THE ANATOMICAL RECORD 294:427–444 (2011)

Toward Comprehensive Analysis of the Galectin Network in Chicken: Unique Diversity of Galectin-3 and Comparison of its Localization Profile in Organs of Adult Animals to the Other Four Members of this Lectin Family

HERBERT KALTNER,[1*] DIETER KÜBLER,[2*] LARA LÓPEZ-MERINO,[3] MICHAELA LOHR,[1] JOACHIM C. MANNING,[1] MARTIN LENSCH,[1] JOERG SEIDLER,[4] WOLF D. LEHMANN,[4] SABINE ANDRÉ,[1] DOLORES SOLÍS,[3,5] AND HANS-JOACHIM GABIUS[1]

[1]Faculty of Veterinary Medicine, Institute of Physiological Chemistry, Ludwig-Maximilians-University, Munich, Germany
[2]Biomolecular Interactions, German Cancer Research Center, Heidelberg, Germany
[3]Instituto de Química Física Rocasolano, CSIC, Madrid, Spain
[4]Molecular Structure Analysis, German Cancer Research Center, Heidelberg, Germany
[5]Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Bunyola, Mallorca, Illes Balears, Spain

ABSTRACT

Characterization of all members of a gene family established by gene divergence is essential to delineate distinct or overlapping expression profiles and functionalities. Their activity as potent modulators of diverse physiological processes directs interest to galectins (endogenous lectins with β-sandwich fold binding β-galactosides and peptide motifs), warranting their study with the long-term aim of a comprehensive analysis. The comparatively low level of complexity of the galectin network in chicken with five members explains the choice of this organism as model. Previously, the three proto-type chicken galectins CG-1A, CG-1B, and CG-2 as well as the tandem-repeat-type CG-8 had been analyzed. Our study fills the remaining gap to determine gene structure, protein characteristics and expression profile of the fifth protein, that is, chimera-type chicken galectin-3 (CG-3). Its gene has a unique potential to generate variants: mRNA production stems from two promoters, alternative splicing of the form from the second transcription start point (tsp) can generate three mRNAs. The protein with functional phosphorylation sites in the N-terminus generated by transcription from the first tsp (tsp1CG-3) is the predominant CG-3 type present in adult tissues. Binding assays with neoglycoproteins and cultured cells disclose marked similarity to properties of human galectin-3. The expression and localization profiles as well as proximal promoter regions have characteristic features distinct from the other four CGs. This information on CG-3 completes the description

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Spanish Ministry of Science and Innovation; Grant number: BFU2009-10052; Grant sponsor: CIBER of Respiratory Diseases (CIBERES) (*ISCII*).

*Correspondence to: Herbert Kaltner, Faculty of Veterinary Medicine, Institute of Physiological Chemistry, Ludwig-Maximilians-University, Veterinärstr. 13, 80539 Munich, Germany. Fax: +49-89-21802508. E-mail: Kaltner@lmu.de or Dieter Kübler, Biomolecular Interactions, German Cancer Research Center, Im Neuenheimer Feld 581, 69120 Heidelberg, Germany. E-mail: d.kuebler@dkfz.de

Received 7 September 2010; Accepted 16 November 2010

DOI 10.1002/ar.21341
Published online 2 February 2011 in Wiley Online Library (wileyonlinelibrary.com).

of the panel of CGs, hereby setting the stage for detailed comparative analysis of the entire CG family, e.g., in embryogenesis. Anat Rec, 294:427–444, 2011. © 2011 Wiley-Liss, Inc.

Key words: epithelium; lectin; macrophages; phylogeny; promoter region

INTRODUCTION

The histochemical analysis of tissue glycoconjugates has revealed a large diversity of different glycan structures, their presentation tightly regulated and associated to features such as cell type and degree of differentiation in a fingerprint-like manner (Spicer and Schulte, 1992; Danguy et al., 1994; Lohr et al., 2010). Owing to the unsurpassed capacity of carbohydrate oligomers for high-density information storage they can thus be likened to biochemical signals (code words), which have a significant bearing on cell sociology (for recent reviews covering biochemical and medical aspects of the sugar code, please see Gabius, 2009). These insights gained from glycan profiling with sugar-specific probes such as plant lectins open the eyes to the physiological potential of carbohydrate-protein (lectin) interactions *in situ*, thus guiding efforts to detect and localize endogenous lectins. They are capable of forming intermolecular contacts (e.g., in adhesion) and to initiate signaling following the recognition of cognate glycan determinants, thereby triggering e.g., cell growth control (Villalobo et al., 2006; Gabius, 2009). On the side of the glycan, especially the spatially accessible branch-end structures are contact sites. Among tissue lectins, the family of galectins plays a prominent role in targeting such epitopes with a β-galactoside core (Barondes, 1984; Gabius, 1987; Kasai and Hirabayashi, 1996; Schwartz-Albiez, 2009).

These proteins share a common sequence signature and the β-sandwich fold, reflecting their phylogenetic relationship and conservation of a set of amino acids essential for ligand binding (Cooper, 2002; Houzelstein et al., 2004; López-Lucendo et al., 2004; please see also Supporting Information, Fig. 1). The process of gene diversification has resulted in more than 10 different family members in mammals divided into three subgroups. To start addressing fundamental questions on structural and functional aspects of intrafamily diversity a model organism with reduced level of complexity would be an attractive study object. This reasoning has led us to focus on the chicken galectins (CGs). They are comprised of a total of five members, that is, three homodimeric proto-type proteins termed CG-1A, CG-1B and CG-2 (Beyer et al., 1980; Oda and Kasai, 1983; Sakakura et al., 1990; Varela et al., 1999; Kaltner et al., 2008; López-Lucendo et al., 2009), one tandem-repeat-type protein most closely related to mammalian galectin-8, thus termed CG-8 (Kaltner et al., 2009), and the chimera-type CG-3. Whereas the other four CGs have already been examined with respect to gene structure and expression profiling, these issues have not yet been addressed in detail for CG-3, which defines the aim of this report.

The existence of CG-3 was first detected in extracts of normal and Rous sarcoma-virus-transformed secondary chicken embryo fibroblasts with an antibody against murine galectin-3 (Crittenden et al., 1984). cDNA cloning from hypertrophic tibial chondrocytes of 14-day-old embryos and from bone-marrow-derived osteoclasts of calcium-deficient chicken led to a respective sequence and two variants referred to as CG-3-TM1/TR1, with different lengths in Northern blotting (starting at 1.3 kb of low abundance and then reaching 1.5 kb for CG-3-TM1 and 2.0 kb for CG-3-TR1) (Nurminskaya and Linsenmayer, 1996; Gorski et al., 2002). The comparison of this information with gene expression of mammalian galectin-3 teaches the lesson that presence of several mRNAs is rather common for the galectin-3 gene. In detail, its transcription can be initiated from more than one site, with marked implications for the tightly regulated selection of splice points defining the length of the first untranslated exon of murine galectin-3 (Cherayil et al., 1989; Voss et al., 1994; Gaudin et al., 1997; Kadrofske et al., 1998). But there is a further cause for mRNA diversity from this gene. Of particular note, the human gene for galectin-3 has a second, spatially separated promoter with its own functional transcription start point (tsp). In addition to the mentioned region upstream of the first untranslated exon (tsp1) the tsp2 lies in the second intron, in the case of the human gene used at low abundance and strictly regulated (Raimond et al., 1995; Guittaut et al., 2001). Two productive open reading frames of the resulting mRNAs with a length of 318 or 291 nucleotides, respectively, are out-of-frame with the regular galectin-3-coding sequence, a third in the fourth exon in-frame but so far not reported to be translated (Guittaut et al., 2001). In view of this unique complexity for mammalian galectin-3 genes, the relation between gene structure and transcript nature needed to be firmly established in chicken.

Mammalian galectin-3 has a unique tridomain structure with the N-terminal region, a collagenase-sensitive section with Gly/Pro-rich tandem repeats and the carbohydrate recognition domain (CRD). Given that the so far reported CG-3 sequences mentioned above lack a hallmark of mammalian galectin-3, that is, the sites for serine phosphorylation in the N-terminal section (Fukumori et al., 2007; Szabo et al., 2009), we surmised that the known CG-3 sequences might represent variants. By strategically merging several experimental techniques we herein resolved this problem. At the outset, we took advantage of the well-characterized genomic organization of mammalian galectin-3 with its six exons of characteristic lengths as reliable guideline. After we revealed that the CG-3 gene maintains this feature, transcription of the CG-3 gene from two tsps (tsp1, tsp2) could be

delineated. The mRNA originating from the tsp2 is then subject to alternative splicing, generating a total of three CG-3 forms (tsp2CG-3I/II/III). In contrast to mammalian galectin-3, the N-terminal sequence of the tsp2CG-3I protein was not phosphorylated. N-Terminal phosphorylation and lectin activity were yet ascertained for tsp1CG-3 protein. It turned out to be the predominant CG-3 protein in adult chicken organs. Expression and localization profiles were thus determined to enable comparisons, between CGs and between avian/mammalian forms of galectin-3.

MATERIALS AND METHODS

Computational Sequence Analyses

Chromosomal environment and genomic organization of CG-3 as well as mammalian and *Danio rerio* genes for galectin-3 were comparatively analyzed using the UCSC Genome Browser Gateway (http://genome.ucsc.edu/cgi-bin/hgGateway), the Ensembl Genome Browser (http://www.ensembl.org/index.html) and/or the NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview). DNA sequences were edited by the EditSeq sequence analysis software version 7.1.0 (DNAstar; 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). The known EST sequence (dbEST Id 37586824, represented on the cDNA level by the GenBank entry EF429082.1) and the further available GenBank entries U50339, AF479564.1/AF479565.1 as well as, independently, a previously defined consensus sequence for CGs (Kaltner et al., 2008) ascertained the location of the single gene. To identify putative binding sites for transcription factors in the proximal promoter regions of the CG-3 gene and its second promoter as well as of the genes for human, mouse and rat galectin-3, the sequences stretching from 2000 base pairs (bp) upstream to 150 bp downstream of the starting point for transcription were subjected to processing by the programs Match and P-Match using distinct presetting and stringency criteria to limit the occurrence of false-positive cases as described previously for the other four CGs (Kaltner et al., 2008, 2009).

Cloning, Protein Expression and Purification

As reported for CG-2 (Kaltner et al., 2008), total RNA from embryonic kidney was isolated using the RNeasy® kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, and 2.5 μg were used as template. Rapid amplification of cDNA ends (RACE) was performed with the reagents of the SMART RACE cDNA amplification kit (Clontech, Heidelberg, Germany) and the company's Advantage™ 2 PCR kit together with two CG-3-gene-specific primers covering the 793-819 bp (3′RACE) and the 894-926 bp (5′RACE) sequence stretch within the GenBank entry U50339. The resulting products of 470 bp (3′RACE) and 926 bp (5′RACE) were sequenced. This information led to the design of primers to obtain CG-3-specific cDNA covering the entire sequence of 726 bp by PCR amplification. The reaction was directed by the sense primer 5′-CGTACGCATATGTCGGACGGTTTCTCTC-3′ with an internal *NdeI* restriction site (underlined) and the antisense primer 5′-CGCTAGGGATCCTTAAATCATGGAGGTCAAAAC-3′ with an internal *BamHI* restriction site (underlined). cDNA specific for the protein tsp2CG-3I was amplified from *v-src*-transformed chicken F6CC-PR9692 embryonic fibroblasts (kindly provided by Jiří Plachý, Institute of Molecular Genetics AS CR, Prague, Czech Republic) using the sense primer 5′-GACATATGCAGGCCATGAAGG-3′ with an internal *NdeI* restriction site (underlined) and the antisense primer 5′-CTGAAGCTTTTAAATCATGGAGGTCAAAAC-3′ with an internal *HindIII* restriction site (underlined). To obtain a cDNA clone coding for CG-3 protein as produced by proteolytic truncation with cleavage of the Thr93/Ala94 peptide bond, that is without the N-terminal stretch and the collagenase-sensitive stalk, CG-3-specific cDNA was used as a template. A sequence stretch ranging from nucleotide number 280 to 726 was amplified with the sense primer 5′-CATATGGCACCGTACTCTGAAGCTCC-3′ containing an *NdeI* restriction site (underlined) and with the antisense primer 5′-GTCGACTTAAATCATGGAGGTCAAAACACTG-3′ containing a *SalI* restriction site (underlined). The reactions were performed with the Expand High Fidelity PCR system as recommended by the manufacturer (Roche, Penzberg, Germany). The amplification products were separated from PCR reagents by gel electrophoresis in 2% agarose. They were separately purified using a gel extraction kit (Qiagen), ligated into the *EcoRV*-linearized pET-Blue-1 AccepTor™ vector (Novagen, Darmstadt, Germany) with single 3′dU overhangs and propagated in this company's *E. coli* strain NovaBlue. Subsequently, the CG-3-specific cDNA was ligated into *NdeI*/*BamHI*-treated expression vector pET-12a (Novagen), the tsp2CG-3I-specific coding sequence into the expression vector pET-26b (Novagen) digested with *NdeI*/*HindIII*, and coding sequence for truncated CG-3 in the expression vector pET-24a pretreated with *NdeI*/*SalI*. The products were used for transformation of the *E. coli* strains BL21(DE3)-pLysS (Novagen; pET12a) and RosettaBlue™(DE3)pLysS (Novagen; pET26b, pET24a), respectively. These plasmids facilitated recombinant production of the protein in the respective *E. coli* strain. Optimal yields of lectin production with 25–30 mg in the case of CG-3, 10–15 mg in the case of tsp2CG-3I and 30–40 mg in the case of the truncated version of CG-3 per liter of culture medium were obtained with TB medium (Roth, Karlsruhe, Germany) at 37°C and induction with 100 μM isopropyl β-thio-D-galactoside after thawing frozen bacteria in 7 mL lysis buffer (20 mM phosphate-buffered saline, pH 7.2, containing 2 mM ethylendiaminetetraacetic acid and 4 mM β-mercaptoethanol) per gram (wet weight). Affinity chromatography using lactosylated Sepharose 4B, prepared by ligand conjugation to divinyl-sulfone-activated resin, was used as crucial step in purification, as described previously for other CGs (Kaltner et al., 2008).

Analytical Procedures

Gel electrophoretic analysis of recombinant proteins was routinely performed in a 4–15% linear gradient system with silver staining to visualize any contaminations. Mass spectrometric fingerprint analysis of the purified protein was carried out either without any chemical treatment or with full reduction of disulfide bonds and alkylation with iodoacetamide (Solís et al., 2010). Samples were dialyzed exhaustively against 50 mM

ammonium bicarbonate, digested with modified porcine trypsin (sequencing grade; Promega, Mannheim, Germany), and analyzed by MALDI-MS using an Ultraflex MALDI-TOF/TOF mass spectrometer (Bruker-Daltonik, Bremen, Germany) as described (López-Lucendo et al., 2009). For analysis of the approximately 10 kDa N-terminal tryptic peptide of CG-3, a matrix solution of 15 g/L α-cyano-4-hydroxycinnamic acid in 50% aqueous acetonitrile and 0.15% trifluoroacetic acid was used, and mass spectra were recorded in linear positive mode at 25 kV acceleration voltage and 1.7 kV in the linear detector by accumulating 800 spectra of single laser shots. The equipment was calibrated employing cytochrome c singly and doubly charged mass signals. The analysis of mass data was performed using the flexAnalysis 2.2 software (Bruker-Daltonik). Phosphorylated, proteolytically truncated and biotinylated proteins were in-gel digested with trypsin or chymotrypsin prior to phosphopeptide enrichment by Ga(III)-immobilized metal ion affinity chromatography, if applicable (Lehmann et al., 2006; Seidler et al., 2009). Then reversed phase nano ultrahigh performance liquid chromatography (UPLC)-MS/MS analysis using a nanoAcquity UPLC system (Waters, Milford, MA) combined with a QTOF instrument type Q-Tof2 (Waters Micromass, Manchester, UK), estimation of degree of phosphorylation at the two sites and databank searches were carried out, as described (Kübler et al., 2008; Seidler et al., 2009; Winter et al., 2009).

Phosphorylation Assays

Enzymatic phosphorylation with three different protein kinases, i.e., casein kinases (CK) CK-1 and -2 (New England Biolabs, Frankfurt, Germany) and PKA (catalytic subunit of the cAMP-dependent protein kinase A; kindly provided by N. König and D. Bossemeyer, German Cancer Research Center, Heidelberg), was performed as described for human galectin-3 (Kübler et al., 2008). Following one-dimensional lithium dodecyl sulfate gel electrophoresis in precast 4-12% NuPAGE Bis-Tris mini gels (Invitrogen, Karlsruhe, Germany) product formation was detected either by autoradiography/phosphoimaging when [γ^{32}P]ATP (~110 TBq/mol; Amersham, Braunschweig, Germany) at 10 μM was used or by the ProQ-Diamond phosphoprotein gel stain (Invitrogen).

Binding Assays

CG-3, its truncated version and human galectin-3 were labeled under activity-preserving conditions with the N-hydroxysuccinimide ester derivative of biotin (Sigma, Munich, Germany) (Gabius et al., 1991) and product analysis was performed by mass spectrometry to determine extent of label incorporation and identification of its sites (Purkrábková et al., 2003; Kübler et al., 2008). Binding of biotinylated lectin to neoglycoproteins, which present carbohydrate ligands covalently attached to the carrier protein bovine serum albumin free of any contamination by natural glycoproteins (Gabius et al., 1988, 1990), was quantitated spectrophotometrically in a solid-phase inhibition assay as described (André et al., 2004). Cell binding was analyzed by flow cytofluorometry using a panel of Chinese hamster ovary cell (CHO) lines with the Pro^{-5} parental line and Lec1/Lec2/Lec4/Lec8 mutant lines with defects in complex-type N-glycosylation, sialylation, β1,6-branching of N-glycans and galactosylation, respectively (kindly provided by P. Stanley, Albert Einstein College of Medicine, Bronx, NY), and with lines genetically engineered to express increased level of α2,6-sialylation (Kaltner et al., 2009) and of α1,2-fucosylation or presence of bisecting N-acetylglucosamine (GlcNAc) residue, respectively. Transfection of wild-type cells with cDNA coding for the respective glycosyltransferase (kindly provided by J. B. Lowe, University of Michigan, Ann Arbor, MI and P. Umana, GLYCART Biotechnology AG, Schlieren-Zürich, Switzerland, respectively) and clone selection using the plant lectins UEA-I or PHA-E as screening tool were carried out as described (André et al., 2006a). In addition, the human Capan-1 pancreatic carcinoma cell line with natural reactivity for galectin-3 regulatable by the tumor suppressor p16^{INK4a} was included (André et al., 2007a). Cell culture and staining, data acquisition and processing were performed exactly as described previously for human galectin-3 and other CGs (André et al., 2007a; Kaltner et al., 2008, 2009). Controls for carbohydrate-specific binding comprised experiments in the presence of cognate sugar and treatment with noncognate sugar (mannose) as osmolarity control. Omission of the incubation step with labeled lectin enabled to measure the lectin-independent background (André et al., 2009a). Comparative analyses were routinely performed with aliquots of cell suspensions from the same batch, the inter-assay variability in percent did not exceed 11.2%.

Expression Profiling by RT-PCR, Western Blotting and Immunohistochemistry

The mRNAs specific for CG-3 and the tsp2CG-3 variants were probed in RT-PCR analyses with cDNA preparations of a panel of tissues and the cultured fibroblasts, first applying the sense primer 5′-CCCGGCG TACCCTGGATA-3′ and the antisense primer 5′-TTAAAT-CATGGAGGTCAAAACAC-3′, resulting in amplification products of 625 bp (CG-3, tsp2CG-3I), 835 bp (tsp2CG-3II) or 1076 bp (tsp2CG-3III). To distinguish CG-3- and tsp2CG-3I/II/III-specific coding sequences the sense primers 5′-ATGTCGGACGGTTTCTCT-3′ in the case of CG-3 and 5′-ATGCAGCCCATGAAGGC-3′ for the tsp2CG-3 forms were applied. Combining the respective sense primer with the antisense primer 5′-TTAAATCATGGAG-GTCAAAACAC-3′, the lengths of amplified cDNAs for either CG-3 (726 bp) or tsp2CG-3I (789 bp) were calculated to be sufficiently different for separation. The loading control with chicken β-actin-specific mRNA was established with the sense primer 5′-GATGATGATAT TGCTGCGC-3′ and the antisense primer 5′-GGTGAA GCTGTAGCCTC-3′. A polyclonal anti-CG-3 antibody preparation free of cross-reactivity against the other four CGs was obtained by immunizing rabbits, monitoring the antibody titer regularly by ELISAs to determine optimal timing for booster injections and blood drawing, purifying the immunoglobulin G (IgG) fraction from serum by protein A-Sepharose 4B affinity chromatography and consecutively removing any material cross-reactive to the other four CGs from the IgG fraction by chromatographic affinity depletion on resin covalently loaded with CG-1A, CG-1B, CG-2, or CG-8 at densities of 7.5–10.2 mg/mL (Kaltner et al., 2008). Complete absence of contaminating cross-reactivity was ascertained by ELISA and Western blot

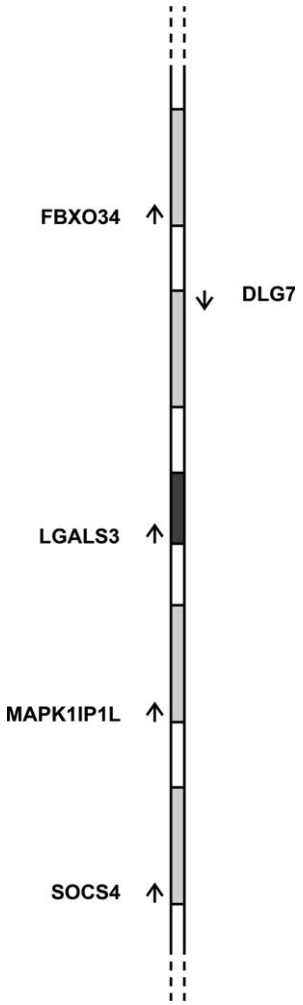

Fig. 1. Chromosomal environment of the CG-3 gene (dark gray box) on chromosome 5. Gene orientation is indicated by arrows; box lengths and spacers are not drawn to scale. Abbreviations: LGALS3, CG-3; DLG7, discs, large homolog 7 (*Drosophila*); FBXO34, F-box only protein 34; MAPK1IP1L, mitogen-activated protein kinase 1-interacting protein 1-like; SOCS4, suppressor of cytokine signaling 4.

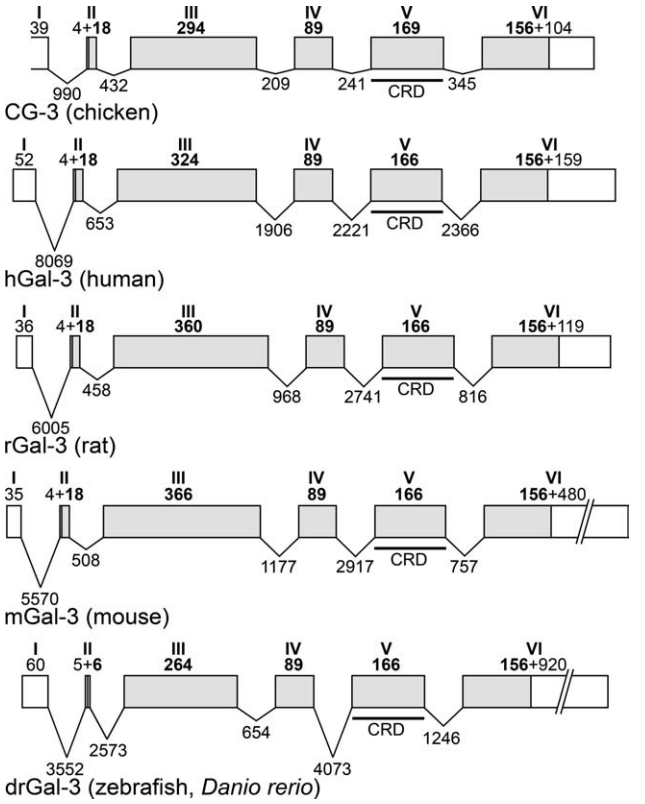

Fig. 2. Organization of the CG-3 gene and the comparison to the respective features of four genes for mammalian/fish galectin-3 (i.e., from the human (hGal-3), rat (rGal-3), mouse (mGal-3) and zebrafish (drGal-3) genomes). All genes are constituted by six exons (Roman numerals), translation of the common form in each case starting within the second 22 (11) bp-long exon and ending within the last exon. Translated exons are drawn as gray-colored boxes. They are connected by lines, representing introns. The exon coding for the CRD is underlined. The dimensions of each box are correlated to exon length (given in Arabic numerals), while the lines for introns are not drawn to scale, numbering at each line providing the information on the precise length in bp. In the scheme for CG-3, the box representing the untranslated exon I depicts the sequence information for the longest CG-3-specific EST clone reported in the database.

analysis (Lensch et al., 2006). This antibody preparation further refined by affinity purification with CG-3 as ligand was applied in immunoprecipitation/Western blot analyses with 2.5 mg total protein per tissue type as described in detail for CG-8 analysis (Kaltner et al., 2009). Immunohistochemical processing of fixed and paraffin-embedded sections of organs of adult (6-month-old) chicken followed an optimized protocol with stringent specificity controls set up previously for the analyses of the prototype and tandem-repeat-type CGs (Kaltner et al., 2008, 2009). Microphotographs were taken using an AxioImager.M1 microscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) equipped with an AxioCam MRc3 and the software Axiovision 4.6.

RESULTS AND DISCUSSION

Gene Structure of CG-3

Systematic searches in the chicken genome for a match separately based on the entries into the GenBank database and the galectin consensus sequence combined with typical galectin-3 features, especially the collagenase-sensitive repeat region, converged on one hit on chromosome 5 (minus strand). As in mammalian genomes (Hughes, 1994), galectin-3 in chicken is encoded by a single gene. The genes neighboring this locus were also identified to enable a comparative analysis (Fig. 1). Identical arrangements were found in the genome of *Homo sapiens* (on chromosome 14), *Bos taurus* (on chromosome 10), *Taeniopygia guttata* (on chromosome 5) and *Anolis carolinensis*, marked similarities in the case of *Mus musculus* (on chromosome 14) or *Rattus norvegicus* (on chromosome 15) and even *Danio rerio* (on chromosome 17, lacking DLG7). Evidently, the pattern of genes in the vicinity of the CG-3 locus is conserved. This finding let us expect maintenance of the structural profile of the gene's organization, a key factor to distinguish the typical CG-3 with its three different structural domains from any variants.

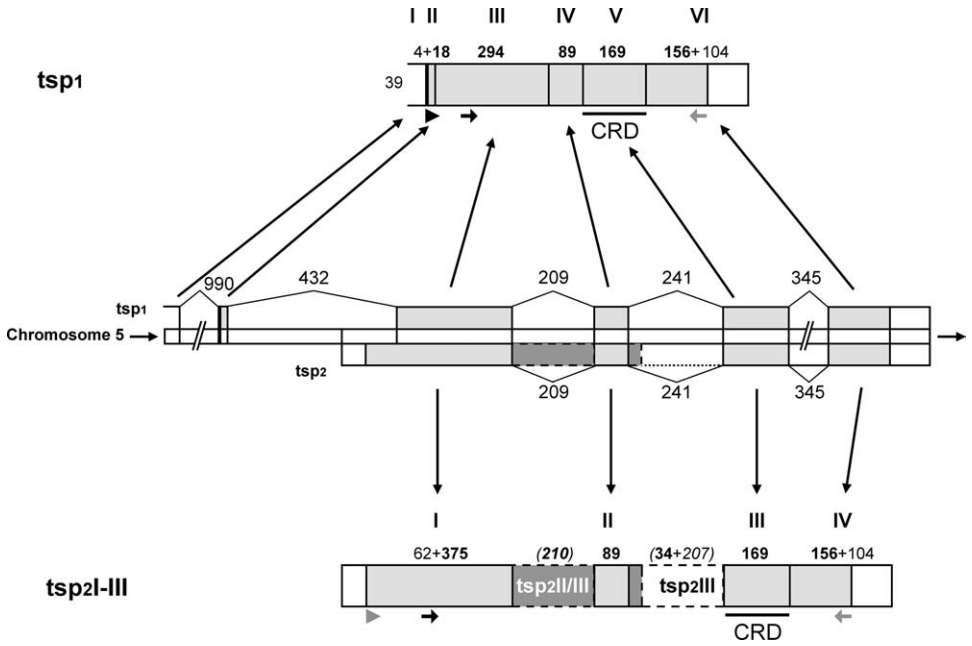

Fig. 3. Architecture of the four different mRNAs derived from the CG-3 gene. Two transcription start points (tsps), that is, tsp1CG-3 (upper part; please see also Fig. 2 for exon/intron display)/tsp2CG-3 (bottom part), are used. Additionally, alternative splicing generates the I–III forms originating from tsp2 (translated exons and the introns turned to an exon to its full extent (tsp2CG-3II) or in part (tsp2CG-3III) are shown as lightly or intensely gray-colored boxes, respectively). Transcription of the single CG-3 gene is thus initiated by two promoters. Whereas no variability in processing is seen for the tsp1-derived product (upper part), 209 bp- and 241 bp-long introns can successively be kept and turned into coding sequences, translation for tsp2CG-3III-specific mRNA will terminate within the second intron (bottom panel) after coding for 10 amino acids. A one-base insertion into the sequence of the first intron (C added in position 184) apparently facilitates completion of this intron to an in-frame open reading frame (Gorski et al., 2002). Sites of complementarity for the primer sets used in discriminatory RT-PCR analyses are depicted as black/gray arrow pair (separating product lengths of tsp1CG-3/tsp2CG-3I-specific mRNAs versus tsp2CG-3II/III mRNAs) and as black/gray arrowheads (forward priming)/gray arrow (backward priming) (separating product lengths on the grounds of the difference in site of transcription initiation, what results in disparities within the 5′-sections of the mRNAs).

The identified gene is composed of six exons, translation starting in the second exon with a 18-bp stretch (Fig. 2). The assumed phylogenetic lineage could be traced by adding information of mammalian and fish genes to Fig. 2. This diagram highlights the conservation of these features (exon number and length) among mammals, fish and birds, even including occurrence of the triplett splitting between exons IV/V. Evidently, a mRNA derived from this exon arrangement will code for the typical galectin-3 product of an organism. Its generation will be initiated at the tsp1, its translated protein will encompass the characteristic tridomain structure of galectin-3 with the N-terminal stretch, the collagenase-sensitive tail and the CRD, the latter encoded by exon V (highlighted in Fig. 2; the predicted amino acid sequence with the positions of the essential residues of the CRD is presented in Supporting Information, Fig. 1). Overall, such a protein can be referred to as CG-3, herewith posing the pertinent question on the origin of the three so far reported forms of cDNA for CG-3.

Sequence alignments and consideration of the lengths of the transcripts given above enabled to resolve this issue. The answer, graphically depicted in Fig. 3, is based on (i) use of a second tsp within the second intron of the CG-3 gene (tsp2) and (ii) alternative splicing of the primary transcript from this tsp. While pre-mRNA processing in the hypertrophic chondrocytes apparently removes the remaining three introns completely, the increased lengths in Northern blotting for the two variants from osteoclasts can be attributed to incomplete splicing (Fig. 3). Explicitly, the 210-bp intron is kept in the first form (formerly Gal-3-TM1), explaining the respective increase in length of the transcript (Gorski et al., 2002). The insertion does not alter the reading frame so that two proteins from the fully spliced pre-mRNA and this first variant form share substantial portions of identity in amino acid sequence with CG-3 (Supporting Information, Fig. 1). The third mRNA from tsp2 (formerly referred to as Gal-3-TR1) maintains the second intron (241 bp), too, together with the split-triplett usage at this site (Fig. 3). Its translation should yet terminate after 10 amino acids due to a stop codon. A resulting protein product will lack the CRD (Gorski et al., 2002; please see Supporting Information, Fig. 1 for the predicted amino acid sequence).

Summarizing the lessons provided by Fig. 3, a protein with properties typical for galectin-3 is likely to be produced. In addition, a second tsp and alternative splicing account for three variants. In contrast to the situation in the human galectin-3 gene, where two entirely different proteins arise from such a variant mRNA (here, translation can start from two initiation sites in the first section of exon III with +1/+2 shifts; Guittaut et al., 2001), and the origin of the tumor suppressor proteins

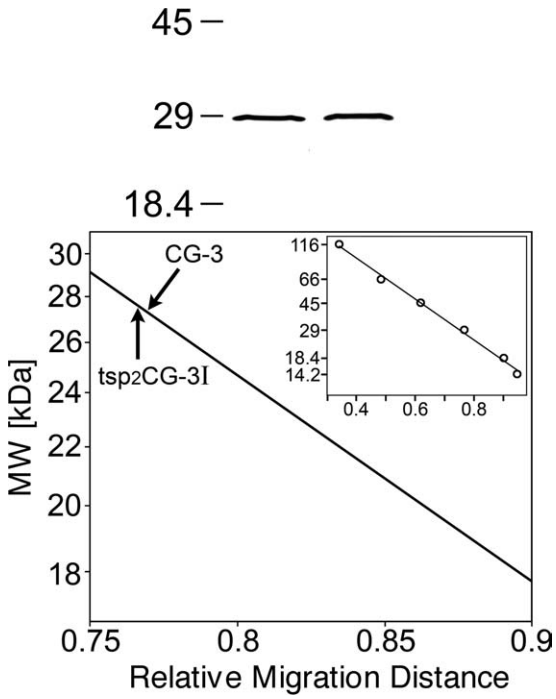

Fig. 4. Gel electrophoretic mobility of two CG-3 proteins. Purified proteins from recombinant production were separated by polyacrylamide gel electrophoresis under denaturing conditions in the presence of β-mercaptoethanol in a 4–15% linear gradient gel. The top panel shows the bands for CG-3 and the tsp2CG-3I variant (50 ng per lane). Molecular masses were determined by the plot given as an inset in the bottom panel. The inset shows the quantitative relationship between the relative migration distance and known molecular mass of the six standard proteins [β-galactosidase (116 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), β-lactoglobulin (18.4 kDa), and lysozyme (14.2 kDa)].

$p16^{INK4a}/p19^{ARF}$ which has set a precedent for protein diversity from one gene (Quelle et al., 1995), in-frame reading is operative. The insertion of one or two introns into the coding sequence derived from tsp2 will not perturb the reading frame of the variants. An event of in-frame alternative splicing had been reported previously in the case of the tandem-repeat-type CG-8, here in the region coding for the linker sequence (Kaltner et al., 2009). Looking at the predicted amino acid sequence of CG-3 (Supporting Information, Fig. 1), lectin activity is expected, and potential for proteolytic truncation by removal of the Gly/Pro-rich tandem-repeat region and for N-terminal phosphorylation is apparent. To prove these three characteristics recombinant production of CG-3 was established.

Protein Characterization

Protein purity was first assessed by gel electrophoresis (Fig. 4). To exclude any sequence deviation or post-translational processing mass spectrometric analysis was performed. A mass of 25,593 Da was determined, close to the calculated mass for a product with acetylated serine at the N-terminus of 25,583.04 Da. N-Terminal sequencing of gel-eluted protein confirmed this assumption and revealed susceptibility of this protein to limited degradation yielding products with ragged N-termini starting at positions 5, 8, 13, 17, and 25, respectively. Tryptic fingerprinting after full reduction and alkylation with iodoacetamide reached sequence coverage of 86%. It confirmed the predicted sequence excluding any post-translational modification (Table 1). However, in addition to the expected S-carbamidomethylation, several cases of apparent N- and O-alkylation were encountered upon chemical treatment of the peptides (+57.05 Da). This type of substitution can occur at aspartic acid, glutamic acid and histidine residues when processing protein digests (Boja and Fales, 2001). All detected peptides with potential for N- and/or O-alkylation contained such modified residues, peptide 1-104 harboring up to six substitutions. To attribute the mass change to the chemical treatment, we omitted the reagents for reduction/alkylation in separate experiments. Indeed, when the analysis of peptide 1-104 was carried out without prior reduction and alkylation of CG-3, a signal at m/z 10029.28, corresponding to the unmodified peptide, was observed. In addition, signals at m/z 10086.62, 10144.02, and 10200.54 were also visible, indicating that up to three N- and/or O-alkylations can occur already during the purification process, where iodoacetamide was present to preclude oxidation of sulfhydryls. With the predicted sequence confirmed and absence of any biochemical modification ascertained, this protein is thus suited for structural and functional analysis, for example, by crystallography or chemical mapping of the binding site to extend previous work with proto-type CGs (Solís et al., 1996; Siebert et al., 1997; Varela et al., 1999; Wu et al., 2001, 2007; López-Lucendo et al., 2009), as well as in assays to probe sensitivity for cleavage of the Pro/Gly-rich stalk and kinase-dependent phosphorylation.

Protein Truncation and Phosphorylation

Sequence comparisons of the Gly/Pro-rich repeat units in the stalk region of CG-3 had indicated pronounced similarity to the mammalian proteins (Cooper, 2002). Proteolytic truncation by collagenase trims mammalian galectin-3 to the CRD, with functional implications due to impairment of pentamerization in the presence of multivalent ligands as presented in microdomains (Kopitz et al., 2001, 2010; Ahmad et al., 2004). The respective activity of matrix metalloproteinase-9 on murine galectin-3 is the molecular event for preventing accumulation of late hypertrophic chondrocytes and increasing osteoclast recruitment during endochondral bone formation (Ortega et al., 2005). In fact, it can be considered as an irreversible regulatory event. This proven functionality warrants respective experiments. Treatment with collagenase under conditions suited for human galectin-3 yielded two closely neighboring bands in electrophoretic analysis at around 15–17 kDa (Fig. 5a). As determined by mass spectrometry, they result from cleavage of the peptide bonds Pro85/Gly86 or Thr93/Ala94, schematically shown in Supporting Information, Fig. 2. Thus, targeted degradation of CG-3, a potential molecular switch in the protein's functionality, is operative.

Turning to the N-terminal phosphorylation, assays using human galectin-3 as positive control and the truncated version of CG-3 as negative control showed strong phosphate incorporation into the human and avian full-length proteins with CK-1 (Fig. 5a,b). Presence of the

TABLE 1. Tryptic peptides of CG-3 detected by fingerprint analysis

m/z	Matching mass	#MC	Modification	Position	Peptide
1,166.58249	1,166.5880	0		180–189	FPFEPGTPFK
1,301.64599	1,301.6385	0		141–151	GQDIAFHFNPR
1,345.64543	1,345.6569	0	Cys_CAM: 194	190–200	LQVLCEGDHFK
1,457.75176	1,457.7396	1		140–151	RGQDIAFHFNPR
1,538.85371	1,538.8399	0		105–118	VPYDLPLPAGLMPR
1,595.87803	1,538.8399 + 57.052	0	O-alkylation: 108	105–118	VPYDLPLPAGLMPR
1,612.91959	1,612.9017	0		119–133	LLITITGTVNSNPNR
1,643.88623	1,643.8652	0		201–214	VAVNDAHLLQFNFR
1,648.7938	1,591.8267 + 57.052	1	O-alkylation: 183	176–189	TAPRFPFEPGTPFK
1,696.8655	1,696.7934	0	Cys_CAM: 162	159–172	VIVCNSMFQNNWGK
1,700.88866	1,643.8652 + 57.052		N- or O-alkylation: 105/107	201–214	VAVNDAHLLQFNFR
1,757.91784	1,643.8652 + 2 × 57.052		N- and O-alkylation: 105,107	201–214	VAVNDAHLLQFNFR
2,110.99464	2,110.9797	1	Cys_CAM: 162	159–175	VIVCNSMFQNNWGKEER
2,350.28462	2,350.2765	1		119–139	LLITITGTVNSNPNRFSLDFK
10,082.1	10,026.0325 + 57.052	0	N-/O-alkylation	1–104	acSDGFSLSDALPAHNPGAPPPQGWNRPPGPGAFPAY PGYPGAYPGAPGPYPGAPGPHHGPPGPYPGGPPGPY PGGPPGPYPGGPPGPYPGGPTAPYSEAPAAPLK
10,141.7	10,026.0325 + 2 × 57.052	0	N-/O-alkylation	1–104	acSDGFSLSDALPAHNPGAPPPQGWNRPPGPGAFPAY PGYPGAYPGAPGPYPGAPGPHHGPPGPYPGGPPGPY PGGPPGPYPGGPPGPYPGGPTAPYSEAPAAPLK
10,198.9	10,026.0325 + 3 × 57.052	0	N-/O-alkylation	1–104	acSDGFSLSDALPAHNPGAPPPQGWNRPPGPGAFPAY PGYPGAYPGAPGPYPGAPGPHHGPPGPYPGGPPGPY PGGPPGPYPGGPPGPYPGGPTAPYSEAPAAPLK
10,255.6	10,026.0325 + 4 × 57.052	0	N-/O-alkylation	1–104	acSDGFSLSDALPAHNPGAPPPQGWNRPPGPGAFPAY PGYPGAYPGAPGPYPGAPGPHHGPPGPYPGGPPGPY PGGPPGPYPGGPPGPYPGGPTAPYSEAPAAPLK
10,313.1	10,026.0325 + 5 × 57.052	0	N-/O-alkylation	1–104	acSDGFSLSDALPAHNPGAPPPQGWNRPPGPGAFPAY PGYPGAYPGAPGPYPGAPGPHHGPPGPYPGGPPGPY PGGPPGPYPGGPPGPYPGGPTAPYSEAPAAPLK
10,369.5	10,026.0325 + 6 × 57.052	0	N-/O-alkylation	1–104	acSDGFSLSDALPAHNPGAPPPQGWNRPPGPGAFPAY PGYPGAYPGAPGPYPGAPGPHHGPPGPYPGGPPGPY PGGPPGPYPGGPPGPYPGGPTAPYSEAPAAPLK

MC: missed cleavages; Cys_CAM: carboxyamidomethyl cysteine.

Mass spectrometric fingerprint analysis of CG-3 was carried out after complete reduction and alkylation with iodoacetamide. Theoretical peptide masses are given as $[M+H]^+$ and were calculated using the PeptideMass tool available at the ExPASy Proteomics Server (www.expasy.ch/tools/peptide-mass.html) and (i) monoisotopic masses of the amino acid residues for peptides below 3,500 Da and (ii) average masses for the ~10 kDa peptide. The complete amino acid sequence of CG-3, showing amino acid numbering and highlighting sequence coverage by peptide mass fingerprinting, is listed below.

```
                10         20         30         40         50         60
  1 acSDGFSLSDAL PAHNPGAPPP QGWNRPPGPG AFPAYPGYPG AYPGAPGPYP GAPGPHHGPP
                70         80         90        100        110        120
 61  GPYPGGPPGP YPGGPPGPYP GGPPGPYPGG PTAPYSEAPA APLKVPYDLP LPAGLMPRLL
               130        140        150        160        170        180
121  ITITGTVNSN PNRFSLDFKR GQDIAFHFNP RFKEDHKRVI VCNSMFQNNW GKEERTAPRF
               190        200        210        220        230        240
181  PFEPGTPFKL QVLCEGDHFK VAVNDAHLLQ FNFREKKLNE ITKLCIAGDI TLTSVLTSMI
```

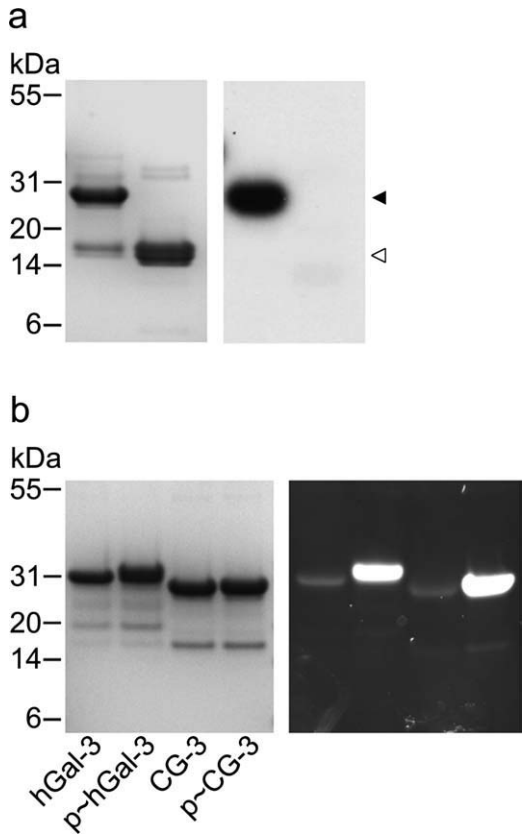

Fig. 5. Gel electrophoretic mobility of the CG-3 protein prior to and after treatment with collagenase (**a**, left) and extent of incorporation of radioactive ^{32}P-label into the two protein preparations in assays for phosphorylation with CK-1 (a, right); arrowheads mark position of full-length CG-3 (◀) or truncated CG-3 (◁), respectively. Comparison of extents of phosphorylation for human galectin-3 (h Gal-3) and CG-3 (5 μg each) after 15 min analyzed by gel electrophoresis and protein stain (**b**, left) or the phospho-stain (b, right). The gel mobility of molecular mass marker proteins are given on the left.

cognate sugar lactose did not affect CG-3's reactivity as CK-1 substrate (not shown). Tests with CK2/PKA were consistently negative. By LC-MS analysis of chymotryptic peptides from CK-1-treated CG-3, acceptor site(s) for phosphorylation was (were) assigned to positions Ser5/Ser7 with mono- and diphosphorylation (Fig. 6). As a consequence, no label incorporation into the truncated protein was detectable (Fig. 5a). Matching the CK1-phosphorylation consensus motif well, Ser5 appears as likely preferential target, this moiety not present in the tsp2-derived variants (please see Supporting Information, Fig. 1). Fittingly, tsp2CG-3I protein, whose purity is documented in Fig. 4, was not a substrate for any of the three kinases tested. N-Terminal phosphorylation, important for the antiapoptotic activity and also for a cooperation with the CRD in selection of glycoprotein ligands (Fukumori et al., 2007; Díez-Revuelta et al., 2010), is thus possible for CG-3 but not the variant sequence. The introduction of this post-translational modification can be considered as a diagnostic test for this aspect of the tridomain structure of a typical galectin-3 protein. Having documented phosphorylation and proteolytic truncation, we proceeded to experiments on ligand binding using labeled lectin. First, mass spectrometric analysis was performed to quantitate label incorporation and identify conjugation sites. It disclosed up to nine positions for biotin conjugation to lysine moieties in the sequence following the cleavage sites for proteolytic truncation, as summarized in Supporting Information, Figure 2. The labeled proteins were tested in solid-phase and cell assays to characterize their ligand-binding properties.

Ligand Binding

In essence, the affinity chromatography step in purification had already indicated CG-3 binding to lactose attached to the resin. Solid-phase assays with surface-presented neoglycoproteins extended the description of the lectin activity for the labeled lectin. Binding was specific depending on the nature of the carbohydrate (mannose and maltose were included as negative controls), saturable and blocked by haptenic sugar. The reactivity to N-acetyllactosamine (LacNAc) and histo-blood group ABH epitopes was comparatively stronger than to α2,3-sialylated LacNAc, as similarly seen for human galectin-3 tested in parallel, confirming this lectin's calorimetric binding data (Bachhawat-Sikder et al., 2001; Ahmad et al., 2002). α2,6-Sialylation of LacNAc impaired reactivity, extending the evidence for similarity to human galectin-3 (data not shown). Since the lectin will physiologically interact with natural glycans presented by N- and O-glycosylated proteins and by glycolipids and the affinity can be notably regulated by local clustering (Dam et al., 2005; Gabius, 2006), binding studies were next performed with cells *in vitro*. Probing into distinct structure/reactivity profiles is made rather convenient by the availability of the panel of CHO glycosylation mutant lines (Patnaik and Stanley, 2006). Quantitative determination of percentage of positive cells and mean fluorescence intensity in comparative measurements under identical conditions will signal reactivity differences due to a distinct change in the glycome profile.

Concentration dependence and sugar inhibition are illustrated for staining of the wild-type CHO cells by CG-3 and its truncated derivative in Fig. 7. The profiles for both proteins were very similar. Proteolytic truncation did not impair the activity of the CRD to target cell surface glycans. These reference wild-type cells lack enzymatic activities for α2,6-sialylation, α1,2/3/4-fucosylations and the synthesis of bisected N-glycans. Four glycosylation mutants and, in addition, three engineered transfectants with overexpression of certain glycosyltransferases enabled to measure the effect of distinct glycan alterations such as sialylation by α2,3/6-linkages on cell surfaces. The wild-type cells and human galectin-3 were the internal references (Fig. 8a). Cell binding was dependent on the presence of galactose (Fig. 8b). Impairment of synthesis of complex-type N-glycans revealed preference for these glycans as ligand (Fig. 8c). Neither O-glycans (mucin-type core 1 structures up to tetrasaccharides and O-fucosylated/glucosylated/mannosylated glycans) nor ganglioside GM3 (α2,3-sialylated lactosylceramide) as major glycolipid appeared to play a remarkable role as binding sites. This result directed further interest to measure the impact of modifications of N-glycans. In contrast to CG-8 (Kaltner et al., 2009),

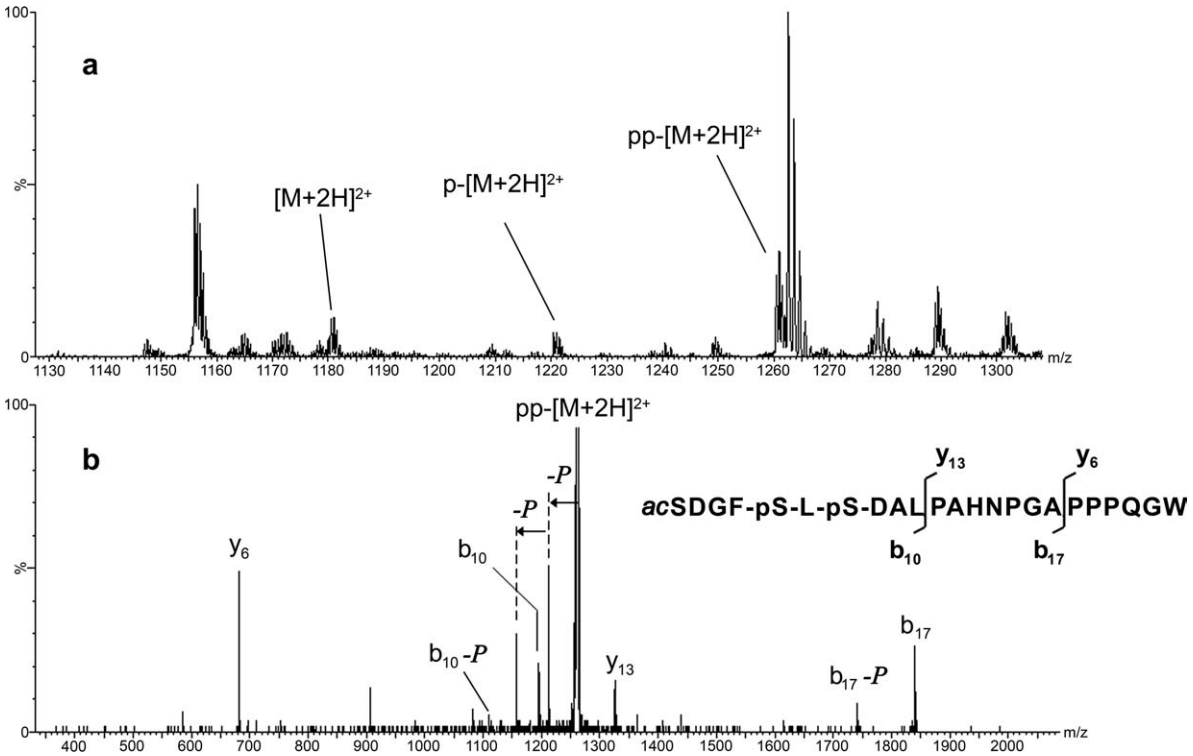

Fig. 6. Identification of the sites for CK-1-dependent CG-3 phosphorylation in the N-terminal peptide (positions 1-23 with the N-acetylated serine at position 1; please see also footnote to Table 1 and Supporting Information, upper part of Fig. 1). Survey LC-MS spectrum showing non- ($[M+2H]^{2+}$), mono- (p-$[M+2H]^{2+}$) and diphosphorylated (pp-$[M+2H]^{2+}$) peptide species eluting between 68.0 and 69.5 min (**a**). The MS/MS spectrum of the diphosphorylated peptide presents the main complementary fragment pairs b_{10}/y_{13} (b_{10}: decapeptide comprising the amino acids at positions 1–10, y_{13}: tridecapeptide comprising the amino acids at positions 11–23) and b_{17}/y_6 (b_{17}: heptadecapeptide comprising the amino acids at positions 1–17, y_6: hexapeptide comprising the amino acids at positions 18–23) confirming phosphorylation at Ser5 and Ser7 (**b**). *–P* indicates loss of phosphoric acid. Estimation of the degree of phosphorylation for the two sites in this peptide yielded 65% for Ser5 and 47% for Ser7.

the loss of α2,3-sialylation increased the fluorescence intensity (Fig. 8d). This result is in full accord with the data from the solid-phase assays, in which CG-8 but not CG-3 proved very reactive with α2,3-sialylated LacNAc (not shown). Introduction of α2,6-sialylation into N-glycans significantly reduced staining intensity, the α1,2-fucosylation augmented staining slightly (Fig. 8e,f), fully in line with the measured neoglycoconjugate reactivity. In both cases, the percentage of positive cells increased. These neo-epitopes for CHO cells apparently conveyed reactivity to cells by these two modes of branch-end tailoring, galectin-3 being reactive with α2,6-sialylated LacNAc repeats in N-glycans (Ahmad et al., 2002). When the degree of branching of N-glycans was increased, human galectin-3 had a slight preference for presence of β1,6-branching in triantennary structures (André et al., 2006b). Fittingly, the loss of β1,6-branching from complex-type N-glycans diminished intensity of staining (Fig. 8g), the introduction of the bisecting GlcNAc moiety into the core-fucosylated N-glycans having a comparatively stronger negative effect on percentage of positive cells (Fig. 8h). The presence of disubstituted core had an especially strong effect on binding of truncated CG-3 (Fig. 8h). Since it will not interact with the CRD, topological factors will matter. In fact, adding this substitution will induce a shift in the conformational equilibrium of the relative presentation of N-glycan antennae, which could underlie the affinity change previously noted in solid-phase experiments with neoglycoproteins presenting N-glycans with mono- and disubstituted cores (Unverzagt et al., 2002; André et al., 2004, 2007b, 2009b). For human galectin-3, presence of both core substitutions reduced the affinity more than twofold relative to the core-fucosylated N-glycan (André et al., 2007b). Overall, these results document the sensitivity of CG-3 binding to various, even subtle changes in the glycomic profile and the similarity to human galectin-3.

The validity of these two conclusions was further underscored by measurements in a human cell system, with a reactivity increase controlled by a tumor suppressor (André et al., 2007a), as shown in Supporting Information, Figure 3. Extent of fluorescent staining with CG-3 was higher than for the three proto-type CGs, CG-1A being the most reactive among them (Kaltner et al., 2008). Of note, in this model system human galectin-3 could functionally compete with the proto-type galectin-1 and hereby interfere with its pro-anoikis activity (Sanchez-Ruderisch et al., 2010), a result inspiring considerations of respective functional divergence among CGs. At this point, the presented data indicated that CG-3 is a lectin that can sense alterations in glycan profiles. It can thus act as effector via the characteristic tridomain structure. Because the gene structure had already

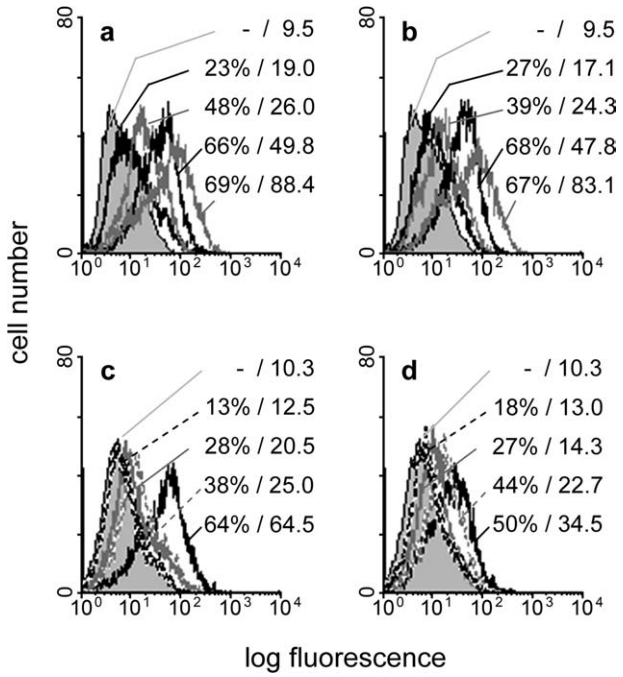

Fig. 7. Fluorescent cell surface staining by labeled full-length CG-3 (**a, c**) and its proteolytically truncated derivative (**b, d**). Semilogarithmic representation of staining profiles of the parental CHO line Pro^{-5} (lacking expression of β1,4-galactosyltransferase VI) documenting dependence of staining on lectin concentration (a, b) and presence of haptenic inhibitor (c, d). Characteristics of background staining when using the fluorescent indicator without the prior incubation step with biotinylated lectin are illustrated as a reference in each panel by the shaded area. Quantitative data on percentage of positive cells (%) and mean fluorescence intensity are presented for each curve in the order of listing from top to bottom. Concentrations of lectin used were 0.25 μg/mL, 1 μg/mL, 2 μg/mL, and 5 μg/mL for CG-3 (a) and 0.25 μg/mL, 0.5 μg/mL, 1 μg/mL, and 2 μg/mL for the truncated protein (b). Concentrations of lactose used were 40 mM, 10 mM and 2 mM with the reference without inhibitor (100%; bottom) in assays with 2 μg/mL CG-3 (c) and 1 μg/mL truncated CG-3 (d).

suggested that this CG-3 protein might be the prevalent form, expression profiling will not only decide the issue on CG-3 presence but also on the relative abundance of the possible forms.

Expression Profiling by RT-PCR/Western Blotting

The detailed gene map of CG-3 given in Fig. 3 led to the design of primer sets discriminating between the different forms. First, we tested a primer set targeting sequences in exons III/V to separate mRNAs specific for CG-3/tsp2CG-3I (amplification product at 625 bp) from those for tsp2CG-3II (835 bp) and tsp2CG-3III (1076 bp), the sites for complementarity with primers shown in Fig. 3 as pair of black/gray arrows. As documented in Fig. 9a, amplification products were obtained in all tested tissues, the intensity of the main band at 625 bp yet varied. Faint signals for variants were restricted to tsp2CG-3III and two organs. To measure the relative contribution of CG-3/tsp2CG-3I to the recorded signals, the next primer sets were designed to target a common

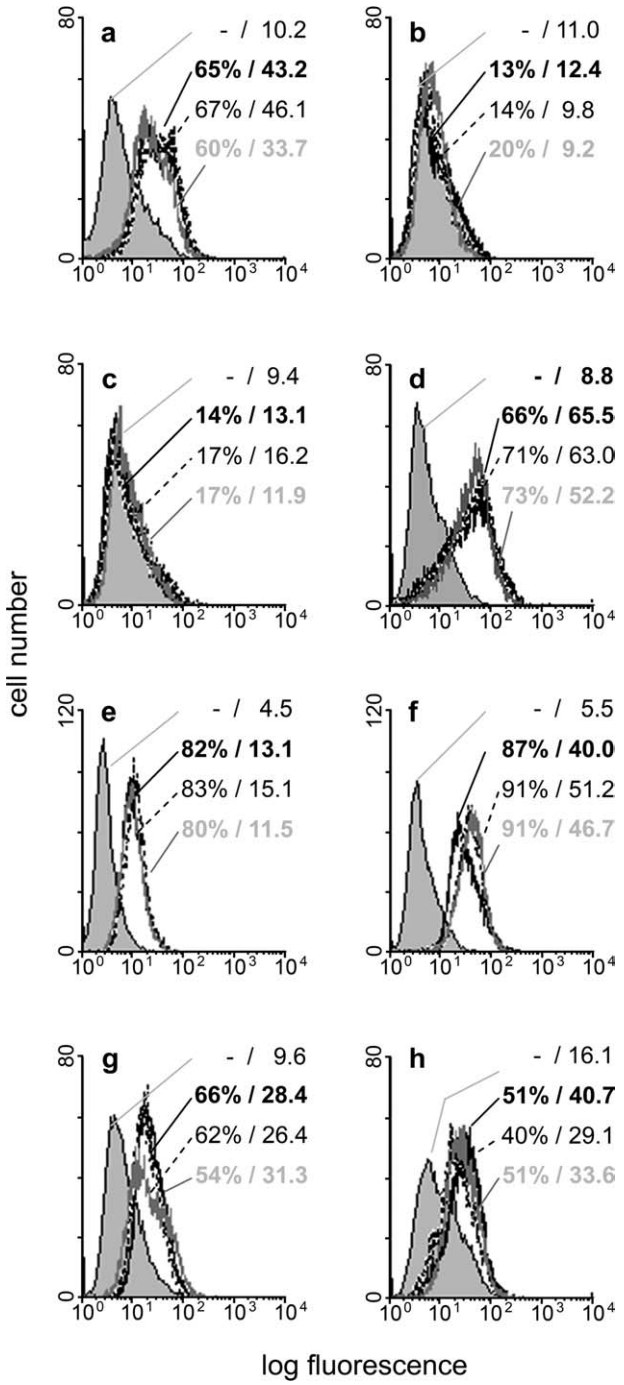

Fig. 8. Fluorescent cell surface staining by labeled full-length CG-3 (2 μg/mL; black), its proteolytically truncated derivative (1 μg/ml; dashed line) and human full-length galectin-3 (2 μg/mL; gray). All quantitative data are presented starting with the background in the given order from top to bottom (for further details, please see legend to Fig. 7). Semilogarithmic representation of staining profiles are given for the parental line (**a**), the Lec8 mutant (**b**; reduced galactosylation), the Lec1 mutant (**c**; impaired generation of complex-type N-glycans), the Lec2 mutant (**d**; reduced sialylation), the transfectant line for α2,6-sialyltransferase I (**e**), the transfectant line for α1,2-fucosyltransferase I (**f**), the Lec4 mutant (**g**; reduced β1,6-branching in N-glycans) and the transfectant line for *N*-acetylglucosaminyltransferase III (**h**; increased addition of bisecting GlcNAc residue).

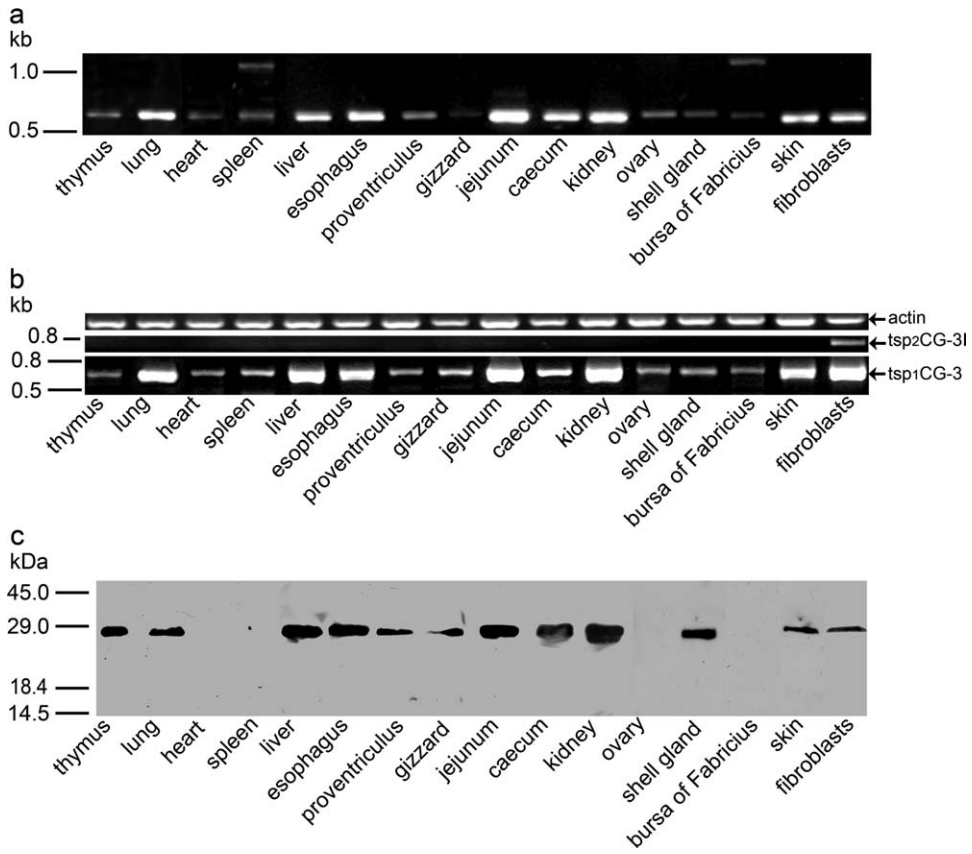

Fig. 9. Expression profiling of tsp1CG-3 and tsp2CG-3I-III by RT-PCR and Western blotting. Presence of mRNAs in extracts from a panel of tissues and transformed embryonic fibroblasts was first tested with the primer set given in Fig. 3 by black/gray arrows and led to direct amplification products of 625 bp (tsp1CG-3/tsp2CG-3I) or 1076 bp (tsp2CG-3III) (**a**). Calibration of cDNA length is given. The following RT-PCR analysis, in which the amplification is driven by two primer sets, distinguishes between transcripts from both tsps (for details, please see Fig. 3). The product of the length of 726 bp (arrow, tsp1CG-3) reveals presence of tsp1CG-3-specific mRNA in 15 tested tissues and the cultured fibroblasts, whereas tsp2CG-3I-specific mRNA (arrow, 789 bp) was detected only in extracts of the fibroblasts (**b**). No evidence for products by amplifying mRNAs for tsp2CG-3II/III was obtained. Actin-specific mRNA was monitored as internal loading control for each tissue type and the cells. Calibration of cDNA length is given. Detection of tsp1CG-3 protein in tissue extracts (2.5 mg of total protein processed by immunoprecipitation) and in fibroblast extracts (50 μg protein) was performed with affinity-purified anti-CG-3 IgG fractions free of any cross-reactivity to the other four CGs (0.5 μg/mL) (**c**). Positions of molecular weight markers are included.

sequence in the final exon but two different regions for the 5′-primer to keep the two start regions apart (Fig. 3). The signals were nearly exclusively confined to the 726 bp product representing CG-3, except for the specimen from the cultured transformed fibroblasts weakly positive for the 789-bp product (Fig. 9b). This sample excluded false-negative results for a tsp2CG-3I-specific transcript. No signal for a tsp2CG-3III-specific mRNA could be picked up in these experiments for spleen and bursa of Fabricius, underlining the very low abundance, if present at all, of this form. In relation to mRNA production from tsp1 the second promoter thus appears conspicuously less active in adult tissue and independently controlled, as previously described for the human tsp2 with rather low abundance of the respective product (Guittaut et al., 2001). What these results clearly demonstrate is that expression of the CG-3 gene mainly uses tsp1, as suggested by the phylogenetic comparison (Fig. 2) and the tridomain structure (Fig. 3), when monitoring adult organs. They also emphasize the requirement to select primer sites accordingly to avoid focusing on certain variants such as tsp2CG-3II/III in array-based studies (Geatrell et al., 2009), and they prompted us to take expression profiling to the level of the protein. After all, human galectin-3 production is known to be under post-transcriptional control (Ramasamy et al., 2007; Sanchez-Ruderisch et al., 2010).

Using the thoroughly characterized recombinant protein, we raised polyclonal antibodies and ensured lack of cross-reactivity to any of the other four CGs. The antibody preparation was reactive with CG-3, its truncated version and also tsp2CG-3I (Supporting Information, Fig. 4). However, the absence of the respective signal in the RT-PCR analysis made it unnecessary to further fractionate the IgG preparation or to exploit the differences in isoelectric points of the two proteins for detection. In line with the RT-PCR analyses presented above protein was detected in 11 from 15 tested tissues (Fig. 9c). Extracts of spleen and bursa of Fabricius, presenting low-intensity bands for CG-3 and a variant (Fig. 9a), were negative. No evidence for natural occurrence of a proteolytically truncated protein in adult organs was

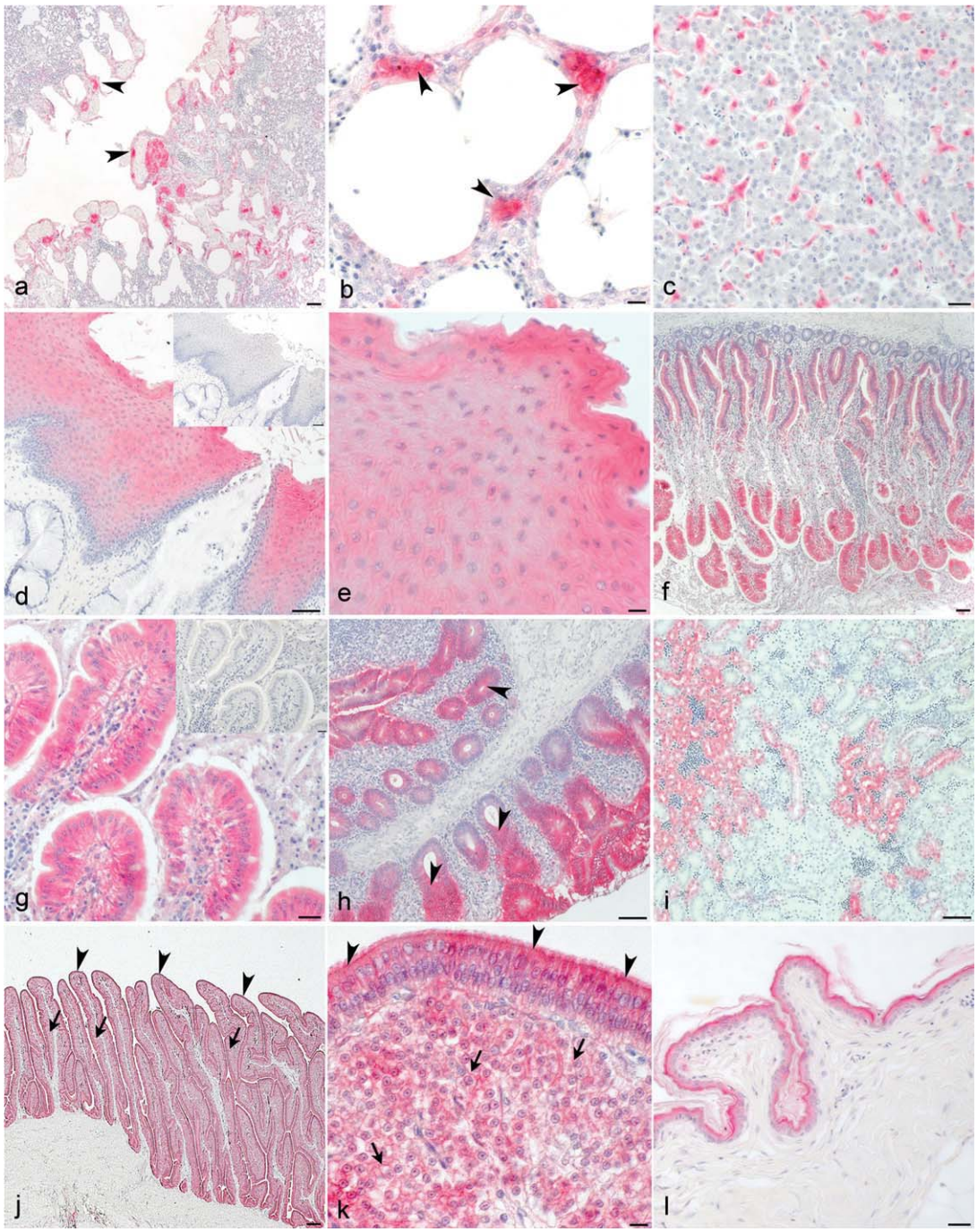

Fig. 10. Localization of CG-3 in tissue sections by immunohistochemistry. Control sections after processing without the incubation step with the CG-3-specific antibodies to ascertain lack of antigen-independent staining are shown in the insets of panels d and g. Microphotographs of cross-sections of the lumen and the wall of a parabronchus as well as connected atria at two levels of magnification (**a, b**). The presence of strongly reactive cells was visualized in atria (a), especially in interatrial septa, characteristic of alveolar macrophages (arrowheads) (b). In liver, marked positivity was confined to stellate-shaped cells in hepatic sinusoids, that is, Kupffer cells (**c**). Staining of the typical stratified squamous epithelium within the middle-third of the esophagus at two levels of magnification (**d, e**). Submucosal glands and its excretory ducts were negative (d). CG-3 presence was restricted to the *stratum corneum* and to the *stratum spinosum*. In both layers, staining was observed extracellularly and within the cytoplasm, comparatively strong intensity in the *stratum corneum* (e). In jejunum (**f, g**), the staining of epithelium along the microvilli was zonally graded, especially intense in the tips of microvilli and decreased in the basal part. No staining was seen in the middle part and in the jejunal crypts (f). On the cellular level, CG-3 reactivity was exclusively cytoplasmic, goblet cells were not reactive (g). In contrast to jejunum, epithelial lining of caecal crypts (arrowheads) presented a strong signal (**h**). In kidney, intense staining was observed in tubules of cortex and medullary cones comprising epithelia of distal tubules and loops of Henle (**i**). In shell gland's mucosa, positivity for CG-3 was found both in the pseudostratified surface epithelium (arrowheads) and the glandular epithelial lining (arrows) (**j**). Showing the tip of a shell gland's mucosal fold at increased level of magnification, CG-3 localization in surface epithelium was confined to the apical cytoplasm (arrows), mostly membrane-associated presence (arrowheads) was observed in glandular epithelial cells (**k**). Immunopositivity for CG-3 in skin was limited to the *stratum intermedium* (equivalent to *stratum spinosum* in mammals) of the epidermis (**l**). The scale bars are 10 μm (b, e, k), 20 μm (c, g, l, insets of d and g), 50 μm (a, d, f, h, i) and 200 μm (j).

TABLE 2. Immunohistochemical profiling of CG-3 presence in various organs of adult chicken

Type of organ	Positivity	Type of organ	Positivity
Cerebrum/cerebellum	–	*Gldd. ventriculares*	
Thymus		Epithelium	+[a]
Thymocytes, macrophages	–	*L. propria mucosae*	–
Hassall's corpuscles	++	Stratum compactum	–
Larynx[b]		Gut[b,c]	
Respiratory epithelium	+[d,a]	Epithelial lining[a]	–
L. propria mucosae	–	Villi	+++[e]
Trachea[b]		Crypts	–[f],+[g],+++[h]
Respiratory epithelium	–/+[d,a]	Goblet cells	–
L. propria mucosae	–	*L. propria mucosae*	–
Lung[b]		Kidney	
Parabronchal wall	+++[i,j]	Glomeruli	–
Atria		Epithelial lining of tubules	
Respiratory epithelium	–	Proximal convolution	–
Interatrial septa	+++[i,j]	Distal convolution	+++[a,k]
Heart	–	Loops of Henle	++[a,l]
Spleen	–	Collecting ducts	–
Liver		Connective tissue	–
Hepatocytes	–	Ovary	–
Kupffer cells	+++[m]	Oviduct[n]	
Esophagus[b]		Surface epithelium	+[a]
Epithelium	++/+++[a]	Glandular epithelial lining	–
L. propria mucosae	–	Uterus (shell gland)[b]	
Proventriculus[b]		Surface epithelium	+++[d,a]
Gldd. proventriculares superficiales		Glandular epithelial lining	++[o]
Epithelium	–/+[a]	Skin[b]	
L. propria mucosae	–	Epidermis	–
Gldd. proventriculares profundae		*Stratum corneum*	–
Epithelium	–	*Stratum intermedium*[p]	+++
L. propria mucosae	–	*Stratum basalis*	–
Gizzard[b]		Dermis	–
Pellicula	–	Subcutis	–

Signal intensity was semiquantitatively grouped into the categories: – negative, + weak but significant, ++ medium, +++ strong;
[a]Exclusively cytoplasmic;
[b]Smooth muscle layers are negative for CG-3;
[c]gut = duodenum, jejunum, ileum, caecum, rectum;
[d]Exclusively apical cytoplasma and cilia;
[e]Jejunum only tips of microvilli;
[f]Jejunum, ileum;
[g]Duodenum;
[h]Caecum, rectum;
[i]Clusters of cells;
[j]Exclusively alveolar macrophages;
[k]Renal cortex only;
[l]Medullary cones only;
[m]Single cells scattered in the parenchyma;
[n]Infundibulum, isthmus, magum, vagina, uterus is listed separately;
[o]Staining membrane-associated;
[p]Equivalent to *stratum spinosum* in mammals. Abbreviations: *gldd.* = *glandulae*, *l* = *lamina*.

obtained when working under conditions of thorough protection against proteolytic degradation. These immunoblotting data provided a guideline for the immunohistochemical localization of CG-3 in fixed tissue sections.

CG-3 Localization

After rigorously excluding antigen-independent staining by respective controls including the absence of any staining in organs negative in blotting, that is, heart, spleen, ovary and bursa of Fabricius, the CG-3-specific IgG preparation was used for comparative monitoring of tissue specimen processed under identical conditions. The positivity of lung and liver extracts was attributed to alveolar macrophages and the Kupffer cells, respectively (Fig. 10a–c). The other prominent feature of CG-3 localization is positivity of epithelial cells, which form a continuous layer covering the surface of the mucosa, in digestive and reproductive tracts, and also encompass the lumen of kidney tubules (Table 2). In addition, thymic (*Hassall's*) corpuscles derived from epithelio-reticular cells were reactive (Table 2). The specificity of the reaction was further illustrated by the distinct staining of the apical zone and cilia in the epithelia of trachea/larynx and of the shell gland (Table 2). A zonal grading within the esophageal stratified squamous epithelium was observed (Fig. 10d,e), the columnar epithelium of the proventriculus being negative. The cuboidal

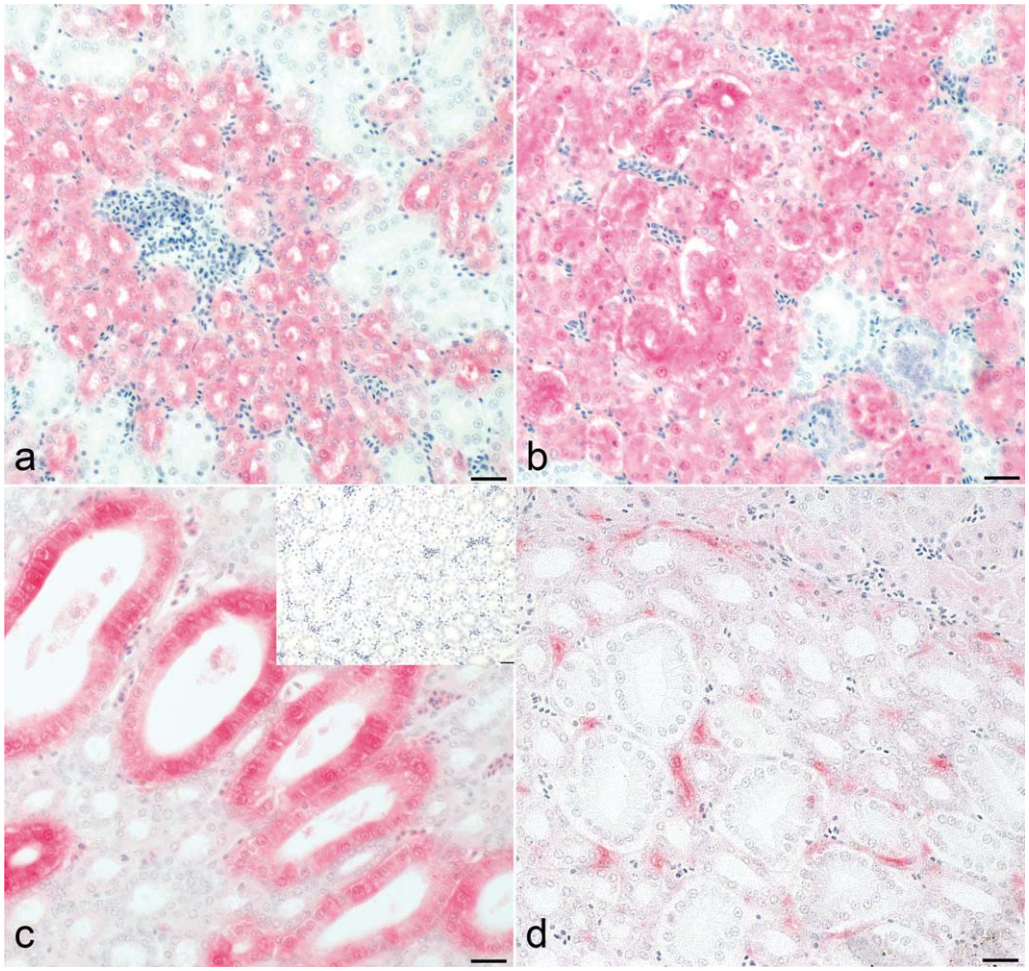

Fig. 11. Localization of CG-3, CG-1A, CG-1B, CG-2 and CG-8 in adult kidney by immunohistochemistry. Microphotographs of cross-sections from adult kidney after processing with antibody preparations specific for CG-3 (**a**), CG-1A (**b**), CG-2 (**c**), CG-8 (**d**) and CG-1B (inset to c). Colocalization of the profiles of presence of CG-3 (a) with CG-1A (b) is seen in epithelia of distal tubules (please see also Fig. 10i for additional documentation of CG-3 localization in kidney). Further positivity for CG-1A is present in epithelia of proximal tubules (b). In contrast, CG-2-specific staining was found in epithelia of medullary collecting ducts (c), the connective tissue between these ducts harboring reactivity for anti-CG-8 (d). No signal was detected for CG-1B (inset to panel c), hereby excluding any antigen-independent staining with IgG preparations. The scale bars are 20 μm (a-d, inset of c).

epithelium of the *glandulae proventriculares* and *ventriculares* was negative (Table 2). Immunopositivity was presented in villi of the jejunum (Fig. 10f,g), extending to crypts in the caecum (Fig. 10h). Goblet cells did not contain CG-3 (Table 2). The pseudostratified surface epithelium, which covers the extensive mucosal fold of the shell gland, was also conspicuously positive (Fig. 10i,j), as was the *stratum intermedium* (equivalent to *stratum spinosum* in mammals) in skin (Fig. 10k, Table 2). Neither immune cells (in thymus, the bronchial or gut-associated lymphoid tissue) nor smooth muscle layers were positive. *Laminae propriae mucosae*, a site of residence of immune cells, too, showed no reactivity (Table 2).

The presented mapping completes the initial monitoring of adult tissue for all five CGs and thus affords the opportunity to start comprehensive comparative analysis. To do so we selected kidney as instructive example with its known staining properties for CG-1A (Stierstorfer et al., 2000), CG-2 (Kaltner et al., 2008) and CG-8 (Kaltner et al., 2009), and the lack of expression of CG-1B (Kaltner et al., 2008). The sections processed with noncross-reactive IgG fractions against the five CGs revealed a partial overlap in localization profiles in epithelia of distal tubules and loops of Henle in the cases of CG-3 (Fig. 11a) and CG-1A (Fig. 11b). CG-1A positivity was also present in proximal tubules (Fig. 11b). CG-2, in epithelial lining of medullary collecting ducts (Fig. 11c), and CG-8, in connective tissue between collecting ducts (Fig. 11d), were present in spatially clearly separated regions, the occurrence of disparate profiles also serving as internal specificity controls. These results, together with further information on the localization profiles of the other four CGs, are summarized in Supporting Information, Table 1. It also includes the data on mouse galectin-3 (also referred to as Mac-2 antigen) to track down cross-species differences, for example in the cases of spleen and ovary (Flotte et al., 1983; Lohr et al., 2008). The emerging differences in intra- and inter-

species analyses intimate diversity on the level of regulatory sequences in the promoter region. As a first step to its characterization, we performed computational searches for putative transcription-factor-binding sites in the proximal promoter regions of both tsps presented in Supporting Information, Table 2. Qualitative differences are delineated, among them the presence of putative target sites for c-Ets-1, GATA-3, and NF-κB, a potent mediator of inflammatory responses, in the promoter of the CG-3 gene (tsp1). These data sets enabled to answer the question on such differences among proximal promoter regions for all CGs. They are summarized in Supporting Information, Table 3 for the tsp1 region and Table 4 for the tsp2 region. In view of the similarities in organization with mammalian genes, we completed the comparisons with setting CG-3 features in relation to characteristics of mammalian promoter regions (Supporting Information, Table 5). Of interest, the two target sequences for Nkx2-5 and RFX1, which are not present in the CG-3 promoter but in mammalian promoters, are found in the proximal promoter region of the CG-8 gene. At any rate, caution should yet be exercised regarding the predictive value of these *in silico* evaluations, which should be considered to give further experimental work a direction.

In summary, the production of mRNA from the avian CG-3 gene has the highest degree of complexity among CGs. In addition to alternative splicing, known from CG-8 (Kaltner et al., 2009), two principal tsps can be used. Adult organs from normal chicken, if positive, mostly express the form with reactive phosphorylation sites in the N-terminus, closely related to mammalian galectin-3. Only RT-PCR analysis of specimen of spleen and bursa of Fabricius yielded a signal for tsp2CG-3III variants. As noted before in the cases of tsp2CG-3I (Nurminskaya and Linsenmayer, 1996) and tsp2CG-3II/III (Gorski et al., 2002), cultured cells can harbor the variants, tsp2CG-3II also detected by RT-PCR in cDNA from whole embryo extract (Geatrell et al., 2009). Whereas the analysis of glycan binding using neoglycoproteins and cell lines proved similar properties to human galectin-3, the immunohistochemical expression profile revealed conspicuous differences to mammalian galectin-3 and also to the other four CGs. Having herewith clarified the inherent complexity of mRNA production for CG-3, the unique chimera-type CG, and defined tsp1CG-3 as predominant form in adult tissues, the last step toward a solid foundation for the comprehensive analysis of all five CGs, which form the complete galectin network in this model organism, has been taken.

ACKNOWLEDGMENTS

We are indebted to Dr. J. Plachý for kindly supplying the *v-src*-transformed chicken embryonic fibroblasts, Dr. Y. Nekcic for helpful advice, A. Helfrich, B. Hofer, and L. Mantel for skilful technical assistance.

LITERATURE CITED

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

Ahmad N, Gabius HJ, Kaltner H, André S, Kuwabara I, Liu FT, Oscarson S, Norberg T, Brewer CF. 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.

André S, Kojima S, Gundel G, Russwurm R, Schratt X, Unverzagt C, Gabius HJ. 2006b. Branching mode in complex-type triantennary N-glycans as regulatory element of their ligand properties. Biochim Biophys Acta 1760:768–782.

André S, Kožár T, Kojima S, Unverzagt C, Gabius HJ. 2009b. From structural to functional glycomics: core substitutions as molecular switches for shape and lectin affinity of N-glycans. Biol Chem 390:557–565.

André S, Kožár T, Schuberth R, Unverzagt C, Kojima S, Gabius HJ. 2007b. Substitutions in the *N*-glycan core as regulators of biorecognition: the case of core-fucose and bisecting GlcNAc moieties. Biochemistry 46:6984–6995.

André S, Pei Z, Siebert HC, Ramström O, Gabius HJ. 2006a. 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, Sanchez-Ruderisch H, Nakagawa H, Buchholz M, Kopitz J, Forberich P, Kemmner W, Böck C, Deguchi K, Detjen KM, Wiedenmann B, von Knebel Doeberitz M, Gress TM, Nishimura SI, Rosewicz S, Gabius HJ. 2007a. Tumor suppressor $p16^{INK4a}$: modulator of glycomic profile and galectin-1 expression to increase susceptibility to carbohydrate-dependent induction of anoikis in pancreatic carcinoma cells. FEBS J 274:3233–3256.

André S, Specker D, Bovin NV, Lensch M, Kaltner H, Gabius HJ, Wittmann V. 2009a. Carbamate-linked lactose: design of clusters and evidence for selectivity to block binding of human lectins to (neo)glycoproteins with increasing degree of branching and to tumor cells. Bioconjug Chem 20:1716–1728.

André S, Unverzagt C, Kojima S, Frank M, Seifert J, Fink C, Kayser K, von der Lieth CW, Gabius HJ. 2004. 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.

Bachhawat-Sikder K, Thomas CJ, Surolia A. 2001. Thermodynamic analysis of the binding of galactose and poly-*N*-acetyllactosamine derivatives to human galectin-3. FEBS Lett 500:75–79.

Barondes SH. 1984. Soluble lectins: a new class of extracellular proteins. Science 223:1259–1264.

Beyer EC, Zweig SE, Barondes SH. 1980. Two lactose binding lectins from chicken tissues. Purified lectin from intestine is different from those in liver and muscle. J Biol Chem 255:4236–4239.

Boja ES, Fales HM. 2001. Overalkylation of a protein digest with iodoacetamide. Anal Chem 73:3576–3582.

Cherayil BJ, Weiner SJ, Pillai S. 1989. The Mac-2 antigen is a galactose-specific lectin that binds IgE. J Exp Med 170:1959–1972.

Cooper DNW. 2002. Galectinomics: finding themes in complexity. Biochim Biophys Acta 1572:209–231.

Crittenden SL, Roff CF, Wang JL. 1984. Carbohydrate-binding protein 35: identification of the galactose-specific lectin in various tissues of mice. Mol Cell Biol 4:1252–1259.

Dam TK, Gabius HJ, André S, Kaltner H, Lensch M, Brewer CF. 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 HJ. 1994. Contribution of carbohydrate histochemistry to glycobiology. Histol Histopathol 9:155–171.

Díez-Revuelta N, Velasco S, André S, Kaltner H, Kübler D, Gabius HJ, Abad-Rodríguez J. 2010. Phosphorylation of adhesion- and growth-regulatory human galectin-3 leads to the induction of axonal branching by local membrane L1 and ERM redistribution. J Cell Sci 123:671–681.

Flotte TJ, Springer TA, Thorbecke GJ. 1983. Dendritic cell and macrophage staining by monoclonal antibodies in tissue sections and epidermal sheets. Am J Pathol 111:112–124.

Fukumori T, Kanayama Ho, Raz A. 2007. The role of galectin-3 in cancer drug resistance. Drug Resist Update 10:101–108.

Gabius HJ. 1987. Endogenous lectins in tumors and the immune system. Cancer Invest 5:39–46.

Gabius HJ. 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 HJ, editor. 2009. The Sugar Code. Fundamentals of glycosciences. Weinheim: Wiley-VCH.

Gabius HJ, Bodanowitz S, Schauer A. 1988. Endogenous sugar-binding proteins in human breast tissue and benign and malignant breast lesions. Cancer 61:1125–1131.

Gabius HJ, Schröter C, Gabius S, Brinck U, Tietze LF. 1990. Binding of T-antigen-bearing neoglycoprotein and peanut agglutinin to cultured tumor cells and breast carcinomas. J Histochem Cytochem 38:1625–1631.

Gabius HJ, 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.

Gaudin JC, Arar C, Monsigny M, Legrand A. 1997. Modulation of the expression of the rabbit galectin-3 gene by p53 and c-Ha-ras proteins and PMA. Glycobiology 7:1089–1098.

Geatrell JC, Gan PMI, Mansergh FC, Kisiswa L, Jarrin M, Williams LA, Evans MJ, Boulton ME, Wride MA. 2009. Apoptosis gene profiling reveals spatio-temporal regulated expression of the p53/Mdm2 pathway during lens development. Exp Eye Res 88: 1137–1151.

Gorski JP, Liu FT, Artigues A, Castagna LF, Osdoby P. 2002. New alternatively spliced form of galectin-3, a member of the β-galactoside-binding animal lectin family, contains a predicted transmembrane-spanning domain and a leucine zipper motif. J Biol Chem 277:18840–18848.

Guittaut M, Charpentier S, Normand T, Dubois M, Raimond J, Legrand A. 2001. Identification of an internal gene to the human galectin-3 gene with two different overlapping reading frames that do not encode galectin-3. J Biol Chem 276:2652–2657.

Houzelstein D, Gonçalves IR, Fadden AJ, Sidhu SS, Cooper DNW, Drickamer K, Leffler H, Poirier F. 2004. Phylogenetic analysis of the vertebrate galectin family. Mol Biol Evol 21:1177–1187.

Hughes RC. 1994. Mac-2: a versatile galactose-binding protein of mammalian tissues. Glycobiology 4:5–12.

Kadrofske MM, Openo KP, Wang JL. 1998. The human LGALS3 (galectin-3) gene: determination of the gene structure and functional characterization of the promoter. Arch Biochem Biophys 349:7–20.

Kaltner H, Solís D, André S, Lensch M, Manning JC, Mürnseer M, Sáiz JL, Gabius HJ. 2009. Unique chicken tandem-repeat-type galectin: implications of alternative splicing and a distinct expression profile compared to those of the three proto-type proteins. Biochemistry 48:4403–4416.

Kaltner H, Solís D, Kopitz J, Lensch M, Lohr M, Manning JC, Mürnseer M, Schnölzer M, André S, Sáiz JL, Gabius HJ. 2008. Prototype chicken galectins revisited: characterization of a third protein with distinctive hydrodynamic behaviour and expression pattern in organs of adult animals. Biochem J 409:591–599.

Kasai Ki, Hirabayashi J. 1996. Galectins: a family of animal lectins that decipher glycocodes. J Biochem 119:1–8.

Kopitz J, Bergmann M, Gabius HJ. 2010. How adhesion/growth-regulatory galectins-1 and -3 attain cell specificity: case study defining their target on neuroblastoma cells (SK-N-MC) and marked affinity regulation by affecting microdomain organization of the membrane. IUBMB Life 62:624–628.

Kopitz J, von Reitzenstein C, André S, Kaltner H, Uhl J, Ehemann V, Cantz M, Gabius HJ. 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.

Kübler D, Hung CW, Dam TK, Kopitz J, André S, Kaltner H, Lohr M, Manning JC, He L, Wang H, Middelberg A, Brewer CF, Reed J, Lehmann WD, Gabius HJ. 2008. Phosphorylated human galectin-3: facile large-scale preparation of active lectin and detection of structural changes by CD spectroscopy. Biochim Biophys Acta 1780:716–722.

Lehmann WD, Wei J, Hung CW, Gabius HJ, Kirsch D, Spengler B, Kübler D. 2006. Effective solvation of alkaline earth ions by proline-rich proteolytic peptides of galectin-3 upon electrospray ionisation. Rapid Commun Mass Spectrom 20:2404–2410.

Lensch M, Lohr M, Russwurm R, Vidal M, Kaltner H, André S, Gabius HJ. 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.

Lohr M, Kaltner H, Lensch M, André S, Sinowatz F, Gabius HJ. 2008. Cell-type-specific expression of murine multifunctional galectin-3 and its association with follicular atresia/luteolysis in contrast to pro-apoptotic galectins-1 and -7. Histochem Cell Biol 130:567–581.

Lohr M, Kaltner H, Schwartz-Albiez R, Sinowatz F, Gabius HJ. 2010. Towards functional glycomics by lectin histochemistry: strategic probe selection to monitor core and branch-end substitutions and detection of cell-type and regional selectivity in adult mouse testis and epididymis. Anat Histol Embryol 39:481–493.

López-Lucendo MF, Solís D, André S, Hirabayashi J, Kasai Ki, Kaltner H, Gabius HJ, Romero A. 2004. Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. J Mol Biol 343:957–970.

López-Lucendo MF, Solís D, Sáiz JL, Kaltner H, Russwurm R, André S, Gabius HJ, Romero A. 2009. Homodimeric chicken galectin CG-1B (C-14): crystal structure and detection of unique redox-dependent shape changes involving inter- and intrasubunit disulfide bridges by gel filtration, ultracentrifugation, site-directed mutagenesis, and peptide mass fingerprinting. J Mol Biol 386:366–378.

Nurminskaya M, Linsenmayer TF. 1996. Identification and characterization of up-regulated genes during chondrocyte hypertrophy. Dev Dyn 206:260–271.

Oda Y, Kasai Ki. 1983. Purification and characterization of β-galactoside-binding lectin from chick embryonic skin. Biochim Biophys Acta 761:237–245.

Ortega N, Behonick DJ, Colnot C, Cooper DNW, Werb Z. 2005. Galectin-3 is a downstream regulator of matrix metalloproteinase-9 function during endochondral bone formation. Mol Biol Cell 16:3028–3039.

Patnaik SK, Stanley P. 2006. Lectin-resistant CHO glycosylation mutants. Methods Enzymol 416:159–182.

Purkrábková T, Smetana K Jr., Dvořánková B, Holíková Z, Böck C, Lensch M, André S, Pytlík R, Liu FT, Klíma J, Smetana K, Motlík J, Gabius HJ. 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.

Quelle DE, Zindy F, Ashmun RA, Sherr CJ. 1995. Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. Cell 83:993–1000.

Raimond J, Rouleux F, Monsigny M, Legrand A. 1995. The second intron of the human galectin-3 gene has a strong promoter activity down-regulated by p53. FEBS Lett 363:165–169.

Ramasamy S, Duraisamy S, Barbashov S, Kawano T, Kharbanda S, Kufe D. 2007. The MUC1 and galectin-3 oncoproteins function in a microRNA-dependent regulatory loop. Mol Cell 27:992–1004.

Sakakura Y, Hirabayashi J, Oda Y, Ohyama Y, Kasai K. 1990. Structure of chicken 16-kDa β-galactoside-binding lectin. Complete amino acid sequence, cloning of cDNA, and production of recombinant lectin. J Biol Chem 265:21573–21579.

Sanchez-Ruderisch H, Fischer C, Detjen KM, Welzel M, Wimmel A, Manning JC, André S, Gabius HJ. 2010. Tumor suppressor $p16^{INK4a}$: downregulation of galectin-3, an endogenous competitor of the pro-anoikis effector galectin-1, in a pancreatic carcinoma model. FEBS J 277:3552–3563.

Schwartz-Albiez R. 2009. Inflammation and glycosciences. In: Gabius HJ, editor. The sugar code. Fundamentals of glycosciences. Weinheim: Wiley-VCH. p 447–467.

Seidler J, Adal M, Kübler D, Bossemeyer D, Lehmann WD. 2009. Analysis of autophosphorylation sites in the recombinant catalytic subunit alpha of cAMP-dependent kinase by nano-UPLC-ESI-MS/MS. Anal Bioanal Chem 395:1713–1720.

Siebert HC, Adar R, Arango R, Burchert M, Kaltner H, Kayser G, Tajkhorshid E, von der Lieth CW, Kaptein R, Sharon N, Vliegenthart JFG, Gabius HJ. 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.

Solís D, Maté MJ, Lohr M, Ribeiro JP, López-Merino L, André S, Buzamet E, Cañada FJ, Kaltner H, Lensch M, Ruiz FM, Haroske G, Wollina U, Kloor M, Kopitz J, Sáiz JL, Menéndez M, Jiménez-Barbero J, Romero A, Gabius HJ. 2010. N-Domain of human adhesion/growth-regulatory galectin-9: preference for distinct conformers and non-sialylated N-glycans and detection of ligand-induced structural changes in crystal and solution. Int J Biochem Cell Biol 42:1019–1029.

Solís D, Romero A, Kaltner H, Gabius HJ, Díaz-Mauriño T. 1996. Different architecture of the combining site of the two chicken galectins revealed by chemical mapping studies with synthetic ligand derivatives. J Biol Chem 271:12744–12748.

Spicer SS, Schulte BA. 1992. Diversity of cell glycoconjugates shown histochemically: a perspective. J Histochem Cytochem 40:1–38.

Stierstorfer B, Kaltner H, Neumüller C, Sinowatz F, Gabius HJ. 2000. Temporal and spatial regulation of expression of two galectins during kidney development of the chicken. Histochem J 32: 325–336.

Szabo P, Dam TK, Smetana K Jr, Dvořánková B, Kübler D, Brewer CF, Gabius HJ. 2009. Phosphorylated human lectin galectin-3: analysis of ligand binding by histochemical monitoring of normal/malignant squamous epithelia and by isothermal titration calorimetry. Anat Histol Embryol 38:68–75.

Unverzagt C, André S, Seifert J, Kojima S, Fink C, Srikrishna G, Freeze HH, Kayser K, Gabius HJ. 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.

Varela PF, Solís D, Díaz-Mauriño T, Kaltner H, Gabius HJ, Romero A. 1999. The 2.15 Å crystal structure of CG-16, the developmentally regulated homodimeric chicken galectin. J Mol Biol 294: 537–549.

Villalobo A, Nogales-González A, Gabius HJ. 2006. A guide to signalling pathways connecting protein-glycan interaction with the emerging versatile effector functionality of mammalian lectins. Trends Glycosci Glycotechnol 18:1–37.

Voss PG, Tsay YG, Wang JL. 1994. Galectin-3: differential accumulation of distinct mRNAs in serum-stimulated mouse 3T3 fibroblasts. Glycoconj J 11:353–362.

Winter D, Pipkorn R, Lehmann WD. 2009. Separation of peptide isomers and conformers by ultra performance liquid chromatography. J Sep Sci 32:1111–1119.

Wu AM, Singh T, Liu JH, Krzeminski M, Russwurm R, Siebert HC, Bonvin AMJJ, André S, Gabius HJ. 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.

Wu AM, Wu JH, Tsai MS, Kaltner H, Gabius HJ. 2001. Carbohydrate specificity of a galectin from chicken liver (CG-16). Biochem J 358:529–538.