CHAPTER I

INTRODUCTION

1.1 Statement of the problem

The extracellular matrix (ECM) not only supports cells and provides architecture characteristic of individual tissues, but also regulates cell behavior by storing and distributing cytokines and growth factors to target cells. Proper functions of the ECM require an elaborate matrix assembly that involves various ECM molecules including collagens, proteoglycans, hyaluronan (HA), and glycoproteins. An alteration of the ECM structure may transmit a signal to cells and change their behavior.

Versican/PG-M (Zimmermann *et al.,* 1989; Shinomura *et al.,* 1993) is a large chondroitin sulfate (CS) proteoglycan of the extracellular matrix (ECM), mainly synthesized by fibroblasts and vascular smooth muscle cells. Its core protein consists of two globular domains G1 and G3 at the N- and C-termini, respectively, and two CS-attachment domains CS-α and CS-β between the two globular domains. Up to 23 CS chains are attached to these domains, and a molecular mass of versican reaches 1,000 kDa. The N-terminal G1 domain consists of A, B, and B′ looped subdomains. The B-B′ stretch binds hyaluronan (HA) and the A subdomain enhances the binding (Matsumoto *et al.*, 2003). The C-terminal G3 domain binds other ECM molecules (Wu *et al.,* 2005) including fibrillins (Isogai *et al.,* 2002), fibulin-1 (Aspberg, Adam et al. 1999), 2

(Olin *et al.,* 2001), tenascins (Aspberg *et al.,* 1995) (Suwiwat *et al.,* 2004) (Tsujii *et al.,* 2006), type I collagen (Yamagata *et al.,* 1986), and fibronectin (Yamagata *et al.,* 1986). By interacting with these molecules, versican is incorporated into the ECM and serves as a structural macromolecule.

Versican exhibits two distinct expression patterns. Whereas it is constitutively expressed in adult tissues such as dermis and blood vessels, and serves as a structural macromolecule of the ECM, it is transiently expressed at high levels in various embryonic tissues and regulates cell behavior such as adhesion (Yamagata and Kimata, 1994; Braunewell *et al.,* 1995; Mazzucato *et al.*, 2002), migration (Landolt *et al.,* 1995; Evanko *et al.,* 1999; Perrissinotto *et al.,* 2000), proliferation (Cattaruzza *et al.,* 2004; Hinek *et al.,* 2004; Sheng *et al.* 2005), and differentiation (Domenzain *et al.,* 2003; Wu *et al.,* 2004). The G3 domain of versican was shown to enhance proliferation of NIH3T3 fibroblasts via epidermal growth factor (EGF)-like motifs (Zhang *et al.,* 1998). The G3 domain without the EGF-motifs enhances interaction of EGF receptor (EGFR) and β1-integrin, impairing growth of U87 astrocytoma cells (Wu *et al.,* 2004). The V1 variant enhances proliferation and inhibits apoptosis of NIH3T3 fibroblasts, while the V2 variant exhibits an opposite activity (Sheng *et al.,* 2005). Although lines of evidence showing direct effects of versican domains have been provides, the precise mechanisms by which versican regulates cell behavior have not been fully understood.

Recently, a knock-in mice (*Cspg2*$^{\Delta 3/\Delta 3}$) whose versican lacks the A subdomain of the G1 domain has been generated. These mice, expressing the mutant versican at a level of ~50% compared with normal versican in wild type (WT) mice, exhibit abnormalities in cardiovascular systems and skin, contrasting to versican-null *hdf* mice which die at stage

E9.5 by severe cardiac defects (Mjaatvedt *et al.,* 1998). As *hdf* heterozygotes, expressing ~50% normal versican, are viable with no obvious phenotype, the abnormalities of *Cspg2*$^{\Delta 3/\Delta 3}$ embryos are likely due to decreased levels of versican deposition in the ECM by the absence of the A subdomain. Thus, analysis of *Cspg2*$^{\Delta 3/\Delta 3}$ mice leads to elucidation of functions of versican in the ECM, separating from those of transiently expressed versican.

To obtain insight into the mechanisms underlying these abnormalities, the embryonic fibroblasts obtained from *Cspg2*$^{\Delta 3/\Delta 3}$ mice had been characterized.

1.2 Literature reviews

1.2.1 The extracellular matrix

Tissues are made up of cells and extracellular space. An extracellular space is largely filled by an intricate network of macromolecules constituting the ECM. This matrix is composed of a variety of proteins and polysaccharides that are assembled into an organized meshwork as demonstrated in figure 1.1. Variations in the relative amounts of the different types of matrix macromolecules and the way in which they are organized give rise diversity of forms, each adapted to the functional requirements of the particular tissue. The matrix can become calcified to form the rock-hard structures of bone or teeth, or it can form the transparent matrix of the cornea, or it can adopt the ropelike organization that gives tendons their enormous tensile strength. Two main classes of extracellular macromolecules make up the matrix: (1) polysaccharide chains of the class called glycosaminoglycans (GAGs), which are usually found covalently linked to protein

in the form of proteoglycans, and (2) fibrous proteins, including collagen, elastin, fibronectin, and laminin, which have both structural and adhesive functions. The members of both classes come in a great variety of shapes and sizes.

Glycosaminoglycans (GAGs) are unbranched polysaccharide chains composed of repeating disaccharide units. They are called GAGs because one of the two sugars in the repeating disaccharide is always an amino sugar (*N*-acetylglucosamine or *N*-acetylgalactosamine), which in most cases is sulfated. The second sugar is usually an uronic acid (glucuronic or iduronic). Because there are sulfate or carboxyl groups on most of their sugars, GAGs are highly negatively charged. Four main groups of GAGs are distinguished according to their sugars, the type of linkage between the sugars, and the number and location of sulfate groups: (1) hyaluronan, (2) chondroitin sulfate and dermatan sulfate, (3) heparan sulfate, and (4) keratan sulfate. Except for HA, all GAGs are found covalently attached to protein in the form of proteoglycans, which are made by most animal cells. The polypeptide chain, or core protein, of a proteoglycan is made on membrane-bound ribosomes and threaded into the lumen of the endoplasmic reticulum. The polysaccharide chains are mainly assembled on this core protein in the Golgi apparatus. First, a special link tetrasaccharide is attached to a serine side chain on the core protein to serve as a primer for polysaccharide growth; then, one sugar at a time is added by specific glycosyltransferases. While still in the Golgi apparatus, many of the polymerized sugars are covalently modified by a sequential and coordinated series of reactions. Epimerizations alter the configuration of the substituents around individual carbon atoms in the sugar molecule; sulfations increase the negative charge. A single type of core protein can vary greatly in the number and types of attached GAG chains.

Moreover, the underlying repeating pattern of disaccharides in each GAG can be modified by a complex pattern of sulfate groups. GAG and proteoglycans associate to form polymeric complexes in the ECM. Molecules of aggrecan, for example, the major proteoglycan in cartilage assembles with hyaluronan in the extracellular space to form aggregates. Moreover, GAGs and proteoglycans associate with fibrous matrix proteins such as collagen and with protein meshwork such as the basal lamina, creating extremely complex structures. Decorin, which binds to collagen fibrils, is essential for collagen fiber formation; mice that can not make decorin have fragile skin that has reduced tensile strength.

Collagen, a family of fibrous proteins, is secreted by connective tissue cells, as well as by a variety of other cell types. They are the most abundant proteins in mammals, constituting 25% of the total protein mass in these animals. After being secreted into the extracellular space, collagen molecules assemble into higher-order polymers called collagen fibrils. Collagen fibrils often aggregate into larger, cable-like bundles, several micrometers in diameter, which can be seen in the light microscope as collagen fibers. Cells interact with the ECM mechanically as well as chemically, with dramatic effects on the architecture of the tissue. Thus, for example, fibroblasts work on the collagen they have secreted, crawling over it and tugging on it-helping to compact it into sheets and drawing it out into cables. When fibroblasts are mixed with a meshwork of randomly oriented collagen fibrils that form a gel in a culture dish, the fibroblasts tug on the meshwork, drawing in collagen from their surroundings and thereby causing the gel to contract to a small fraction of its initial volume. By similar activities, a cluster of

fibroblasts surrounds itself with a capsule of densely packed and circumferentially oriented collagen fibers.

Elastin, a highly hydrophobic protein which likes collagen, is unusually rich in proline and glycine but, unlike collagen, is not glycosylated and contains some hydroxy-proline but no hydroxylysine. Soluble tropoelastin (the biosynthetic precursor of elastin) is secreted into the extracellular space and assembled into elastic fibers close to the plasma membrane, generally in cell-surface. After secretion, the tropoelastin molecules become highly cross-linked to one another, generating an extensive network of elastin fibers and sheets.

Fibronectin is another fibrous protein other than collagen, contributing in ECM. It is a dimer composed of two very large subunits joined by disulfide bonds at one end. Each subunit is folded into a series of functionally distinct domains separated by regions of flexible polypeptide chain. The domains in turn consist of smaller modules, each of which are serially repeated and usually encoded by a separate exon, suggesting that the fibronectin gene, like the collagen genes, evolved by multiple exon duplications. All forms of fibronectin are encoded by a single large gene that contains about 50 exons of similar size. Transcription produces a single large RNA molecule that can be alternatively spliced to produce the various isoforms of fibronectin. Fibronectin is important not only for cell adhesion to the matrix but also for guiding cell migrations in vertebrate embryos.

Laminin-1 is a large, flexible protein composed of three very long polypeptide chains (α, β, and γ) arranged in the shape of an asymmetric cross and held together by disulfide bonds. Several isoforms of each type of chain can associate in different

combinations to form a large family of laminins. The laminin in basement membranes consists of several functional domains: one binds to perlecan, one to nidogen, and two or more to laminin receptor proteins on the surface of cells.

Most cells need to attach to the extracellular matrix to grow and proliferate and, in many cases, even to survive. This dependence of cell growth, proliferation, and survival on attachment to a substratum is known as anchorage dependence, and it is mediated mainly by integrins and the intracellular signals they generate. The physical spreading of a cell on the matrix also has a strong influence on intracellular events. Cells that are forced to spread over a large surface area survive better and proliferate faster than cells that are not so spread out, even if in both cases the cells have the same area making contact with the matrix directly.

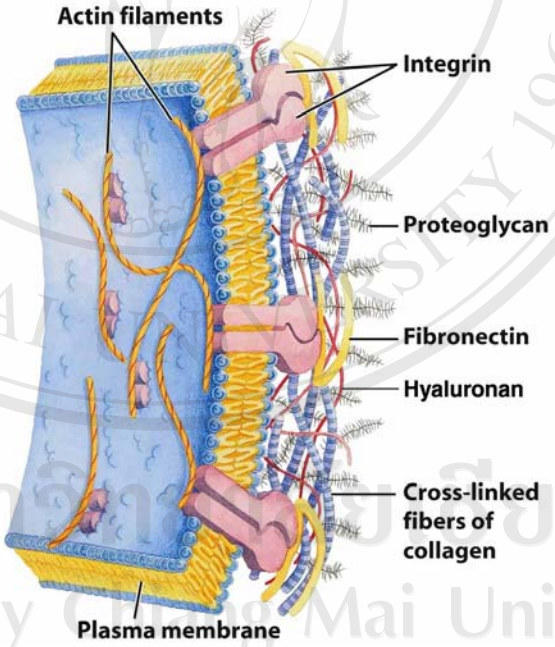

Figure 1.1 Structure of extracellular matrix composes of several extracellular molecules including proteoglycan, fibronectin, hyaluronan, collagen fibers, attaching to receptor on cell membrane (Nelson and Cox, 2005).

1.2.2 Hyaluronan binding proteins

There is protein family which specifically interacts with hyaluronan, called hyaluronan binding proteins (HABPs), or hyadherin. HABPs contain a common structural domain of ~ 100 amino acids in length, termed a link module, which is involved in HA binding (Day, 1999). However, a number of HABPs lack of link module also found as shown in figure 1.2. HABPs family proteins are possibly classified into two groups; link and non-link module containing HABPs.

The link module (also referred as a proteoglycan tandem repeat (Kohda *et al.,* 1996) was first identified in the link protein isolated from cartilage, it was recently found in brain as a brain specific link protein (BRAL1). The link protein is comprised of immunoglobulin (Ig) domain and two contiguous link modules. This molecular arrangement is also found in the G1 domains of aggrecan, versican, neurocan, and brevican. Besides, link module found in the hyaluronan receptor; CD44, existing in numerous isoforms because of alternative splicing of 10 variant exons in different combination. Each isoform contains a single link module close to the N terminus of protein (Lesley *et al.,* 2000). Apart from CD44, the only other member of the link module HABP is a LYVE-1. This molecule is restricted in its expression to lymph vessel endothelium and appears to be involved in hyaluronan degradation (Prevo *et al.,* 2001). The protein product of tumor necrosis factor-stimulated gene-6 (TSG-6) is another single link module containing molecule. It is secreted in response to inflammatory stimuli (Wisniewski *et al.,* 1996; Wisniewski *et al.,* 2005). The link module can be divided into three subgroups according to the expected size of their HA-binding domains. For instance, type A domains comprise a single, independently folding link module that can

support high affinity binding, as found in TSG-6. Type B domain, require N- and C-terminal extensions flanking the link module for correct folding and functional activity, and these are typified by the ~150 amino acid HA-binding domain of CD44. The third class, type C domains, contains N-terminal Ig domain followed by two contiguous link modules as found in cartilage link protein and members of the chondroitin sulfate proteoglycan family, aggrecan, versican, neurocan, and brevican.

Inter-α-inhibitor (IαI), a serine protease inhibitor plentiful in serum, was one of the first proteins found to associate with HA (Bost *et al.,* 1998). It is a non-link module HABP. IαI is an unusual proteoglycan with a chondroitin 4-sulfate chain linked to bikunin (containing two Kunitz inhibitor domains) and two heavy chains attach to the GAG by ester bonds via their C-terminal aspartic acid residues (Enghild *et al.,* 1999). IαI is essential during ovulation, acting to stabilize the hyaluronan-rich cumulus extracellular matrix with which it forms a covalent complex (Zhao *et al.,* 1995). CD38, a type II membrane glycoprotein is another non-link module HABP (Nishina *et al.,* 1994). This protein functions as an enzyme with NADase activity, and this property has been studied much more extensively than its hyaluronan binding function. Recently two related non-link module HABP, IMP-50 and SPACRCAN have been isolated from human retina. IMP-50 is expressed by cone and rod photoreceptor cells and is presented in the interphotoreceptor matrix. SPACRCAN (a chondroitin sulfate proteoglycan) is likely to be a receptor on photoreceptors and pinealocytes (Acharya *et al.,* 2000). The last example of non-link module HABP is the receptor for hyaluronan-mediated motility (RHAMM). It mediates cell migration and proliferation in normal and cancer cells (Turley *et al.,* 2002).

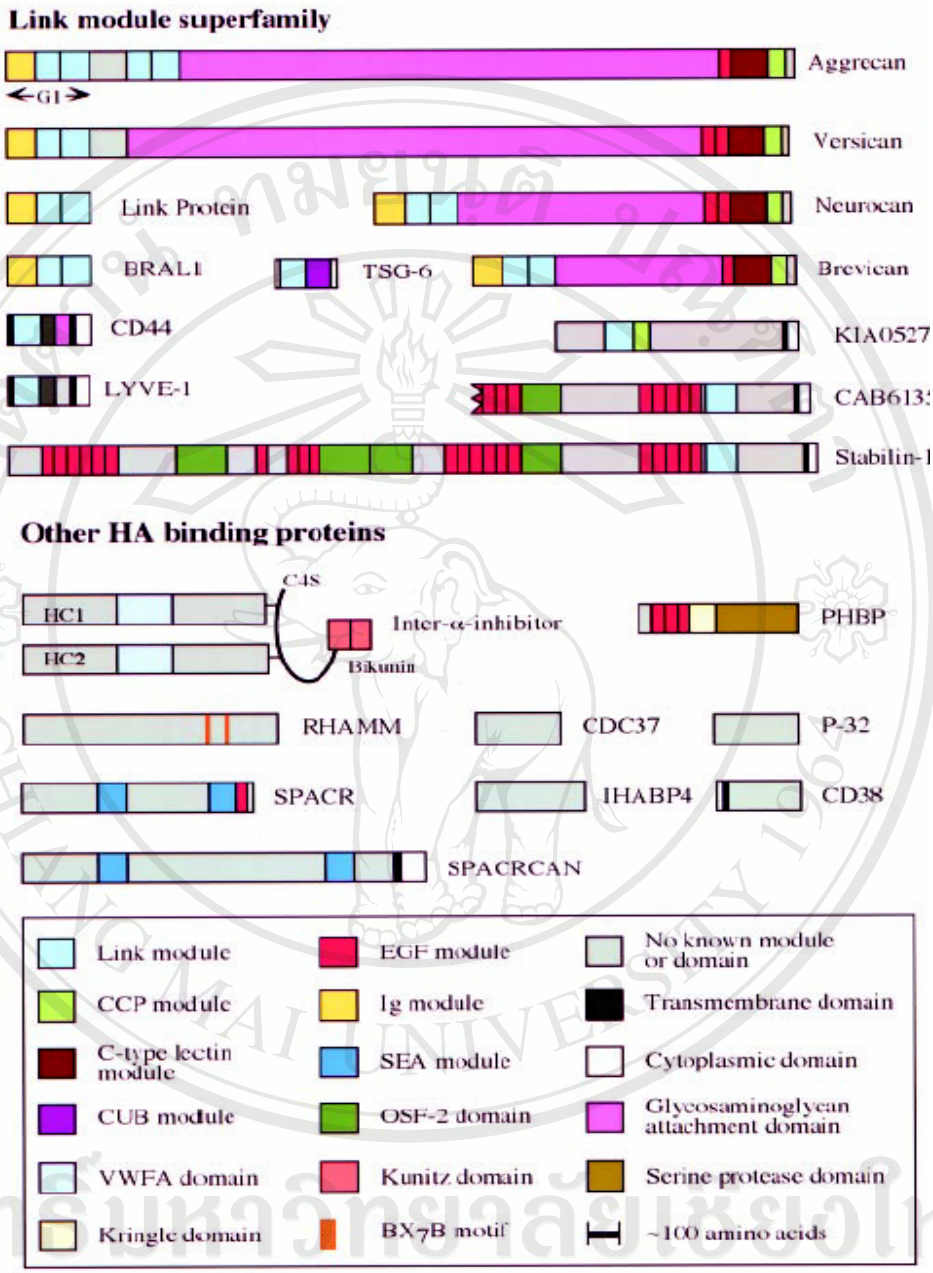

Figure 1.2 The modular organization of Hyaluronan Binding Proteins (Day and Prestwich, 2002).

1.2.3 Versican Gene and Structure

A complete genomic structure of the human versican (chondroitin sulfate proteoglycan2; CSPG2) gene was first deciphered by Micheal F. Naso (Naso *et al.,* 1995). As present in Figure 1.3, versican gene locates in human chromosome 5q12-14 (Iozzo *et al.,* 1992), and composing of 15 exons that spanning more than 90 kb of genomic DNA. These exons encode five domains protein: hyaluronan-binding region (HBR), GAG-α, GAG-β, EGF-like, lectin-like and complement regulatory protein (CRP)-like which gather to be a versican. The exons ranges in size from 76 to 5262 base pairs (bp), and all followed the AG/GT rule for splicing junction. The versican promoter harbors a typical TATA box locates approximately 16 bp upstream of the transcription start site, and binding sites for a number of transcription factors involved in regulation of gene expression. Stepwise 5′ deletions have identified a strong enhancer element between -209 and -445 bp, and a strong negative element between -445 and -632 bp (Wu *et al.,* 2005).

Signal Peptide- following exon 1, which encoded 284 bp of the 5′-untranslated region, the versican gene contained a 76 bp (exon 2) encoding the signal peptide and the first few amino acids of the mature protein core.

Domain I: the hyaluronan-binding region. Exons 3-6 coded for the amino-terminal hyaluronan-binding region of the first globular domain (G1). The corresponding protein module binds specifically to HA. Within this domain, there was an Ig-like fold encoded by exon 3 (375 bp), also within this hyaluronan binding domain of versican, there were two repeats of about 100 amino acid residues each, known to mediate binding

to hyaluronan (Goetinck *et al.,* 1987); the first was encoded by exon 4 (175 bp) and 5 (128 bp), whereas the second was encoded by exon 6 (294 bp).

Domain II: the GAG-α-binding region. This region presents in an alternatively splicing gene of versican (Dours-Zimmermann and Zimmermann, 1994). It is a contiguous stretch of genomic DNA containing an exon of 2961 bp (exon 7). This region lacked cysteine residues, but contained several serine-glycine dipeptides flanked by acidic amino acid residues that could be the attachment site of chondroitin sulfate chains.

Domain III: the GAG-β-binding region. This domain was encoded by a uniquely large, single exon of 5262 bp (exon 8). It contains numerous potential attachment sites for chondroitin sulfate side chains through the serine-glycine or glycine-serine residue (Zimmermann and Ruoslahti, 1989). This exon is an alternatively splicing gene also.

Domain IV: homology to selectins. The carboxyl terminal region of the second globular domain (G3), which contains two EGF-like repeats motif, a lectin-like motif, and a CRP-like motif, was encoded by exons 9-14. Interestingly, these motifs are found in the family of leukocyte homing and cell adhesion molecules known as selectins (Bevilacqua and Nelson, 1993). These motifs are thought to participate in cell-cell or cell-matrix interactions by binding carbohydrate moieties on glycoproteins or other glycosylated molecules (Stoolman, 1989). This domain can be divided into subdomains called IVa, IVb, and IVc because of the distinct structural properties manifests. Subdomain IVa: the epidermal growth factor-like repeats; the two EGF-like repeats were encoded by two exons of identical size (114 bp each), exon 9 and 10. These EGF repeats are thought to mediate protein-protein interactions in protein containing these motifs.

Subdomain IVb: the lectin-like motif; it was encoded by three distinct exons (exon 11-13) of 159, 83, and 145 bp respectively. These exons contain a number of highly conserved residues, which are probably very important for folding and/or carbohydrate recognition and binding. Subdomain IVc: the complement regulatory protein-like motif; it was encoded by a single exon (exon 14) of 183 bp, this exon coded for a structural motif found in a number of complement regulatory proteins.

The carboxyl-terminal end and 3′-untranslated region- exon 15 encoded 42 amino acid residues of the carboxyl terminus, the translation stop codon, and the 3′-untranslated region.

The genomic organization of versican is shared by homologous proteins including aggrecan, link protein, CD44, P-selectin, and E-selectin. The similarities between the versican gene and genes of homologous proteins suggest that versican may have evolved from primordial gene through the process of exon shuffling and intron recombination. Various domains of versican may have a function analogous to other proteins harboring similar motifs and sharing the same genomic organization.

Considering to protein structure of versican core protein encoded from versican gene (Wu *et al.,* 2005), It consists of an N-terminal G1 domain, a C-terminal G3 domain, and CS chain binding regions between G1 and G3, as illustrated in figure 1.4. The G1 domain is composed of an Ig-like motif, followed by two proteoglycan tandem repeats that are known as hyaluronan-binding region. The G3 domain of versican consists of two EGF-like motifs, and a complement regulatory protein-like motif (Wight, 2002). RNA splicing occurs in the two large exons encoding the GAG attachment sites of versican mRNA. Four mRNA transcripts arise from alternative splicing, giving rise to V0, V1, V2

and V3. V0 possesses GAG-α and GAG-β attachment region, V1 possesses GAG-β, but not GAG-α, V2 possesses GAG-α but not GAG-β and V3 possesses neither GAG-β nor GAG-α. The four isoforms of versican including V0, V1, V2, and V3 present molecular weight of the core proteins of about 370 kDa, 263 kDa, 180 kDa, and 74 kDa, respectively. However, these theoretical values are significantly lower than the core protein sizes deduced from SDS-PAGE, where versican V0, V1, and V2 migrate after chondroitinase ABC digestion around 550, 500, and 400 kDa, respectively (Dours-Zimmermann and Zimmermann, 1994; Schmalfeldt *et al.,* 1998). This difference may originate from a high content of N- and O-linked oligosaccharides or from the very low isoelectric points of the versican core proteins (calculated between 4 and 5), which may lead to aberrant gel migration properties as a result of reduced SDS-binding. The molecular weights of the intact proteoglycan isoforms of versican are even more difficult to estimate because of the heterogeneity of the GAG chains. Sizes between 6×10^5 and 1.5×10^6 for the intact proteoglycan variants of versican (V2 to V0) are plausible. Versican contains different lengths of the GAG binding region with an accompanying variation in number of attached GAG chains. The number of GAG side chains per length of the GAG-binding regions is rather constant. Estimated numbers of GAG chains are 17-23, 12-15, and 5-8 in versican isoforms V0, V1, and V2, respectively (Dours-Zimmermann and Zimmermann, 1994). Due to the absence of both central domains, versican V3 is most likely devoid of GAG side chains and therefore may not be a proteoglycan (Zako *et al.,* 1995). Most of the above numbers are estimates based on putative carbohydrate attachment sites in the primary structure of the core proteins. However, the number, size, and composition of the carbohydrate substitution in not only

influenced by the length of the GAG chain binding regions, but it may also be affected by G1 and G3 domains and the cell and tissue types (Yang *et al.,* 2000). Versican's GAG profile varies amongst tissues and specific molecules. GAG sulfation patterns change both markedly and precisely during chick brain development (Kitagawa *et al.,* 1997), indicating the significance of specific carbohydrate information. GAG chains of versican are long, repeating disaccharides of uronic acid, and N-acetylgalactosamine. Three possible sulfation sites exist, leading to at least five specific chondroitin sulfate subtypes. Of these, chondroitin 4-sulfate (C-4S) and chondroitin 6-sulfate (C-6S) are important mediators of molecular interaction. GAG chains are composed of approximately 40 repeating units. These chains repel each other, and within themselves, thus assuming long, extend shapes. Versican's structure is complex, and its observed effects *in vitro* and *in vivo* are as well. It has been shown to be involved in diverse cell functions, such as adhesion, migration, proliferation, and apoptosis. Reports are sometime contradictory. The seemingly confused observations may be reconciled by an appreciation of the structure and expression diversity of versican. On the other hand, versican does not act by itself. Its many binding partners play important roles in determining the behaviors that involve versican.

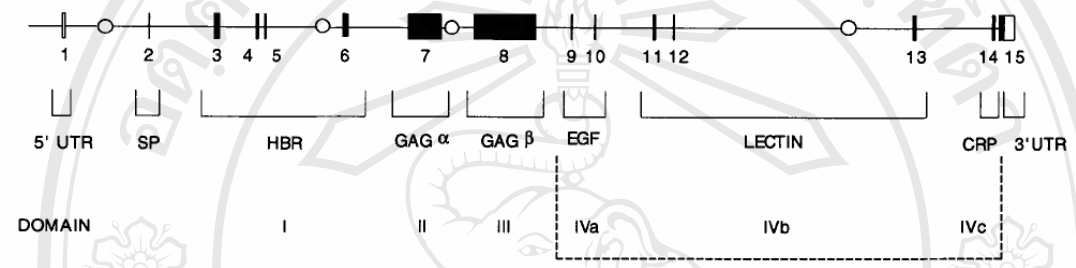

Figure 1.3 Schematic representation of the human versican gene and intron/exon organization. Exons are represented by filled rectangles and introns by thin lines. Empty boxes signify untranslated regions. Open circles in introns indicate that the exact size of that intron is not known. The four major domains are indicated by roman numerals. SP = signal peptide; HBR = hyaluronan-binding region; GAG-α = glycosaminoglycan α-binding domain; GAG-β = glycosaminoglycan β-binding domain; EGF = epidermal growth factor repeat like motif; CRP = complement regulatory protein like motif; UTR = untranslated region (Naso *et al.,* 1995).

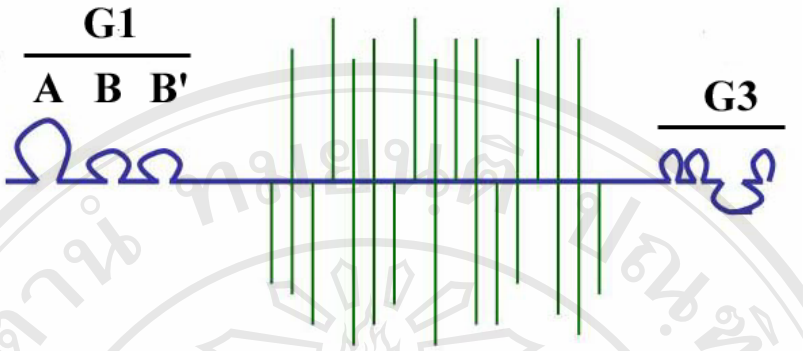

Figure 1.4 A schematic of versican structure, composing of core protein (G1, CS chain binding region, and G3 domain). Blue line represents core protein and Green lines show GAG attaching to CS chain binding region.

1.2.4 Biological function of versican

Considerable biomechanical role of versican, it is one of the main components of the ECM where it provides hygroscopic properties to create a loose and hydrated matrix that is necessary to support key events in development and disease. Through direct or indirect interactions with cells and molecules, versican is able to regulate cell adhesion and survival, cell proliferation, cell migration, and ECM assembly (Wight, 2002).

Regard to cell adhesion property, early studies show that versican was anti-adhesive, and this activity appeared to reside in the G1 domain of versican (Yamagata *et al.,* 1989; Yamagata *et al.,* 1993; Yamagata and Kimata, 1994). However, the carboxyl-terminal domain of versican interacts with the β1 integrin of glioma cells, activating focal adhesion kinase (FAK), promoting cell adhesion and preventing apoptosis in this cell

type (Wu *et al.,* 2002). The pro-adhesive property of the G3 domain of versican raises the possibility that different breakdown products of versican might differentially affect cell adhesion in different ways. Versican also binds to adhesion molecules on the surface of inflammatory leukocytes. For example, versican binds to L- and P-selectins through specific oversulfated sequences in the chondroitin sulfate chains (Atarashi *et al.,* 2002). In addition, versican binds CD44 through chondroitin sulfate chains (Kawashima *et al.,* 2000) and this interaction might stabilize HA-versican aggregates associated with the cell surface (Bajorath *et al.,* 1998). Versican interacts to CD44 through the link module domain present on the amino-terminal end of CD44. Thus, one could envision an ECM in which multiple binding sites for inflammatory cells exist. Once the leukocytes bind, versican could influence the availability of inflammatory chemokines involved in the recruitment of mononuclear leukocytes (Hirose *et al.*, 2001). The chondroitin sulfate chains of versican mediate this binding. Such interactions may sustain the inflammatory response, and thus could put versican at the head of the list of molecules to target for controlling inflammation.

Versican also is involved in cell proliferation. For example, mitogens such as platelet-derived growth factor (PDGF) upregulate versican expression in arterial smooth muscle cells, and contribute to the expansion of the pericellular ECM that is required for the proliferation and migration of these cells (Schonherr *et al.,* 1997; Evanko *et al.,* 1999; Evanko *et al.,* 2001). Thus, versican expression is associated with a proliferative cell phenotype and is often found in tissues exhibiting elevated proliferation, such as in development and in a variety of tumors, including melanoma (Paulus *et al.,* 1996; Ricciardelli *et al.,* 1998; Touab *et al.,* 2002). In fact, levels of versican present in the

stroma of prostate correlate positively with progression in early-stage prostate cancer (Ricciardelli *et al.,* 1998). Moreover, cell culture studies suggest that elevated levels of versican result from the prostate tumor cells inducing host stem cells to increase the synthesis of versican via a paracrine mechanism involving transforming growth factor beta 1 (TGF-β1) (Sakko *et al.,* 2001). The signaling pathways that are involved in regulating versican synthesis are beginning to emerge. For example, it is clear that endogenous tyrosine kinase activity of the synthesis at both the transcript and protein levels in vascular smooth muscle cells (Schonherr *et al.,* 1997). The PDGF effect on versican synthesis in vascular smooth muscle cells is quite specific, since mRNA transcripts for decorin and biglycan are either not affected or are decreased. Molecules that associate with versican such as HA and the HA receptor (CD44) are also upregulated by PDGF in arterial smooth muscle cells, suggesting a need for a higher-ordered versican complex for events associated with PDGF stimulation of smooth muscle cell proliferation (Evanko *et al.,* 2001). In addition to transcriptional/post-transcriptional regulation of versican synthesis, PDGF increases the length of the CS chain attached to the versican core protein (Schonherr *et al.,* 1991). The increase in versican and associated proteins in response to growth factors such as PDGF and TGF-β1 causes increasing in the pericellular matrix of the cells and expansion of the ECM. Pericellular matrix expansion involves the interaction of versican with several binding proteins, such as HA and CD44. These complexes increase the viscoelastic nature of the pericellular matrix, creating a highly malleable extracellular environment that supports a cell-shape change necessary for cell proliferation and migration (Lee *et al.,* 1993). Furthermore, these versican enriched macromolecular complexes may have a dramatic effect on the tension exerted

on the cells themselves and the traction forces generated by the cell. Such mechanical changes could impact mechanically coupled signaling (Chicurel *et al.,* 1998; Chicurel *et al.,* 1998). Thus, the versican-HA complex that surrounds cells serves as an important, but infrequently considered, mechanism for controlling cell shape and cell division. In fact, inhibiting the formation of this pericellular coat blocks the proliferation of arterial smooth muscle cells in response to PDGF (Evanko *et al.,* 1999). Another mechanism by which versican could influence proliferation is by acting as a mitogen itself, through the EGF sequences in the G3 domain of the molecule. For example, expression of G3 minigenes in NIH3T3 cells enhances cell proliferation, and the effect can be blocked by deletion of the EGF domain in the G3 construct (Zhang *et al.,* 1998). Maximal growth-promoting activity is achieved in NIH3T3 cells and chondrocytes with both G1 and G3 minigene constructs, supporting the concept that versican regulates proliferation by binding directly to a growth factor receptor and by interfering with cell adhesion (Yang *et al.,* 1999; Zhang *et al.,* 1999).

Versican is expressed along neural crest pathways and influences neural cell migration (Perissinotto *et al.,* 2000). A number of other studies suggest that versican blocks neural crest migration because cells do not enter tissues that express versican (Landolt *et al.,* 1995; Henderson *et al.,* 1997). Pax3 is a transcription factor associated with defective neural cell migration. Splotch mice are characterized by mutations in the Pax3 gene and exhibit neural crest related abnormalities, including the failure of neural crest cells to colonize target tissues. However, neural crest cells derived from these mutant mice migrate as controls *in vitro*, so it has been suggested that the defect may not reside in the neural crest cell themselves, but rather in the ECM environment through

which they migrate. Indeed, earlier studies (Henderson *et al.,* 1997) demonstrated that versican was markedly overexpressed in Splotch mutant in neural crest cell migration pathways, suggesting that versican may be responsible for defective cell migration in this species. Recent studies show that overexpression of Pax3 in a medulloblastoma cell line causes upregulation of the V2 splice variant of versican and a downregulation of the V3 variants (Mayanil *et al.,* 2001). Such differential regulation of the versican isoforms may explain, in part, the migratory defect in the Splotch mouse. It is of interest that the V3 isoform lacks CS chains, which should reduce the exclusionary properties of the ECM. Versican also appears to plays a role in the migration of embryonic cells in the development of the heart (Henderson and Copp, 1998). Versican gene expression occurs at high levels during the development of the heart. Versican is expressed in a chamber-specific manner, with high levels in trabeculations of the right ventricle. In addition, versican is expressed in the endocardial cushion of the atrioventricular, semilunar and venous valves. That versican plays an essential role in the development of the heart has been demonstrated by the identification of an insertional transgene mutation in the versican gene in the heart-defect (*hdf*) mouse (Mjaatvedt *et al.,* 1998). The loss of versican expression in the homozygous *hdf* is associated with the failure of the endocardial cushion cells to migrate. It is of interest that this endocardial cushion phenotype also resembles the phenotype in the hyaluronan synthase 2 knockout mouse, suggesting that the interaction of versican with hyaluronan is critical to cell movement in this tissue (Camenisch *et al.,* 2000). Versican influences the migration of a variety of other cell types, and this activity appears to be mostly associated with the anti-adhesive activities involving the G1 domain of the molecule. In the nervous system and in axonal

growth, the V2 splice variant inhibits axonal outgrowth and migration (Schmalfeldt *et al.,* 2000). This inhibiting activity of versican can be reduced, but not eliminated, by removing CS chains, indicating that multiple domains of versican are involved in controlling axon regeneration. Although the V2 isoform is widely present in the CNS, it is predominately localized to the myelinated fiber tracts. Oligodendrocyte cells are the likely source of V2 (Milev *et al.,* 1998; Asher *et al.,* 2002). How versican inhibits axonal growth remains open. The finding that both the GAGs and core protein domains of the molecule are involved in the inhibitory activity suggests a direct interaction with the cells, or modification of the surrounding matrix to form exclusionary boundaries. The fact that a versican play a fundamental role in axonal migration is highlighted by recent studies that show upregulation of versican following central nervous system injury (Asher *et al.,* 2002). These changes have been associated with the failure of nerves to regenerate. The importance of versican in preventing nerve regeneration is highlighted by recent studies that show that degradation of CS chains by chondroitinase ABC lyase treatment following spinal cord injury in experimental animals promotes regeneration of both ascending and descending corticospinal-tract axons (Bradbury *et al.,* 2002). Such results suggest that manipulating versican synthesis in spinal cord injury may be a useful intervention for therapeutic treatment of this condition. Failure of axons to regenerate is also characteristic of multiple sclerosis, and versican appears to increase in plaques present in the white matter of the brain from patients with multiple sclerosis (Sobel and Ahmed, 2001).

Versican interacts with several different ECM molecules and in part, plays a central role in ECM assembly. The domain structure of versican lends itself to multiple types of

interactions through either protein-protein or protein-carbohydrate interactions. Perhaps the best known of these interactions involves a specific interaction between the amino-terminal domain of versican (G1) and hyaluronan. The binding of versican to hyaluronan involves a tandem double-loop sequence in the G1 domain of versican and a stretch of five repeat disaccharides in HA. This interaction is stabilized by another protein-link protein-which exhibits selective binding specificity for both HA and versican. In addition to HA, versican interacts with other ECM molecules such as tenascin-R (Aspberg *et al.,* 1995). Versican interacts with tenascin-R through the lectin-binding domain of versican and involves protein–carbohydrate interactions. The lectin-binding domain participates in other ligand interactions as well. For example, versican interacts with fibulin-1 and fibulin-2 (Aspberg *et al.,* 1999; Olin *et al.,* 2001), a growing family of ECM proteins that are expressed in particularly high levels in the developing heart valve. In adults, however, fibulin-1 and -2 are found associated with microfibrils that are part of elastic fibers. Versican also can interact with proteins associated with elastin in elastic fibers. For example, versican interacts with the elastic fiber-associated protein; fibrillin (Isogai *et al.,* 2002), and versican has been shown to co-localize with elastic fibers in skin (Zimmermann *et al.,* 1994). Furthermore, fibrillins bind fibulin-2, and fibulin is preferentially localized to the elastin/microfibril interface in some tissue, but not in others (Reinhardt *et al.,* 1996). It may be that fibulin serves as a bridge between versican and fibrillin, forming high-ordered multi-molecular structures important in the assembly of elastic fibers. The relationship of versican to elastic fiber assembly is interesting and unusual. For example, rat pup arterial smooth muscle cells have high levels of tropoelastin expression, but no detectable levels of versican synthesis, while the opposite

is true for adult arterial smooth muscle cells (Lemire *et al.,* 1996). Furthermore, it is known that chondroitin sulfate inhibits the formation of elastic fibers, through a mechanism involving interference with the binding of the elastin receptor to the surface of arterial smooth muscle cells (Hinek *et al.,* 1991). Recent studies have shed light on the relationship of versican and elastic fiber assembly. For example, it was found that overexpressing of the versican splice variant that lacks CS chains (V3) dramatically alters arterial smooth muscle cell phenotype by enhancing cell adhesion, decreasing growth and migration (Lemire *et al.,* 2002), and upregulating tropoelastin expression (Merrilees *et al.,* 2002). Placement of V3-transduced arterial smooth muscle cells into injured blood vessels results in the formation of multiple elastic laminae during injury repair (Merrilees *et al.,* 2002). Thus, it may be that overexpressing of versican that lacks CS chains competes for binding sites with versican that contain CS chains associated with HA on the cell surface. This would allow the elastin-binding protein to associate with the cell surface and promote elastic fiber assembly (Hinek *et al.,* 1991; Hinek and Wilson, 2000). Versican plays a role in tissue remodeling of the ECM that is associated with developmental events as well. For example, during pregnancy and involution an extensive remodeling of the human cervical connective tissue occurs. Versican increases in this tissue during pregnancy and achieves its highest concentration in the ripe cervix immediately after vaginal delivery, but falls dramatically during involution (Westergren-Thorsson *et al.,* 1998). These changes in versican are opposite to what was found for collagen types I and III and the small leucine-rich proteoglycans decorin, biglycan and fibromodulin. Thus, versican characterizes a more provisional and loosely organized ECM that accommodates particular development events.

1.2.5 The interactions among versican with extracellular matrix molecules and cell surface proteins.

Versican does not stand alone in ECM, but it interacts with several biomolecules including HA, fibronectin, fibulins, fibrillin, tenascin, selectins, CD44, integrin, and EGFR as demonstrated in Figure 1.5. These multiple binding interactions play important roles in cell and tissue behavior (Wu *et al.,* 2005). Versican has been found to be co-localized with HA, CD44, and tenascin in the pericellular matrix of cultured fibroblasts (Yamagata *et al.,* 1993), and in epidermal keratinocyte tumors, levels of HA, CD44 and versican correlated with the aggressiveness of the disease (Karvinen *et al.,* 2003). Versican's binding to HA is well known, occurring through its tandem repeats. The association of versican with HA is, at least in part, mediated by link protein, and indeed it has recently been shown that both link protein and HA bind to versican's G1 domain (Matsumoto *et al.,* 2003) as shown in figure 1.6. Using different versican constructs, the results showed that both HBR domains were required and sufficient for *in vitro* binding, but that binding affinity was increased when the Ig-like domain was included. Link protein contains two HBR domains which appear to interact with the homologous structures of versican, and this interaction did not affect versican's binding to HA. Overexpression of versican G1 domain can enhance cell proliferation and reduce cell adhesion in different cell types (Ang *et al.,* 1999). Taken together, studies have suggested that versican might enhance tumor behavior through interaction with HA and subsequent activation of CD44. Interestingly, studies have demonstrated that versican may interact directly with CD44 through its GAG chains, independent of HA

(Kawashima *et al.,* 2000). It is important to know if this interaction also induces CD44 activation and transduces signals.

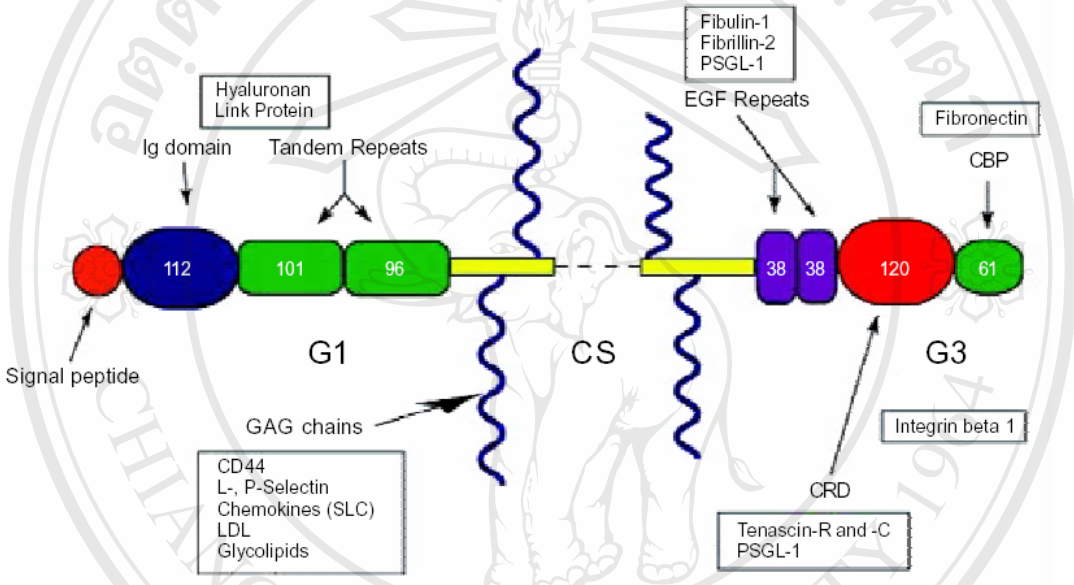

Figure 1.5 The interaction of versican with other molecules. The locations of versican motifs that interact with other molecules are shown (Wu *et al.,* 2005).

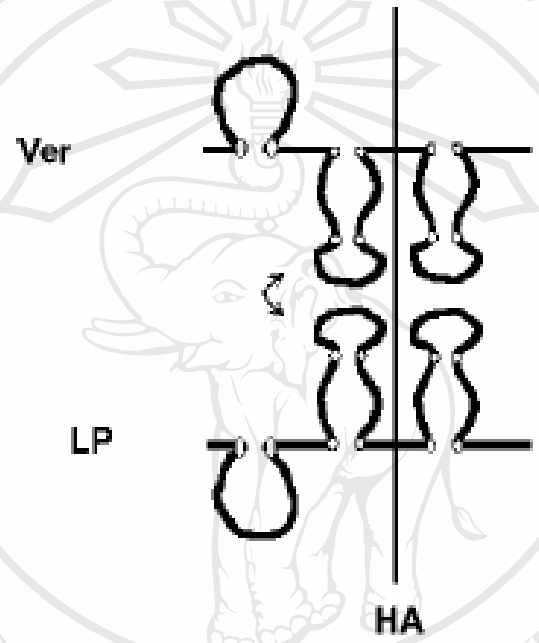

Figure 1.6 Schematic diagram of interaction of the G1 domain of versican (Ver), link protein (LP), and hyaluronan (HA). Versican G1 domain binds to both HA and LP at the BB′ stretches (Matsumoto *et al.,* 2003).

At C-terminal lectin-like motif in versican core protein, it binds to tenascin-R through protein-protein interaction. In addition, tenascin binds to many other members of the HABP family, such as brevican, neurocan, phosphocan, and syndecan through lectin-like motif (Grumet *et al.,* 1994; Aspberg *et al.,* 1997). The tenascin family of large ECM glycoproteins comprises five distinct genes coding for (in a chronological order) TN-C, TN-R, TN-X, TN-Y and TN-W (Erickson, 1993). Of these, only TN-R and TN-C are expressed in the central nervous system. The structural homology shared by TN-R and TN-C has raised the logical question about similarities in function. Indeed, TN-R and TN-C both have been found to interfere with β1 integrin-dependent cell adhesion and neurite outgrowth on fibronectin. The defined spatiotemporal expression of chondroitin sulfate proteoglycans in the ECM or at the cell membrane of glial and neuronal cells is believed to play a role in biological phenomena as diverse as neural cell adhesion and migration; axon path finding, synaptogenesis and plasticity; growth factor and cytokine action; neuronal survival; and structural organization of the ECM (Margolis and Margolis, 1997). Similar in distribution to chondroitin sulfate proteoglycans in the brain, TN-R is also implicated in such phenomena and therefore likely to interfere with chondroitin sulfate proteoglycan action or vice versa. TN-R has been suggested to participate in the macromolecular organization of perineuronal nets, along with HA and HABP, based on its divalent cat ion-dependent homophilic binding properties (Pesheva *et al.,* 1991). This large HA matrix may create a local physical barrier, preventing other cells from entering the microenvironment. TN-R binds to the C-terminal lectin-like motif of versican but surprisingly, this association is not

mediated by carbohydrate moieties of TN-R (Aspberg *et al.,* 1997). Instead, protein-protein interactions, involving at least two fibronectin III domains of TN-R, are responsible. TN-R gene knockout mice show abnormal perineuronal net formation, although the anatomy of all major brain areas and the formation and structure of myelin appear normal (Forsberg *et al.,* 1996). These results suggest that in the brain ECM, there may be many molecules, some of which may also interact with versican which are functionally complementary.

Fibulin-1 and fibulin-2 are modular ECM proteins, expressed in extracellular fibrils, basement membranes, and elastic fibres (Timpl *et al.,* 2003). Versican binds to fibulin-2 and fibrillin-1 through its C-terminal lectin-like motif in a calcium-dependent manner. As with the unexpected observations with TN-R, versican appears to bind to fibulin-1, even following the latter's deglycosylation, again indicating protein-protein interaction (Aspberg *et al.,* 1999). The binding of fibulin-1 and fibulin-2 to versican has been mapped to their calcium-binding epidermal growth factor (cbEGF)-like domains. In addition to versican, aggrecan and brevican also interact with fibulin-2 by their lectin-like motif, whereas neurocan does not (Olin *et al.,* 2001). Electron microscopy studies have confirmed the mapping and demonstrated that HA-aggrecan complexes can be cross-linked by the fibulins. The binding site for versican in microfibrils is most likely within a region of fibrillin-1, one of the major building blocks of microfibrils, between its cbEGF domains 11 and 21 (Isogai *et al.,* 2002). Mutations in this region of the human gene, which potentially disrupt versican or other protein binding, can result in severe forms of Marfan syndrome ("neonatal" Marfan syndrome), with significant cardiovascular disease. The connection between versican

and fibrillin microfibrils may thus be functionally significant, particularly in cardiovascular tissues. Fibulin may serve as a bridge between versican and fibrillin, forming highly ordered multimolecular structures important in the assembly of elastic fibers. This organization may be particularly meaningful in blood vessels and skin, where these molecules are co-localized (Wight, 2002).

Studies have shown versican to interact with fibronectin, as well as collagen type I (Yamagata *et al.,* 1986) and that these interactions are responsible for reduced cell adhesion in melanoma cells (Touab *et al.,* 2002). Fibronectin and collagen are integrin ligands and play roles in enhancing cell adhesion. Versican prevents cell binding to fibronectin, but not collagen type I, in several central nervous system cell types (Braunewell *et al.,* 1995). Conditioned medium from prostate cancer fibroblasts containing versican V0 and V1 can result in reduced adhesion of prostate cancer cell lines to fibronectin, but not to laminin (Sakko *et al.,* 2003). Adding RGD peptide, a classical integrin-binding sequence, to the V0/V1 containing conditioned medium yielded increased cell attachment to fibronectin. This may mean that versican can sequester fibronectin through fibronectin's RGD domain, preventing it from binding to cells. Treatment of the conditioned medium with chondroitinase ABC decreased versican's capacity to reduce adhesion suggesting that GAG side chains are involved. Recently, it was demonstrated that versican G3 domain enhances tumor growth and angiogenesis, perhaps through enhancement of fibronectin expression (Zheng *et al.,* 2004). Furthermore, versican G3 domain can form complexes with fibronectin and vascular endothelial growth factor. The complex was found to stimulate endothelial cell adhesion, proliferation and migration. Removal of this complex with anti-

fibronectin antibody reversed G3's enhancing effects on endothelial cell activities (Wijelath *et al.,* 2002).

Versican appears to interact with selectins via its GAGs as shown in figure 1.5, and may only do so when they are oversulfated. Specific tetrasaccharides involved in the interaction were identified. In addition, versican is also found to bind certain chemokines and regulate chemokine function (Hirose *et al.,* 2001). In particular, the secondary lymphoid tissue chemokine (SLC) was shown to bind to dermatan sulfate with strong specificity. SLC assists in the binding of integrin α4β7, on leukocytes, with its ligand, MAdCAM-1, found on the endothelium (Pachynski *et al.,* 1998). Versican, through its GAG chains, may thus prevent leukocyte adhesion to vessel walls through SLC sequestration. Importantly, SLC binding to its receptor was not affected following GAG addition, suggesting that versican's negative modulation occurs downstream in the signal transduction pathway. Although no consensus sequence has been identified in versican-binding chemokines, it appears that versican generally interacts with those that attract mononuclear lymphocytes. A specific layer of binding specificity is the particular CS chain that is expressed by versican, either through isoform expression or carbohydrate modification. As the same GAG chains also bind L-selectin, and as some versican populations do not bind to this leukocyte cell surface adhesion molecule, versican expression patterns may be important in certain *in vivo* situations such as in inflammatory responses, where L-selectin and chemokines are concomitantly involved. Versican has been found to be upregulated in many inflammatory conditions, such as arthritis, asthma, and granulomatous lung diseases (Johnson, 2001). Therefore,

versican might be involved in inflammation by regulating interaction with selectins and chemokine activities.

As indicated above, GAG chains of versican bind to CD44, and competitive inhibition assays showed that versican, HA, and GAG chains all bind to the link module of CD44, with HA showing a substantially higher binding affinity. Although, as mentioned, versican GAG oversulfation is critical for binding to L- and P-selectins, GAG modifications do not appear to be important in CD44 binding. Interestingly, the binding of either L- or P-selectin to versican was inhibited with the addition of the solubilized other, but CD44 binding was not affected in the presence of selectins. This suggests that different GAG sites are responsible for these interactions (Kawashima *et al.,* 2002). As mentioned above, versican has been found to be co-localized with HA, CD44 and tenascin in the pericellular matrix of cultured fibroblasts (Yamagata *et al.,* 1993). As CD44 binds to both hyaluronan and versican, it is possible that the three may form complexes together. However, as versican and CD44 bind following hyaluronidase treatment, the complex association does not depend on the presence of HA. The fact that GAG chains alone can bind the proteins mention here, without any assistance from versican core proteins, suggests that other proteoglycans that contain the proper side chains may also be involved in complexes.

Versican interacts with integrin β1 through its C-terminal G3 domain. This interaction of these molecules expresses modulated cell behavior. For example, astrocytoma cells expressing a construct containing versican G3 domains exhibited enhanced spreading and adhesion, with increased phosphorylation of FAK and reduced H_2O_2-induced apoptosis. Exogenous addition of purified versican G3 product

(containing two EGF-like motifs, lectin-like motif and CRP motif) was also able to induce FAK phosphorylation and increased cell adhesion (Wu *et al.,* 2002).

EGFR, another important cell surface receptor, mediates many cellular responses in both normal biological processes and pathological states. Like all receptor tyrosine kinases (RTKs), the EGFR family members comprise an extracellular domain containing a ligand-binding site, a single hydrophobic transmembrane helix, and a cytosolic domain that includes a conserved protein tyrosine kinase core, flanked by regulatory sequences. The ligands of EGFR, such as EGF, bind to their receptors and induce conformational changes in the receptor monomers that promote their activation. Integrins are able to form clusters with EGF receptors, and this physical interaction frequently affects the intensity of EGFR-induced down-stream signal to extracellular signal-regulated kinase (ERK). Growing evidence indicates that collaborative pathways derived from these two signals are crucial in regulating a range of cell activities, such as proliferation, differentiation, apoptosis, adhesion, and migration (Yamada and Even-Ram, 2002). Astrocytoma cells expressing the C-terminal domain of versican exhibit reduced EGFR activity, represented by lower levels of EGF-induced EGFR autophosphorylation and accelerated reduction of the protein expression level after phosphorylation (Wu *et al.,* 2004). With enhanced integrin β1-mediated adhesion signals, the cells exhibited increased integrin β1-EGFR interaction, accompanied by increased FAK phosphorylation (Wu *et al.,* 2002). This increased cell adhesion and decreased EGF-induced EGFR phosphorylation was followed by reduced cell migration. The effects of versican on integrin and EGFR activation were also demonstrated in versican V1-transfected PC12 cells, a phaeochromocytoma cell line.

Expression of versican V1 in PC12 cells resulted in upregulation of EGFR and integrin expression and induced NGF-independent PC12 cell neuronal differentiation and neurite outgrowth. Blockade of integrin β1, EGFR, or Src, a downstream signal transduction kinase, significantly blocked this differentiation. This suggested that the differentiation requires signals mediated by versican derived from both integrin and EGFR (Wu *et al.,* 2004). Versican appears to not only present or recruit molecules to the cell surface, but also to modulate the expression levels of genes and coordinate complex signal pathways.

P-selectin glycoprotein ligand-1 (PSGL-1) is a homodimeric glycoprotein held together by disulfide-bonds expressed on the cell surface of leukocytes and mediates leukocyte rolling on the vascular endothelium (Pouyani and Seed, 1995). Leukocyte rolling from the blood stream into tissues in response to inflammatory stimuli is a critical component of immunoresponse, which involves the activities of a range of adhesion and signaling molecules. The first events in this process are leukocyte adhesion and leukocyte rolling on the endothelial surface under vascular shear flow, mediated by selectins. PSGL-1 is expressed by essentially all blood leukocytes including lymphocytes, monocytes, neutrophils, and platelets, and has been shown to mediate the rolling of human neutrophils on selectins (Moore *et al.,* 1995). The N-terminal fragment of PSGL-1 is extensively glycosylated, which is critical for binding to the selectins. PSGL-1 binds to selectins and mediates leukocyte rolling on the vascular endothelium. PSGL-1 also binds to the C-terminal (G3 domain) of versican, cells transfected with PSGL-1 or a shorter form of PSGL-1 containing the G3-binding site, or cells expressing endogenous PSGL-1, aggregate in the presence of versican or

G3 product. The aggregation appears to be induced by G3 multimers that bind to PSGL-1 and form a network. Endogenous versican and/or G3-containing fragments also bind to PSGL-1 in human plasma. Removal of the endogenous G3-containing fragments reduces the effect of plasma on leukocyte aggregation. It reveals a physiologically relevant role for PSGL-1/versican binding and may have implications in the immunoresponse (Zheng *et al.,* 2004).

Beside extracellular matrix and cell surface protein, versican also bind to molecule outside tissue. Studies showed that versican, along with other vascular proteoglycans, bind to and retain low density lipoproteins (LDL) in vascular walls, perhaps contributing to atherosclerotic formations (Olsson *et al.,* 1997). Strong co-localization of LDL and versican has been observed. Lipoprotein lipase (LPL) has domains that bind to both LDL and proteoglycans, while oxidized LDL does not appear to efficiently interact with proteoglycans, including versican. LPL may act as an efficient bridge between these molecules (Olin *et al.,* 1999). It appears that the negatively charged GAG chains of versican interact with positively charged clusters of amino acids on apolipoproteins. Increased LDL binding affinity is seen with longer GAG side chains, greater levels of sulfation, and increased ratios of C-6S to C-4S (or CS C to CS A) (Srinivasan *et al.,* 1995). The number and composition of GAG chains in versican isoforms should have a significant effect on these molecular interactions. The interaction of versican with HA, which generates a very high local density of GAG chains, would also magnify this effect.

1.2.6 Hyaluronan structure, biosynthesis, degradation, and distribution.

The structure of hyaluronan was determined in the l950s. By chemical and enzymatic methods it could be established that HA is a linear polymer built from repeating disaccharide units with the structure [D-glucuronic acid (1-β-3) N-acetyl-D-glucosamine (1-β-4)] as illustrated in figure 1.7. HA preparations extracted from tissues are polydisperse with regard to molecular mass, and the average M_r is usually several million. When hyaluronan is visualized in the electron microscope by the Kleinschmidt technique, i.e., after binding of basic cytochrome c molecules to the negative charges on the polysaccharide, the molecules appear as single chains. The contour length of a chain of M_r 4 x 10^6 is 10 μm. The HA molecule is stabilized by hydrogen bonds parallel with the chain axis. The polymer consequently takes up a stiffened helical configuration, which gives the molecule an overall expanded coil structure in solution. The radius of gyration is about 200 nm, i.e., the coil can be regarded as a highly hydrated sphere containing approximately 1000-fold more water than polymer. The main part of the water is mechanically immobilized within the coil and not chemically bound to the polysaccharide (Laurent and Fraser, 1992).

Biosynthesis of hyaluronan was regulated by three isoforms of hyaluronan synthases (HAS) isoenzymes including HAS1, HAS2, and HAS3. All three isoforms have been first characterized in the bacterium *Streptococcus pyogenes* (DeAngelis *et al.,* 1993), and also have been found in human, mice, and chicken (Meyer and Kreil, 1996; Spicer and McDonald, 1998). To date, most of our knowledge of how this family of HAS genes produces HA derives from studies of the bacterial enzyme. The *Streptococcus pyogenes* HAS a 48-kDa transmembrane protein that passes through the plasma membrane 4 times

and requires the association of about 16 molecules of cardiolipin for full activity (Tlapak-Simmons *et al.,* 1999). Its cytoplasmic UDP-*N*-acetylglucosamine and UDP-glucuronic acid transferase sites add the alternating monosaccharide, most probably to the reducing end of the growing HA chain with continuous extrusion through the plasma membrane using a pore provided by the enzyme itself. Northern blots indicate that the production each HAS is differentially scheduled during the embryonic development and that there are tissue and cell-specific variations in their expression. Genetic deletion of the three HAS genes in mice indicates that only HAS2 is vital to development, resulting in death day 10 because of a failure in the development of the heart (Camenisch *et al.,* 2000). The specific roles that HAS1 and HAS3 play are not yet as well documented. However, overexpression of HAS1, 2, or -3 in several cell types indicates that there are distinct differences in their requirement for cellular UDP-*N*-acetylglucosamine and UDP-glucuronic acid, in elongation rates, and in the final polymer size (Meyer and Kreil, 1996). For instance, HAS3 produces lower molecular weight HA than HAS2 (Brinck and Heldin, 1999). Given increasing evidence that lower molecular mass HA (<200,000 Da) more efficiently activates intracellular signaling pathways via cell associated HABPs than high molecular weight HA, the differential regulation of the HAS may have important consequences for the modulation of cell behavior. HAS2, for instance, may be important in contributing high molecular weight HA that is required for formation of extracellular proteoglycan complexes, particularly abundant in cartilage. Because the synthesis rate and product size of HA appear to depend on the cellular background, HAS are also likely to be subjected to regulation by other proteins (Itano *et al.,* 1999). Overexpression of HAS genes in Chinese hamster ovary cells resulted in greater than

1000 folds enhancement of HA production, inhibited cell migration, and reduced the expression of cell surface CD44, consistent with a study that showed down-regulation of HA receptors with the addition of high levels of exogenous HA. These results indicate that an optimal size and amount of HA is required to promote cell motility, and this response may be cell background-dependent. An altered balance in the ratio of synthesis to catabolism is strongly associated with neoplastic progression (Tammi *et al.,* 2002).

The removal of HA appears to be as important as its synthesis in both morphogenesis and tissue homeostasis. It has been estimated that about a third of HA in the human body is removed and replaced each day. In many instances, removal is achieved by endocytic uptake, either within the tissue where it is made or in lymph nodes and the liver (Tammi *et al.,* 2002). After endocytosis hyaluronan is transported to lysosomes that contain hyaluronidase, β-glucuronidase and β-N-acetylglucosaminidase. The degradation products are glucuronic acid and N-acetylglucosamine. The monosaccharide transferred to the cytosol where the glucuronic acid is degraded via its regular pathway. The N-acetylglucosamine becomes phosphorylated to N-acetylglucosamine 6-phosphate, which is specifically deacetylated. The label of HA tagged in the acetyl group shows up as acetate in cultures of liver endothelial cells. *In vivo,* the acetate becomes oxidized to carbon dioxide and water (e.g., in the hepatocytes). Thus tritium in the acetyl moiety appears in water. The glucosamine 6-phosphate is deaminated to fructose phosphate, which enters the glycolytic pathway (Laurent and Fraser, 1992). The catabolic rate of HA greatly varies between tissues. Labeled HA in the epidermal compartment of human skin organ cultures disappears with a half-life of about 1 day, in contrast to ~20 and ~70 days in the cartilage and vitreous body,

respectively (Fraser *et al.,* 1997). CD44, which plays a major part in the formation of cell-associated matrices can, under certain conditions, also mediate HA endocytosis (for instance during morphogenesis of tissues such as the lung and skin, during long bone growth, and in adult tissues such as cartilage) (Hua *et al.,* 1993; Toole, 1997). The mechanisms of HA uptake by cells appear unique because clathrin-coated pits and caveolae, the most common vehicles for wrapping and detaching receptor bound cargo into endosomal vesicles, seem not to be operative. Mechanisms by which internalized HA is broken down have not been studied in detail but likely involve endosomal and lysosomal hyaluronidases. In some circumstances HA is degraded outside the cell by extracellular hyaluronidases. For example, a hyaluronidase, PH20, is attached to the sperm surface via a glycophosphatidylinositol anchor and is required for sperm penetration through the HA-rich matrix surrounding the oocyte. Thus, it is required for successful fertilization (Lin *et al.,* 1994).

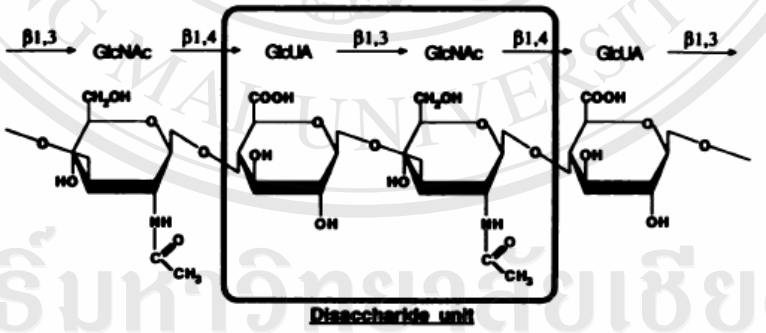

Figure 1.7 The chemical structure of hyaluronan. The polymer is built from alternating units of glucoronic acid (GlcUA) and N-acetylglucosamine (GlcNAc) (Laurent and Fraser, 1992).

Hyaluronan is present in the ECM, on the cell surface, and inside the cell. It is useful therefore to broadly divide the functions of HA into those associated with the organization of the ECM, those associated with a formation of a HA coat on the cell surface, those associated with receptor-mediated signaling, and those associated with the intracellular presence of HA. Extracellular HA is found in tissues that are comprised primarily of ECM, for example cartilage, where its role as an important structural component of the matrix (Knudson *et al.,* 1999). Concentration of hyaluronan in tissues and tissue fluids is depended on type of tissue as shown in Table 1.1. For a HA-rich pericellular matrix to form, it is necessary for it to be anchored to the cell surface. In many instances this is mediated by CD44, but nascent chains of HA that remain attached to the HAS machinery can also contribute to coat formation (Bono *et al.,* 2001). Retention of hyaluronan as a coat at the cell surface allows capture and incorporation of extracellular HABPs, such as aggrecan, into the immediate environment of the cell (Nishida *et al.,* 1999). Evanko *et al.* (Evanko *et al.,* 1999) have recently shown that a HA coat incorporating versican is required for the proliferation and migration of smooth muscle cells *in vitro*. There is growing evidence for the presence of intracellular HA. Intracellular HA has been detected in the cytoplasm of vascular smooth muscle cells during late prophase/early prometaphase of mitosis and in key subcellular compartments such as the nucleus and lamellae during cell locomotion and following serum stimulation. Intracellular HA can be derived from either the extracellular environment or from an as yet unidentified intracellular source and may be involved in nuclear function, chromosomal rearrangement, and other events associated with cell proliferation and motility (Collis *et al.,* 1998; Evanko and Wight, 1999).

Table 1.1 Concentration of hyaluronan in tissues and tissue fluids (Laurent and Fraser, 1992).

Tissue or fluid	Concentration, mg/l
Rooster comb	7500
Human umbilical cord	4100
Human synovial fluid	1420–3600
Bovine nasal cartilage	1200
Human vitreous body	140–338
Human dermis	200
Rabbit brain	65
Rabbit muscle	27
Human thoracic lymph	8.5–18
Human urine	0.1–0.5
Human serum	0.01–0.1

1.2.7 Physical and biological functions of hyaluronan.

Hyaluronan plays several critical roles in organisms both physical and biological aspects. For physical functions, HA plays a structural role in cartilage and other tissues. The cartilage proteoglycan, aggrecan, is bound specifically to HA chains, the bond being stabilized by so-called link proteins. The aggregates formed have masses on the order of 108 Da and are deposited within the collagen framework. Without this interaction the proteoglycans would not be retained in cartilage. In addition, at physiological concentrations, the HA molecules entangle and form a random network of chains. Such networks interact sterically with other macromolecular components. HA excludes other macromolecules, especially large ones that can not find space in the network. It also retards the diffusion of other substances that can not penetrate the network. By these

properties HA and other polysaccharides regulate the distribution and transport of plasma proteins in the tissues. The extraordinary rheological properties that HA solutions exhibit, (e.g., in synovial fluid) have led to speculations about its role in the lubrication of joints and tissues. Joint fluid and concentrated (0.1-1%) solutions of HA show visco-elastic behavior and the viscosity is strongly shear-dependent. These properties are ideal for a lubricant (Laurent and Fraser, 1992).

For biological functions of HA, It is well known from many studies that HA production is high during cell proliferation. Brecht et al. (Brecht *et al.,* 1986) have suggested that the polysaccharide has a function during the mitotic process. The synthesis of HA seems to be maximal during mitosis. A possible function would be that HA, growing out from the cell surface, causes a detachment of the cell from its supporting matrix so that it can divide more easily. In some early studies it was shown that low concentrations of HA could aggregate certain cells such as lymphoma cell lines, virally transformed cell lines, and lung macrophages. This phenomenon is mediated by HABPs on the cell surface. Underhill (Underhill, 1992) subsequently isolated and described a HA-binding receptor from SV-3T3 cells. He found that the receptor was a glycoprotein with a molecular weight of 85,000 Da, and with the aid of a monoclonal antibody he could identify the receptor histochemically in various tissues. It was verified that CD44 can bind HA, a schematic of CD44-HA binding is demonstrated in figure 1.8, and that CD44 apparently was identical with the HA receptor characterized by Underhill (Culty *et al.,* 1990). HA has been connected with various processes during morphogenesis and differentiation. HA concentration often increases in compartments where cellular migration is going to take place-one example is the neural crest cell

migration-and it has been suggested by Toole *et al.* (Toole, 1997; Toole, 2001) that the polysaccharide opens up migration paths. More recent investigations by Turley (Turley *et al.,* 1990) suggest that the polysaccharide actually promotes locomotion. Turley isolated a HABP from the media of chick fibroblast cultures. This protein can be inserted into the cell membrane of urea-pretreated cells and gives the cells an appearance similar to that of virally transformed cells. The protein has been purified by affinity chromatography and contains at least three immunologically related components (56-70 kDa), which can form large aggregates. The HABP co-locates on the cell membrane with actin filaments. Hyaluronan induces metabolic changes in the cells via the binding protein, e.g., stimulation of protein kinase activity. In a recent experiment with Ras-transformed cells it was shown that increased expression of the HABP stimulates locomotion 6- to 10-fold, and that this is dependent on the presence of free extracellular hyaluronan. An important aspect of the role of hyaluronan in locomotion is tumor invasion. Tumors are often enriched in hyaluronan and in some tumors elevated levels are correlated with invasiveness. Tumor cells do not always make HA themselves but can stimulate other cells to produce the polysaccharide. In general, concentrated (viscous) HA solutions seem to inhibit various cell activities whereas dilute solutions in some instances have been reported to be stimulatory. There are reports among others on inhibition of the response to mitogen in lymphocyte stimulation, phagocytosis, locomotion, neutrophil adhesion, production migration inhibitory factor, formation of active E rosettes by T lymphocytes, and rejection of skin allografts. HA stimulation of phagocytosis and other cellular activities has been described for monocytes and in

granulocytes. Subcutaneous injections of small amounts of HA also activate granulocytes *in vivo* (Laurent and Fraser, 1992).

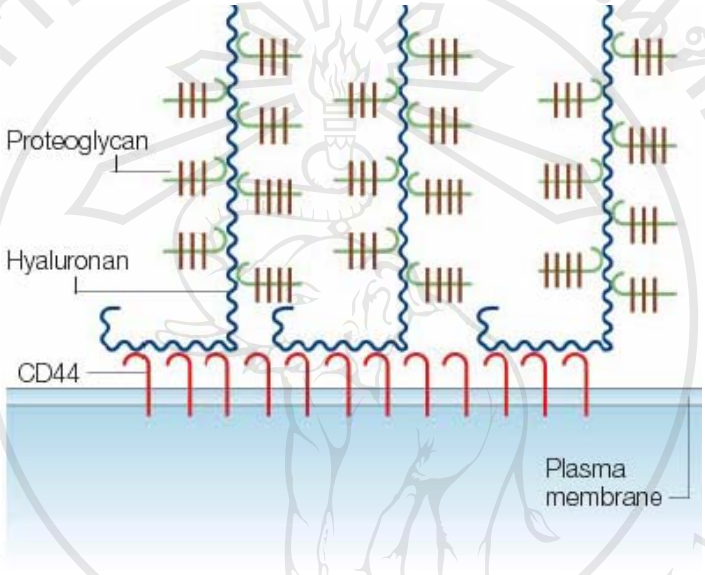

Figure 1.8 Schematic of CD44-HA interactions. HA binds to CD44 on cell surface, and to proteoglycan in extracellular matrix (Toole, 2004).

1.2.8 Hyaluronan-induced signal transduction (Toole, 2004).

Hyaluronan interacts with many proteins, several of which are known or potential cell-surface receptors. Of these, CD44 and RHAMM are established signal-transducing receptors that influence cell proliferation, survival and motility, and are known to be relevant to cancer. Other cell-surface HABPs, such as lymphatic-vessel endothelial

hyaluronan receptor 1 (LYVE1) and TOLL4, might also have roles in cancer pathogenesis. CD44 is a cell-surface glycoprotein that contains an ectodomain, a transmembrane domain and a cytoplasmic domain. The ectodomain includes an amino terminal HA-binding domain that is related to the link modules of HA-binding proteoglycans and link proteins. The region of the CD44 gene that encodes the ectodomain contains a site into which many exon products are spliced in numerous combinations. Although HA is the main ligand for CD44, several other molecules interacting with this protein, many of which bind to carbohydrate side groups that are attached to the spliced-in regions. Among these other ligands, fibroblast growth factors, osteopontin and matrix metalloproteinases (MMPs) are particularly important in terms of relevance to cancer. In response to HA binding, and depending on the cellular context, the cytoplasmic tail of CD44 interacts with many regulatory and adaptor molecules, such as SRC kinases, RHO GTPases, VAV2, GAB1, ankyrin and ezrin. CD44 also mediates the cellular uptake and degradation of HA, which affects growth regulation and tissue integrity. RHAMM is alternatively spliced and the different forms of the resulting protein are found both on cell surfaces and inside cells. Although RHAMM mRNA does not contain a recognizable leader sequence, the protein is transported to the cell surface, where it binds HA and like CD44 transduces signals that influence growth and motility. There is no link module domain in RHAMM, but it does include a HA-binding motif that is presented in several HABPs and contains the sequence B[X7]B (where 'B' represents arginine or lysine, and 'X' represents any non-acidic amino acid). Intracellular RHAMM interacts with several signaling proteins and cytoskeletal components, including SRC, ERK1, actins and microtubules. Interactions of HA with CD44 and RHAMM lead to

numerous cellular responses, including those that involve tyrosine kinases, protein kinase C, FAK, phosphatidylinositol 3-kinase (PI3K), mitogen-activated protein kinase (MAPK), nuclear factor-β and Ras, as well as cytoskeletal components. Although it is clear that CD44 and RHAMM can participate independently in proliferative and migratory phenomena, their relative contributions to any given event have not been fully resolved in most cases, and it is likely that they have redundant or overlapping functions in some situations. In general, the interactions of HA with CD44 and RHAMM are of obvious physiological importance, and their normal activities seem to be disrupted in cancer cells. The other ligands of CD44 and RHAMM compound this complexity. The interactions of HA with these receptors, particularly with CD44, regulate two specific cellular functions that are especially important for tumorigenesis cell survival and ERbB-family signaling.

One of the best studied hallmarks of the behavior of malignant cells is the ability to survive under conditions that would lead to growth arrest or apoptosis in normal cells. This is reflected by the observation that cancer cells are usually able to grow as colonies in soft agar or in suspension-anchorage independent conditions in which normal cells, especially epithelial cells, would undergo apoptosis (a phenomenon that is sometimes known as anoikis). Mechanistically, the phenomenon of anchorage dependent growth in normal epithelia is now known to be due to the requirement for cell-survival signals from ECM components, as well as from growth factors. These signals are usually transduced through interactions of matrix molecules-such as fibronectins, laminins and collagens-with integrins. Cooperative interactions between growth-factor receptors and integrins initiate complex signaling pathways that regulate cell survival and proliferation. Escape

from the requirement for properly regulated anchorage to the ECM presumably allows malignant cells to grow outside normal matrix microenvironments and to travel though the circulation to sites of metastasis. Recent studies have shown that HA strongly promotes anchorage-independent growth and that the resistance of cancer cells to growth arrest and apoptosis under anchorage-independent conditions is dependent on constitutive interactions between HA and CD44. Consistent with these results, several groups have shown that hyaluronan activates the PI3K–AKT signaling pathway, which promotes cell survival. Hyaluronan also stimulates the phosphorylation of FAK and BAD, which also promotes cell survival. These effects are reversed when constitutive HA-CD44 interactions are inhibited. Interestingly, FAK activation can be induced by the interaction of HA with either CD44 or RHAMM. Recent studies have shown that HA promotes the interaction of the cytoplasmic domain of CD44 with the p110 subunit of PI3K through the adaptor protein GAB1 and activates this pathway, providing a direct mechanism by which HA-CD44 interactions regulate cell survival. RHAMM interacts with and activates ERK1, and consequently maintain cell survival. So, the interaction of HA with CD44 or RHAMM can promote cell survival.

It was recently shown that perturbing constitutive HA-CD44 interactions with HA oligomers, soluble HABPs or siRNAs directed against CD44 dissociates a signaling complex that contains CD44, activated ERbB2, ezrin and PI3K, and therefore inhibits downstream ERbB2 signaling. It therefore seems that HA-CD44 interactions regulate constitutive ERbB2 signaling in cancer cells. A similar relation might occur between HA and other receptor kinases that regulate cell behaviors. For example, the interaction of HA with CD44 elicits high affinity binding between the cytoplasmic domains of CD44

and transforming growth factor β-1 receptor (TGFβR-1), leading to increased SMAD2/SMAD3 signaling and other downstream events. In addition, CD44 is required for c-MET signaling in response to the interaction of c-MET with its ligand, hepatocyte growth factor. When it is associated with the plasma membrane, the oncoprotein Tpr-MET stimulates PI3K activity, which in turn induces HA and CD44 production and leads to cell transformation.

1.2.9 Senescence.

Senescence is a terminally arrested cellular growth state. It differs from the non-proliferative state of terminal differentiation and, since senescent cells remain viable, senescence should not be confused with cell death processes such as apoptosis or necrosis. Senescent cells are arrested at the G1/S phase of the cell cycle and are thus also distinct from non-dividing G0-arrested quiescent cells (Sherwood *et al.,* 1988). Human cells are defined as senescent when they fail to respond to mitogens and the population does not divide in a set time period (Smith and Pereira-Smith, 1990). Senescence was first identified as the final stage in the lifespan of normal cells, but it is now recognized that a senescent-like state can be induced in young cells by a variety of factors (Reddel, 1998). Phenotypic changes associated with senescence include increased cell size, multi-nucleation, cytoplasmic vacuolation, and decreased membrane fluidity. The rate of protein, DNA and RNA synthesis is reduced in senescent cells. Senescent cells accumulate altered macromolecules and exhibit DNA alterations such as shortened telomeres, decreased chromatin condensation, increased karyotypic abnormalities and decreased methylation (Goldstein, 1990). Acidic *β*-galactosidase activity is widely used

as a biomarker for senescent cells. There is evidence that senescent cells have altered signaling pathways however, since the expression of many genes in senescent cells remains mitogen inducible, their failure to synthesize DNA following mitogen stimulation is not due to a global signal transduction block (Cristofalo *et al.,* 1989). The expression and activity of many other genes and proteins also change when cells enter senescence. Senescent cells may accumulate in aged tissue and could, by both their altered pattern of gene expression and their non-proliferative state, compromise tissue function and thus contribute to the aged phenotype (Faragher and Kipling, 1998).

The terminal growth arrest of cultured diploid mammalian cells which follows a defined number of Population Doubling (PD) is generally referred to as replicative senescence (also known as the Hayflick limit). This occurs in a variety of different human cell types, as well as in other vertebrate species, indicating that it is a general phenomenon (Cristofalo and Pignolo, 1993). The limited replicative potential of normal human cells is genetically determined and is dominant over the immortal phenotype of tumor-derived cells (Pereira-Smith and Smith, 1983). A senescent-like state can also be induced prematurely by various stimuli including oxidative stress, irradiation, DNA demethylating agents, accumulation of double stranded DNA breaks, histone deacetylase inhibitors and DNA topoisomerase inhibitors. Inappropriate mitogenic signaling also induces premature senescence. Inhibition of PI3K or constitutive MAPK signaling via overexpression of oncogenic Ras, Raf or MEK induces premature senescence in human diploid fibroblasts. Induction of premature senescence by ceramide may involve alterations in phospholipaseD/diacylglycerol signaling. Thus, in addition to preventing tumor formation by limiting replicative potential, senescence also functions as an

alternative mechanism to apoptosis to prevent proliferation of damaged, potentially cancer-forming cells (Duncan *et al.,* 2000).

For oncogene activation-induced senescence such as senescence in response to the overexpression of the Ras oncogene or of downstream effectors such as Raf, activated MAPK and the promyelocytic (PML) oncoprotein. This response has been suggested to represent a tumor-suppressive mechanism, by which cells prevent uncontrolled proliferation in response to the aberrant activation of proliferation-driving oncogenes. It seems, however, that extremely high levels of Ras expression and of its effectors are required for senescence to occur, indeed levels that may not be present in the majority of spontaneously arising human tumors, even those carrying Ras mutations. In a recently described mouse tumor model carrying a single mutated copy of K-Ras, no signs of senescence were observed (Tuveson *et al.,* 2004). However, there are examples of human tumors where senescence may occur in response to excessive Ras signaling. An interesting example is that of Spitz nevi, a benign melanocytic lesion that is considered a precursor for melanoma. Some of these nevi carry a mutated form of H-Ras in multiple copies due to gene amplification. Cells in these lesions express high levels of p16, and are growth arrested (Maldonado *et al.,* 2004).

There are evidences that senescence is regulated by various tumor suppressors such as p53 or p16, presumably it could function as a natural barrier to tumorigenesis. This hypothesis could be directly tested by subjecting a normal cell to potential oncogenic or mitogenic stimuli. Indeed, activated H-Ras (V12) was found to induce premature senescence in primary rodent and human cells (Serrano *et al.,* 1997). Depending on the cellular context, induction of senescence by oncogenic signals such as activated H-Ras

depends on either or both p53 and p16INK4a tumor suppressor proteins. Since Ras signaling involves the Raf-MEK-ERK pathway, it is conceivable that other components of this pathway can also induce premature senescence. Indeed, oncogenic Raf and constitutive expression of MAPK mimic Ras-induced premature senescence in IMR90 fibroblasts (Lin *et al.,* 1998). Furthermore, during oncogenic Ras-induced premature senescence, the Raf-MEK-ERK pathway activates p38 MAPK, and inhibition of p38 activity results in a failure to induce premature senescence by activated Ras. Constitutively active MKK3 and MKK6, which activate p38 MAPK by phosphorylation, can also induce premature senescence by upregulating p53 and p16INK4a in human fibroblasts (Wang *et al.,* 2002). Apart from p53, PML, ARF, and p16INK4a, other proteins are also likely to mediate the H-Ras response. For example, recently, a genetic suppressor element screen identified Seladin-1 as a target gene that is involved in H-Ras induced premature senescence. Interestingly, Seladin-1 encodes an oxidoreductase enzyme involved in cholesterol metabolism. Further studies on Seladin-1 demonstrated that it is an effector of Ras-induced reactive oxygen species (ROS) signaling (Wu *et al.,* 2004). A careful analysis of oncogenic, mitogenic, and other hyperproliferative signals is likely to reveal more cases of premature senescence induction by such signals in primary cells, which in all likelihood represents a failsafe mechanism. For example, similar to activated H-Ras expression, E2F1 overexpression, a potent mitogenic signal, leads to premature senescence in normal human fibroblasts. Premature senescence induction by E2F1 depends on the p53 status of the cells and is mediated by transcriptional induction of p14ARF by E2F1 (Dimri *et al.,* 2000). Constitutive overexpression of E2F3 also results in induction of senescence in transgenic mouse model and cultured mouse

embryonic fibroblasts (MEFs). It was found that a sustained E2F activity, which provides a hyperproliferative signal, induced senescence-like features in mouse pituitary gland (Lazzerini *et al.,* 2005). This report is significant because it indicates that senescence induction by hyperproliferative signals is not merely an *in vitro* phenomenon. Recently, it was shown that the overexpression of oncogenic ERbB2 also upregulates p21 and induces premature senescence in MCF-7 cells. The induction of ERbB2 in this setting causes p53-independent p21 upregulation and premature senescence, which can be reversed by the inhibition of p38 MAPK or functional inactivation of p21 by antisense oligonucleotides (Trost *et al.*, 2005).

1.2.10 Oncogene-induced senescence.

There are evidences that cellular senescence suppresses the development of cancer by arresting the proliferation of damaged or stressed cells that are at risk for malignant transformation (Campisi, 2005). The pivotal integrator and effector of senescent phenomenon is p53, a tumor-suppressor protein that controls the expression of numerous genes. Activity of p53 is regulated at multiple levels, including degradation by H/MDM2, stabilization by ARF, and phosphorylation by the DNA-damage response kinases. Many activated or overexpressed oncogenes, especially those that deliver strong mitogenic signals from plasma membrane receptors, also trigger p53-dependent senescence both in culture and *in vivo* (Braig and Schmitt, 2006). These oncogenes, of which (activated) H-Ras (V12) is the prototype, also trigger MAPK pathway as demonstrated in figure 1.9. Components of the MAPK cascade also increase the expression of the tumor suppressor ARF and initiate a negative feedback loop that

ultimately inhibits H/MDM2, resulting in p53 stabilization (Courtois-Cox *et al.,* 2006). Thus, oncogene induced senescence appears to operate through known mechanisms that converge on p53. Sun *et al.* (Sun *et al.,* 2007) identify a downstream component of the MAPK pathway, MAPK-activated protein kinase 5 (PRAK) that is important for oncogene-induced senescence. Interestingly, PRAK-unlike other MAPK components, such as Raf, immediately downstream of Ras did not cause senescence when oncogenically activated or overexpressed. Rather, PRAK overexpression alone only slightly suppressed cell proliferation. However, it substantially enhanced Ras-induced senescence, increasing both the kinetics and robustness of the response. In fact, PRAK was required for the increase in p53 transcriptional activity that is stimulated by oncogenic Ras. Sun *et al.* (Sun *et al.,* 2007) also show that PRAK directly phosphorylates p53 on a residue that is distinct from those targeted by the DNA-damage response kinases. These new findings provide insights into how some oncogenes engage the senescence program to ensure suppression, rather than promotion of tumorigenesis. They indicate that oncogene-induced senescence and the DNA-damage response have both common and unique modes of action, and that both processes activate p53. In many cases, oncogene-induced senescence is p53 dependent. Oncogenic Raf, for example, induces senescence in both human fibroblasts and mammary epithelial cells independent of p53 function (Olsen *et al.,* 2002). In fibroblasts, oncogenic Raf increases the expression of p16INK4A (Zhu *et al.,* 1998), a cyclin-dependent kinase inhibitor that prevents the phosphorylation and inactivation of the Retinoblastoma protein (pRb). pRb lies at the heart of a second major tumor-suppressor pathway that regulates the senescence response. In mammary epithelial cells, though, Raf induces senescence

independently of both p53 and p16INK4a (Olsen *et al.,* 2002). Given the well-established and crucial role that p53 plays in mediating DNA-damage responses, these findings suggest that oncogenic Ras, but not oncogenic Raf, drives unbalanced DNA replication and the subsequent DNA-damage response-despite the fact that Raf is immediately downstream of Ras (Wahl and Carr, 2001). Moreover, they suggest that there are multiple pathways to the final growth-arrested senescent state (Campisi, 2005). Additionally, there may well be both cell type- and oncogene-specific differences in how oncogenes induce senescence. Moreover, each component of signaling cascades, such as those governed by Ras, may have unique effects not shared by downstream components. For example, Ras and similar GTPases generate ROS as signaling molecules, which in turn activate or inactivate redox-sensitive mediator proteins (Finkel, 2003). It is therefore possible that signaling by oncogenic Ras, but not all components of the MAPK pathway, initiates unique processes, such as the increased expression of genes that result in unbalanced DNA replication. Further, the high level of ROS produced by an overexpressed Ras oncogene might damage nuclear DNA, leading directly to a p53-dependent DNA damage response or might also exacerbate unbalanced DNA replication because of increased oxidative lesions that can retard or stall replication fork progression. Although p53 controls multiple responses to stress and damage, including the permanent senescent growth arrest, a transient cell-cycle arrest, and cell death, p53 is dispensable at least for Raf-induced senescence of mammary epithelial cells (Olsen *et al.,* 2002). The pRb pathway is assumed to be involved in the ultimate mechanisms that establish and maintain the senescent growth arrest (Campisi, 2005), which can be indirectly activated by p53 via transcriptional upregulation of the p21 cyclin-dependent kinase inhibitor.

Although p16 INK4A does not appear to be required for the senescence growth arrest, other means of activating the pRb tumor suppressor pathway may be crucial. Yet, p21 is induced similarly whether cells undergo a transient or senescent growth arrest, and Raf induced senescence of mammary epithelial cells occurs even in the presence of viral oncogenes that inactivate pRb. Thus, the available data can not rule out the possibility of a senescence pathway that is independent of both p53 and pRb (Yaswen and Campisi, 2007). The mechanisms that regulate and implement oncogene-induced senescence, and their relationship to other inducers of senescence, may well be as myriad and complex as the data to date suggest in figure 1.9.

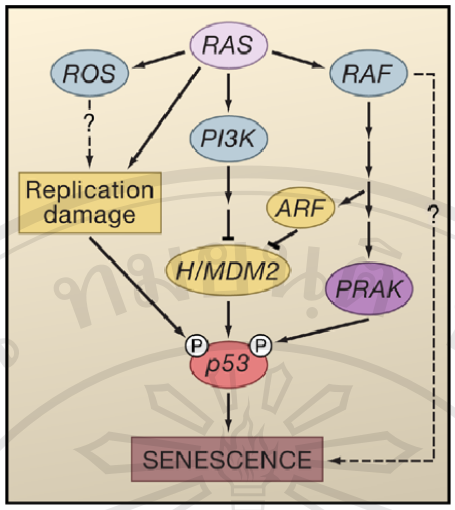

Figure 1.9 Multiple pathways mediate oncogene-induced senescence. Active Ras initiates two protein kinase cascades: the MAPK pathway and the PI3K pathway. The MAPK pathway includes the Raf and PRAK kinases and increases the activity of p53, by two distinct mechanisms. First, MAPK components induce the expression of ARF, and second is PRAK, a downstream MAPK component. In addition, active Raf induces the senescence of some cells independent of p53 function. The PI3K pathway stabilizes p53 through the inactivation of H/MDM2. In addition to activating these kinase cascades, and possibly working through them, oncogenic Ras causes unbalanced DNA replication, resulting in DNA damage. The DNA damage response initiates yet another protein kinase cascade that activates p53, but in this case phosphorylation occurs on sites distinct from that phosphorylated by PRAK. Finally, ROS are an intermediary in Ras signaling and may damage DNA directly (Yaswen and Campisi, 2007).

1.2.11 Senescence and tumorigenesis

Recent studies from Weinberg and Hahn (Boehm and Hahn, 2005) demonstrated that the first step in creating an *in vitro* model of human cancer involves the abrogation of cellular senescence. These studies show that a combination of SV40 large T, human teromerase reverse transcriptase (hTERT), and H-Ras is able to transform a variety of normal human cell types such as fibroblasts, embryonic kidney cells, mammary epithelial cells, ovarian epithelial cells, and endothelial cells. Abrogation of senescence can be achieved by SV40 large T, a combination of HPV oncoproteins E6 and E7, E1A and MDM2 co-expression, or the use of short inhibitory RNA against pRb and p53. Since the INK4a/ARF locus is an upstream regulator of both pRb and p53, various mutations in this locus presumably can also substitute SV40 large T function in transformation assays. Similarly, a combined knockdown of p16INK4a and p53 by the RNAi approach, together with SV40 small t, hTERT, and H-Ras, causes transformation of normal human fibroblasts (Voorhoeve and Agami, 2003). Compared to human cells, murine cells are clearly less rigid in terms of requirements for different genetic mutations for transformation. Nevertheless, bypass of senescence is also essential for murine cell transformation. Established immortal MEF cell lines have often lost p53 or p19ARF. MEFs deficient in p19ARF are highly susceptible to oncogenic transformation and resistant to H-Ras-induced senescence (Sharpless *et al.,* 2004). The H-Ras induced senescence can also be abrogated by other potential oncogenes, such as hDRIL and BCL6, which indirectly affects pRb and p53 pathways of senescence. For example, BCL6 overrides p53 pathway by inducing cyclin D1 expression, while hDRIL targets pRb pathway by binding to E2F1 (Peeper *et al.,* 2002). These potential oncogenes

strongly cooperate with H-Ras to transform mouse and human cells. Similarly, downregulation of Seladin-1 expression by short hairpin RNA, together with hTERT and H-Ras, not only overcomes Ras-induced senescence, but also transforms human fibroblasts (Wu *et al.,* 2004). Deficiency of transcription factor DMP-1, works upstream of ARF, also promotes H-Ras-induced transformation in MEFs. Thus, malfunction of senescence and tumor suppressor pathways facilitate transformation by oncogenes in human and mouse cells (Inoue *et al.,* 2000). So far, most studies have suggested that the failure of senescence-induction pathways in conjunction with activated oncogenes possibly leads to tumor formation *in vivo*. Several recent studies have shown that chemotherapeutic drugs and radiation can induce senescence in tumor cells.

1.2.12 CD44 structure and interactions (Goodison *et al.,* 1999)**.**

The human CD44 gene has been mapped to the chromosomal locus 11p13 and is composed of two groups of exons. One group, comprising exons 1–5 and 16–20, are spliced together to form a transcript that encodes the ubiquitously expressed standard isoform as shown in figure 1.10. The 10 variable exons 6–15 (also known as v1–10) can be alternatively spliced and included within the standard exons at an insertion site between exons 5 and 16. Molecules containing the variable exons or their peptide products are designated CD44v and, in theory, alternative splicing would allow more than a thousand different CD44 variants to be generated. Further complexity of expression can be achieved by the alternative splicing of some of the standard exons and by the use of cryptic splice sites. In the human gene, exon 6 (v1) contains a stop codon at the 17th

amino acid and so, unlike in mice, this exon is not normally included in processed transcripts.

The most abundant standard isoform of human CD44 protein (CD44s) contains 363 amino acids and has a theoretical molecular mass of 37 kDa. The protein consists of three regions, a 72 amino acids C-terminal cytoplasmic domain, a 21 amino acids transmembrane domain, and a 270 amino acids extracellular domain as shown in figure 1.11. The cytoplasmic region is encoded by part of exon 18 and by exons 19 and 20. This highly conserved cytoplasmic tail can exist as a short or a more prevalent long form by the inclusion of the C-terminal exon. The hydrophobic transmembrane domain is encoded by exon 18 and is 100% conserved between mammalian species. The extracellular domain can be subdivided further into conserved and non-conserved regions. The N-terminal ectodomain, encoded by exons 1 to 5 is highly conserved (~ 85%) between mammalian species and is thought to fold into a globular tertiary structure by the formation of disulfide bonds between three pairs of cysteine residues. Also present in this section is a 100 amino acids region of homology with other HABPs. The variable region is the point at which up to 381 amino acids encoded by the 10 alternatively spliced variant exons are inserted at a site between exons 5 and 16 of the RNA transcript, corresponding to amino acid position 223. The membrane proximal region of the extracellular domain, encoded by exons 16 and 17 is less conserved (35%) between mammalian species and includes several carbohydrate modification sites. The apparent molecular mass of the CD44s protein, as estimated by gel electrophoresis, is ~80 kDa and the largest possible protein, containing peptides from all variant exons can be over 200 kDa. This is far greater than the expected value as calculated from amino

acid residues and is the result of the extensive post-translational modification of CD44 isoforms. There are at least five conserved consensus N-glycosylation sites in the N-terminal ectodomain and two CS attachment sites on the exon 5 product. Several potential O-linked glycosylation sites reside in the membrane proximal extracellular region and there are also consensus attachment sites for heparan sulfate, keratin sulfate, and sialic acid residues on the standard extracellular region. The alternatively spliced, variable exon products also have extensive potential modification sites, including serine/threonine rich regions for O-glycosylation, 10 glycosaminoglycan attachment consensus SGXG motifs, and tyrosine sulfation. The degree of glycosylation can affect the ligand binding characteristics of the protein and therefore alter its function. Thus, the regulation of the amount and the type of posttranslational modification can add further diversity to the range of potential functions of CD44 isoforms.

The principal ligand of CD44 is HA, and there are at least three sites for HA binding on the CD44 molecule: one in the link domain encoded by exon 2 and the other two overlap in the region encoded by exon 5. The HA binding sites consist of basic amino acid clusters, with specific arginine residues being essentially required, as shown by mutation studies. However, although all CD44 isoforms contain the HA recognition site, not all cells expressing CD44 bind the HA ligand constitutively. Cells can express CD44 in an active, an inducible, or an inactive state with respect to HA binding. The differences in the HA binding state of CD44 are cell specific and have been shown to be related to post-translational modification patterns. Inhibition of N-glycosylation enhances HA binding, and abrogation of glycosylation by mutation at specific sites can convert the inducible form to the constitutively active form. In certain cell types, inactive

CD44 surface protein can be rapidly induced to bind HA by interaction with specific antibodies. The induced binding state can also be achieved more slowly by incubation of the cells with stimulating agents such as Phorbol esters or interleukins. Detailed mutational analysis of basic amino acid clusters in the cytoplasmic domain has defined specific arginine and lysine residues through which protein kinase C stimulating reagents differentially regulate the binding of CD44 to HA.

The attachment of chondroitin sulfate to sites on the CD44 protein (exon 5) can lead to a number of indirect ECM binding interactions. Collagen, fibronectin, and laminin can bind to CD44 in this way. Serglycin, a family of proteoglycans named for its Ser-Gly dipeptide repeats, has also been reported to bind to CD44 indirectly. Chondroitin sulfate modified serglycin is secreted from haemopoietic cells and its interaction with CD44 leads to the aggregation, adherence, and activation of lymphoid cells. CD44 isoforms containing the alternatively spliced exon v3 can be modified with heparan sulfate at a Ser-Gly-X-Gly motif and are then able to bind basic fibroblast growth factor and heparin binding epidermal growth factor. The manipulation of glycosylation patterns by the alternative splicing of CD44 is an important mechanism for the achievement of multifunctional products from a single gene.

The status of the cytoplasmic domain of CD44 can influence the binding of extracellular HA, presumably by membrane localization, involving clustering or dimerization. However, the main role for the cytoplasmic domain is to transduce signals from extracellular stimuli. The cytoplasmic domain has been shown to bind to a number of intracellular proteins directly. An ankyrin binding domain has been identified and binding appears to be enhanced by palmitoylation of CD44 amino acid residues. A

correlation between the ability to bind ankyrin and HA has been proposed and, although it is not demonstrable in all cell types, it is clear that dynamic intracellular and extracellular interactions can be facilitated by CD44 proteins. As an example of this, it has been reported that a CD44v3 isoform, present in a metastatic breast cell line, can interact with ankyrin via the cytoplasmic domain and is found closely associated with an active MMP in a complex within pseudopodia structures during cell migration. CD44 interactions with the ERM family (ezrin, radixin, and moesin) of proteins have also been described. CD44 has been shown to co-precipitate and co-localize with ERM proteins, and a region of basic amino acids in the cytoplasmic tail has been identified as essential for the binding of moesin and ezrin. Ankyrin and ERM proteins connect elements of the plasma membrane with the actin filament network of the cell and thus a direct link between CD44 and intracellular scaffold structures can be envisioned, facilitating cell motility and migration, as well as membrane localization of CD44.

1.2.13 CD44 functions and relating signals.

The multiple functions of the CD44 family of proteins are centered on the binding of HA and, to a lesser extent, other extracellular molecules. Functional diversity is achieved by the alternative splicing of the pre-mRNA, and the fine tuning of ligand binding can be achieved by post-translational modification and dynamic interaction with other cellular factors. The CD44s protein is involved primarily in the maintenance of three dimensional organ/tissue structures. HA accumulates in angiogenesis, wound healing, and embryonic cell migration (Jain *et al.,* 1996; Trochon *et al.,* 1996). HA appears to serve as a ligand support for CD44 mediated cell movement. CD44 molecules can also

mediate the aggregation of cells. This can be achieved via multivalent HA binding by CD44 on adjacent cells or via inter-CD44 binding via attached glycosylation moieties. HA dependent binding can cause aggregation of macrophages, lymphocytes, and fibroblasts (Underhill and Dorfman, 1978; Green *et al.,* 1988; Cooper and Dougherty, 1995). Moreover, CD44 can serve as a co-receptor, physically linked to other classical signaling receptors such as c-Met, members of the ErbB family of receptor tyrosine kinases (Orian-Rousseau *et al.,* 2002) and TGF-βR1 (Ito *et al.,* 2004; Ito *et al.,* 2004), and in the process, facilitate the association of intracellular mediators of signal transduction. This interaction with CD44v6 is not due to heparin sulfate substation. Downstream signaling of c-Met (MEK and ERK phosphorylation) requires the presence of an intact CD44 cytoplasmic domain (Orian-Rousseau *et al.,* 2002). CD44 also provide a support stage for binding basic growth factors (Bennett *et al.,* 1995). CD44 also capable to mediate internalization of aggrecan fragments that remain bound to HA (Jiang *et al.,* 2002; Embry and Knudson, 2003).

CD44 can mediate cell signals through itself directly or via other interacted proteins. In signal transduction, however the CD44 cytoplasmic tail exhibits no inherent receptor kinase or phosphatase, a simple view of CD44-mediated signaling is similar to that of integrin signaling; that is, indirect transfer of ECM information via associated signaling proteins that become linked to the cytoplasmic domain of CD44. These functional signaling complexes are organized primarily by anchoring proteins and elements of the cytoskeleton. In nucleated blood-borne cells, as well as migrating embryonic, endothelial or malignant cells, unoccupied CD44 receptors are activated by the binding of HA, initiating extracellular clustering of CD44. The extracellular clustering of CD44 results

in intracellular events including the activation of kinases such as c-Src and FAK as well as Rho and Rac, leading to enhanced association of CD44 in to actin cytoskeleton complexes and the recruitment and activation of additional signaling partners that lead to cell migration (Bourguignon *et al.,* 2001; Murai *et al.,* 2004). HA-CD44 interactions also augment the phosphorylation of MAPK and Akt pathways (Sohara *et al.,* 2001). In other tissues where HA is more ubiquitous and in abundance, clustered CD44 receptors occupied in a multivalent fashion with high-molecular-mass HA represent the quiescent state of the cells. For these cells, CD44 signaling is initiated by disruption of HA-CD44 interactions. This disruption may occur by degradation of the HA (Nishida *et al.,* 2003), by the presence of soluble CD44 acting as a competitor for HA (Peterson *et al.,* 2000), by cleavage of the ectodomain of CD44 (Sugahara *et al.,* 2003) or by the presence and competition of small HA oligosaccharides (Ghatak *et al.,* 2002). In these instances the release of extracellular constraints or clustering imposed by bound macromolecular HA is the likely signal transduction. The association of CD44 with the cytoskeleton ERM (ezrin/radixin/moesin) and/or ankyrin adaptor proteins is regulated in part by CD44 phosphorylation (Ponta *et al.,* 2003). Morrison *et al.,* (Morrison *et al.,* 2001) proposed a mechanism whereby dynamic associations of CD44 with cytoskeletal complexes could be regulated by change in the phosphorylation state of ERM proteins and of another adaptor protein, merlin. Unlike ERM however, merlin has little capacity to engage with the actin cytoskeleton and as such functions as a competitive inhibitor of ERM-mediated CD44 tethering to the actin cytoskeleton. Phosphorylated ezrin bind CD44, by dephosphorylation of merlin results in its binding to CD44. Thus, the regulated phosphorylation/dephosphorylation of merlin could function as a regulatory switch by

which CD44-HA interactions impact cell proliferation. Several reports have documented a new mechanism for direct CD44-mediated signaling: a two steps process beginning with an extracellular domain cleavage, followed by γ-secretase cleavage of the transmembrane domain releases the CD44 intracellular domain (ICD) (Okamoto *et al.,* 2001; Lammich *et al.,* 2002), which exhibits nuclear translocation and functions as a transcription factor in a manner similar to the Notch receptor (De Strooper *et al.,* 1999). The CD44-ICD potentiates transactivation with the p300/CREB-binding protein to activate transcription, and one of the target genes identified was CD44 itself (Okamoto *et al.,* 2001).

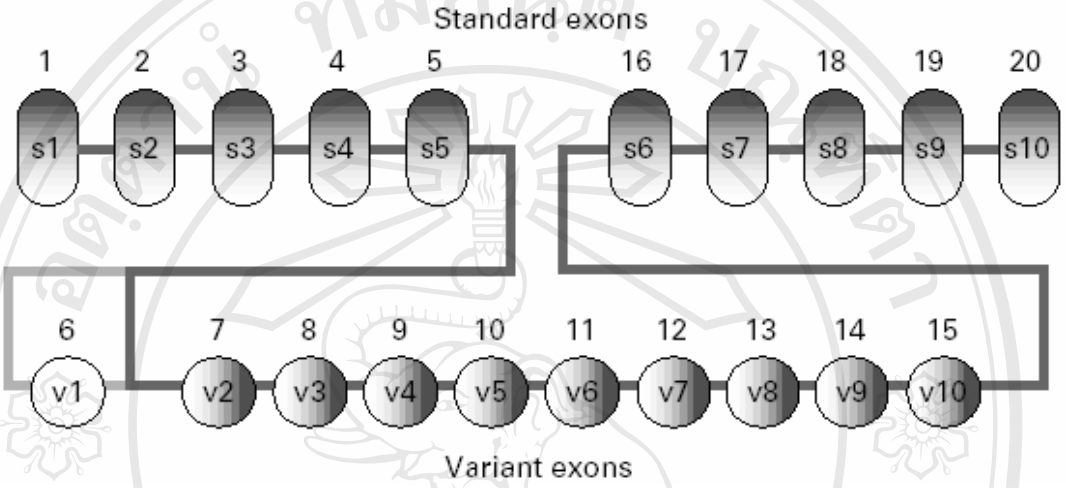

Figure 1.10 Schematic diagram of the structure of the CD44 gene. The standard exons (s1–10) encode the ubiquitously expressed standard protein isoform, CD44s. Combinations of the variant exons (v1–10) can be alternatively spliced between s5 and s6 to encode variant protein isoforms, CD44v. Unlike the mouse gene, exon 6 (v1) of the human gene contains a stop codon and is not normally included in human CD44 mRNA (Goodison et al., 1999).

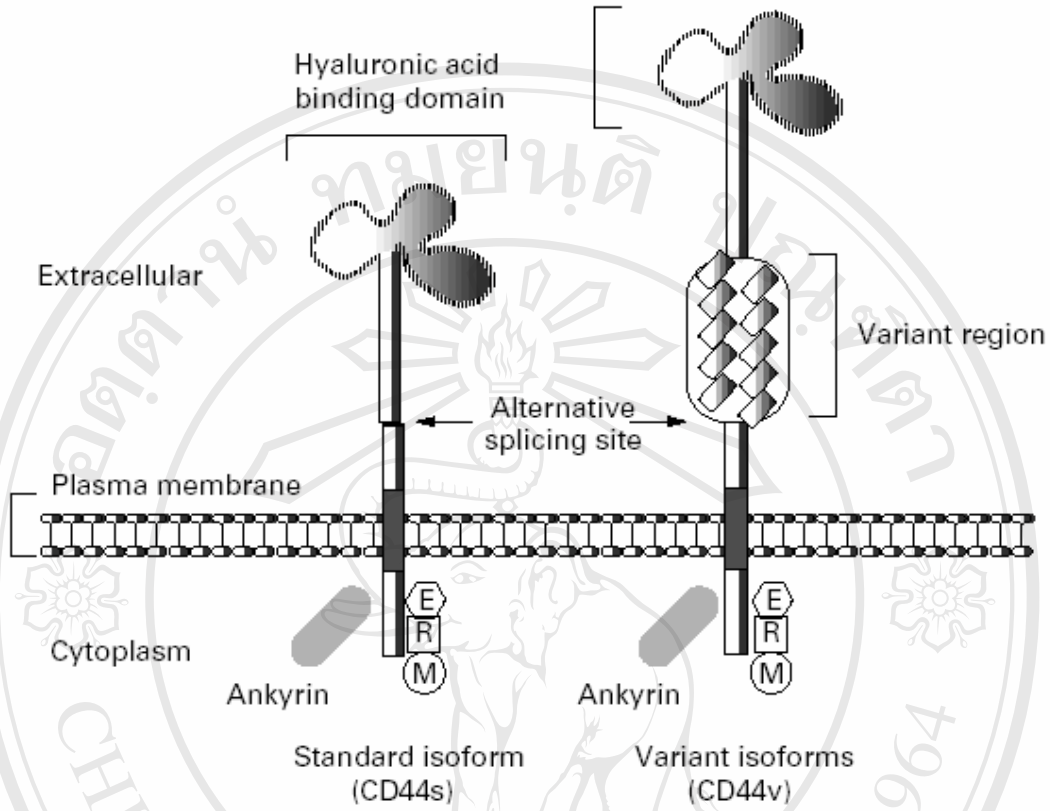

Figure 1.11 CD44 protein structure. The standard isoform binds its principal ligand, HA at the N-terminal, distal extracellular domain. The inclusion of combinations of the variant exons (v1–10) within the extracellular domain can alter the binding affinity for HA and confer interaction with alternative ligands. The molecule interacts with the cytoskeleton through the binding of ankyrin and the ERM family (ezrin, radixin, moesin) to the cytoplasmic domain (Goodison et al., 1999).

1.3 Objective

To understand mechanism or interaction of versican in controlling cell behavior and organ development.