Original Article

Murine Homodimeric Adhesion/Growth-Regulatory Galectins-1, -2 and -7: Comparative Profiling of Gene/Promoter Sequences by Database Mining, of Expression by RT-PCR/Immunohistochemistry and of Contact Sites for Carbohydrate Ligands by Computational Chemistry

(crypts / epithelium / lectin / modelling / ovary)

M. LOHR[1], M. LENSCH[1], S. ANDRÉ[1], H. KALTNER[1], H.-C. SIEBERT[1], K. SMETANA JR.[2], F. SINOWATZ[3], H.-J. GABIUS[1]

[1]Institute of Physiological Chemistry, [3]Institute of Veterinary Anatomy, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Munich, Germany
[2]Institute of Anatomy, Charles University in Prague, First Faculty of Medicine, Prague, Czech Republic

Abstract. Following the detection of individual members of the family of galectins it is an obvious challenge to define the extent of functional overlap/divergence among these proteins. As a step to address this issue a comparative profiling has been started in the mouse as a model organism, combining sequence analysis, expression patterns and structural features in the cases of the homodimeric galectins-1, -2 and -7. Close relationship was apparent at the level of global gene organization. Scrutiny of the proximal promoter regions for putative transcription-factor-binding sites by two search algorithms uncovered qualitative and quantitative differences with potential to influence the combinatorial functionality of regulatory sequences. RT-PCR mapping with samples from an array of 17 organs revealed significant differences, separating rather ubiquitous gene expression of galectin-1 from the more restricted individual patterns of galectins-2 and -7. Using specific antisera obtained by affinity depletion including stringent controls to ascertain lack of cross-reactivity these results were corroborated at the level of galectin localization in fixed tissue sections. Nuclear presence was seen in the case of galectin-1. In addition to non-identical expression profiles the mapping of the carbohydrate recognition domains of galectins-1 and -7 by homology modelling and docking of naturally occurring complex tetra- and pentasaccharides disclosed a series of sequence deviations which may underlie disparate affinities for cell surface glycans/glycomimetic peptides. In view of applicability the presented data can serve as useful reference to delineate changes with respect to disease and in genetically engineered models. To enable more general conclusions on the galectin network it is warranted to further pursue this combined approach within this lectin family.

Introduction

Lectins are defined as carbohydrate-binding proteins without enzymatic activity on the ligand, explicitly excluding the immunoglobulins. Such proteins from plants and invertebrates are also called haemagglutinins. They have attained the status of convenient probes for glycoconjugate characterization and of popular cyto- and histochemical tools, which were crucial e.g. for defining the chemical nature of ABH blood group epitopes (Spicer and Schulte, 1992; Danguy et al., 1994; Roth, 1996; Cummings, 1997; Watkins, 1999; Rüdiger and Gabius, 2001). Their application as glycan-selective sensor proteins will visualize structural consequences of the fine-tuned mechanisms of regulation of glycosylation, which involves orchestrating the activities of diverse glycosyltransferases (Brockhausen and Schachter, 1997; Pavelka, 1997). The hereby described aspects of glycomic profiles (the complete pattern of carbohydrate epitopes in cellular glycoconjugates) will depend on the cell type and also on developmental and disease parameters such as malignant transformation (Caselitz, 1987; Spicer and Schulte, 1992; Brinck et al., 1995, 1996, 1998; Reuter and Gabius, 1999; Holíková et al., 2002). These intricacies of glycan chains, which comprise branch-end/core

Received May 15, 2007. Accepted June 1, 2007.

This study was supported by an EC Marie Curie Research Training Network Grant (contract No. MRTN-CT-2005-019561).

Corresponding author: Herbert Kaltner, Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Veterinärstr. 13, D-80539 Munich, Germany. Phone: (+49) 89 2180 3984; e-mail: kaltner@lmu.de

Abbreviations: bp – base pairs, Ig – immunoglobulin, PBS – phosphate-buffered saline.

substitutions, degree of branching and variations in anomeric and linkage-point positioning, inspired investigators to envision more than structural roles for the products of glycosylation, and – indeed – when attached to inert carriers, glycans were capable to track down specific binding sites in cells and tissues (Gabius et al., 1993; Danguy et al., 1995; Lee and Lee, 1997; Gabius, 2001). These results are strong evidence for an operative information transfer via protein-carbohydrate interaction. It involves endogenous lectins and embodies an important aspect of the concept of the sugar code with carbohydrates as high-density coding units (Gabius, 2000; Solís et al., 2001; Gabius et al., 2002, 2004). The endogenous lectins will then serve as versatile effectors to elicit cellular processes as diverse as adhesion, growth control, mediator release and cell activation, migration and routing as well as tissue invasion and to contribute to maintain immune homeostasis (Gabius, 1987, 1997, 2006; Buzás et al., 2006; Villalobo et al., 2006). Toward this end, distinct glycan targets that convey the relevant information are selected by the fitting endogenous lectins. They will "read" the respective signals, not responding to the presence of a wide variety of other epitopes. Glycan diversity, it turns out, is matched by the emergence of a variety of lectin families, and this interplay at the level of complementarity between carbohydrate determinant and carbohydrate recognition domain with its functional consequences explains the interest to study the presence of animal lectins systematically.

Among the lectin families a particularly wide spectrum of cell functions has so far been uncovered for the galectins. They home in on β-galactosides present in glycoconjugates, with branch-end and core substitutions as well as local glycan density affecting the affinity (André et al., 1997, 2004a, b, 2006a; Unverzagt et al., 2002; Wu et al., 2002, 2004, 2006; Siebert et al., 2003). Of note, cell surface binding can be restricted to only few glycans. Fibronectin, laminin and carcinoembryonic antigen as well as integrins such as the fibronectin receptor of carcinoma cells have received attention in this respect (Ohannesian et al., 1995; André et al., 1999; Fischer et al., 2005). The inhibition of neuroblastoma cell growth depends on the pentasaccharide of ganglioside GM_1 as key contact point for galectins, and functional divergence among galectins is in this case likely due to differences in propensity for cross-linking surface ligands (Kopitz et al., 1998, 2001, 2003; Ahmad et al., 2004; André et al., 2005a). The dissection of induction of apoptosis in activated T-cells by galectins-1, -2 and -7 resulted in the following conclusion: functional divergence is due to differences in glycan selection underlying the characteristic profiles of caspase activation (Sturm et al., 2004). Beyond such productive protein-carbohydrate interactions galectins are also engaged in equally specific recognition of peptide motifs. They render members of this lectin family capable to exert intracellular functions, e.g. relevant for growth control by associating with Bcl-2, β-catenin or oncogenic H-ras (Liu et al., 2002; Rotblat et al., 2004; Shimura et al., 2005; Smetana et al., 2006; André et al., 2007). As a consequence of sequence variation among galectins, distinct properties can well be established due to an impact on these two types of interactions, then setting members of the family apart in certain aspects of function. In this case, the homologous proteins will not simply share redundant functional profiles. However, the aspect to define individual features within the galectin network is not yet well resolved at different levels. Addressing this issue clearly justifies respective investigations, especially with focus on this protein family in a selected organism.

Toward this end, here we start a comparative profiling of the three groups of galectins in the mouse. This initial part comprises the proto-type homodimeric family members galectin-1, -2 and -7. Although the three corresponding human proteins crystallographically share the same β-sandwich folding (Lobsanov et al., 1993; Leonidas et al., 1998; López-Lucendo et al., 2004), each protein is known to harbour unique features, as already mentioned above for induction of T-cell apoptosis. These inter-galectin differences are further highlighted by two cases as follows: galectin-2 is involved in lymphotoxin-α secretion with implications for myocardial infarction and galectin-7 is referred to as p53-induced gene 1 based on the conspicuous upregulation of gene expression before the onset of apoptosis of DLD-1 colon carcinoma cells triggered by the tumour suppressor (Polyak et al., 1997; Ozaki et al., 2004; Ozaki and Tanaka, 2005). In our stepwise approach, we first document the gene structures of galectins-1, -2 and -7 including close scrutiny of the proximal promoter regions by two different search algorithms for putative transcription-factor-binding sites. We next report the gene expression profiles by RT-PCR and proceed to present a map of protein presence by immunohistochemistry. Rigorous controls ensured that the applied antibodies were free of cross-reactivity. Finally, we compile the amino acid sequences, generate structural models of the galectins, as far as was possible, and computationally map key contact sites with low-energy conformations of two natural glycans to help spot sequence deviations potentially relevant to modulate fine-specificity of the carbohydrate recognition domains. The availability of this detailed information in a common model organism will establish a reference for several purposes: a.) to define inter-galectin differences in global expression and intra-/extracellular presence, in the course of development and in diseases including malignancy, b.) to turn attention to distinct organs for study of knockout models, c.) to provide a guideline for selection of negative/positive controls, d.) to help relate the presence

of distinct sequence motifs in the proximal promoter regions uncovered by *in silico* monitoring to gene activity patterns, and e.) to track down functionally relevant sequence deviations. In addition to its use as reference a comparative data collection revealing non-uniform properties gives further research on dissecting the galectin network a clear direction.

Material and Methods

Sequence analysis

Information about the nucleic acid sequences of the three galectin genes, the putative transcription start sites and the genomic sequences up- and downstream of these sites were downloaded from the respective GenBank entries available at the database of the National Center for Biotechnology Information (NCBI, Bethesda, MD; http://www.ncbi.nlm.nih.gov/Genbank/index.html). The amino acid sequences of murine galectins-1 (P16045), -2 (Q9CQW5) and -7 (O54974) were obtained from the Swiss-Prot protein sequence database (http://www.expasy.ch/sprot/). DNA and amino acid sequences were edited using the EditSeq sequence analysis software version 4.0 (DNAstar Inc., Madison, WI) as well as the Reformat and Map algorithms included in the GCG (Genetics Computer Group Inc. Sequence Analysis Software Package) programs available from the HUSAR biocomputing service of the German Cancer Research Center (Heidelberg, Germany; http://genius.embnet.dkfz-heidelberg.de/). Amino acid alignments were performed using the program Multalin version 5.4.1 (Corpet, 1988) (http://bioinfo.genopole-toulouse.prd.fr/multalin/multalin.html). Homology scores were calculated via the program EMBOSS (with the "needle" algorithm for global pairwise alignments), available on the homepage of the European Bioinformatics Institute (Wellcome Trust Genome Campus, Cambridge, UK; http://www.ebi.ac.uk/emboss/align/). To predict putative transcription-factor-binding sites in the proximal promoter regions, the online versions of the two programs Match™ (Kel et al., 2003) and P-Match™ (Chekmenev et al., 2005) were used, both available on the website of the BIOBASE Biologische Datenbanken GmbH (Wolfenbüttel, Germany; http://www.gene-regulation.com/pub/programs.html). In contrast to other web-based resources, these two programs process data input of the latest update of the TRANSFAC® transcription factor database (accessible on the same site) as a source for weight matrices and consensus sequences. On this basis the employed algorithms calculate the probability with which each transcription factor listed in the database will bind a distinct part of a given DNA sequence (on any of the two strands) and present the results as so-called "scores". The binding motifs are further classified into entire matrix and core sequences, both of which receive scores. Presettings were chosen to include "low-quality" matrices (matrices that yield a higher than average number of false-positive hits, if applied to intron 3 sequences) in order to detect sites for transcription factors such as Sp1 for which only "low-quality" weight matrices exist in the database. To counterbalance this selection of a parameter with low level of stringency the option to customize the "cut-off", i.e. the minimal value a score must reach to be considered as a significant hit, was used, and the cut-off was set to the most restricted value in order to minimize the number of false-positive hits.

RT-PCR analysis of murine galectins-1, -2 and -7

Total RNA was isolated from organs of the mice strains NMRI (outbred strain) and C57BL/6 (inbred strain) using RNAeasy mini columns (Qiagen, Hilden, Germany) as described previously (Saal et al., 2005). Briefly, the frozen organ samples were weighed, placed in sample tubes, 600 µl RLT lysis buffer (containing guanidine isothiocyanate) were added per 30 mg tissue, and the tissue specimens were immediately homogenized for 10-20 s with the Ultra-Turrax® system TP18/10 (IKA-Werke, Staufen, Germany). Subsequently, the detailed instructions of the manufacturer's protocol were followed. Ethanol was added to the lysate, creating conditions that promote selective binding of RNA to the silica-based membrane of the mini-spin column. The sample was then applied to the spin column and total RNA bound to the membrane, whereas contaminants were effectively washed away. Total RNA was subsequently eluted in a small volume of RNase-free water. The level of integrity of the purified RNA was controlled in agarose gel electrophoresis. Prior to cDNA production total RNA preparations were incubated for 30 min at 37°C with RNase-free DNase I (Roche Diagnostics, Mannheim, Germany) to digest any contaminating genomic DNA. The reaction was stopped by incubation for 10 min at 75°C. As a template for first-strand synthesis of cDNA 2.5 µg of total RNA were used. Reverse transcription was performed for 50 min at 42°C in a reaction volume of 20 µl containing first-strand synthesis buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl and 15 mM $MgCl_2$; provided by the supplier), dNTPs (1 mM each), oligo$(dT)_{12\text{-}18}$ primer (0.5 µg), 200 units Superscript™ II RNase H^- Reverse Transcriptase (Invitrogen, Karlsruhe, Germany) and diethylpyrocarbonate-treated water. The reaction was stopped by raising the incubation temperature to 70°C for 15 min, and an aliquot of 2 µl of each reaction mixture was used in the following PCR amplification protocol, already applied for mapping expression of human galectin genes in tumour cell lines (Lahm et al., 2001). To produce specific fragments of defined size from the galectin cDNAs, suitable oligonucleotide primers were designed using gene sequences in GenBank accessible under Nos. X15896, AK007364 and AF038652. PCR amplification of mouse

galectin-1 cDNA was directed by the sense primer 5′-ATCATGGCCTGTGGTCTGGTC-3′ and the antisense primer 5′-TTCACTCAAAGGCCACGCACT-3′, of mouse galectin-2 by the sense primer 5′-GTCACCATGTCGGAGAAATTTG-3′ and the antisense primer 5′-CTAAGGTCTTCTGAGGTGCCG-3′ and of mouse galectin-7 by the sense primer 5′-GCCCG-CCATGTCTGCTACCCA-3′ and the antisense primer 5′-CTTGGGTCCTTTGCTTAGAAG-3′. The expected fragment lengths were 410 base pairs (bp) for mouse galectin-1, 420 bp for mouse galectin-2 and 432 bp for mouse galectin-7. Quality of cDNA preparations was confirmed by PCR amplification with the β-actin-specific primers 5′-GGCATCGTGATGGACTCCG-3′ (β-actin sense) and 5′-GTCGGAAGGTGGACAGCGA-3′ (β-actin antisense) as controls. The reaction mix for the PCR reaction was prepared as recommended by the distributor of the *Taq* DNA polymerase (Qiagen). In general, the analyses were carried out in a volume of 20 µl containing 2 µl cDNA, 0.5 U *Taq* polymerase, 50 µM dNTPs, 1 × PCR-buffer (commercial mixture of Tris-HCl, KCl, $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, pH 8.7 at 20°C), 1 × Q-Solution and 0.1 µM of sense and antisense primers. Amplification of the galectin-specific cDNAs was started with an initial denaturation step at 94°C for 4 min, then 35 consecutive cycles of the following series of steps were performed: denaturation at 94°C for 45 s, annealing at 60°C for 45 s and extension at 72°C for 1 min. The final extension step was carried out at 72°C for 10 min. PCR products and reagents were separated in 2% agarose gel electrophoresis at 80 V for 40 min.

Preparation of galectins and their polyclonal antibodies including quality controls

Recombinant production of human galectins and their purification by affinity chromatography on lactosylated Sepharose 4B, obtained by divinyl sulphone activation, as a crucial step resulted in material suited for immunization, the purity carefully ascertained by one- and two-dimensional gel electrophoresis, gel filtration and nano-electrospray ionization mass spectrometry and the activity determined by haemagglutination as well as solid-phase and cell-binding assays (Gabius, 1990; André et al., 1999, 2005b, 2006b; Kopitz et al., 2003; Purkrábková et al., 2003; Chovanec et al., 2004). Polyclonal antibodies were raised in rabbits using repeated booster injections under the control of ELISA assays to monitor the course of the titre, and the immunoglobulin G (IgG) fractions were purified from serum by affinity chromatography on protein-A Sepharose 4B resin (Amersham Biotech, Freiburg i. Br., Germany) (Kaltner et al., 1997, 2002). These preparations were routinely subjected to specificity controls by ELISA and Western blots using these three galectins and members of the chimera- and tandem-repeat-type groups to reduce risk of cross-reactivity and ensure specific interspecies reactivity. In cases of positivity, e.g. anti-galectin-2 binding to galectins-1 and -3, chromatographic affinity depletion was performed using bead-immobilized galectins, prepared by divinyl sulphone activation to yield protein density of 5–7 mg/ml, as ligand for cross-reactive antibodies (Gabius et al., 1991; Saal et al., 2005; Lensch et al., 2006). Quality controls by ELISA and Western blots reported on successful removal of cross-reactive material.

Western blot analysis of tissue extracts

Tissue samples from murine lung, stomach and urinary bladder were routinely homogenized in extraction buffer (20 mM phosphate-buffered saline (PBS), pH 7.2, containing 1% Triton X-100, 0.1% sodium deoxycholate, 0.05 M lactose, 2 mM dithiothreitol, 2 mM ethylene diamine tetraacetic acid, and the protease inhibitors pefabloc® at 1 mM (aprotinin at 2 µg/ml and leupeptin at 5 µg/ml). Protein content was routinely quantitated using the Bradford assay with bovine serum albumin (BSA) as a standard (Bio-Rad, München, Germany). Western blotting followed an optimized procedure (Gabius et al., 1986; Kaltner et al., 2002), first separating a total of 50 µg extracted protein on a discontinuous SDS polyacrylamide gel (4% stacking gel and 14.5% running gel), followed by electrophoretic transfer of proteins to a nitrocellulose membrane. Residual binding sites on the membrane were then blocked (5% dry milk and 0.5% Tween 20 in Tris-buffered saline, pH 7.5), and membranes were routinely incubated overnight at 4 °C with an IgG-containing solution at a concentration of 1 µg/ml, the antibody always freshly added to the blocking solution. In the next step, free IgG was completely removed by three washing steps with Tris-buffered saline, pH 7.5, containing 0.05% Tween 20, the membranes were then placed in a solution containing the indicator conjugate, i.e. horseradish-peroxidase-labelled goat anti-rabbit IgG (0.5µg/ml; Sigma-Aldrich, München, Germany) for 1 h at room temperature. Finally, membranes were placed in a darkroom in a mixture of 2 ml 0.1 M Tris-HCl, pH 8.6, and 1.25 mM luminol sodium salt, 0.2 ml of a solution of 6.7 mM p-coumaric acid in dimethyl sulphoxide and 0.6 µl H_2O_2 (30% v/v) for 2 min at room temperature. Exposition periods with CL-X Posure™ X-ray film (Pierce, Bonn, Germany) for signal visualization and subsequent film development were set to yield optimal signal intensity and minimal background.

Immunohistochemical processing

Organs were immediately fixed by immersion in either methanol with 30% acetic acid, buffered paraformaldehyde (4% w/v, pH 7.2) or Bouin's solution for 24 h at 4°C. Tissue specimens were then dehydrated using an ascending series of ethanol (70%, 80%, 96% and 100%), embedded in paraffin wax at 61°C, cut into serial sections of 3.5 µm in thickness, and these sections

were mounted on Superfrost slides (Menzel, Braunschweig, Germany). A selection was processed immunohistochemically to determine optimal conditions. Routine processing was performed with sections from specimen of fixation with Bouin's solution, as described (Kaltner et al., 2002; Manning et al., 2004; Siebert et al., 2006). In brief, the sections were routinely deparaffinized, hydrated, incubated with 1% H_2O_2 in distilled water to quench endogenous peroxidase activity, washed with 10 mM PBS (pH 7.2–7.4) and treated with 1% (w/v) solution of BSA (Sigma-Aldrich) in PBS with 5% (v/v) normal swine serum (DakoCytomation GmbH, Hamburg, Germany) to saturate non-specific protein-binding sites, hereby contributing to exclusion of antigen-independent staining. Sections were then incubated in a humid chamber overnight at 4°C with a solution containing one of the IgG fractions raised against galectin-1 (2 µg/ml), galectin-2 (5 µg/ml) or galectin-7 (1 µg/ml), respectively, dissolved in PBS with 1% BSA. Following an incubation period of 16 h the sections were carefully rinsed in PBS and then incubated at room temperature with a solution of biotinylated secondary antibody (swine anti-rabbit Ig, 1.25 µg/ml; DakoCytomation GmbH). After 1 h the sections were thoroughly rinsed again in PBS to remove unbound protein and then incubated at room temperature with Vectastain® Elite ABC Kit PK-6100 reagent for peroxidase or with Vectastain® ABC Kit AK-5000 for alkaline phosphatase (Axxora, Grünberg, Germany). Localization profiles of the three galectins via the marker conjugate were visualized using the Vector® DAB Substrate Kit SK-4100 for peroxidase (Axxora), applied to the sections under visual control with a maximum period of time of 10 min, or the Vector® Red alkaline phosphatase Substrate Kit SK-5100 for 20–30 min in the dark. Later on, some sections were counterstained with Mayer's haemalaun. Following dehydration sections were finally mounted in Eukitt® (Kindler, Freiburg, Germany). Negative controls included the use of pre-immune rabbit serum which replaced the galectin-specific IgG fraction to exclude staining by antigen-independent protein-protein interactions such as Fc-dependent binding of any antibody and the omission of the primary antibody in the first incubation step of routine processing to ascertain antigen-dependent staining, hereby excluding e.g. carbohydrate-mediated binding of the glycoprotein horseradish peroxidase to mannose-specific receptors in tissues (Kuchler et al., 1990). Photographs were taken with a Leica Microscope (type 020-525.719) equipped with a digital camera (Leica DFC 320; Leica Camera, Solms, Germany).

Molecular modelling of galectin-ligand complexes

Based on crystallographic information for the corresponding human galectins (PDB codes 1GZW, 1HLC and 1BKZ) homology modelling was performed using previously described and internally validated tools and settings (André et al., 2005a). Ligands were docked into the contact sites using either the topology of the bound-state conformation of the pentasaccharide of ganglioside GM_1 to human galectin-1 in solution (Siebert et al., 2003) or the low-energy conformation of the histo-blood group A-tetrasaccharide (Wu et al., 2007), the strictly conserved tryptophan residue, which is indispensable for association to the galactose residue, serving as internal reference (Siebert et al., 1997).

Results

Gene organization and promoter sequences

This stepwise analysis leads from comparing gene structures including putative regulatory elements to expression profiles determined by RT-PCR/immunohistochemistry, finally mapping the amino acid substitutions in the vicinity of the crucial contact points to carbohydrate ligands *in silico*. With the full genome information at hand, the occurrence of single genes was detected for each galectin. The genes for galectins-1 and -2 are both located on chromosome 15, close to each other (approximately 50 kbp apart) but on opposite strands. In contrast, the galectin-7 gene is found on chromosome 7. In the first step of the comparative program the distribution of introns in the genes for the three homodimeric murine galectins was evaluated drawing on publicly accessible sequence information. Whereas the number of introns was identical and their positioning equivalent, the lengths of the intron sequences and 5′/3′-untranslated regions varied, up to 9-fold in the case of the first intron (Fig. 1). The distribution of coding regions in exons 3 and 4 showed a dissimilarity between galectins-1/-2 and galectin-7 which in effect was compensatory (Fig. 1). By aligning the coding sequences of 408 nucleotides (galectin-1), 393 nucleotides (galectin-2) and 411 nucleotides (galectin-7) an identity score of 51.8% was calculated in the galectin-1/-2 comparison. The scores in the galectin-1/-7 and the galectin-2/-7 comparison were 49.3% and 49.9%, respectively. Using triplet-to-triplet matching, the following percentages of deviations between the three sequences were calculated: 74.5% of the coding triplets differ between galectins-1 and -2 (in at least one position), but 26.7% of these alterations are "silent" (cause no change in the amino acid sequence). The respective numbers for the other two comparisons are 87.9% differences, 17.7% of them being "silent" for galectin-1 vs -7 and 80.1 % differences/15% being "silent" for galectin-2 vs -7.

The availability of full genomic sequences next enabled us to systematically sift through the proximal promoter regions (-1000 bp to +100 bp relative to the initiation of transcription). We deliberately applied two independent search algorithms and selected settings to include putative target sequences for common factors,

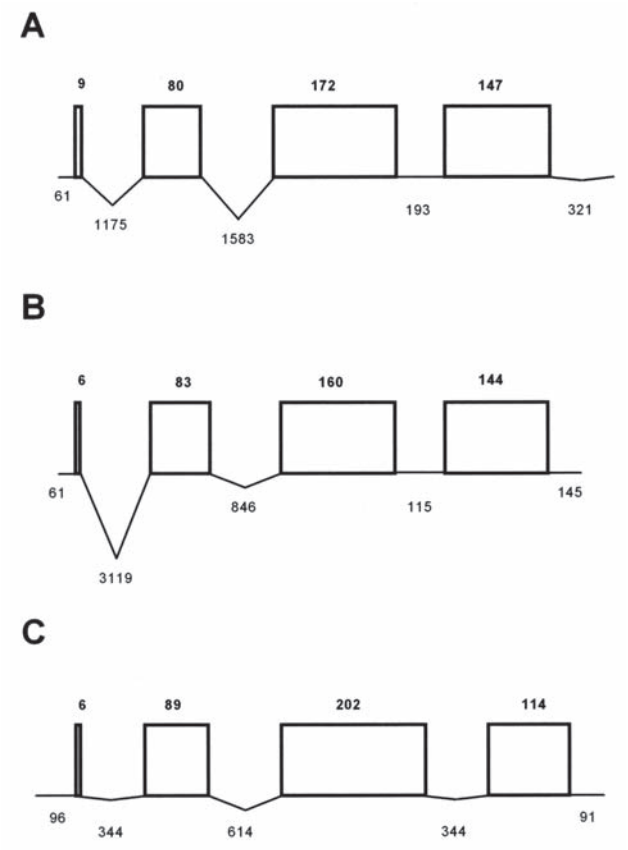

Fig. 1. Organization of the primary transcripts of the genes for murine galectins-1 (**A**), -2 (**B**) and -7 (**C**). In each case, the four exons are given as open boxes connected by lines representing 5'/3'- untranslated sequence stretches and introns, respectively. Box sizes are correlated to exon length (given in numbers of base pairs (bp)), while the lines are not drawn to scale, numbering for each line providing the information on the precise length in bp.

while maintaining stringent selectivity by a restrictive cut-off value. As summarized in Table 1, a series of sites for putative interaction with transcription factors was found. In addition to this overview we epitomize the detected qualitative differences, i.e. presence/absence of a distinct sequence for a putative contact, in Table 2. Besides these qualitative differences, the occurrence of sequence hits also shows quantitative alterations that can be relevant for the combinatorial functionality of regulatory units (Table 1). The developmental regulator En-1, for example, showed up in the sequence list only once for galectin-1 compared to five and three hits for galectins-2 and -7, respectively. A similar variability was found for the frequency of presence of sites that at least one search algorithm assigned to the transcription factor AML-1a (Runx) or the developmental activator HOX A3 (Table 1). On average, this region of the galectin-2 gene is fairly dense in putative sites for the negative regulator in brain and heart AML-1a. The case of galectin-7 presents with the largest number of unique sites for regulation. Evidently, a series of inter-galectin changes satisfying the criteria of the search algorithms were thus detectable in the proximal promoter regions. These results prompted us to proceed to mapping actual gene expression profiles, first at the highest level of sensitivity to avoid false-negative results, i.e. RT-PCR profiling.

Expression profiling by RT-PCR

The sufficiently low degree of homology of these galectin genes noted above enabled us to design fully discriminatory primer sets. Initial validation of technical aspects was carried out with organ samples that were known to produce galectin-1 in several mammals, i.e. lung and heart. Having successfully run the controls, cDNA specimens were prepared with aliquots of the same set of reagents from various organs. Two panels originating from different strains to answer the additional question on inter-strain variability were processed in parallel. After compiling the results of the analysis, routinely obtaining amplification products of the expected length, an abundant presence of the galectin-1 signal was seen (Fig. 2). In comparison, gene transcription for galectins-2 and -7, determined at this level of sensitivity, was less frequent, and the inter-strain comparison yielded rather similar patterns (Fig. 2). Due to this absence of conspicuous differences, further work could thus be focussed on the C57BL/6 strain, and these data are presented in Table 3 along with the data of immunohistochemistry. Because the parameters of amplification were chosen to exclude false-negative results, even spurious presence of mRNA can lead to a positive signal. To assess the level of protein presence and its distribution in an organ, immunohistochemical monitoring is required. Toward this end antibodies against the three lectins were raised, which had to pass stringent quality controls to avoid cross-reactivity among the three proteins and also with other members of the galectin family.

Expression profiling by immunohistochemistry

The three different antibody preparations were routinely tested by ELISA and Western blotting for any cross-reactivity, followed by chromatographic processing to completely remove such contaminating activity in the case of positivity. Besides the panel of the tested galectins we included members of the other two groups of the galectin family to avoid ambiguities in interpretation of staining. A weak reactivity of the IgG fraction against galectin-2 was measured with galectins-1 and -3, as were low-intensity signals obtained with the anti-galectin-7 preparation with galectins-1 and -4. Using the mentioned galectins as ligands in affinity chromatography the IgG preparations were fractionated and carefully retested. The lack of cross-reactivity of these refined preparations when testing the same amounts of the three galectins is illustrated in Fig. 3A, the positive control

Table 1. Compilation of putative transcription-factor-binding sites in the proximal promoter regions of the genes for murine galectins-1, -2 and -7

	mGal-1					mGal-2					mGal-7				
Binding Factor	Pos. (Ori.)	Sequence	Core Score	Matrix Score	**Program**	Pos. (Ori.)	Sequence	Core Score	Matrix Score	**Program**	Pos. (Ori.)	Sequence	Core Score	Matrix Score	**Program**
AML-1a (Runx)	-958 (+)	tGCGGT	1.0	1.0	**PM**	-897 (-)	GCCACa	1.0	1.0	**PM**	-828 (+)	gGTGGT	1.0	1.0	**PM**
	-560 (-)	GCCACa	1.0	1.0	**PM**	-865 (-)	ACCACt	1.0	1.0	**PM**	-786 (-)	ATCACa	1.0	1.0	**PM**
	-513 (+)	tGTGGT	1.0	1.0	**M, PM**	-857 (-)	ACCACc	1.0	1.0	**PM**	-709 (-)	ACCACt	1.0	1.0	**PM**
	-460 (-)	GCCACa	1.0	1.0	**PM**	-677 (+)	gGTGGT	1.0	1.0	**PM**	-458 (-)	ACCACt	1.0	1.0	**PM**
	-259 (-)	GCCACa	1.0	1.0	**PM**	-610 (-)	ACCACt	1.0	1.0	**PM**	+21 (-)	ACCACt	1.0	1.0	**PM**
	-108 (+)	tCTGGT	1.0	1.0	**PM**	-337 (+)	tGTGGT	1.0	1.0	**M, PM**	+65 (-)	GCCACa	1.0	1.0	**PM**
						-277 (+)	tGTGGT	1.0	1.0	**M, PM**					
						-220 (-)	ACCACg	1.0	1.0	**PM**					
						-55 (+)	gGTGGT	1.0	1.0	**PM**					
						+33 (+)	tGTGGC	1.0	1.0	**PM**					
AP-1		not found					not found				-974 (-)	tatgAGTCAcc	1.0	0.993	**M***
Cdx-1	-131 (+)	CTTTAta	0.986/1.0	0.973/1.0	**M, PM**	-459 (+)	aATTATa	0.984	0.987	**M**	-977 (+)	CTTTAtg	0.986	0.987	**M**
	+64 (-)	ctTCAAT	1.0	1.0	**PM**	-69 (+)	aTTTATg/ATTTAtg	1.0	0.994/1.0	**M**	-864 (+)	TATTGga	1.0	1.0	**PM**
											-496 (-)	cATTATt	0.984	0.981	**M**
											-492 (+)	ATTTGac	1.0	1.0	**PM**
C/EBPα, β; CHOP-10	-951 (-)	ttATTTCttaattc	0.996	0.983	**M**	-563 (-)	ctgtaaTCCCAgc	1.0	1.0	**PM**	-638 (-)	tctttggGCAAAtt	0.984	0.967	**M**
	-923 (-)	ttctatCCAAAcc	1.0	0.969	**M**	-320 (-)	ggctcaGCAAAga	0.972	0.965	**M**					
	-577 (+)	agcTGCACtgagc	1.0	0.998	**PM**	-136 (+)	tgTTTGTtgagtg	0.965	0.96	**M**					
COMP1		not found				-723 (-)	agttggttctttcCTTCCatcatg	0.964	0.919	**PM**	-183 (+)	tgaaagGGGAGgtgagaagagcac	0.929	0.905	**PM**
											+4 (+)	ggaaagGGAGGagctgaaccacta	0.929	0.891	**PM**
CUTL-1		not found				-175 (-)	gctcACGATc	1.0	1.0	**PM**		not found			
En-1	-127 (+)	ATAATcc	1.0	1.0	**PM**	-930 (-)	caAGCTC	1.0	1.0	**PM**	-646 (+)	GATCTtg	1.0	1.0	**PM**
						-838 (+)	ATAGTtg	1.0	1.0	**PM**	-467 (-)	ccATTAC	1.0	0.996	**M**
						-336 (+)	GTGGTtg	1.0	1.0	**PM**	+22 (-)	ccACTAC	0.996	0.992	**M**
						-292 (-)	caAGATC	1.0	1.0	**PM**					
						-32 (-)	ggATTAT	1.0	1.0	**PM**					
Evi-1		not found				-626 (+)	ccctAGATGtg	0.988	0.983	**PM***	-845 (-)	gTGTTTcatttgatc	0.782	0.904	**PM**
						-582 (-)	caGATATagtg	0.972	0.982	**PM***	-230 (+)	cactAGATCtc	0.988	0.99	**PM***
						-100 (+)	tgatattagGAGATa	0.826	0.822	**M***	-137 (-)	gTATTAaaggttgtc	0.93	0.924	**PM**
											-120 (-)	taTCTCTaagc	0.988	0.981	**PM***
GATA-3	-326 (-)	ccCTATCcc	1.0	0.995	**M**		not found				-123 (-)	ccCTATCtc	1.0	1.0	**M**
GKLF	-810 (+)	aatgaatggGAAGG	0.933	0.9	**M**	-432 (+)	ggatgagagAAGGG	1.0	0.926	**M**	-704 (+)	tgagaagagAGAGG	0.949	0.9	**M**
											-703 (+)	gagaagagaGAGGG	0.965	0.907	**M**
											-399 (+)	ataggataaAAAGG	0.968	0.911	**M**
											-105 (+)	gaaaaggggAGAGG	0.949	0.92	**M**
											-104 (+)	aaaaggggaGAGGG	0.965	0.932	**M**
											+28 (-)	CCTTGcctttactt	0.918	0.907	**M**
HNF-4α2, COUP-TF		not found				-86 (+)	aaaggtcCAAAGtccttat	1.0	0.912	**M***	-644 (-)	tcttgatCTTTGggcaaat	1.0	0.941	**M***

HOX A3	-820 (+)	ATAACtgct	1.0	1.0	**PM**	-959 (+)	TGCACgttt	1.0	1.0	**PM**	-563 (+)	GGTAAggtt	1.0	1.0	**PM**
	-740 (-)	caccCTGGA	1.0	0.995	**PM**	-796 (+)	TTCAGttat	1.0	0.995	**PM**	-360 (+)	TTGAGttat	1.0	0.995	**PM**
	-313 (-)	actcTTAGG	1.0	0.953/1.0	**M, PM**	-786 (+)	CTTATttgt	0.95	0.96	**M**					
	-96 (-)	catcCTTGT	1.0	0.995	**PM**	-130 (+)	TTGAGtgag	1.0	1.0	**PM**					
Lmo2-complex		not found					not found				-123 (-)	cccTATCTc	1.0	1.0	**PM**
Msx-1		not found				-115 (-)	cATATAcag	0.936	0.956	**M**	-511 (+)	ctgTTAGTg	0.9	0.931	**M**
											-482 (+)	ccgTTACTg	0.9	0.931	**M**
MZF-1	-839 (+)	acagtAGGGGctc	1.0	0.984	**PM**	-742 (+)	agaGGGGA	1.0	1.0	**PM**	-853 (-)	TCCCCaat	1.0	1.0	**PM**
	-385 (-)	atcCCCCTccccc	1.0	0.96	**M**	-14 (+)	agaGGGGA	1.0	1.0	**PM**	-195 (+)	gcggtAGGGGtat	1.0	0.992	**PM**
	-358 (+)	gacctAGGGGctt	1.0	0.983	**PM**						+46 (-)	TCCCCtcc	1.0	1.0	**PM**
	-330 (-)	cagGCCCTatccc	0.987	0.987	**PM**										
	-191 (+)	gagggAGGGGcta	1.0	0.934/0.98	**M, PM**										
	-100 (-)	CCCCCatc	1.0	1.0	**PM**										
Pax-2	-745 (-)	tggcacacccTGGACtgcc	0.98	0.963	**PM**	-356 (-)	cagagccagcCTGTCtttt	0.972	0.962	**PM**	-757 (+)	ctcaGTTCCaaggaggagc	0.963	0.96	**PM**
	-378 (+)	tcccCCACCagctgcgctg	0.992	0.975	**PM**	-310 (+)	agagGTTGGatggcctgcc	0.974	0.962	**PM**	-275 (-)	tttcagaaagGGTGGaggc	0.992	0.964	**PM**
	-301 (-)	aactcatcctTAAACtatt	1.0	0.963	**PM**	-163 (-)	gtcagagtgaGTGAGgagc	0.966	0.966	**PM**	-205 (+)	gtctGACAGggcggtaggg	0.972	0.965	**PM**
	-101 (-)	gcccccatccTTGTCctga	0.98	0.968	**PM**	-31 (-)	gattataagcCTGACcaag	0.98	0.968	**PM**	-58 (-)	tcgcccggccTTGACagct	0.989	0.987	**PM**
	-96 (-)	catccttgtcCTGACatgc	0.98	0.984	**PM**	+43 (-)	agctgtcaccATGTCggtg	0.992	0.974	**PM**	+51 (-)	tcccacagcgACTGGccac	1.0	0.969	**PM**
	-88 (+)	tcctGACATgcaattggct	0.992	0.977	**PM**										
Pax-6	-304 (-)	tttaactcatcCTTAAactat	0.792	0.836	**M***		not found					not found			
Sox-5		not found					not found				+68 (-)	acATTGTtgc	1.0	1.0	**PM**
Sp1	-192 (+)	tgagGGAGGggcta	0.956	0.942	**M**	-854 (-)	accCAGCCct	0.962	0.954	**M**	-78 (+)	ggGGCAGggc	1.0	1.0	**PM**
	-63 (+)	ccggGGAGGggcgg	0.956	0.954	**M**	-59 (+)	tcGGGGTggt	1.0	1.0	**PM**					
	-58 (+)	gaggGGCGGgactc	1.0	0.942	**M**										
SRY	-631 (+)	AAGGAca	1.0	1.0	**PM**	-718 (-)	gtTCTTT	1.0	1.0	**PM**	-988 (-)	tcTCTTT	1.0	1.0	**PM**
	-302 (+)	TAACTca	1.0	1.0	**PM**	-464 (+)	AAACAaa	1.0	1.0	**M, PM**	-846 (-)	tgTGTTT	1.0	1.0	**PM**
											-359 (-)	tgAGTTA	1.0	1.0	**PM**
											-279 (-)	ctCCTTT	1.0	1.0	**PM**
STAT1		not found					not found				-291 (-)	ttacAGGAA	1.0	1.0	**PM***
v-Myb		not found					not found				-484 (-)	ctCCGTTact	1.0	0.967	**M***
ZEB (AREB6)		not found				-489 (+)	ctGTTTCac	1.0	1.0	**PM**	-845 (+)	gtGTTTCat	1.0	1.0	**PM**

Putative transcription-factor-binding sites found in the proximal promoter regions (from -1000 bp upstream of the putative transcription start site to +100 bp downstream) of the genes for murine galectins-1, -2 and -7. Pos. (Ori.): Position (relative to the putative start site of transcription) and orientation of the putative binding site ((+): sense strand; (-): antisense strand). Sequence: DNA sequence of the putative binding site; core sequence is set in uppercase. Core Score: score for the matrix core sequence; Matrix Score: score for the entire matrix sequence; Program: the algorithm yielding the given hit - M: Match™; PM: P-Match™. If a putative binding site is predicted by both algorithms, different scores and slightly deviating matrix sequences for the same hit are possible between the output of the programs - in these cases, both values are given. Two score values are also given when several different weight matrices exist for the same binding factor and both yield scores surpassing the cut-off but deviating from each other. Hits which were detected by "high quality" matrices are marked by an asterisk after M/PM, respectively. Binding Factor: AML-1a: acute myeloid leukemia gene 1a (Runx/runt homology domain); AP-1: activator protein 1; Cdx-1: chicken homeobox gene cdxA product; C/EBP: CCAAT/enhancer binding protein; CHOP-10: C/EBP homologous protein 10; COMP1: cooperates with myogenic proteins 1; CUTL-1: cut-like homeodomain protein 1; En-1: engrailed 1; Evi-1: ectopic viral integration site 1-encoded factor; GATA: GATA-box binding factor; GKLF: gut-enriched Krueppel-like factor; HNF: hepatic nuclear factor; COUP-TF: chicken ovalbumin upstream promoter-transcription factor; HOX A3: homeobox cluster protein A3; Lmo2: LIM-only protein 2; Msx-1: msh-like homeobox protein 1; MZF-1: myeloid zinc finger protein 1; Pax: paired box gene product; Sox: SRY-related HMG-box gene product; Sp1: stimulatory protein 1; SRY: sex-determining region Y gene product; STAT1: signal transducer and activator of transcription 1, p91; v-Myb: viral myoblastoma oncoprotein; ZEB: zinc finger E-box-binding protein or Atp1a1 regulatory-element-binding protein 6 (AREB6).

Table 2. Unique features of proximal promoter regions in galectin genes defined by the transcription factor for the target sequence[a]

	presence of target sequence	absence of target sequence
mGal-1	Pax-6	COMP-1, Evi-1, HNF-4α, Msx-1, ZEB (AREB6)
mGal-2	CUTL-1	GATA-3
mGal-7	AP-1, Lmo2-complex, Sox-5, STAT1, v-Myb	–

[a]the cases with restriction of the absence/presence of a putative target site for the given transcription factor to the proximal promoter region of only one gene are singled out from the listing in Table 1.

deliberately generating a very strong signal to pick up low-level contaminations. That the antiserum fractions without exception did not bind to other components in the extracts is shown in Fig. 3B by Western blot data using different organ specimens. When probing extracts of heart, lung, stomach, jejunum, colon and urinary bladder, the galectin-1 presence was invariably observed, whereas the signals for galectins-2 and -7 showed a relatively restricted occurrence, the positive cases given in Fig. 3B. In addition to running these biochemical controls, the antigen-dependent staining was ascertained by immunohistochemical controls, as illustrated exemplarily in Fig. 4A, B. Also, the optimal conditions for fixation in terms of degree of preservation of tissue integrity and of signal intensity had been defined in preliminary testing.

Having verified the quality of the reagents and the protocol, the patterns of galectin presence and localiza-

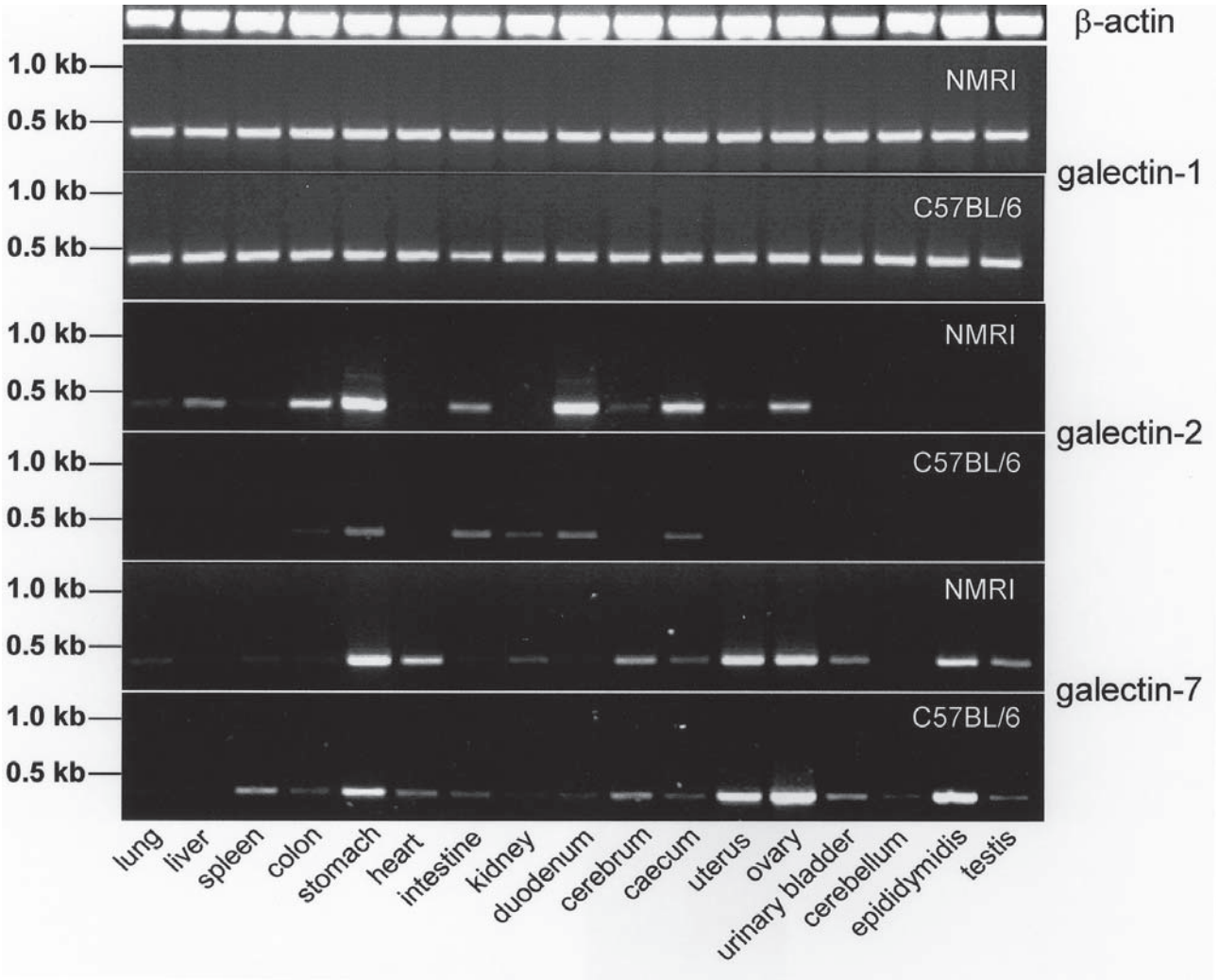

Fig. 2. Profiling of gene transcription for murine galectins -1, -2 and -7 in various tissue types of two mouse strains. The presence of galectin-gene-specific transcripts was determined performing RT-PCR with discriminatory primer sets in cDNA preparations of organ samples, yielding fragment lengths of 410 bp for galectin-1, 420 bp for galectin-2 and 432 bp for galectin-7. Amplification of β-actin cDNA with a β-actin-specific primer set (predicted product size: 622 bp) served as loading and quality control.

Table 3. Summary of semi-quantitative detection of galectins (gene expression/protein presence) by RT-PCR/immunohistochemistry (IHC)[a]

	galectin-1		galectin-2		galectin-7	
	RT-PCR	IHC	RT-PCR	IHC	RT-PCR	IHC
nervous system						
cerebrum	++	(+)	+	–	+	–
cerebellum	++	(+)	–	–	+/–	–
digestive tract						
stomach/oesophagus	++	(+)/(+)	+	+/–	++	–/++
duodenum	++	(+)	+/–	(+)	–	–
jejunum	++	(+)	+	(+)	+/–	–
colon	++	(+)	+/–	(+)	+/–	++[d]
caecum	++	(+)	+/–	(+)	+/–	–
respiratory and circulatory tract						
lung	++	(+)	–	–	–	–
heart	++	(+)	–	–	+	–
urogenital tract						
kidney	++	(+)/++[b]	+/–	–	–	–
urinary bladder	++	(+)	–	–	+	++[e]
ovary	++	(+)	–	–	++	++[f]
testis	++	(+)/++[c]	–	–	++	–
epididymidis	++	(+)	–	–	+/–	–
uterus	++	(+)	–	–	++	–
miscellaneous						
liver	++	(+)	–	–	–	–
spleen	++	(+)	–	–	+	–
skin	n. t.[g]	(+)	n. t.[g]	-	n. t.[g]	++

[a]signal intensity was semi-quantitatively grouped into the categories: "++" (strong), "+" (medium), "(+)" (weak but significant), "(+/-)" (very weak); the cases b-f define staining intensity of distinct cell types: [b]transitional epithelium of the renal pelvis; [c]tubuli seminiferi contorti; [d]epithelial lining of villi/crypts; [e]transitional epithelium of the urinary bladder; [f]ovarian surface epithelium; [g]not tested

tion were systematically determined immunohistochemically. The results are summarized in Table 3. In principle, there is a close correlation between the two data sets, when considering the high level of sensitivity of the RT-PCR approach. Strictly speaking, a spurious presence of transcripts might not translate into sufficiently high protein production for immunohistochemical detection. In equal accord with the blot analyses, the galectin-1 presence was seen ubiquitously, examples documenting the staining profile of fixed sections in urinary bladder (vessels, Fig. 4C) and testicular tissue (Fig. 4D). In the latter case, the terminal plugs of the *tubuli seminiferi contorti* were strongly positive as was the transitional epithelium of the renal pelvis. Nuclear reactivity appeared in a substantial percentage of cells. On the face of it, the range of organ positivity was considerably restricted for galectins-2 and -7, as was the cellular staining to cytoplasmic positivity. The digestive tract, especially the apical glands of the stomach and the apical/luminal parts of the jejunum and the colon, was the site with galectin-2 presence (Table 3; Fig. 4E, F), epithelium in different organs, conspicuously seen e.g. in oesophagus, colon, ovary, skin and urinary bladder, the site with galectin-7 presence (Table 3; Fig. 5A-E). A comparison of urinary bladder positivity between galectins-1 and -7 is given in Fig. 5E, F to illustrate the differences in the cell-type positivity in an organ. Evidently, the immunohistochemical profiling revealed characteristic expression patterns that intimate non-redundant characteristics. Further evidence for non-redundant properties can be provided by examining the proteins' sequences. Deviations in the coding regions mentioned in the first paragraph let differences in the amino acid sequences become expectable, and the knowledge on position of contact sites for carbohydrate ligands renders *in silico* work

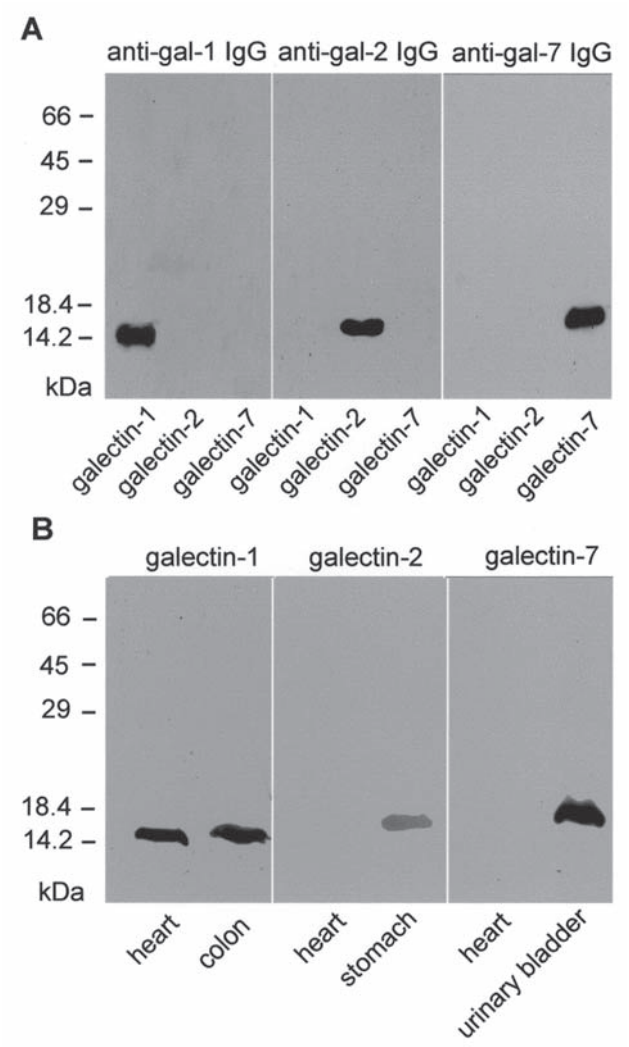

Fig. 3. Western blot analyses for excluding any cross-reactivity (A) or antigen-independent reactivity (B) in tissue extracts for the three immunoglobulin G fractions against galectins -1, -2 and -7, respectively. The galectins were tested at 100 ng, samples from organ extracts at 50 µg protein for heart, colon and urinary bladder or 100 µg for stomach. Samples were subjected to SDS-PAGE gel electrophoresis, and the proteins were transferred thereafter to the nitrocellulose membrane in a tank blotting procedure. The period of film exposure after immunostaining was 20 s for detection of cross-reactivity and 1-2 (stomach) min for extract processing.

feasible to relate sequence alterations to the carbohydrate recognition domain.

Structural profiling by computational chemistry

Multiple alignment revealed a sequence identity of 41.7% (57.6% similarity) between galectins-1 and -2, 27.9% (45% similarity) between galectins-1 and -7, and 30.1% (51.5% similarity) between galectins-2 and -7. The detailed illustration of the positioning of identities marks especially amino acids crucial for direct contact to carbohydrate ligands centred around the sole tryptophan residue, i.e. the equivalents of His44, Asn46, Arg48, Lys63, Asn61, Trp68, Glu71 and Arg73 in human galectin-1 (Fig. 6A). The way these residues can be involved in ligand binding was visualized by a docking analysis, selecting naturally occurring oligosaccharides as binding partners to facilitate a view on the extended binding domain. The low-energy conformations of the tetra- and pentasaccharides are given in Fig. 6B. Of note, computational model building for structures of homodimeric murine galectins was reliably feasible for galectins-1 and -7 but not for galectin-2. Considering the sequences and the architecture of the carbohydrate recognition domains, two types of disparities with potential impact on ligand binding were discerned: the occurrence of sequence gaps (e.g. absence of the equivalent of His53 in galectin-7; see Figs. 6C, D) and of substitutions (e.g. Asp124/Gly125 in Fig. 6C, D and in the immediate vicinity of key moieties such as Thr71/Arg72 or His73/Glu74).

Discussion

This study deals with the intra-family diversity of galectins, potent extra- and intracellular regulators of diverse cellular activities. It addresses the fundamental question on the extent of divergence at different levels, reporting on gene/promoter sequences, expression profiles and structural features of carbohydrate recognition domains. As a step toward a comprehensive analysis of the proteins in the galectin network we herein comparatively mapped the respective characteristics of three homodimeric proteins, i.e. murine galectins-1, -2 and -7. In order to obtain valid conclusions it was imperative to keep experimental/computational parameters identical in all the study parts, e.g. search algorithms or tissue fixation.

The gene organization obviously reflected common ancestry. Despite the maintained exon/intron organization the relation between exon length and placement of sequence stretches in the β-strands of the two antiparallel β-sheets had been interpreted to suggest a separation of galectin-7 from the branch with galectins-1 and -2 in phylogenesis (Houzelstein et al., 2004). Changes in the length of the exon encoding the central tryptophan residue (see Fig. 1 for details and Fig. 6 for its involvement in contact to carbohydrate ligands) underlie the variable contribution of this exon to the F4 β-strand, this shift implying a certain evolutionary distance. At the level of putative sites for interaction with transcription factors in the proximal promoter regions, a series of sequence hits was found in each case. In general, the quantitative and qualitative differences that we extracted from extensive database mining are indicative of the potential for non-uniform regulation of gene activities. Of note, the overall transcriptional rate will likely be the sum of combinatorial events and not yet readily predictable. As it

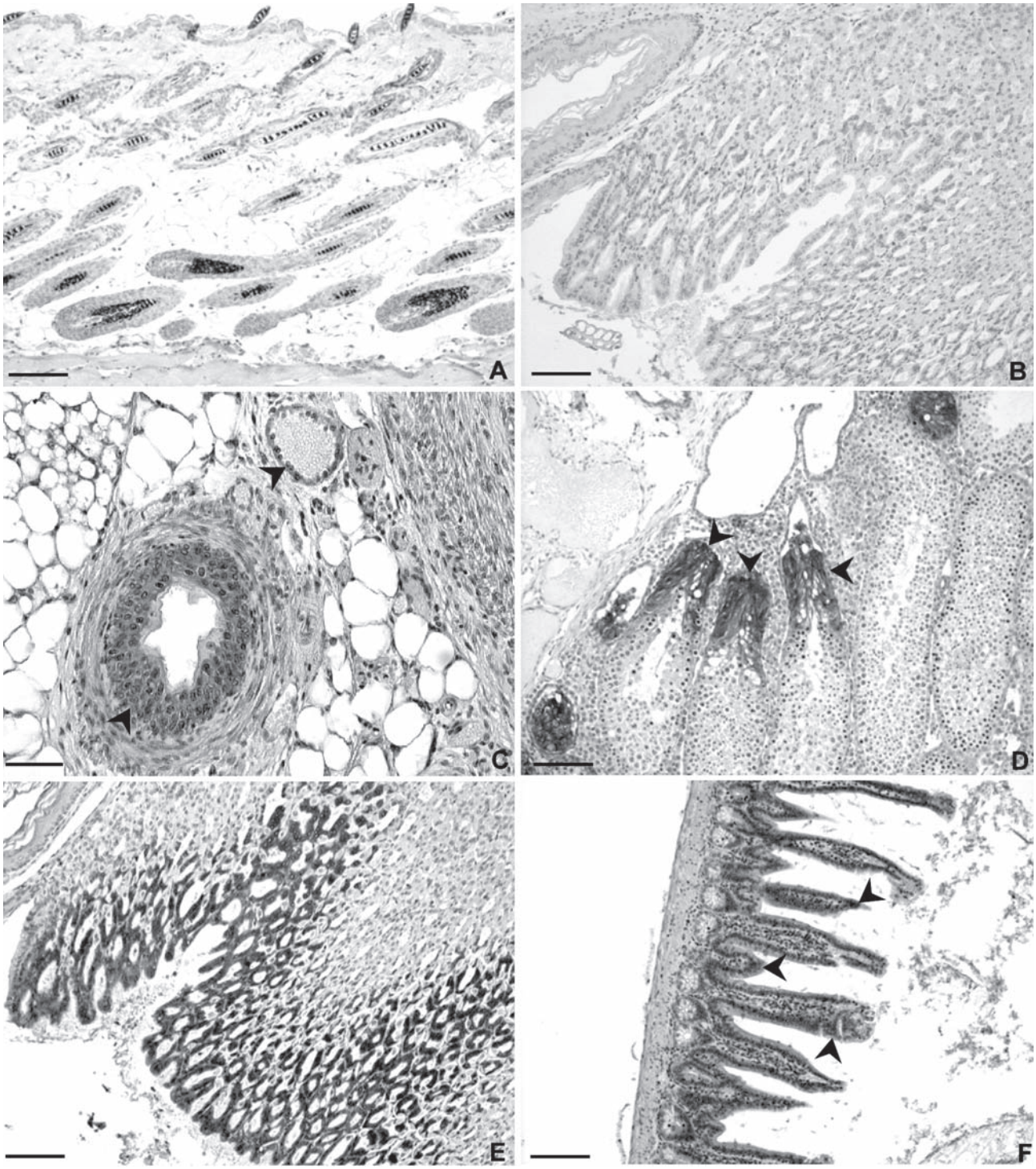

Fig. 4. Illustration of galectin localization with anti-galectin-1 and anti-galectin-2 in fixed tissue sections of different murine organs. Controls by omission of the primary antibody from the staining protocol (section of skin; A) and by processing pre-immune rabbit serum (section of stomach; B) showed no positive reaction of the tissue. Dark granules in panel A are composed of melanin. Processing with anti-galectin-1 resulted in positivity of vessels (C; arrowheads) and smooth muscle cells, as demonstrated in a section of urinary bladder (C). In testis, a strong signal was detected in the terminal plugs of the *tubuli seminiferi contorti* (D; arrowheads). Regarding anti-galectin-2 the luminal parts of the glands of stomach presented a positive reaction (E), for comparison to the control using pre-immune rabbit serum see panel B. Positivity in jejunum and colon was found in the epithelium of villi (F; arrowheads) but not in the epithelium of crypts. Bar: 100 µm (A, B, D-F), bar: 50 µm (C).

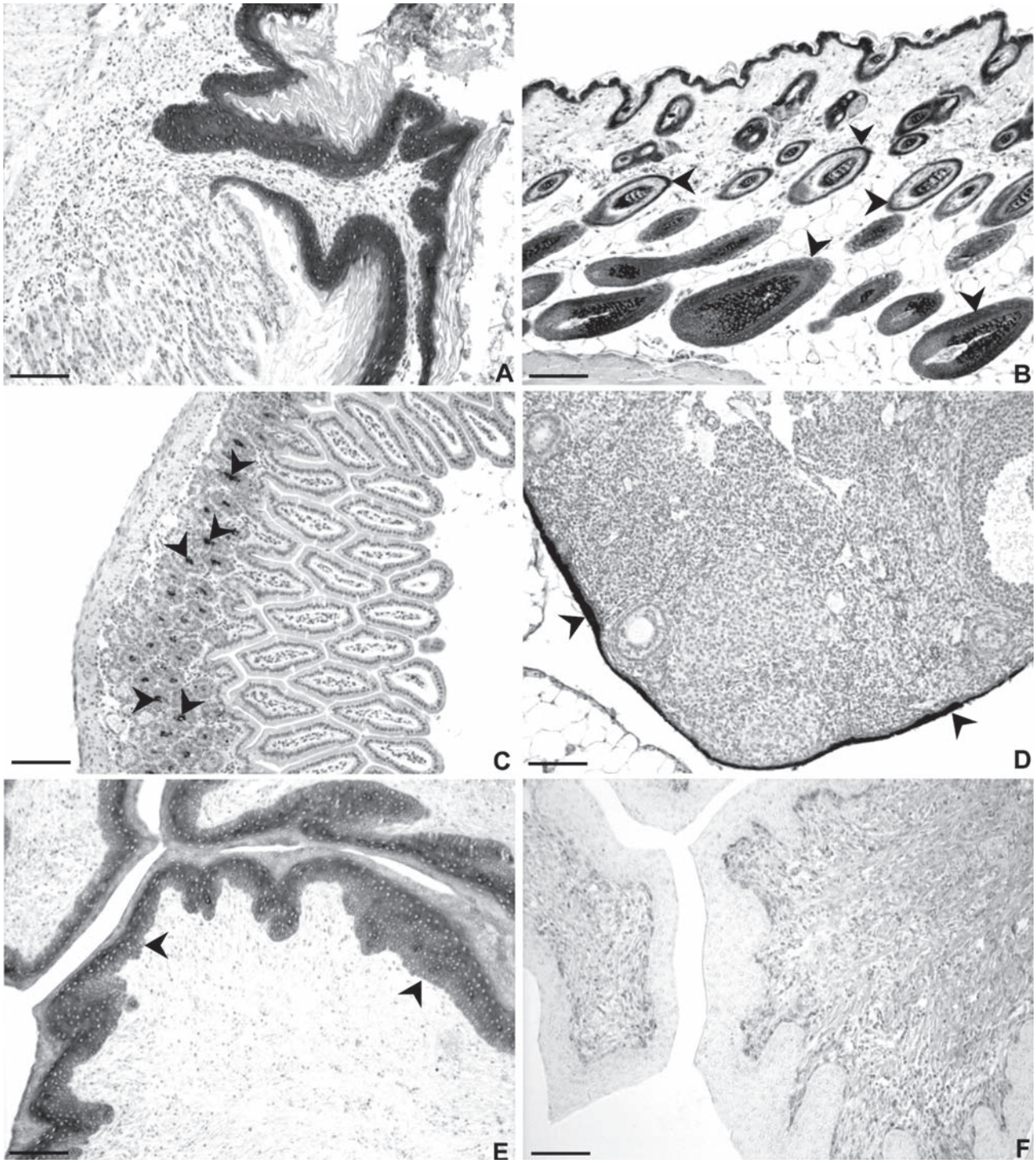

Fig. 5. Illustration of galectin localization with anti-galectin-7 (A-E) and anti-galectin-1 (F) in fixed tissue sections of different murine organs. Galectin-7 was localized in stratified epithelia of, e.g., oesophagus (A) and skin (B), for comparison to processing without the primary antibody see panel A of Fig. 4. Positivity in skin also included the outer root sheath (B; arrowheads). Epithelial lining of the apical parts of the crypts in colon was strongly positive (C, arrowheads). In ovary, the surface epithelium (D, arrowheads) was strongly reactive with the galectin-7-specific IgG preparation. Cell-type selectivity of the galectin presence was documented in the case of urinary bladder for galectin-7, exclusively present in the transitional epithelium (E, arrowheads), and for galectin-1, weakly present in smooth muscle cells and connective tissue (F). No counterstaining was performed in this case. Bar: 100 μm (A-F).

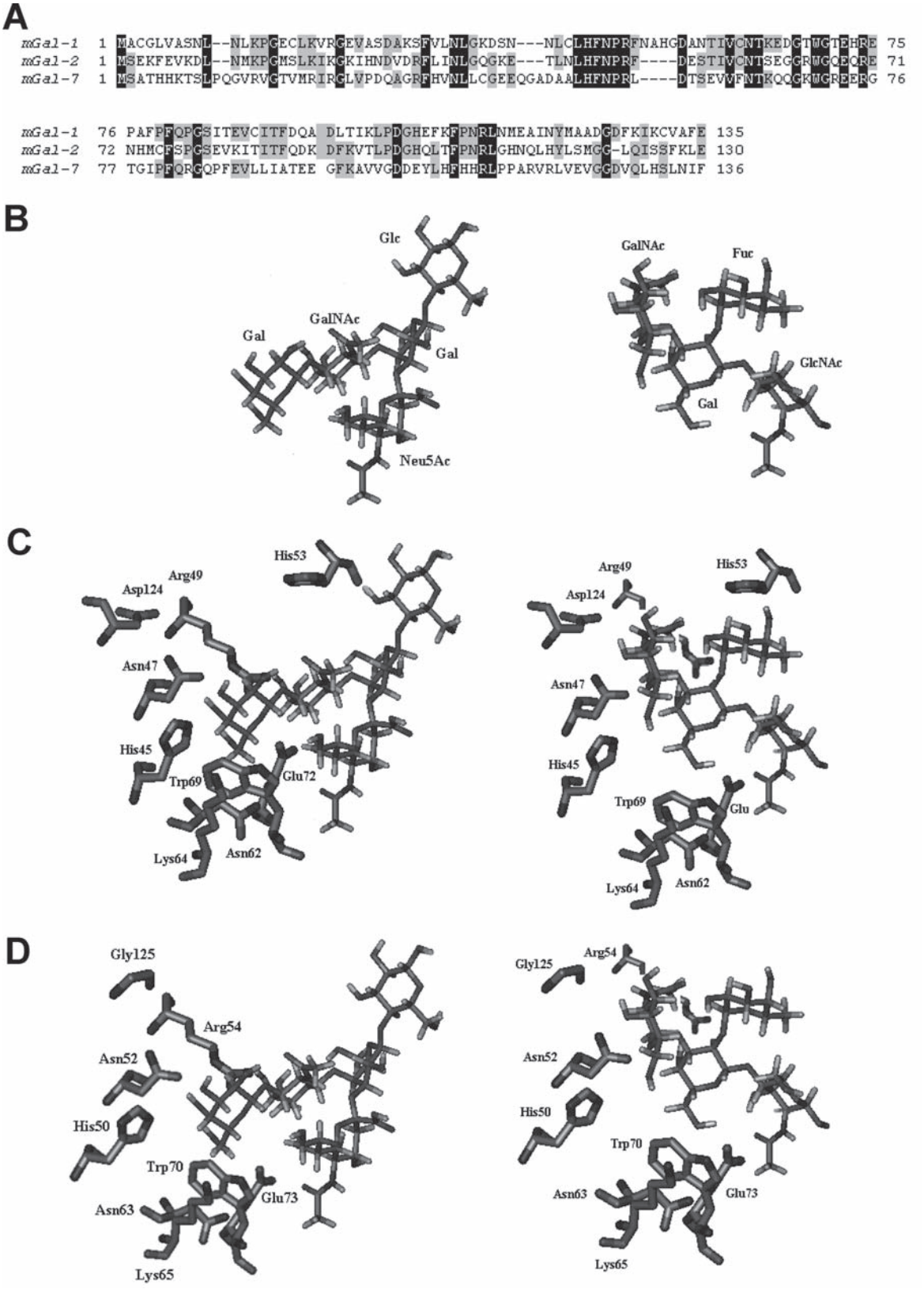

Fig. 6. Graphical illustration of sequence comparison by multiple alignment in the cases of murine galectins-1, -2 and -7 and of architecture of carbohydrate recognition domains when accommodating natural glycans. Identical residues between all three protein sequences are shown as white letters on black background, while those moieties identical between two of the sequences are highlighted by grey background (A). The low-energy conformations of the two carbohydrate ligands used in docking, i.e. the pentasaccharide of ganglioside GM_1 (left) and the histo-blood group A tetrasaccharide (right), are shown in panel B. The relative positions of the two glycans and key sites of two galectin structures derived from homology modelling (galectin-1: C; galectin-7: D) are given.

stands, our analysis markedly extends previous monitoring of promoter regions of galectins-1 and -2 in the case of the human genes, in which the main attention was directed to Sp1/Ap1 sites (Gitt et al., 1992).

Regarding the technical side of this approach it should be mentioned that applying the two search algorithms independently did not lead to identical profiles of target sequences that meet the particular criteria. Although it would thus be premature to draw definitive conclusions on contextual functionality of each sequence *in vivo*, it could, though, be assumed that non-identical expression patterns will emerge. Determination of haemagglutinating activity, as performed for extracts of mouse organs (Levi and Teichberg, 1984), and of cell reactivity to neoglycoconjugates with non-discriminatory ligands, as performed with cells and tissue sections (Gabius and Bardosi, 1991; Gabius and Gabius, 1992), may thus not simply indicate presence of a single lectin but additive contributions of several proteins. Our results on three proteins under identical conditions identify such cases of overlap. They would remain undetected when exclusively studying an individual protein, as e.g. in the studies with panels of human, rat or rabbit organs and galectin-1 (Harrison et al., 1984; Wasano et al., 1990; Allen et al., 1991). The rather ubiquitous presence of galectin-1 in murine tissue matches previous experience by *in situ* hybridization or immunohistochemistry on the lectin presence in midgestation embryo and adult muscle (Poirier et al., 1992; Poirier and Robertson, 1993), uterus except luminal and glandular epithelium (Phillips et al., 1996), ovarian stroma and *corpus luteum* (Choe et al., 1997; Nio and Iwanaga, 2007), hair follicles (Wollina et al., 2000), interstitium and seminiferous epithelium of testis (Timmons et al., 2002) as well as *lamina propria mucosae* and the muscle layer of the digestive tract (Nio et al., 2005).

At the same time, we performed the corresponding analysis with galectins-2 and -7, using antisera in immunohistochemistry that we rigorously controlled for specificity (Holmseth et al., 2006) and subjected to chromatographic affinity depletion to eliminate cross-reactivity among the galectin family. Compared to the galectin-1 features the detection of galectins-2 and -7 revealed a more restricted range of organ positivity with galectin-type-dependent differences. The low degree of overlap constituted an inherent quality control, as did presentation of nuclear staining only for galectin-1. Fittingly, this lectin is a component of spliceosomal complexes (Smetana et al., 2006; Wang et al., 2006). Staining (in the case of galectin-2) confined to the digestive tract is evocative of the situation in rat and in line with results on mouse tissues by *in situ* hybridization, hereby rendering a major impact of negative translational control or substantial transport from the site of production to the site of final residence unlikely (Oka et al., 1999; Nio et al., 2005). Galectin-7, a marker of stratified epithelia (Magnaldo et al., 1998; Timmons et al., 1999; Sato et al., 2002), served to underscore the inter-galectin differences, also when sharing sites of presence with galectin-1. The staining profile of the ovarian surface epithelium, corroborated by mRNA detection (Nio and Iwanaga, 2007), is homogeneous for galectin-7 and heterogeneous for galectin-1. Whereas in this case the same cell type harbours two tested galectins, the urinary bladder presents a case with cell-type selectivity, as documented in Fig. 5E, F. With the presence of galectins in murine tumours having initially been proven biochemically (Gabius et al., 1984) and the given results underlining the tissue presence of more than one lectin in certain normal tissues, it is evidently promising to monitor malignancy and other diseases with a set of different antibodies. Ovarian/urothelial tumours or gastric cancer, where expression of the galectin-2 gene was found to be downregulated by chip analysis of *Helicobacter*-induced tumours (Takaishi and Wang, 2007), are attractive study objects in view of the data reported herein. Investigations on human material, e.g. pulmonary/head and neck cancer or muscle biopsies have already attested the potential of localizing these proteins or galectin-binding sites as markers (Bardosi et al., 1989; Kayser et al., 2003a, b, 2005; Plzák et al., 2004; Saussez et al., 2006). Flanking the galectin mapping with syntactic structure analysis and application of other markers may then help shed light on effector pathways of the lectins *in vivo*, e.g. by correlating growth patterns to the presence of a galectin (Kayser and Gabius, 1997, 1999; Kayser et al., 2001; Sheikholeslam-Zadeh et al., 2001; Chovanec et al., 2005) or to the absence of a functional galectin ligand such as glycoprotein CD7 in Sézary syndrome (Rappl et al., 2002).

Intuitively, the detection of disparities in regulation and in cell-type patterns implies distinct structural features. At the level of carbohydrate recognition, affinity to the typical branch-end N-acetyllactosamine is comparatively low for galectin-7 (Ahmad et al., 2002; Dam et al., 2005). Frontal affinity chromatography and isothermal titration calorimetry uncovered further differences including toleration of α2,6-sialylation in structures with an N-acetyllactosamine repeat by galectin-7 but not galectin-1 (Ahmad et al., 2002; Hirabayashi et al., 2002). Based on experimental and computational description of extended binding sites for natural glycans (Siebert et al., 2003; Wu et al., 2007) we set out to spot sequence differences in the vicinity of crucial contact points. Model building was feasible in two cases, and sequence changes of assumed relevance were delineated. Since long-range effects by substitutions (i.e. Cys2Ser and Arg111His) and the impact of ligand binding on the gyration radius are known from the case of galectin-1 (He et al., 2003; López-Lucendo et al., 2004), the actual relevance of these alterations will definitely

need to be tested experimentally by site-directed mutagenesis.

Taken together, the presented evidence for the extent of divergence among the homodimeric galectins-1, -2 and -7 prompts us to pursue the systematic profiling. Structurally, the chimera-type galectin-3, which harbours a collagenase-sensitive domain and an N-terminal section susceptible to serine phosphorylation, and tandem-repeat-type proteins such as galectins-4/-6, -8 and -9 are readily separated from the homodimeric proteins. Corresponding data for these family members and their correlation to the results presented here are expected to contribute to enhance our understanding of the galectin network.

References

Ahmad, N., Gabius, H.-J., Kaltner, H., André, S., Kuwabara, I., Liu, F.-T., Oscarson, S., Norberg, T., Brewer, C. F. (2002) Thermodynamic binding studies of cell surface carbohydrate epitopes to galectins-1, -3, and -7: evidence for differential binding specificities. *Can. J. Chem.* **80**, 1096-1104.

Ahmad, N., Gabius, H.-J., André, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F., Brewer, C. F. (2004) Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *J. Biol. Chem.* **279**, 10841-10847.

Allen, H. J., Sucato, D., Gottstine, S., Kisailus, E., Nava, H., Petrelli, N., Castillo, N., Wilson, D. (1991) Localization of endogenous β-galactoside-binding lectin in human cells and tissues. *Tumor Biol.* **12**, 52-60.

André, S., Unverzagt, C., Kojima, S., Dong, X., Fink, C., Kayser, K., Gabius, H.-J. (1997) Neoglycoproteins with the synthetic complex biantennary nonasaccharide or its α2,3/α2,6-sialylated derivatives: their preparation, assessment of their ligand properties for purified lectins, for tumor cells *in vitro* and in tissue sections, and their biodistribution in tumor-bearing mice. *Bioconjugate Chem.* **8**, 845-855.

André, S., Kojima, S., Yamazaki, N., Fink, H., Kaltner, H., Kayser, K., Gabius, H.-J. (1999) Galectins-1 and -3 and their ligands in tumor biology. Non-uniform properties in cell surface presentation and modulation of adhesion to matrix glycoproteins for various tumor cell lines, in biodistribution of free and liposome-bound galectins and in their expression by breast and colorectal carcinomas with/without metastatic propensity. *J. Cancer Res. Clin. Oncol.* **125**, 461-474.

André, S., Kaltner, H., Furuike, T., Nishimura, S.-I., Gabius, H.-J. (2004a) Persubstituted cyclodextrin-based glycoclusters as inhibitors of protein-carbohydrate recognition using purified plant and mammalian lectins and wild-type and lectin-gene-transfected tumor cells as targets. *Bioconjugate Chem.* **15**, 87-98.

André, S., Unverzagt, C., Kojima, S., Frank, M., Seifert, J., Fink, C., Kayser, K., von der Lieth, C.-W., Gabius, H.-J. (2004b) Determination of modulation of ligand properties of synthetic complex-type biantennary N-glycans by introduction of bisecting GlcNAc *in silico*, *in vitro* and *in vivo*. *Eur. J. Biochem.* **271**, 118-134.

André, S., Kaltner, H., Lensch, M., Russwurm, R., Siebert, H.-C., Fallsehr, C., Tajkhorshid, E., Heck, A. J. R., von Knebel-Döberitz, M., Gabius, H.-J., Kopitz, J. (2005a) Determination of structural and functional overlap/divergence of five proto-type galectins by analysis of the growth-regulatory interaction with ganglioside GM_1 *in silico* and *in vitro* on human neuroblastoma cells. *Int. J. Cancer* **114**, 46-57.

André, S., Kojima, S., Prahl, I., Lensch, M., Unverzagt, C., Gabius, H.-J. (2005b) Introduction of extended LEC14-type branching into core-fucosylated biantennary N-glycan. Glycoengineering for enhanced cell binding and serum clearance of the neoglycoprotein. *FEBS J.* **272**, 1986-1998.

André, S., Kojima, S., Gundel, G., Russwurm, R., Schratt, X., Unverzagt, C., Gabius, H.-J. (2006a) Branching mode in complex-type triantennary N-glycans as regulatory element of their ligand properties. *Biochim. Biophys. Acta* **1760**, 768-782.

André, S., Pei, Z., Siebert, H.-C., Ramström, O., Gabius, H.-J. (2006b) Glycosyldisulfides from dynamic combinatorial libraries as O-glycoside mimetics for plant and endogenous lectins: their reactivities in solid-phase and cell assays and conformational analysis by molecular dynamics simulations. *Bioorg. Med. Chem.* **14**, 6314-6326.

André, S., Maljaars, C. E. P., Halkes, K. M., Gabius, H.-J., Kamerling, J. P. (2007) Discovery of galectin ligands in fully randomized combinatorial one-bead-one-compound (glyco)peptide libraries. *Bioorg. Med. Chem. Lett.* **17**, 793-798.

Bardosi, A., Dimitri, T., Wosgien, B., Gabius, H.-J. (1989) Expression of endogenous receptors for neoglycoproteins, especially lectins, that allow fiber typing on formaldehyde-fixed, paraffin-embedded muscle biopsy specimens. A glycohistochemical, immunohistochemical, and glycobiochemical study. *J. Histochem. Cytochem.* **37**, 989-998.

Brinck, U., Bosbach, R., Korabiowska, M., Schauer, A., Gabius, H.-J. (1995) Lectin-binding sites in the epithelium of normal human appendix vermiformis and in acute appendicitis. *Histol. Histopathol.* **10**, 61-70.

Brinck, U., Bosbach, R., Korabiowska, M., Schauer, A., Gabius, H.-J. (1996) Histochemical study of expression of lectin-reactive carbohydrate epitopes and glycoligand-binding sites in normal human appendix vermiformis, colonic mucosa, acute appendicitis and colonic adenoma. *Histol. Histopathol.* **11**, 919-930.

Brinck, U., Korabiowska, M., Bosbach, R., Gabius, H.-J. (1998) Detection of inflammation- and neoplasia-associated alterations in human large intestine using plant/invertebrate lectins, galectin-1 and neoglycoproteins. *Acta Anat.* **161**, 219-233.

Brockhausen, I., Schachter, H. (1997) Glycosyltransferases involved in N- and O-glycan biosynthesis. In: *Glycosciences: Status and Perspectives*, eds. Gabius H.-J., Gabius S., pp 79-113. Chapman & Hall, Weinheim, London.

Buzás, E., György, B., Pásztói, M., Jelinek, I., Falus, A., Gabius H.-J. (2006) Carbohydrate recognition systems in autoimmunity. *Autoimmunity* **39**, 691-704.

Caselitz, J. (1987) Lectins and blood group substances as tumor markers. *Curr. Top. Pathol.* **77**, 245-278.

Chekmenev, D. S., Haid, C., Kel, A. E. (2005) P-Match: transcription factor binding site search by combining patterns and weight matrices. *Nucleic Acids Res.* **33,** W432-W437.

Choe, Y. S., Shim, C., Choi, D., Lee, C. S., Lee, K.-K., Kim, K. (1997) Expression of galectin-1 mRNA in the mouse uterus is under the control of ovarian steroids during blastocyst implantation. *Mol. Reprod. Dev.* **48**, 261-266.

Chovanec, M., Smetana, K. Jr., Dvořánková, B., Plzáková, Z., André, S., Gabius, H.-J. (2004) Decrease of nuclear reactivity to growth-regulatory galectin-1 in senescent human keratinocytes and detection of non-uniform staining profile alterations upon prolonged culture for galectins-1 and -3. *Anat. Histol. Embryol.* **33**, 348-354.

Chovanec, M., Smetana, K. Jr., Betka, J., Plzák, J., Brabec, J., Moya-Álvarez, V., André, S., Kodet, R., Gabius, H.-J. (2005) Correlation of expression of nuclear proteins pKi67 and p63 with lectin histochemical features in head and neck cancer. *Int. J. Oncol.* **27**, 409-415.

Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881-10890.

Cummings, R. D. (1997) Lectins as tools for glycoconjugate purification and characterization. In: *Glycosciences: Status and Perspectives*, eds. Gabius H.-J., Gabius S., pp 191-199. Chapman & Hall, Weinheim, London.

Dam, T. K., Gabius, H.-J., André, S., Kaltner, H., Lensch, M. and Brewer, C. F. (2005) Galectins bind to the multivalent glycoprotein asialofetuin with enhanced affinities and a gradient of decreasing binding constants. *Biochemistry* **44**, 12564-12571.

Danguy, A., Akif, F., Pajak, B., Gabius, H.-J. (1994) Contribution of carbohydrate histochemistry to glycobiology. *Histol. Histopathol.* **9**, 155-171.

Danguy, A., Kayser, K., Bovin, N. V., Gabius, H.-J. (1995) The relevance of neoglycoconjugates for histology and pathology. *Trends Glycosci. Glycotechnol.* **7**, 261-275.

Fischer, C., Sanchez-Ruderisch, H., Welzel, M., Wiedenmann, B., Sakai, T., André, S., Gabius, H.-J., Khachigian, L., Detjen, K. M., Rosewicz, S. (2005) Galectin-1 interacts with the α5β1 fibronectin receptor to restrict carcinoma cell growth via induction of p21 and p27. *J. Biol. Chem.* **280**, 37266–37277.

Gabius, H.-J. (1987) Endogenous lectins in tumors and the immune system. *Cancer Invest.* **5**, 39-46.

Gabius, H.-J. (1990) Influence of type of linkage and spacer on the interaction of β-galactoside-binding proteins with immobilized affinity ligands. *Anal. Biochem.* **189**, 91-94.

Gabius, H.-J. (1997) Animal lectins. *Eur. J. Biochem.* **243**, 543-576.

Gabius, H.-J. (2000) Biological information transfer beyond the genetic code: the sugar code. *Naturwissenschaften* **87**, 108-121.

Gabius, H.-J. (2001) Glycohistochemistry: the why and how of detection and localization of endogenous lectins. *Anat. Histol. Embryol.* **30**, 3-31.

Gabius, H.-J. (2006) Cell surface glycans: the why and how of their functionality as biochemical signals in lectin-mediated information transfer. *Crit. Rev. Immunol.* **26**, 43-79.

Gabius, H.-J., Bardosi, A. (1991) Neoglycoproteins as tools in glycohistochemistry. *Progr. Histochem. Cytochem.* **22(2)**, 1-66.

Gabius, H.-J., Gabius, S. (1992) Chemical and biochemical strategies for the preparation of glycohistochemical probes and their application in lectinology. *Adv. Lectin Res.* **5**, 123-157.

Gabius, H.-J., Engelhardt, R., Rehm, S., Cramer, F. (1984) Biochemical characterization of endogenous carbohydrate-binding proteins from spontaneous murine rhabdomyosarcoma, mammary adenocarcinoma, and ovarian teratoma. *J. Natl. Cancer Inst.* **73**, 1349-1357.

Gabius, H.-J., Engelhardt, R., Rehm, S., Barondes, S. H., Cramer, F. (1986) Presence and relative distribution of three endogenous β-galactoside-specific lectins in different tumor types of rat. *Cancer J.* **1**, 19-22.

Gabius, H.-J., Wosgien, B., Hendrys, M., Bardosi, A. (1991) Lectin localization in human nerve by biochemically defined lectin-binding glycoproteins, neoglycoprotein and lectin-specific antibody. *Histochemistry* **95**, 269-277.

Gabius, H.-J., Gabius, S., Zemlyanukhina, T. V., Bovin, N. V., Brinck, U., Danguy, A., Joshi, S. S., Kayser, K., Schottelius, J., Sinowatz, F., Tietze, L. F., Vidal-Vanaclocha, F., Zanetta, J.-P. (1993) Reverse lectin histochemistry: design and application of glycoligands for detection of cell and tissue lectins. *Histol. Histopathol.* **8**, 369-383.

Gabius, H.-J., André, S., Kaltner, H., Siebert, H.-C. (2002) The sugar code: functional lectinomics. *Biochim. Biophys. Acta* **1572**, 165-177.

Gabius, H.-J., Siebert, H.-C., André, S., Jiménez-Barbero, J., Rüdiger, H. (2004) Chemical biology of the sugar code. *ChemBioChem* **5**, 740-764.

Gitt, M. A., Massa, S. M., Leffler, H., Barondes S. H. (1992) Isolation and expression of a gene encoding L-14-II, a new human soluble lactose-binding lectin. *J. Biol. Chem.* **267**, 10601-10606.

Harrison, F. L., Fitzgerald, J. E., Catt, J. W. (1984) Endogenous β-galactoside specific lectins in rabbit tissues. *J. Cell Sci.* **72**, 147-162.

He, L., André, S., Siebert, H.-C., Helmholz, H., Niemeyer, B., Gabius, H.-J. (2003) Detection of ligand- and solvent-induced shape alterations of cell-growth-regulatory human lectin galectin-1 in solution by small angle neutron and X-ray scattering. *Biophys. J.* **85**, 511-524.

Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Müller, W. E. G., Yagi, F., Kasai, K. I. (2002) Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim. Biophys. Acta* **1572**, 232-254.

Holíková, Z., Hrdlicková-Cela, E., Plzák, J., Smetana, K. Jr., Betka, J., Dvořánková, B., Esner, W., Wasano, K., André, S., Kaltner, H., Motlík, J., Hercogová, J., Kodet, R., Gabius, H.-J. (2002) Defining the glycophenotype of squamous epithelia by plant and mammalian lectins. *APMIS* **110**, 845-856.

Holmseth, S., Lehre, K. P., Danbolt, N. C. (2006) Specificity controls for immunocytochemistry. *Anat. Embryol.* **211**, 257-266.

Houzelstein, D., Gonçalves, I. R., Fadden, A. J., Sidhu, S. S., Cooper, D. N. W., Drickamer, K., Leffler, H., Poirier, F. (2004) Phylogenetic analysis of the galectin family. *Mol. Biol. Evol.* **21**, 1177-1187.

Kaltner, H., Lips, K. S., Reuter, G., Lippert, S., Sinowatz, F., Gabius, H.-J. (1997). Quantitation and histochemical localization of galectin-1 and galectin-1-reactive glyco-

conjugates in fetal development of bovine organs. *Histol. Histopathol.* **12**, 945-960.

Kaltner, H., Seyrek, K., Heck, A., Sinowatz, F., Gabius, H.-J. (2002) Galectin-1 and galectin-3 in fetal development of bovine respiratory and digestive tracts. *Cell Tissue Res.* **307**, 35-46.

Kayser, K., Gabius, H.-J. (1997) Graph theory and the entropy concept in histochemistry. Theoretical considerations, application in histopathology and the combination with receptor-specific approaches. *Progr. Histochem, Cytochem.* **32**, 1-106.

Kayser, K., Gabius, H.-J. (1999) The application of thermodynamic principles to histochemical and morphometric tissue research: principles and practical outline with focus on the glycosciences. *Cell Tissue Res.* **296**, 443-455.

Kayser, K., Böhm, G., Blum, S., Beyer, M., Zink, S., André, S., Gabius, H.-J. (2001) Glyco- and immunohistochemical refinement of the differential diagnosis between mesothelioma and metastatic carcinoma and survival analysis of patients. *J. Pathol.* **193**, 175-108.

Kayser, K., Dünnwald, D., Kazmierczak, B., Bullerdiek, J., Kaltner, H., Zick, Y., André, S., Gabius, H.-J. (2003a) Chromosomal aberrations, profiles of expression of growth-related markers including galectins and environmental hazards in relation to the incidence of chondroid pulmonary hamartomas. *Pathol. Res. Pract.* **199**, 589-598.

Kayser, K., Hoeft, D., Hufnagl, P., Caselitz, J., Zick, Y., André, S., Kaltner, H., Gabius, H.-J. (2003b) Combined analysis of tumor growth pattern and expression of endogenous lectins as prognostic tool in primary testicular cancer and its lung metastases. *Histol. Histopathol.* **18**, 771-779.

Kayser, K., Trott, J., Böhm, G., Huber, M., Kaltner, H., André, S., Gabius, H.-J. (2005) Localized fibrous tumors (LFTs) of the pleura: clinical data, asbestos burden, and syntactic structure analysis applied to newly defined angiogenic/growth-regulatory effectors. *Pathol. Res. Pract.* **201**, 791-801.

Kel, A. E., Gößling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., Wingender, E. (2003) MATCH™: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* **31**, 3576-3579.

Kopitz, J., von Reitzenstein, C., Burchert, M., Cantz, M., Gabius, H.-J. (1998) Galectin-1 is a major receptor for ganglioside GM_1, a product of the growth-controlling activity of a cell surface ganglioside sialidase, on human neuroblastoma cells in culture. *J. Biol. Chem.* **273**, 11205-11211.

Kopitz, J., von Reitzenstein, C., André, S., Kaltner, H., Uhl, J., Ehemann, V., Cantz, M., Gabius, H.-J. (2001) Negative regulation of neuroblastoma cell growth by carbohydrate-dependent surface binding of galectin-1 and functional divergence from galectin-3. *J. Biol. Chem.* **276**, 35917-35923.

Kopitz, J., André, S., von Reitzenstein, C., Versluis, K., Kaltner, H., Pieters, R. J., Wasano, K., Kuwabara, I., Liu, F.-T., Cantz, M., Heck, A. J. R., Gabius, H.-J. (2003) Homodimeric galectin-7 (p53-induced gene 1) is a negative regulator for human neuroblastoma cells. *Oncogene* **22**, 6277-6288.

Kuchler, S., Zanetta, J.-P., Vincendon, G., Gabius, H.-J. (1990) Detection of binding sites for biotinylated neoglycoproteins and heparin (endogenous lectins) during cerebellar ontogenesis in the rat. *Eur. J. Cell Biol.* **52**, 87-97.

Lahm, H., André, S., Höflich, A., Fischer, J. R., Sordat, B., Kaltner, H., Wolf, E., Gabius H.-J. (2001) 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**, 375-386.

Lee, R. T., Lee, Y. C. (1997) Neoglycoconjugates. In: *Glycosciences: Status and Perspectives*, eds. Gabius H.-J., Gabius S., pp 55-77. Chapman & Hall, Weinheim, London.

Lensch, M., Lohr, M., Russwurm, R., Vidal, M., Kaltner, H., André, S., Gabius, H.-J. (2006) Unique sequence and expression profiles of rat galectins-5 and -9 as a result of species-specific gene divergence. *Int. J. Biochem. Cell Biol.* **38**, 1741-1758.

Leonidas, D. D., Vatzaki, E. H., Vorum, H., Celis, J. E., Madsen, P., Acharya, K. R. (1998) Structural basis for the recognition of carbohydrates by human galectin-7. *Biochemistry* **37**, 13930-13940.

Levi, G., Teichberg, V. I. (1984) The distribution of electrolectin in mouse: genetic and ontogenic variations. *Biochem. Biophys. Res. Commun.* **119**, 801-806.

Liu, F.-T., Patterson, R. J., Wang, J. L. (2002) Intracellular functions of galectins. *Biochim. Biophys. Acta* **1572**, 263-273.

Lobsanov, Y. D., Gitt, M. A., Leffler, H., Barondes, S. H., Rini, J. M. (1993) X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-Å resolution. *J. Biol. Chem.* **268**, 27034-27038.

López-Lucendo, M. F., Solís, D., André, S., Hirabayashi, J., Kasai, K. I., Kaltner, H., Gabius, H.-J., Romero, A. (2004) Growth-regulatory human galectin-1: crystallographic characterization of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *J. Mol. Biol.* **343**, 957-970.

Magnaldo, T., Fowlis, D., Darmon, M. (1998) Galectin-7, a marker of all types of stratified epithelia. *Differentiation* **63**, 159-168.

Manning, J. C., Seyrek, K., Kaltner, H., André, S., Sinowatz, F., Gabius, H.-J. (2004) Glycomic profiling of developmental changes in bovine testis by lectin histochemistry and further analysis of the most prominent alteration on the level of the glycoproteome by lectin blotting and lectin affinity chromatography. *Histol. Histopathol.* **19**, 1043-1060.

Nio, J., Iwanaga, T. (2007) Galectins in the mouse ovary: concomitant expression of galectin-3 and progesterone degradation enzyme (20α-HSD) in the corpus luteum. *J. Histochem. Cytochem.* **55**, 423-432.

Nio, J., Kon, Y., Iwanaga, T. (2005) Differential cellular expression of galectin family mRNAs in the epithelial cells of the mouse digestive tract. *J. Histochem. Cytochem.* **53**, 1323-1334.

Ohannesian, D. W., Lotan, D., Thomas, P., Jessup, J. M., Fukuda, M., Gabius, H.-J., Lotan, R. (1995) Carcinoembryonic antigen and other glycoconjugates act as ligands for galectin-3 in human colon carcinoma cells. *Cancer Res.* **55**, 2191-2199.

Oka, T., Murakami, S., Arata, Y., Hirabayashi, J., Kasai, K. I., Wada, Y., Futai, M. (1999) Identification and cloning of rat

galectin-2: expression is predominantly in epithelial cells of the stomach. *Arch. Biochem. Biophys.* **361**, 195-201.

Ozaki, K., Tanaka, T. (2005) Genome-wide association study to identify SNPs conferring risk of myocardial infarction and their functional analyses. *Cell. Mol. Life Sci.* **62**, 1804-1813.

Ozaki, K., Inoue, K., Sato, H., Iida, A., Ohnishi, Y., Sekine, A., Sato, H., Odashiro, K., Nobuyoshi, M., Hori, M., Nakamura, Y., Tanaka, T. (2004) Functional variation in LGALS2 confers risk of myocardial infarction and regulates lymphotoxin-α secretion *in vitro*. *Nature* **429**, 72-75.

Pavelka, M. (1997) Topology of glycosylation: a histochemist's view. In: *Glycosciences: Status and Perspectives*, eds. Gabius H.-J., Gabius S., pp 115-120. Chapman & Hall, Weinheim, London.

Phillips, B., Knisley, K., Weitlauf, K. D., Dorsett, J., Lee, V., Weitlauf, H. (1996) Differential expression of two β-galactoside-binding lectins in the reproductive tracts of pregnant mice. *Biol. Reprod.* **55**, 548-558.

Plzák, J., Betka, J., Smetana, K. Jr., Chovanec, M., Kaltner, H., André, S., Kodet, R., Gabius, H.-J. (2004) Galectin-3: an emerging prognostic indicator in advanced head and neck carcinoma. *Eur. J. Cancer* **40**, 2324-2330.

Poirier, F., Robertson, E. J. (1993) Normal development of mice carrying a null mutation in the gene encoding the L14 S-type lectin. *Development* **119**, 1229-1236.

Poirier, F., Timmons, P. M., Chan, C.-T. J., Guénet, J.-L., Rigby, P. W. J. (1992) Expression of the L14 lectin during mouse embryogenesis suggests multiple roles during pre- and post-implantation development. *Development* **115**, 143-155.

Polyak, K., Xia, Y., Zweier, J. L., Kinzler, K. W., Vogelstein, B. (1997) A model for p53-induced apoptosis. *Nature* **389**, 300-305.

Purkrábková, T., Smetana, K. Jr., Dvořánková, B., Holíková, Z., Böck, C., Lensch, M., André, S., Pytlík, R., Liu, F.-T., Klíma, J., Smetana, K., Motlík, J., Gabius, H.-J. (2003) New aspects of galectin functionality in nuclei of cultured bone marrow stromal and epidermal cells: biotinylated galectins as tool to detect specific binding sites. *Biol. Cell* **95**, 535-545.

Rappl, G., Abken, H., Muche, J. M., Sterry, W., Tilgen, W., André, S., Kaltner, H., Ugurel, S., Gabius, H.-J., Reinhold, U. (2002) $CD4^+CD7^-$ leukemic T-cells from patients with Sézary syndrome are protected from galectin-1-triggered T-cell death. *Leukemia* **16**, 840-845.

Reuter, G., Gabius, H.-J. (1999) Eukaryotic glycosylation: whim of nature or multipurpose tool? *Cell. Mol. Life Sci.* **55**, 368-422.

Rotblat, B., Niv, H., André, S., Kaltner, H., Gabius, H.-J., Kloog, Y. (2004) Galectin-1(L11A) predicted from a computed galectin-1 farnesyl-binding pocket selectively inhibits Ras-GTP. *Cancer Res.* **64**, 3112-3118.

Roth, J. (1996) Protein glycosylation in the endoplasmic reticulum and the Golgi apparatus and cell-type specificity of cell surface glycoconjugates expression: analysis by protein A-gold and lectin-gold techniques. *Histochem. Cell Biol.* **106**, 79-92.

Rüdiger, H., Gabius, H.-J. (2001) Plant lectins: occurrence, biochemistry, function and application. *Glycoconjugate J.* **18**, 589-613.

Saal, I., Nagy, N., Lensch, M., Lohr, M., Manning, J. C., Decaestecker, C., André, S., Kiss, R., Salmon, I., Gabius, H.-J. (2005) Human galectin-2: expression profiling by RT-PCR/immunohistochemistry and its introduction as a histochemical tool for ligand localization. *Histol. Histopathol.* **20**, 1191-1208.

Sato, M., Nishi, N., Shoji, H., Kumagai, M., Imaizumi, T., Hata, Y., Hirashima, M., Suzuki, S., Nakamura, T. (2002) Quantification of galectin-7 and its localization in adult mouse tissues. *J. Biochem.* **131**, 255-260.

Saussez, S., Cucu, D.-R., Decaestecker, C., Chevalier, D., Kaltner, H., André, S., Wacreniez, A., Toubeau, G., Camby, I., Gabius, H.-J., Kiss, R. (2006) Galectin 7 (p53-induced gene 1): a new prognostic predictor of recurrence and survival in stage IV hypopharyngeal cancer. *Ann. Surg. Oncol.* **13**, 999-1009.

Sheikholeslam-Zadeh, R., Decaestecker, C., Delbrouck, C., Danguy, A., Salmon, I., Zick, Y., Kaltner, H., Hassid, S., Gabius, H.-J., Kiss, R., Choufani, G. (2001) The levels of expression of galectin-3, but not of galectins-1 and -8, correlate with apoptosis in human cholesteatoma. *Laryngoscope* **111**, 1042-1047.

Shimura, T., Takenaka, Y., Fukumori, T., Tsutsumi, S., Okada, K., Hogan, V., Kikuchi, A., Kuwano, H., Raz, A. (2005) Implication of galectin-3 in Wnt signaling. *Cancer Res.* **65**, 3535-3537.

Siebert, H.-C., Adar, R., Arango, R., Burchert, M., Kaltner, H., Kayser, G., Tajkhorshid, E., von der Lieth, C.-W., Kaptein, R., Sharon, N., Vliegenthart, J. F. G., Gabius, H.-J. (1997) Involvement of laser photo-CIDNP (chemically induced dynamic nuclear polarization)-reactive amino acid side chains in ligand binding by galactoside-specific lectins in solution. *Eur. J. Biochem.* **249**, 27-38.

Siebert, H.-C., André, S., Lu, S.-Y., Frank, M., Kaltner, H., van Kuik, J. A., Korchagina, E. Y., Bovin, N. V., Tajkhorshid, E., Kaptein, R., Vliegenthart, J. F. G., von der Lieth, C.-W., Jiménez-Barbero, J., Kopitz, J., Gabius, H.-J. (2003) Unique conformer selection of human growth regulatory lectin galectin-1 for ganglioside GM_1 versus bacterial toxins. *Biochemistry* **42**, 14762-14773.

Siebert, H.-C., Rosen, J., Seyrek, K., Kaltner, H., André, S., Bovin, N. V., Nyholm, P. G., Sinowatz, F., Gabius, H.-J. (2006) α2,3/α2,6-Sialylation of N-glycans: non-synonymous signals with marked developmental regulation in bovine reproductive tracts. *Biochimie* **88**, 399-410.

Smetana, K. Jr., Dvořánková, B., Chovanec, M., Bouček, J., Klíma, J., Motlík, J., Lensch, M., Kaltner, H., André, S., Gabius, H.-J. (2006) Nuclear presence of adhesion-/growth-regulatory galectins in normal/malignant cells of squamous epithelial origin. *Histochem. Cell Biol.* **125**, 171-182.

Solís, D., Jiménez-Barbero, J., Kaltner, H., Romero, A., Siebert, H.-C., von der Lieth, C.-W., Gabius, H.-J. (2001) Towards defining the role of glycans as hardware in information storage and transfer: basic principles, experimental approaches and recent progress. *Cells Tissues Organs* **168**, 5-23.

Spicer, S. S., Schulte, B. A. (1992) Diversity of cell glycoconjugates shown histochemically: a perspective. *J. Histochem. Cytochem.* **40**, 1-38.

Sturm, A., Lensch, M., André, S., Kaltner, H., Wiedenmann, B., Rosewicz, S., Dignass, A. U., Gabius, H.-J. (2004) Human galectin-2: novel inducer of T cell apoptosis with distinct profile of caspase activation. *J. Immunol.* **173**, 3825-3837.

Takaishi, S., Wang, T. C. (2007) Gene expression profiling in a mouse model of *Helicobacter*-induced gastric cancer. *Cancer Sci.* **98**, 284-293.

Timmons, P. M., Colnot, C., Cail, I., Poirier, F., Magnaldo, T. (1999) Expression of galectin-7 during epithelial development coincides with the onset of stratification. *Int. J. Dev. Biol.* **43**, 229-235.

Timmons, P. M., Rigby, P. W. J., Poirier, F. (2002) The murine seminiferous epithelial cycle is pre-figured in the Sertoli cells of the embryonic testis. *Development* **129**, 635-647.

Unverzagt, C., André, S., Seifert, J., Kojima, S., Fink, C., Srikrishna, G., Freeze, H., Kayser, K., Gabius, H.-J. (2002) Structure-activity profiles of complex biantennary glycans with core fucosylation and with/without additional α2,3/α2,6 sialylation: synthesis of neoglycoproteins and their properties in lectin assays, cell binding, and organ uptake. *J. Med. Chem.* **45**, 478-491.

Villalobo, A., Nogales-González, A., Gabius, H.-J. (2006) A guide to signaling pathways connecting protein-glycan interaction with the emerging versatile effector functionality of mammalian lectins. *Trends Glycosci. Glycotechnol.* **18**, 1-37.

Wang, W., Park, J. W., Wang, J. L., Patterson, R. J. (2006) Immunoprecipitation of spliceosomal RNAs by antisera to galectin-1 and galectin-3. *Nucleic Acids Res.* **34**, 5166-5174.

Wasano, K., Hirakawa, Y., Yamamoto, T. (1990) Immunohistochemical localization of 14 kDa β-galactoside-binding lectin in various organs of rat. *Cell Tissue Res.* **259**, 43-49.

Watkins, W. M. (1999) A half century of blood-group antigen research: some personal recollections. *Trends Glycosci. Glycotechnol.* **11**, 391-411.

Wollina, U., Lange, D., Paus, R., Burchert, M., Gabius, H.-J. (2000) Expression of galectins-1 and -3 and of accessible binding sites during murine hair cycle. *Histol. Histopathol.* **15**, 85-94.

Wu, A. M., Wu, J. H., Tsai, M.-S., Liu, J.-H., André, S., Wasano, K., Kaltner, H., Gabius, H.-J. (2002) Fine specificity of domain-I of recombinant tandem-repeat-type galectin-4 from rat gastrointestinal tract (G4-N). *Biochem. J.* **367**, 653-664.

Wu, A. M., Wu, J. H., Liu, J.-H., Singh, T., André, S., Kaltner, H., Gabius, H.-J. (2004) Effects of polyvalency of glycotopes and natural modifications of human blood group ABH/Lewis sugars at the Galβ1-terminated core saccharides on the binding of domain-I of recombinant tandem-repeat-type galectin-4 from rat gastrointestinal tract (G4-N). *Biochimie* **86**, 317-326.

Wu, A. M., Singh, T., Wu, J. H., Lensch, M., André, S., Gabius, H.-J. (2006) Interaction profile of galectin-5 with free saccharides and mammalian glycoproteins: probing its fine specificity and the effect of naturally clustered ligand presentation. *Glycobiology* **16**, 524-537.

Wu, A. M., Singh, T., Liu, J.-H., Krzeminski, M., Russwurm, R., Siebert, H.-C., Bonvin A. M. J. J., André, S., Gabius, H.-J. (2007) Activity-structure correlations in divergent lectin evolution: fine specificity of chicken galectin CG-14 and computational analysis of flexible ligand docking for CG-14 and the closely related CG-16. *Glycobiology* **17**, 165-184.