Glycoconjugate Journal 20, 247–255, 2004


# Galectin-4 in normal tissues and cancer

Margaret E. Huflejt[1] and Hakon Leffler[2]

[1] *Sidney Kimmel Cancer Center, 10835 Altman Row, San Diego, CA 92121, USA,* [2] *Section MIG (Microbiology, Immunology, Glycobiology), Inst Laboratory Medicine, Lund University, Sweden*

**Galectin-4 belongs to a subfamily of galectins composed of two carbohydrate recognition domains within the same peptide chain. The two domains have all the conserved galectin signature amino acids, but their overall sequences are only approximately 40% identical. Both domains bind lactose with a similar affinity as other galectins, but their respective preferences for other disaccharides, and larger saccharides, are distinctly different. Thus galectin-4 has a property of a natural cross-linker, but in a modified sense since each domain prefers a different subset of ligands. Similarly to other galectins, galectin-4 is synthesized as a cytosolic protein, but can be externalized. During development and in adult normal tissues, galectin-4 is expressed only in the alimentary tract, from the tongue to the large intestine. It is often found in relatively insoluble complexes, as a component of either adherens junctions or lipid rafts in the microvillus membrane, and it has been proposed to stabilize these structures. Strong expression of galectin-4 can be induced, however, in cancers from other tissues including breast and liver. Within a collection of human epithelial cancer cell lines, galectin-4 is overexpressed and soluble in those forming highly differentiated polarized monolayers, but absent in less differentiated ones. In cultured cells, intracellular galectin-4 may promote resistance to nutrient starvation, whereas—as an extracellular protein—it can mediate cell adhesion. Because of its distinct induction in breast and other cancers, it may be a valuable diagnostic marker and target for the development of inhibitory carbohydrate-based drugs.**
***Published in 2004.***

***Keywords:*** **galectin-4, galectins, carbohydrate specificity, promoter, breast cancer, epithelial cells**

## Introduction: Early discoveries, basic properties, and expression in normal tissues

Galectins are defined as proteins containing a canonical carbohydrate recognition domain (CRD) with affinity for $\beta$-galactosides [1–3]. The first discovered protein in the family, now known as galectin-1, is a non-covalent dimer composed of a subunit with one CRD, and the later discovered galectin-3 has one CRD linked to a long N-terminal repetitive sequence. Galectin-4, and an independently discovered 32 kDa galectin from *C. elegans* [4], were the first galectins defining a subfamily with two CRDs in one polypeptide chain. Other presently known mammalian galectins with two CRDs are galectins-8, -9, and -12, and galectin-6, most likely a recent duplication of galectin-4 only found in mouse [2,3], and orthologues of the mammalian bi-CRD galectins have recently been found in other vertebrates [5,6]. The bi-CRD galectins in more distantly related species *e.g. C.elegans* and *Drosophila*, are too different to permit identification as orthologues but demonstrate that bi-CRD galectins are ancient members of the galectin family [5,6].

Rat and human galectin-4 consist of CRDs of 133aa and 130aa, respectively, connected by a link peptide of 34aa and preceded by 17 aa [7,8], and these proportions are similar for other species. The sequences of the two CRDs are about as related to each other (about 40% identical) as they are to the CRDs of galectin-3 and many other galectins, but they are less similar to galectin-1 (about 20% identical). Composed of two CRDs, galectin-4 may be functionally divalent, but, as will be described in detail below, the carbohydrate binding specificities of the two CRDs are quite different and would be expected to show preference for different sets of ligands. The link peptide is homologous to the proline- and glycine-rich repeating domain of galectin-3. There is no evidence for variation in linker length due to alternative splicing as found for galectins-8 [9] and -9 [2,10]. High content of basic amino acids in galectin-4 renders the protein quite basic, especially the C-terminal CRD, with pI of approximately 9.

The early discoveries of galectin-4 highlight two of its properties: firstly, it tends to be associated with relatively insoluble tissue components, and, secondly, the link peptide is sensitive to proteolysis. Only a fragment containing the C-terminal CRD

To whom correspondence should be addressed: Margaret E. Huflejt, Sidney Kimmel Cancer Center, 10835 Altman Row, San Diego, CA 92121, USA. Tel: 858-450-5990; Fax: 858-450-3251; E-mail: mhuflejt@skcc.org

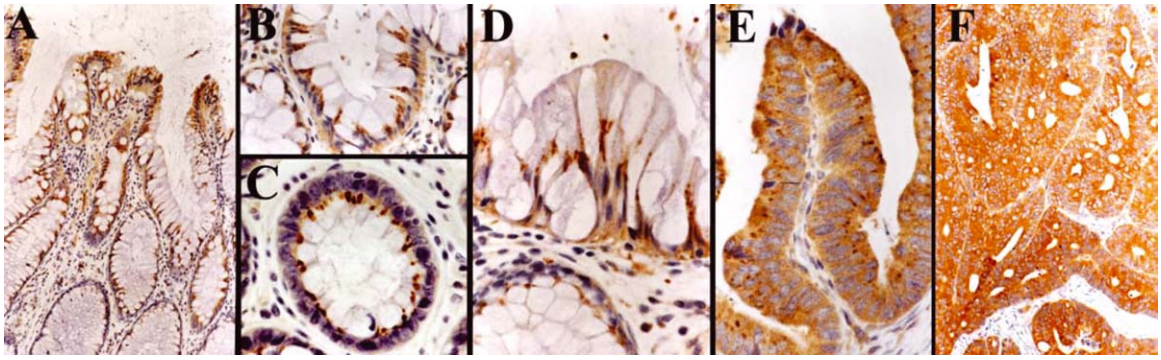


**Figure 1.** Immunohistochemical detection of galectin-4 in normal human colonic mucosa (panel A–D), an adenomatous polyp (panel E), and in colon adenocarcinoma (panel F). Five $\mu$m-thick sections of formalin-fixed and paraffin-embedded tissues were deparaffinized, hydrated, and treated using an antigen retrieval procedure. Galectin-4 was localized using rabbit antiserum raised against the C-terminal CRD of rat galectin-4 (RI-H, [7]) followed by peroxidase-conjugated goat anti-rabbit IgG, and development with diaminobenzidine (amber to deep brown). Nuclei were counterstained with hematoxylin (blue). In the normal colonic mucosa, galectin-4 is found in dense supra-nuclear formations in crypt cells (panels A–D, and also distributed as diffusely cytosolic in some cells closer to the lumen (upper part of panel A). In an adenomatous polyp proximal to an adenocarcinoma the supra-nuclear formation of galectin-4 are still present, but increased diffuse cytosolic expression is apparent (panel E). In an adenocarcinoma the supra-nuclear formation of galectin-4 are absent and instead there is strong diffuse cytosolic expression (panel F). Original magnification: A: 100×; B: 200×; C: 400×; D: 400×; E: 400×; F: 100×.

of rat galectin-4 (named RI-H at the time) was initially isolated by affinity chromatography of intestinal extracts on lactosyl-Sepharose [11]; later the N-terminal CRD was also isolated from rat small intestine [12]. Independently, Chiu *et al.* isolated, by a harsh extraction procedure, a protein tightly associated with adherens junctions in pig tongue epithelium [13], which was then shown to be full length galectin-4 [14]. Full length galectin-4 was also identified in detergent-insoluble complexes from pig small intestine [15]. Localization with antibodies showed galectin-4 also in dense aggregates in esophageal squamous epithelium [16] and small and large intestinal epithelium (Figure 1).

The relative insolubility in tissues described above may be related to the biochemical properties of intact galectin-4 being relatively hydrophobic (as pointed out, [14]) and poorly soluble (precipitates in buffers of physiological ionic strength at approximately 2 mg/ml), but may also be due to its tight association(s) with endogenous ligands [17]. The fact that only individual CRDs are found during affinity purification from tissue extracts suggests cleavage of the link peptide by tissue proteases. Whether this occurs *in vivo* and has functional significance, however, remains to be demonstrated.

Full-length galectin-4 has only been isolated in soluble form from human cancer cell lines with a differentiated polarizing epithelial phenotype (panels A–C in Figure 2, [8]) as discussed in detail below. Even in this case, however, galectin-4 was targeted to specific subcellular domains as demonstrated by immunofluorescence. Galectin-4 also displayed a strikingly different localization from that of galectin-3 expressed by the same cells [8]. Such a difference has also been revealed by immunohistochemistry in normal large and small intestinal tissues (Huflejt and Leffler, unpublished). These observations suggest that the two galectins will preferentially interact with different ligands.

Galectin-4 mRNA has been identified in human [8,18], pig [14], rat [7] and mouse alimentary tract, and in rabbit bladder [19]. Human galectin-4 EST clones have been isolated mainly from intestinal cDNA libraries, with the highest prevalence in colon (1.5–4%), but galectin-4 EST clones have also been found as rare species in libraries from other tissues such as breast, ovary and blastocyst (as searched by UniGene under http://www.ncbi.nlm.nih.gov/entrez/, Nov. 16, 2003).

The analysis of mouse galectin-4 was complicated by the concomitant discovery of galectin-6 [20]. Galectin-6 is closely related to galectin-4, so far has only been found in mouse, and probably arose from a recent gene duplication. It has a linker domain shorter by 24aa as compared with galectin-4, but otherwise the sequences of these two galectins are 93% identical at the nucleotide level and 83% identical at the amino acid level, with small differences scattered throughout. Because of this similarity, it has been difficult to distinguish galectin-4 from galectin -6, and they are detected together in most antibody-based and nucleotide hybridization assays in mice. During embryogenesis, *in situ* hybridization showed expression of their mRNAs only in the developing alimentary system starting from day 13.5 [20], in stark contrast to the more wide and different expression patterns for galectins-1, -3, and -7 [21]. In adult mouse tissues galectin-4/-6 was found by Northern and Western blots again only in gastrointestinal tissues but not in brain, kidney, skeletal muscle, heart, liver, lung [20]. An RNAse protection assay, permitting their separate detection, indicated high expression of galectin-4 in small and large intestine but much lower in stomach, whereas galectin-6 was equally expressed also in stomach. Western blot supported this finding, although the detection of galectin-6 was

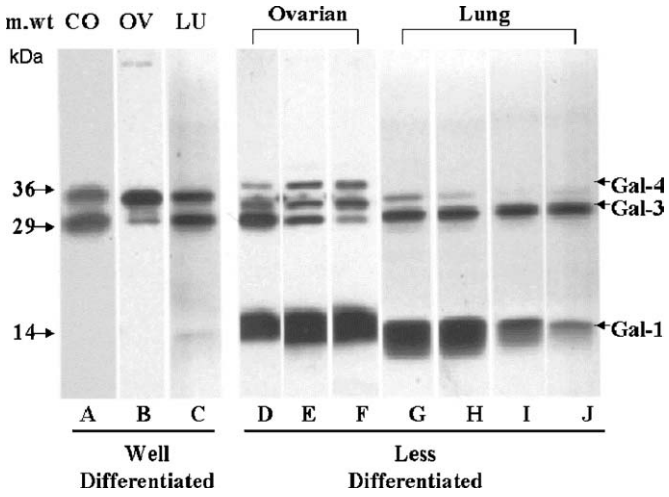


**Figure 2.** Galectin expression profiles in human epithelial cancer cell lines. Galectins were isolated by affinity chromatography on lactosyl-Sepharose from lysates of cells metabolically labeled with $^{35}$S Met/Cys, resolved by SDS-PAGE and visualized by autoradiography [8,42,62]. Lanes A–C: galectins expressed by well-differentiated cell lines: T84 colon adenocarcinoma (CO, lane A), Mls ovarian adenocarcinoma (OV, lane B), and Calu-3 lung adenocarcinoma (LU, lane C). Lanes D–F: galectins expressed by poorly differentiated ovarian carcinoma cell lines: SRS (lane D), SKA (lane E), OW-1 (lane F); Lanes G–J: galectins expressed by poorly differentiated lung carcinoma cell lines: SK-Mes-1 (lane G); Calu-1 (lane H); SK-Lu-1 (lane I); A549 (lane J). The bands at about 14 and 29 kDa were identified as galectins-1 and -3, respectively, by Western blotting with specific antibodies. The band at about 36 kDa in the well-differentiated cells was identified as galectin-4 by Western blotting and by mass spectrometric mapping of tryptic peptides. The other bands slightly above or below 36 kDa in the less differentiated cell lines have not been identified but they likely correspond to other bi-CRD galectin(s).

very weak [20]. There are many mouse galectin-4 ESTs, but very few, if any, of galectin-6 (as searched by UniGene under http://www.ncbi.nlm.nih.gov/entrez/, Nov. 16, 2003). This indicates that of the two—galectin-4 is the one mainly expressed in mouse and is the orthologue of galectin-4 of other mammals, whereas galectin-6 is a minor component which may have both overlapping and different functions.

## Carbohydrate-binding specificity and cross linking activity of galectin-4 and its individual domains

Both CRDs of galectin-4 have similar affinities for lactose ($K_d$ about 0.5–1 mM) to those of galectins-1 and -3, consistent with their full complement of the galectin signature amino acids that interact with bound disaccharide in the core binding site, with identity of six and a conservative substitution of the seventh aa (Lys for Arg) [7]. Moreover, the X-ray crystal structure of the N-terminal domain of rat galectin-4 (Kayden, Lobsanov, Leffler, Rini, unpublished) confirms that it adheres closely to the canonical galectin structure. This justifies interpreting galectin-4 specificity based on the general model for galectins, with subsites A-E, presented in the Introduction to the special issue [3]). The core $\beta$-Gal residue is bound in subsite C and interacts with most of the signature aa residues. However, the two CRDs of galectin-4 differ significantly, both from each other and from other galectins, in preference for other disaccharides (extensions at the reducing side of the Gal into subsite D) and fine specificity for larger saccharides (extensions at the non-reducing side of the Gal into subsite B and A).

For the N-terminal CRD of rat galectin-4, Gal$\beta$1-4GlcNAc (LacNAc) is a significantly worse ligand than lactose, in contrast to the case of galectin-1 and -3 for which LacNAc is the preferred and most common ligand. The affinity of galectin-4-N for LacNAc, is, in fact, about 50 fold lower as compared to galectins-1 and -3. Instead, 3-linked saccharides at the reducing side of Gal (in subsite D), *i.e.* Gal$\beta$1-3GlcNAc and Gal$\beta$1-3GalNAc (T-antigen) were better ligands than LacNAc [7]. These results were confirmed and extended in a more recent study with a larger number of glycoproteins and saccharides [22]. Of the saccharides used in this study, Gal$\beta$1-3GlcNAc$\beta$1-4Gal$\beta$1-4Glc was the best ligand/inhibitor. However, it is not known in which of two possible alternative ways it binds the galectin: with the terminal Gal$\beta$1-3GlcNAc in the core binding site (C–D, [3]) and the Gal$\beta$1-4Glc in site E, or with the internal Gal$\beta$1-4Glc in the core site with Gal$\beta$1-3GlcNAc in sites A–B. Except for possibly the latter case, no specific preference for extension at the non-reducing side of the core Gal into subsite B was found so far [7,22].

For the C-terminal domain of rat galectin-4, LacNAc is also a weaker ligand than Lac, but so are Gal$\beta$1-3GlcNAc and Gal$\beta$1-3GalNAc [7]. At the non-reducing side of the core Gal (in subsite B) the C-terminal domain shows a preference for GalNAc$\alpha$1-3 (part of a blood group A determinant) almost as strong as galectin-3, whereas this is not the case for the N-terminal CRD [7,22].

When the two domains of rat galectin-4 were used as histochemical probes on intestinal tissue sections, they showed strikingly different staining patterns indicating preferences for different endogenous ligands [23]. Most recently distinct specificity differences between the two CRDs of human galectins-4 have been confirmed by fluorescence polarization analysis using a panel of probes [24]. In another study, a preference of full-length human galectin-4 for 3-O-sulfated Gal was found, but the role of each domain was not determined [25]. It, therefore, may be suggested that the N-terminal domain was involved in this interaction since the best saccharide inhibitor had the 3-O-sulfate linked to Gal$\beta$1-3GalNAc, a disaccharide preferred by the N-terminal domain as discussed above.

From the specificities described above, it is clear that galectin-4 is bivalent in a modified sense: it cross-links ligands, but these ligands will likely be different for each CRD. If and how this fact influences the biological activities of galectins-4 remains to be elucidated. Differences in carbohydrate-binding specificities of individual CRDs are also an important question

in case of other bi-CRD galectins, as discussed for galectins-9 [10].

The full-length galectin-4 and the two domains behave as monomers upon size exclusion chromatography [7,11]. However, this does not rule out that under certain conditions such as high concentration, or encounter with appropriate ligands, dimers or oligomers of higher order can be formed, as has been the case for mono-CRD galectins-3, -5 and -7 [3,26]. Such aggregation might explain why the N-terminal domain of rat galectin-4 appears to interact particularly avidly with certain mucins carrying multiple Gal$\beta$-containing saccharides [22]. However, as these assays are complex and require direct binding to immobilized mucin or inhibition by soluble mucin, and the mucins are large (over 1 million Da), heterogeneous and only partially defined, other explanations are also possible.

Recently synthetic inhibitors of galectin-4 have begun to be developed. Screening of a panel of 66 compounds, constructed by aromatic addition at C3 of the core Gal [27], revealed those compounds that inhibited either domain, as well as both [28]. In another line of investigation lactulose-amines have been found to bind to the full-length galectin-4 in a BIAcore assay, with di-lactulose hexamethylene di-amine (L2) showing higher binding affinity than mono-lactulose amines [29]. The L2 compound also retarded growth of breast tumors in mice transgenic for Her-2/*neu*, as described in detail below. It is currently under investigation whether galectin-4 or other galectin(s) in murine tumors are targeted by L2.

## Galectin-4 in human malignancies

Over 200 reports describe tumor-dependent changes in localization and expression levels of galectins 1 and -3, and various tumor-promoting activities of these proteins have been demonstrated [30–32]. More recent evidence shows that expression of many other galectins, including galectin-4, is altered in human malignancies [9,10,18,29,33–38]. Galectin-4 was also identified as a target for autologous antibodies in patients with colon cancer [39], a phenomenon also observed for galectin-9 in Hodgkin' lymphoma [40]. Galectin-8 has also been found to be immunogenic in cancer tissue, as it was identified as a major target for tumor specific monoclonal antibodies raised against prostate and lung cancers [9,41].

Most of the studies on tumor-related galectin-4 expression report only changes at the mRNA level. In colon cancers galectin-4 mRNA expression was found to be much lower than in normal colon tissues [18]. In contrast, it was higher in hepatocellular carcinomas [35] and in gastric cancer cells with increased metastatic potential [36], as compared to the low level in the corresponding normal tissues. To understand the biological significance of such changes it will be necessary to also analyze galectin-4 protein expression and localization in the same tissues.

Immunohistochemical analysis of galectin-4 protein expression and localization is the subject of an on-going prospective study [29,37]. Examples of colon and breast tissues are shown in Figures 1 and 3. In colon, the dominating feature during malignant transformation appears to be the progressive loss of the dense supra-nuclear galectin-4 aggregates, typical for the normal crypt and upper crypt epithelial, cells (panels A–D), and increase in the concentration of cytosolic galectin-4 showing more diffuse distribution (Figure 1E–F). In another study an increased percentage of cells expressing galectin-4 (immunohistochemical labeling index) was found to correlate with a poor prognosis of colon carcinoma [38].

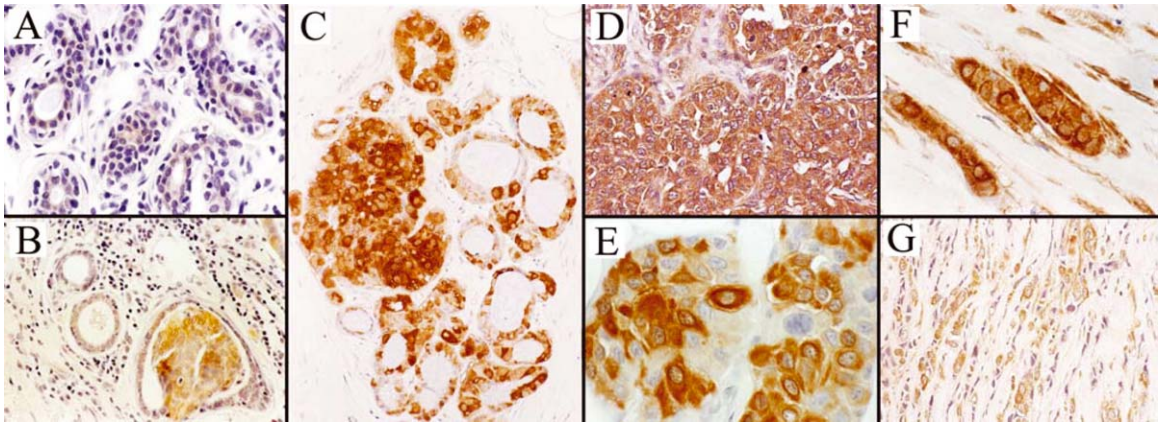


**Figure 3.** Immunohistochemical detection of galectin-4 in normal human breast (panel A), benign breast disease (panel B), ductal carcinoma *in situ* (DCIS) (panel C), ductal carcinoma (panels D–F) and lobular carcinoma (panel G). Breast tissues were processed and immunostained as in Figure 2, omitting the antigen retrieval step. Normal reduction mammoplasty with no staining for galectin-4 (panel A), hyperplasia without atypical component with weak to moderate staining for galectin-4 (panel B), ductal carcinoma *in situ* (DCIS) with galectin-4 expression in enlarged individual cells (panel C), infiltrating ductal carcinoma with high galectin-4 in cancer cells but not surrounding cells (panel D–F), infiltrating lobular carcinoma with galectin-4 expression in interspersed cancer cells (panel G). Original magnification: A: 400×; B: 200×; C: 400; D: 400×; E: 600×; F: 600×; G: 200×.

Galectin-4 expression patterns indicated during malignant transformation of the breast were very different from the patterns found in colon. Tissues from reduction mammoplasties without cytopathological abnormalities and normal tissues surrounding the malignant component, in most cases showed minimal or no galectin-4 expression (Figure 3A) [37]. Weak induction of galectin-4 was clearly associated with epithelial hyperproliferation (Figure 3B), and the atypical component in benign breast biopsies often showed very high intracellular expression. The highest levels of galectin-4 expression were found in the ductal carcinoma *in situ* cases (DCIS, Figure 3C) and in a subset of infiltrating ductal carcinomas (Figure D–E). The high intracellular galectin-4 in breast cancer cells was diffusely cytosolic as in colon cancer. However, the nuclear localization of galectin-4 was found much more frequently in the breast tissues as compared with the colon tissues.

To explore galectin-4 protein expression in cultured cancer cells and to identify a suitable cell culture system for functional studies, galectin profiles of several human cancer cell lines were generated. Galectins were isolated from lysates of metabolically labeled cells by affinity chromatography on lactosyl-Sepharose and separated by SDS-PAGE (Figure 2) [8,42]. This study revealed, that galectin-4 was only expressed by highly differentiated cell lines, which form polarized monolayers and are capable of developing high transepithelial resistance: colon adenocarcinoma T84, lung adenocarcinoma Calu-3, and ovarian adenocarcinoma Mls. The presence of galectin-4 in T84 cells might reflect the fact that this galectin is expressed in normal colon epithelium. However, the other two cell lines were derived from cancers arising in lung and ovary, where expression of galectin-4 has not been previously reported. As mentioned above, the galectins-3 and -4, occupy distinctly different subcellular and membrane-associated domains in T84 cells [8].

The poorly differentiated epithelial cell lines derived from ovarian and lung cancers, did not form high transepithelial resistance when grown on filters, but instead had an increased ability to migrate through a model basement membrane [42]. These cells lacked galectin-4, and instead expressed high levels of galectin-1 protein. Galectin-3 appeared to be equally expressed in all cell lines tested. Poorly differentiated epithelial cancer cell lines often show multiple features of mesenchymal cells. Galectin-1 is absent in normal epithelia, but is often found in normal cells and tissues of mesenchymal origin [21], and therefore its presence in epithelial cancer cells might be symptomatic of advanced stage of malignant transformation.

**Human galectin-4 gene regulation**

Mammalian genes encoding galectins are named *LGALS* (lectin, galactoside-binding, soluble), and numbered consistently with the proteins [43,44]. Murine *Lgals4* and *Lgals6* were mapped cytogenetically to a site near ApoE on chromosome 7 [45], which is syntenic with the q13.1–2 region of human chromosome 19, where the human galectin-4 gene, *LGALS4*, was later found during sequencing of the genome. In the current release of the mouse genome DNA sequence, however, *Lgals4/6* like genes are found both on chromosome 7 and 3, suggesting that further analysis is necessary to ascertain their location.

To identify the transcriptional start site(s) and the upstream regulatory elements in the human galectin-4 encoding gene *LGALS4*, mRNA from the human colon adenocarcinoma T-84 cell was used in primer extension and RNase protection assays [46,47]. The main transcriptional start site was found at position –55 nt, which is 33 bases downstream from a near consensus TATA box, and therefore it appears that in T-84 cells, a TATA promoter primarily regulates galectin-4 gene expression. Low levels of transcription may also occur from the sites at –65 and −58 nt, and from another promoter beyond –184 nt upstream of the TATA box.

Analysis of about 1.5 Kb of 5′ non-coding sequence upstream of the transcription start sites in *LGALS4* revealed the presence of putative binding sites for a number of transcription factors associated with epithelial development, differentiation, and malignant transformation (Figure 4), including HNF-4, MyoD, c-Rel, HNF-3$\beta$, CAAT enhancer binding protein (C/EBP) and HFH-2.

HNF-4, HNF-3$\beta$ and HFH-2 are members of the Hepatocyte Nuclear Factor 3 (HNF-3)/fork family of transcription factors important for liver-specific gene expression as well as in epithelial cell type specific gene expression in adult tissues derived from gut endoderm [48,49]. They could, therefore, contribute to the neoplastic transformation-related increases in galectin-4 mRNA expression in liver [35]. Galectin-4 mRNA levels were substantially higher in hepatocellular carcinoma tissues as compared with matched non-tumorous liver tissues in the same patients. In the HuH-7 and HepG2 cell lines derived from hepatocellular carcinoma, the expression of galectin-4 mRNA was dependent on the cell density and serum concentration: undetectable in rapidly proliferating and subconfluent cells grown in 10% serum, high in dense and overconfluent cultures grown in

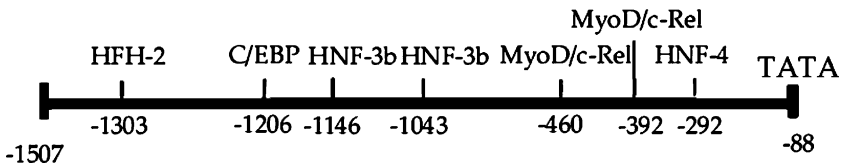


**Figure 4.** Schematic of the human galectin-4 gene (*LGALS4*) upstream regulatory elements [46,47].

10% serum, and induced to high levels in cultures grown at 0.1% serum [35]. These hepatocellular neoplasia-related changes in galectin-4 mRNA expression are consistent with our observations on increases in galectin-4 protein expression levels during malignant transformation of breast, colon and ovarian tissues, described in the previous section.

CAAT enhancer binding proteins, C/EBPs, have been found highly expressed in differentiated cells of the liver, gut epithelial cells, and during differentiation of keratinocytes and squamous epithelium [50,51]. One of them, MyoD, has been implicated in the regulation of gene expression in proliferating and differentiating epithelial cells as well as other tissues [52].

Constitutive expression of galectin-4 in gastrointestinal mucosa and the presence of two binding sites for c-Rel, a subunit of NF$\kappa$-B, in the promoter region of galectin-4 suggests that galectin-4 might be involved in NF$\kappa$-B -mediated gastrointestinal inflammatory responses. Activated during microbial infection NF$\kappa$-B induces multiple target genes which play a role in the host's resistance to pathogenic microorganisms and in host gastrointestinal mucosal defense [53,54]. It is therefore conceivable that in response to microbial infection, galectin-4 is further overexpressed and secreted to interact with microorganisms as suggested for other galectins [55–57].

Aberrant constitutive activation of NF$\kappa$-B and high expression levels of nuclear c-Rel have been recently reported in human breast cancers, as well as in other solid and hematopoietic malignancies [58–60]. To study the role of c-Rel factor in breast tumorigenesis, mice transgenic for murine c-Rel were generated, in which overexpression of c-Rel was driven by the hormone-responsive mouse mammary tumor virus long terminal repeat (MMTV-LTR) promoter [61]. The expression of c-Rel in the mammary gland of these animals was increased, and the animals eventually developed one or more mammary tumors. The tumors expressed increased levels of nuclear NF$\kappa$-B, and aberrant expression of multiple NF$\kappa$-B subunits c-Rel, p50, p52, RelA and RelB was observed. Increased expression of NF$\kappa$-B target genes cyclin D1, c-myc and bcl-xl in these tumors was also found. The presence of c-Rel binding sites in the promoter region of *LGALS4*, and induction/overexpression of galectin-4 in human breast, lung, colon and ovarian malignancies, suggest the interesting possibility that galectin-4 is a downstream component of the c-Rel and NF$\kappa$-B-driven tumor promoting biology.

The upstream sequence of mouse *Lgals6* was also analyzed and transcription start sites identified by primer extension [45]. Two TATA–box-containing promoters were indicated, together with a number of potential binding sites for transcription factors. Notable among these were 6E-boxes from which one was similar to MycMax and the others to MyoD binding sites. There was also present a 19 bp stretch implicated in intestine-specific expression. The proximal TATA box corresponds to the one at −88 in human *LGALS4* with similar surrounding sequence including one downstream E-box. Otherwise the major upstream similarity was in a sequence between −290 and −400 in the *LGALS4*, which includes one conserved E-box.

## Possible function of galectin-4 in normal and cancer tissues

The data presented above support a number of possible functions for galectin-4. The localization and relative insolubility in normal epithelia suggest roles in stabilization of cellular junctions and membranes. This is also supported by the localization of galectin-4 near the junctional complex in EDTA treated cultured T84 cells [8]. The association with detergent insoluble fractions from microvillar membranes, suggests a function in stabilizing certain lipid rafts. Similarly to other galectins, galectin-4 can mediate cell adhesion [8], and induce intracellular signaling (oxidative burst in neutrophils, Almquist, Karlsson, Leffler, unpublished), but the latter have not been explored in great detail.

Effects of intracellular galectin-4 on cell growth are likely since this protein is expressed mainly intracellularly with cytosolic or both cytosolic and nuclear distribution in cancer cells (see above), and it has been shown that galectin-3 and other galectins have such effects. To model the induction of galectin-4 in differentiated epithelial cells, seen in early stages of malignant transformation, we used MDCK cells which form well differentiated epithelial monolayers, and express galectin-3 but not galectin-4 [8,62]. Wild type and mock transfected MDCK cells were unable to grow in serum free medium and became morphologically apoptotic after 7–10 days, whereas MDCK cells transfected with full-length human galectin-4 cDNA and expressing this protein, survived in serum-free medium for 7–8 weeks (Figure 5). Most interestingly, these cells started to proliferate within 4–6 hours after the addition of 10% serum-containing medium. Thus, upon expression of galectin-4, epithelial cells acquired a phenotype characterized by the ability to survive lack of nutrients and growth factors for the prolonged period of time. Such a cellular phenotype is likely to be advantageous in hyperplastic tissues of pre-malignant and malignant tumors.

## Galectin-4 as a diagnostic marker and target for anti-cancer drug development

Many studies have explored the possible use of galectins—in particular galectin-3—as diagnostic cancer markers. Although some promising results have been reported, the picture is often confounded by the rather high expression of galectin-3 in many tissues, and its induction also in inflammatory conditions [9,31,32]. The much more restricted distribution of galectin-4 in normal tissues, but clear induction in early stages of breast and other cancers, therefore, makes it a particularly promising diagnostic and prognostic marker [29,37]. Galectin-4 could be detected using immunohistochemistry as in our ongoing studies, or by tagged ligands for *e.g. in vivo* imaging. In the latter case, the different and high specificities of the two CRDs of galectin-4 described above, suggest that heterodivalent ligands

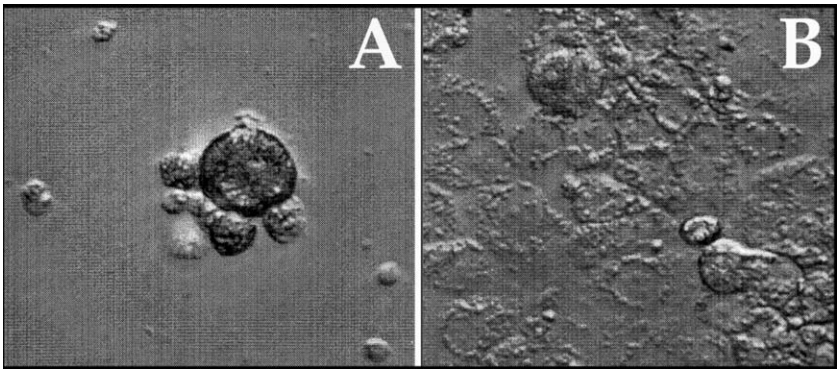


**Figure 5.** Growth of mock-transfected (panel A) and human full-length galectin-4 cDNA-transfected (panel B) MDCK cells. Mock transfected MDCK cells (panel A) stopped growing in medium without serum and only single cells were seen after 7–10 days (wild type MDCK cells display identical behavior (not shown), whereas galectin-4-transfected MDCK cells (panel B) survided in serum free medium for 7–8 weeks. Images were taken by phase contrast/Nomarski optics. Original magnification: 600×.

targeting both domains should provide highly specific imaging probes.

If galectin-4 is not only a marker of early malignancy but also promotes cancer growth as suggested by the functional studies described above, inhibiting its activity should be an important component of anti-cancer therapies. Therapeutic approaches using saccharide derivatives that inhibit galectins *in vitro* have begun. The first galectin inhibitor used in clinical setting was modified citrus pectin, a large and complex polysaccharide rich in anhydrogalacturonic acid, galactose and arabinose, developed by Raz and colleagues [31]. This non-toxic compound has shown strong anti-metastatic activities in rat prostate cancer model [63], and is now in dose-escalation clinical trials (GBC-590, Phase I/II, Glycogenesys, Boston MA). The first trial results presented during the ASCO 2001 meeting reported no toxicity in either colorectal or pancreatic cancer patients (Springgate CF *et al.*, Abstr. 2226, ASCO 2001; Stuart KE *et al.*, Abstr. 2312, ASCO 2001). Glycosylamines have also been shown to inhibit cancer growth in animal models [64,65], and more recently remarkable effects of lactulose and a divalent lactulose amine have been demonstrated [29]. However, in neither of the above therapeutic applications were the substances very specific and it is uncertain which galectin—or galectins—they bind and inhibit *in vivo*. To specifically address tumor-promoting activity of galectin-3, a dominant negative peptide inhibitor of galectin-3, which contains galectin-3 CRD, has been generated. This recombinant protein inhibited primary tumor growth and metastasis in a mouse model of human breast cancer [66]. However, although this peptide was designed as a specific dominant negative inhibitor of galectin-3, it is not certain, whether *in vivo* this was its only function. All the above results lead to a general conclusion that by blocking galectin activities it is possible to reduce tumor growth and metastasis. However, because of the multi-functional nature of the galectin family members, individual galectin-specific inhibitors are needed to address the activities of these proteins in the patho-biological complexities of tumor-bearing organisms.

Attempts at constructing more specific and more potent monovalent or multivalent galectin inhibitors are ongoing [27,28,67]. These projects are encouraged by very promising *in vivo* effects, and by the apparent lack of toxicity of tested inhibitors. Galectin-4 null mutant mice are not yet available, however, galectin-1, galectin-3 and galectin-1/-3 null mutant mice appear healthy. It will therefore be important to determine whether blocking activities of galectin-4, as well as other galectins overexpressed in human cancers, would deliver important therapeutic effects, with no or minimal adverse effects.

## Acknowledgments

The research was funded by the University of California Breast Cancer Research Program (UC BCRP, Award no. 7IB-0039) and Susan G. Komen Breast Cancer Research Foundation (Grant no. BCTR 0100899) to MEH, and by the Swedish Research Council (Grant 12165) and Swedish Funds for Health-Care System -Academic Science Integration, ALF, to HL.

## References

1 Barondes SH, Cooper DN, Gitt MA, Leffler H, Galectins, Structure and function of a large family of animal lectins, *J Biol Chem* **269**, 20807–10 (1994).
2 Cooper DN, Galectinomics: Finding themes in complexity, *Biochim Biophys Acta* **1572**, 209–31 (2002).
3 Leffler H, Carlsson S, Hedlund M, Qian Y, Poirier F, Introduction to galectins, *Glycoconjugate Journal* **19**, 433–40 (2004).
4 Hirabayashi J, Satoh M, Kasai K, Evidence that Caenorhabditis elegans 32-kDa beta-galactoside-binding protein is homologous to vertebrate beta-galactoside-binding lectins, cDNA cloning and deduced amino acid sequence, *J Biol Chem* **267**, 15485–90 (1992).
5 Shoji H, Nishi N, Hirashima M, Nakamura T, Characterization of the Xenopus galectin family, Three structurally different types as in mammals and regulated expression during embryogenesis, *J Biol Chem* **278**, 12285–93 (2003).

6 Houzelstein D, Goncalves I, Sidhu SS, Drickamer K, Cooper D, Leffler H, Poirier F, Phylogenetic analysis of the galectin family, *Molecular Biology and Evolution* **in press** (2004).

7 Oda Y, Herrmann J, Gitt MA, Turck CW, Burlingame AL, Barondes SH, Leffler H, Soluble lactose-binding lectin from rat intestine with two different carbohydrate-binding domains in the same peptide chain, *J Biol Chem* **268**, 5929–39 (1993).

8 Huflejt ME, Jordan ET, Gitt MA, Barondes SH, Leffler H, Strikingly different localization of galectin-3 and galectin-4 in human colon adenocarcinoma T84 cells, Galectin-4 is localized at sites of cell adhesion, *J Biol Chem* **272**, 14294–303 (1997).

9 Bidon-Wagner N, Le Pennec J-P, Human Galectin-8 isoforms and cancer, *Glycoconjugate Journal* **19**, 557–63 (2004).

10 Hirashima M, Kashio Y, Nishi N, Yamauchi A, Imaizumi T-a, Kageshita T, Saita N, Nakamura T, Galectin-9 in physiological and pathological conditions, *Glycoconjugate Journal* **19**, 593–600 (2004).

11 Leffler H, Masiarz FR, Barondes SH, Soluble lactose-binding vertebrate lectins: A growing family, *Biochemistry* **28**, 9222–9 (1989).

12 Tardy F, Deviller P, Louisot P, Martin A, Purification and characterization of the N-terminal domain of galectin-4 from rat small intestine, *FEBS Lett* **359**, 169–72 (1995).

13 Chiu ML, Jones JC, O'Keefe EJ, Restricted tissue distribution of a 37-kD possible adherens junction protein, *J Cell Biol* **119**, 1689–700 (1992).

14 Chiu ML, Parry DA, Feldman SR, Klapper DG, O'Keefe EJ, An adherens junction protein is a member of the family of lactose-binding lectins, *J Biol Chem* **269**, 31770–6 (1994).

15 Danielsen EM, van Deurs B, Galectin-4 and small intestinal brush border enzymes form clusters, *Mol Biol Cell* **8**, 2241–51 (1997).

16 Wasano K, Hirakawa Y, Rat intestinal galactoside-binding lectin L-36 functions as a structural protein in the superficial squamous cells of the esophageal epithelium, *Cell Tissue Res* **281**, 77–83 (1995).

17 Braccia A, Villani M, Immerdal L, Niels-Christiansen LL, Nystrom BT, Hansen GH, Danielsen EM, Microvillar membrane microdomains exist at physiological temperature, Role of galectin-4 as lipid raft stabilizer revealed by "superrafts", *J Biol Chem* **278**, 15679–84 (2003).

18 Rechreche H, Mallo GV, Montalto G, Dagorn JC, Iovanna JL, Cloning and expression of the mRNA of human galectin-4, an S-type lectin down-regulated in colorectal cancer, *Eur J Biochem* **248**, 225–30 (1997).

19 Jiang W, Puch S, Guo X, Bhavanandan VP, Signature sequences for the galectin-4 subfamily, *IUBMB Life* **48**, 601–5 (1999).

20 Gitt MA, Colnot C, Poirier F, Nani KJ, Barondes SH, Leffler H, Galectin-4 and galectin-6 are two closely related lectins expressed in mouse gastrointestinal tract, *J Biol Chem* **273**, 2954–60 (1998).

21 Poirier F, Roles of galectins *in vivo*, *Biochem Soc Symp* 95–103 (2002).

22 Wu AM, Wu JH, Tsai MS, Liu JH, Andre S, Wasano K, Kaltner H, Gabius HJ, Fine specificity of domain-I of recombinant tandem-repeat-type galectin-4 from rat gastrointestinal tract (G4-N), *Biochem J* **367**, 653–64 (2002).

23 Wasano K, Hirakawa Y, Two domains of rat galectin-4 bind to distinct structures of the intercellular borders of colorectal epithelia, *J Histochem Cytochem* **47**, 75–82 (1999).

24 Sörme P, Kahl-Knutsson B, Huflejt M, Nilsson UJ, Leffler H, Fluorescence polarization as an analytic tool to evaluate galectin-ligand interactions, Manuscript (submitted).

25 Ideo H, Seko A, Ohkura T, Matta KL, Yamashita K, High-affinity binding of recombinant human galectin-4 to SO(3)(−)→3Galbeta1→3GalNAc pyranoside, *Glycobiology* **12**, 199–208 (2002).

26 Kopitz J, Andre S, von Reitzenstein C, Versluis K, Kaltner H, Pieters RJ, Wasano K, Kuwabara I, Liu FT, Cantz M, Heck AJ, Gabius HJ, Homodimeric galectin-7 (p53-induced gene 1) is a negative growth regulator for human neuroblastoma cells, *Oncogene* **22**, 6277–88 (2003).

27 Sorme P, Qian Y, Nyholm PG, Leffler H, Nilsson UJ, Low micromolar inhibitors of galectin-3 based on 3′-derivatization of N-acetyllactosamine, *Chembiochem* **3**, 183–9 (2002).

28 Sörme P, Kahl-Knutsson B, Huflejt M, Nilsson UJ, Leffler H, Common and selective high affinity inhibitors for galectins-1, -3 and -4, Manuscript (submitted).

29 Huflejt ME, Mossine VV, Naidenko O, Jazayeri M, Rogers P, Tinari N, Iacobelli SME, Lustgarten JMC, Synthetic lactulose amines bind tumor-promoting galectins-1 and -4, and inhibit breast cancers in Her-2/neu transgenic mice, *24th Annual San Antonio Breast Cancer Symposium* **abstract** (2001).

30 Grassadonia A, Tinari N, Iurisci I, Piccolo E, Cumashi A, Innominato P, D'Egidio M, Natoli C, Piantelli M, Iacobelli S, 90 K (Mac-2 BP) and galectins in tumor progression and metastasis, *Glycoconjugate Journal* **19**, 551–6 (2004).

31 Takenaka Y, Fukumori T, Raz A, galectin-3 and metastasis, *Glycoconjugate Journal* **19**, 543–9 (2004).

32 van den Brule F, Califice S, Castronovo V, Expression of galectins in cancer: A critical review, *Glycoconjugate Journal* **19**, 537–42 (2004).

33 Wollina U, Graefe T, Feldrappe S, Andre S, Wasano K, Kaltner H, Zick Y, Gabius HJ, Galectin fingerprinting by immuno- and lectin histochemistry in cutaneous lymphoma, *J Cancer Res Clin Oncol* **128**, 103–10 (2002).

34 Lahm H, Andre S, Hoeflich A, Fischer JR, Sordat B, Kaltner H, Wolf E, Gabius HJ, Comprehensive galectin fingerprinting in a panel of 61 human tumor cell lines by RT-PCR and its implications for diagnostic and therapeutic procedures, *J Cancer Res Clin Oncol* **127**, 375–86 (2001).

35 Kondoh N, Wakatsuki T, Ryo A, Hada A, Aihara T, Horiuchi S, Goseki N, Matsubara O, Takenaka K, Shichita M, Tanaka K, Shuda M, Yamamoto M, Identification and characterization of genes associated with human hepatocellular carcinogenesis, *Cancer Res* **59**, 4990–6 (1999).

36 Hippo Y, Yashiro M, Ishii M, Taniguchi H, Tsutsumi S, Hirakawa K, Kodama T, Aburatani H, Differential gene expression profiles of scirrhous gastric cancer cells with high metastatic potential to peritoneum or lymph nodes, *Cancer Res* **61**, 889–95 (2001).

37 Huflejt ME, Geradts J, Elliott ML, Leffler H, Liu F-T, Galectin-4 is induced in human breast tumors and is localized to sites of cell adhesion in cultured cells, *Eighty-Eighth Annual Meeting, American Association for Cancer Research, Proc Am Assoc Canc Res* **38**, 267 (1997).

38 Nagy N, Legendre H, Engels O, Andre S, Kaltner H, Wasano K, Zick Y, Pector JC, Decaestecker C, Gabius HJ, Salmon I, Kiss R, Refined prognostic evaluation in colon carcinoma using

immunohistochemical galectin fingerprinting, *Cancer* **97**, 1849–58 (2003).
39 Scanlan MJ, Chen YT, Williamson B, Gure AO, Stockert E, Gordan JD, Tureci O, Sahin U, Pfreundschuh M, Old LJ, Characterization of human colon cancer antigens recognized by autologous antibodies, *Int J Cancer* **76**, 652–8 (1998).
40 Tureci O, Schmitt H, Fadle N, Pfreundschuh M, Sahin U, Molecular definition of a novel human galectin which is immunogenic in patients with Hodgkin's disease, *J Biol Chem* **272**, 6416–22 (1997).
41 Su ZZ, Lin J, Shen R, Fisher PE, Goldstein NI, Fisher PB, Surface-epitope masking and expression cloning identifies the human prostate carcinoma tumor antigen gene PCTA-1 a member of the galectin gene family, *Proc Natl Acad Sci USA* **93**, 7252–7 (1996).
42 Huflejt ME, Krtolica A, Ludlow JW, Leffler H, Barondes SH, Galectin-4 expression in human adenocarcinomas is correlated with a highly differentiated phenotype, *Keystone Symposium on Molecular & Cellular Biology, Conference on Cancer Cell Invasion and Motility, J Cell Biochem* **19B**, 20 (1995).
43 Mehrabian M, Gitt MA, Sparkes RS, Leffler H, Barondes SH, Lusis AJ, Two members of the S-lac lectin gene family, LGALS1 and LGALS2, reside in close proximity on human chromosome 22q12–q13, *Genomics* **15**, 418–20 (1993).
44 Barondes SH, Castronovo V, Cooper DN, Cummings RD, Drickamer K, Feizi T, Gitt MA, Hirabayashi J, Hughes C, Kasai K *et al.*, Galectins: A family of animal beta-galactoside-binding lectins, *Cell* **76**, 597–8 (1994).
45 Gitt MA, Xia YR, Atchison RE, Lusis AJ, Barondes SH, Leffler H, Sequence, structure, and chromosomal mapping of the mouse Lgals6 gene, encoding galectin-6, *J Biol Chem* **273**, 2961–70 (1998).
46 Ponce AM, *Analysis of Transcriptional Expression and Initiation of Galectin-4, a -galactoside Binding Lectin*, M Sci Thesis; University of California, San Diego (1999).
47 Ponce AM, Huflejt ME, Transcriptional regulation of human galectin-4, *Manuscript in Preparation*.
48 Ye H, Kelly TF, Samadani U, Lim L, Rubio S, Overdier DG, Roebuck KA, Costa RH, Hepatocyte nuclear factor 3/fork head homolog 11 is expressed in proliferating epithelial and mesenchymal cells of embryonic and adult tissues, *Mol Cell Biol* **17**, 1626–41 (1997).
49 Ogura K, Choudhuri S, Klaassen CD, Full-length cDNA cloning and genomic organization of the mouse liver-specific organic anion transporter-1 (lst-1), *Biochem Biophys Res Commun* **272**, 563–70 (2000).
50 Jin L, Yang GY, Auborn K, Differences in C/EBPs in normal tissue and papillomas of the larynx, *Cell Prolif* **31**, 127–38 (1998).
51 Birkenmeier EH, Gwynn B, Howard S, Jerry J, Gordon JI, Landschulz WH, McKnight SL, Tissue-specific expression, developmental regulation, and genetic mapping of the gene encoding CCAAT/enhancer binding protein, *Genes Dev* **3**, 1146–56 (1989).
52 Hurlin PJ, Foley KP, Ayer DE, Eisenman RN, Hanahan D, Arbeit JM, Regulation of Myc and Mad during epidermal differentiation and HPV-associated tumorigenesis, *Oncogene* **11**, 2487–501 (1995).
53 Egan LJ, de Lecea A, Lehrman ED, Myhre GM, Eckmann L, Kagnoff MF, Nuclear factor-kappaB activation promotes restitution of wounded intestinal epithelial monolayers, *Am J Physiol Cell Physiol* **285**, C1028–35 (2003).
54 Kagnoff MF, Eckmann L, Epithelial cells as sensors for microbial infection, *J Clin Invest* **100**, 6–10 (1997).
55 Almkvist J, Karlsson A, Galectins as inflammatory mediators, *Glycoconjugate Journal* **19**, 575–81 (2004).
56 Sato S, Nieminen J, Seeing strangers or announcing "danger": Galectin-3 in two models of innate immunity, *Glycoconjugate Journal* **19**, 583–91 (2004).
57 Mandrell RE, Apicella MA, Lindstedt R, Leffler H, Possible interaction between animal lectins and bacterial carbohydrates, *Methods Enzymol* **236**, 231–54 (1994).
58 Cogswell PC, Guttridge DC, Funkhouser WK, Baldwin Jr, AS, Selective activation of NF-kappa B subunits in human breast cancer: Potential roles for NF-kappa B2/p52 and for Bcl-3, *Oncogene* **19**, 1123–31 (2000).
59 Nakshatri H, Bhat-Nakshatri P, Martin DA, Goulet Jr, RJ, Sledge Jr, GW, Constitutive activation of NF-_B during progression of breast cancer to hormone-independent growth, *Mol Cell Biol* **17**, 3629–39 (1997).
60 Sovak MA, Bellas RE, Kim DW, Zanieski GJ, Rogers AE, Traish AM, Sonenshein GE, Aberrant nuclear factor-_B/Rel expression and the pathogenesis of breast cancer, *J Clin Investig* **100**, 2952–60 (1997).
61 Romieu-Mourez R, Kim DW, Shin SM, Demicco EG, Landesman-Bollag E, Seldin DC, Cardiff RD, Sonenshein GE, Mouse mammary tumor virus c-rel transgenic mice develop mammary tumors, *Mol Cell Biol* **23**, 5738–54 (2003).
62 Huflejt ME, Turck CW, Lindstedt R, Barondes SH, Leffler H, L-29, a soluble lactose-binding lectin, is phosphorylated on serine 6 and serine 12 *in vivo* and by casein kinase I, *J Biol Chem* **268**, 26712–8 (1993).
63 Pienta KJ, Naik H, Akhtar A, Yamazaki K, Replogle TS, Lehr J, Donat TL, Tait L, Hogan V, Raz A, Inhibition of spontaneous metastasis in a rat prostate cancer model by oral administration of modified citrus pectin, *J Natl Cancer Inst* **87**, 348–53 (1995).
64 Glinsky GV, Price JE, Glinsky VV, Mossine VV, Kiriakova G, Metcalf JB, Inhibition of human breast cancer metastasis in nude mice by synthetic glycoamines, *Cancer Res* **56**, 5319–24 (1996).
65 Glinsky VV, Glinsky GV, Glinskii OV, Huxley VH, Turk JR, Mossine VV, Deutscher SL, Pienta KJ, Quinn TP, Intravascular metastatic cancer cell homotypic aggregation at the sites of primary attachment to the endothelium, *Cancer Res* **63**, 3805–11 (2003).
66 John CM, Leffler H, Kahl-Knutsson B, Svensson I, Jarvis GA, Truncated galectin-3 inhibits tumor growth and metastasis in orthotopic nude mouse model of human breast cancer, *Clin Cancer Res* **9**, 2374–83 (2003).
67 Andre S, Pieters RJ, Vrasidas I, Kaltner H, Kuwabara I, Liu FT, Liskamp RM, Gabius HJ, Wedgelike glycodendrimers as inhibitors of binding of mammalian galectins to glycoproteins, lactose maxiclusters, and cell surface glycoconjugates, *Chembiochem* **2**, 822–30 (2001).

Accepted 24 January 2004