BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

*Biochemistry* **2009,** *48,* 4403–4416 4403
DOI:10.1021/bi900083q

# Unique Chicken Tandem-Repeat-Type Galectin: Implications of Alternative Splicing and a Distinct Expression Profile Compared to Those of the Three Proto-Type Proteins[†,‡]

Herbert Kaltner,*[,§] Dolores Solís,[∥,⊥] Sabine André,[§] Martin Lensch,[§] Joachim C. Manning,[§] Michael Mürnseer,[§] José Luis Sáiz,[∥,⊥] and Hans-Joachim Gabius[§]

[§]*Institut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität München, Veterinärstrasse 13, D-80539 München, Germany,* [∥]*Instituto de Química Física Rocasolano, CSIC, Serrano 119, E-28006 Madrid, Spain, and* [⊥]*Ciber de Enfermedades Respiratorias (CIBERES), Bunyola, Mallorca, Illes Balears, Spain*

*Received January 20, 2009; Revised Manuscript Received April 1, 2009*

ABSTRACT: Animal galectins (lectins with specificity for $\beta$-galactosides of glycan chains) are potent effectors in diverse aspects of cell sociology. Gene divergence has led to different groups and a marked interspecies variability in the number of members per group. Since the suitability of a model for studying functionality in the galectin network will be distinguished by a rather simple degree of complexity, we have focused on chicken galectins (CGs). Starting from partial expression sequence tag information, we here report on cloning of full-length cDNA for the first avian tandem-repeat-type galectin. It is termed CG-8 on the basis of its sequence similarity to galectin-8 from mammals. Systematic sequence searches revealed its unique character among CGs. Detection of two mature mRNA species points to production of isoforms. Alternative splicing affecting exon V generates the two proteins with linkers of either 9 (CG-8I) or 28 amino acids (CG-8II). Both proteins form monomers with a shape comparable to that of the proto-type proteins CG-1A/B in solution, act as cross-linkers in hemagglutination, and bind cells with a strict dependence on galactose. Western blotting revealed the presence of either CG-8II or the mixture in organ extracts. No evidence of a truncated form was obtained. Preparation of a specific antibody also enabled immunohistochemical localization. Prominent sites of its presence were defense cells in the *l. propria mucosae*, in addition to immune cells in distinct organs such as alveolar macrophages and thymocytes. Overall, we extend the network of CGs to a tandem-repeat-type protein and provide a detailed characterization from gene and protein structures to expression.

The aim of studying products of intrafamily gene diversification is to delineate distinct assignments of individual family members. To be able to do so, an essential task is the identification and characterization of all related proteins forming a particular group. Because of their considerable impact on cell sociology, e.g., regulating cell adhesion, migration, and proliferation, animal lectins are attracting an increasing level of interest in this context (*1*). In fact, they act as potent sensors for alterations in cell surface glycosylation and contribute to ensuring cell homeostasis by an intricate network of protein−carbohydrate and protein−protein interactions (*1*, *2*). In parallel with investigation of hallmarks of the lectin activity of known proteins, the advances in making genomic sequences available account for progress in tracing so far unknown gene products, which share the same phylogenetic origin. This work opens up the attractive opportunity to select organisms with a rather low degree of network complexity as a model for detailed study. With a focus on galectins, defined by their $\beta$-sandwich folding and exquisite specificity for certain $\beta$-galactosides in glycan chains, the established phylogenetic trees reflect a nonuniform degree of intrafamily complexity in different organisms (*3*−*6*).

As a consequence, the presence of only three proto-type galectins in chicken (CG-1A, CG-1B, and CG-2) and one chimera-type protein (CG-3) explains our selection of this organism for detailed study of galectin expression, structure, and functions. The already detected differences in promoter and coding sequences, in hydrodynamic behavior, even disclosing a redox-dependent switch in one case, in carbohydrate fine specificity, and in expression profiles, with CG-1A dominating in liver (parenchyma), CG-1B in lung (respiratory epithelium) and skin, and CG-2 in gut epithelial lining, intimate acquisition of distinct functions (*7*−*16*). As important as gaining new insight into the intragroup analysis are gaps in the complete description of all members of this family that need to be closed. After all, an

[†]This work has been generously supported by Grant BFU2006-10288 from the Spanish Ministry, CIBER de Enfermedades Respiratorias (CIBERES), an EC Marie Curie Research Training Network grant (Contract MRTN-CT-2005-019561), and the research initiative LMU*excellent*. CIBERES is an initiative from *ISCIII*.

[‡]Nucleotide sequences for CG-8I and CG-8II have been deposited in GenBank (entries DQ443719 and AY840215, respectively; later NM_001010843).

*To whom correspondence should be addressed. Telephone: +49-(0)89-2180-3984. Fax: +49-(0)89-2180-2508. E-mail: kaltner@lmu.de.


Published on Web 4/3/2009

available expression sequence tag (EST)[1] of 494 bp (AJ447244), satisfying the criteria for encoding a carbohydrate recognition domain (CRD) connected to a linker-like stretch and a further short sequence elongation, suggests the existence of a tandem-repeat-type galectin in chicken with similarity to mammalian galectin-8 and a respective protein from *Xenopus laevis* (xgalectin-VIIIa), tentatively termed CG-8 (*4*, *6*, *17*). In mammals and *Xenopus*, this type of lectin gene is rather broadly expressed. Marked levels of mRNA production are invariably seen in lung, liver, and kidney, and of note is the fact that processing of its transcript in mammals yields isoforms (*17*−*21*). This knowledge prompts a series of questions about the complete cDNA and also the genomic sequences, including the promoter region, about the actual number of genes for tandem-repeat-type galectins in chicken, about the occurrence of splicing events during transcript maturation as well as about the quaternary structure and hydrodynamic behavior, about lectin activity, and finally about the expression profile on the levels of mRNA/protein and cells. These questions will be answered consecutively, whereby the two forms of CG-8 with differences in linker length (CG-8I and II) are characterized as the chicken tandem-repeat-type galectin.

## EXPERIMENTAL PROCEDURES

*Cloning, Protein Expression, and Purification*. Total RNA from chicken jejunum was isolated using the RNeasy kit (Qiagen, Heidelberg, Germany) following the manufacturer's instructions, and 1 $\mu$g was used as a template to yield full-length cDNA for CG-8. 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 the gene-specific primer covering the 241−267 bp sequence stretch within the known EST. The resulting product of ~1500 bp was sequenced, and this information enabled the design of primers that yielded CG-8-specific cDNA(s) covering the entire sequence by PCR amplification. Amplification of respective sequences from 1 $\mu$g of total RNA was directed by the sense primer 5′-CATATGATGTCCTTGGATGGACCG-3′ with an internal NdeI restriction site (underlined) and the antisense primer 5′-GGATCCCTACCAGCTCCTCAC-3′ with an internal BamHI restriction site (underlined). The reaction was performed with the Expand High Fidelity PCR system as recommended by the manufacturer (Roche, Penzberg, Germany). The amplification product was separated from the PCR reagents by gel electrophoresis in 3% agarose yielding two bands. 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 *Escherichia coli* strain NovaBlue. Subsequently, these two CG-8-specific cDNAs were ligated into NdeI- and BamHI-treated expression vector pET-12a (Novagen), and the products were used for transformation of *E. coli* strain BL21(DE3)pLysS (Novagen). Optimal yields of lectin production with ~25−30 mg of CG-8/L of culture in both cases were obtained with TB medium (Roth, Karlsruhe, Germany) at 22 °C and induction with 100 $\mu$M isopropyl $\beta$-thio-D-galactoside. After frozen bacteria had been thawed in 7 mL of lysis buffer [20 mM phosphate-buffered saline (PBS) (pH 7.2) containing 2 mM ethylenediaminetetraacetic acid and 4 mM $\beta$-mercaptoethanol] per gram (wet weight) and centrifugation, the extracts were processed by affinity chromatographic purification, which was based on lactosylated Sepharose 4B after ligand conjugation to divinyl-sulfone-activated resin, as described previously (*22*, *23*).

*Analytical Procedures*. Gel electrophoretic analysis was carried out on SDS−polyacrylamide gels with 4% stacking and 12.5% running gel parts and subsequent silver staining. Mass spectrometric fingerprinting was preceded by processing the purified proteins in the presence of 9 M urea by incubation with 10 mM dithiothreitol for 30 min at 20 °C in a nitrogen atmosphere for full reduction. Free sulfhydryl groups were subsequently alkylated by incubation for 1 h at 20 °C in the dark with 80 mM iodoacetamide. Samples were then dialyzed exhaustively against 50 mM ammonium bicarbonate, digested with modified porcine trypsin (sequencing grade, from Promega, Madison, WI), and analyzed by MALDI-MS using an Ultraflex time-of-flight mass spectrometer (Bruker-Daltonik, Bremen, Germany) as described previously (*16*). Gel filtration was performed with a Superose 12 HR 10/30 column (void volume, 8.3 mL) (GE Healthcare, Munich, Germany) equilibrated with 5 mM sodium phosphate buffer (pH 7.2) with 0.2 M NaCl and 0.02% $NaN_3$, containing 0.1 M lactose to prevent interactions of CG-8I and CG-8II with the agarose-based matrix and 4 mM $\beta$-mercaptoethanol. The flow rate was 0.5 mL/min and the elution was monitored at 280 nm. Blue dextran (2000 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome *c* (12.4 kDa) were fractionated for calibration. Sedimentation equilibrium experiments were conducted via centrifugation of 80 $\mu$L samples adjusted to different protein concentrations, at 15000, 18000, and 25000 rpm and 20 °C for 12 h in an Optima XL-A analytical ultracentrifuge (Beckman Coulter, Krefeld, Germany) with an AN50-Ti rotor, and sedimentation velocity experiments were conducted at 45000 rpm and 20 °C for 4 h using 400 $\mu$L samples as described previously (*24*). The solvent density and viscosity at 20 °C were calculated using Sednterp. The partial specific volume and degree

[1]Abbreviations: ACTN2, actinin α2; Ahr:Arnt, aryl hydrocarbon receptor/nuclear translator heterodimer; AP-2, activator protein 2; Barbie-Box, barbiturate-inducible element; BSAP, B-cell-specific activator protein; c-Myc:Max, cMyc (cellular counterpart of viral myelocytomatosis oncogene):Max heterodimer; cap, capping signal; CD, cluster of differentiation; cDNA, complementary DNA; CG, chicken galectin; CHO, Chinese hamster ovary; CRD, carbohydrate recognition domain; CP2, CCAAT-binding protein 2; CUTL-1, cut-like homeodomain protein 1; Cys_CAM, carboxyamidomethyl cysteine; Elk-1 (TCF-A), *ets*-like factor 1 (transcription factor A); EDARADD, ectodysplasin A1 receptor-associated death domain protein; ERO1LB, endoplasmic reticulum oxidoreductin 1-like $\beta$ (*Saccharomyces cerevisiae*); EST, expressed sequence tag; *f*, experimental frictional coefficient; $f_0$, minimum frictional coefficient of a compact anhydrous sphere with the mass and volume of the proteins; FoxD3, fork head box D3; *gldd*., *glandulae*; HEATR1, HEAT repeat containing 1; HNF, hepatic nuclear factor; IgG, immunoglobulin G; GATA-3, GATA-box-binding factor 3; IP, immunoprecipitation; *l*., lamina; Lgals8, CG-8/galectin-8; MC, missed cleavages; MSO, methionine sulfoxide; MyoD, myoblast determination gene product; NF-E2, nuclear factor erythroid 2 p45; NF-$\kappa$B (p50), nuclear factor $\kappa$B (p50 subunit); Nkx2−5 (Csx), NK homeobox family member 2−5 (cardiac-specific homeobox protein, tinman homologue); NRF-2, nuclear respiratory factor 2; p300, E1A-associated 300 kDa protein; p53, 53 kDa tumor suppressor protein; Pax, paired box gene product; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; POU2F1/2 (Oct-1), POU domain class 2 transcription factor 1/2 (octamer-binding factor 1); RACE, rapid amplification of cDNA ends; RFX-1, X-box-binding factor 1 (enhancer factor c); RT-PCR, reverse transcription polymerase chain reaction; *s*, sedimentation coefficient; $s^0_{20,w}$, sedimentation coefficient corrected to standard conditions (20 °C in water) and extrapolated to zero concentration; STAT, signal transducer and activator of transcription; v-Maf, avian musculoaponeurotic fibrosarcoma virus AS42 nuclear oncoprotein; YY1, Yin and Yang 1 (nuclear factor E1, delta factor).

of hydration of the proteins were calculated from the amino acid composition. Hemagglutination assays with trypsin-treated, glutaraldehyde-fixed rabbit erythrocytes in 2-fold serial dilutions were conducted in the absence and presence of lactose as an inhibitor in 96-well microtiter plates as described previously (*25*), and lectin binding to neoglycoproteins adsorbed to the surface of plastic wells was assayed with CG-8 proteins labeled under activity-preserving conditions, and also by monitoring the pH sensitivity by covering the pH range from 4 to 11 with different buffer systems (*23*, *26*, *27*). Carbohydrate-inhibitable binding was used for calculation of the dissociation constant by processing data with the Scatchard method (*27*). CG-8-dependent staining of CHO cells [the $Pro^{-5}$ parental line and the Lec2/Lec8 mutants with deficiencies in either sialylation or galactosylation/sialylation due to mutations in the transporters for the respective activated sugars (*28*), kindly provided by P. Stanley (Albert Einstein College of Medicine, Bronx, NY)] and human Capan-1 pancreatic carcinoma cells with reconstituted expression of tumor suppressor $p16^{INK4a}$ [kindly provided by K. M. Detjen (Medizinische Klinik mit Schwerpunkt Hepatologie and Gastroenterologie, Charité-Universitätsmedizin Berlin, Berlin, Germany)] was analyzed by flow cytofluorimetry using streptavidin/(*R*)-phycoerythrin as fluorescent marker (1:40; Sigma, Munich, Germany) as described previously (*29*). A CHO cell clone with ectopic expression of α2,6-sialyltransferase I and an increased level of cell surface α2,6-sialylation was generated by transfection using the pcDNA3.1 vector system (Invitrogen, Karlsruhe, Germany) and the cloned cDNA sequence for this enzyme and selected by screening with biotinylated *Sambucus nigra* agglutinin as described previously (*30*). Enzymatic α2,3-desialylation was performed with the specific neuraminidase from *Salmonella typhimurium* (25 units/50 μL) (New England BioLabs, Frankfurt, Germany) for 1 h at 37 °C using $4 \times 10^5$ cells. The binding parameters (percentage of positive cells and median fluorescence intensity) were computed with the FACScan instrument's software (Becton-Dickinson, Heidelberg, Germany). Carbohydrate-independent binding was minimized by the presence of 100 μg/mL sugar-free bovine serum albumin, and the extent of carbohydrate-inhibitable binding was probed with increasing concentrations of lactose and a mixture of lactose (75 mM) and the glycoprotein asialofetuin (1 mg/mL).

*Expression Profiling by RT-PCR, Western Blotting, and Immunohistochemistry*. Preparations of cDNAs of tested tissues were obtained as described previously (*15*). To harvest sufficient yields of total RNA in the cases of hyaline cartilage (joint surface), long bone (tibia), and skeletal muscle, 100 mg tissue samples were homogenized in 1 mL of TriFast (Peqlab, Erlangen, Germany) using a Precellys Homogenizer (Peqlab) and 2.0 mL screw-cap tubes prefilled with steel beads (diameter of 2.8 mm). Eight runs, each of a period of 10 s, with an interval of 20 s between runs to cool samples on ice were performed at a speed of lysis motion of 6500 rpm. Then samples were kept at room temperature for 5 min. After addition of 0.2 mL chloroform, the samples were vigorously shaken for 15 s and then kept at room temperature for 10 min. During centrifugation at 12000*g*, the mixture separated into the lower phenol/chloroform phase, the interface, and the upper aqueous phase, used for routine RNA preparation. The processed cartilage had been carefully removed from joint surfaces, thereby preventing contaminations from joint capsules in the preparation. Bone marrow was obtained from long bones (tibiae) by thorough washing and centrifugation (2000*g* for 30 s at 4 °C), resulting in complete separation. RT-PCR analyses used the sense primer 5′-ATGATGTCCTTGGATGGAC-3′ and the antisense primer 5′-GGAACTCCAAATTGTAAAACC-3′ and for the loading control established by chicken β-actin-specific mRNA the sense primer 5′-GATGATGATATTGCTGCGC-3′ and the antisense primer 5′-GGTGAAGCTGTAGCCTC-3′. Products 474 and 531 bp in length were obtained after 35 cycles for CG-8I and CG-8II, respectively, with a 468 bp sequence in the case of the β-actin control. The presence of the CG-8II-specific cDNA harboring the sequence of exon V for the linker extension was independently probed with a suitable primer set established by the sense primer 5′-ATCACAGAATTAGCCTTGAA-3′ annealing in the 57 bp sequence stretch exclusively present in the coding sequence of the CG-8II-specific mRNA and the antisense primer 5′-TGTGTTTATCTTTGTTACTCCT-3′, resulting in an amplification product of 123 bp.

A polyclonal anti-CG-8 antibody preparation free of cross-reactivity against the other CGs (CG-1A, -1B, -2, and -3) was obtained after a rabbit had been immunized and any cross-reactive material had been removed from the immunoglobulin G (IgG) fraction by consecutive affinity chromatographical depletion steps using four different batches of affinity resin (Affi-Gel 10/15; Bio-Rad, Munich, Germany), to which one of the other CGs was covalently conjugated at levels between 7.5 and 10.2 mg/mL (*31*, *32*). Finally, following removal of cross-reactive material, CG-8-reactive IgG was purified by affinity chromatography on CG-8-presenting resin. Successful processing was verified by Western blots and ELISA tests (*32*). The affinity-purified preparation of anti-CG-8 IgG facilitated the lectin's detection by combined immunoprecipitation and Western blotting in extracts obtained by tissue/cell homogenization in 50 mM Tris-HCl buffer (pH 8.0) containing 150 mM NaCl and 1% NP40 using 2.5 mg of total protein per assay and by Western blotting from extracts of mononuclear cells isolated from spleens and bursae of Fabricius by standard Ficoll-Paque-based centrifugation steps (GE Healthcare) and extract preparation, using 50 μg of protein/lane. Extraction of protein from cartilage, long bone, and heart muscle followed the protocol for RNA preparation with the buffer given above. The immunoprecipitation protocol started with removal of any proteins binding nonspecifically to the used beads after a 1 h incubation step by centrifugation, and then supernatants were mixed with a solution containing 2.5 μg of affinity-purified anti-CG-8 IgG. Following a 2 h incubation step on ice and then addition of a 25 μL solution with TrueBlot anti-rabbit IgG IP beads (eBioscience, San Diego, CA) and incubation overnight at 4 °C, the samples were centrifuged at 10000*g* for 1 min. Supernatants were removed completely, and pelleted beads were carefully washed three times with 250 μL of the homogenization buffer. Proteins were released from the beads when the suspension was heated to 95 °C for 10 min in 50 μL of 25 mM Tris-HCl buffer (pH 6.5) containing 6% sodium dodecyl sulfate, 10% glycerol, 50 mM dithiothreitol, and bromophenol blue and separated from beads by centrifugation at 10000*g* for 3 min. Gel electrophoresis, electrophoretic transfer to nitrocellulose, probing with antibody, and signal generation by chemiluminescence were performed as described elsewhere (*15*, *33*). As deviation from the cited protocol, washing steps were conducted for 1 h at room temperature with 10 changes of buffer [50 mM Tris-HCl (pH 7.5) containing 150 mM NaCl and 0.1% Tween 20], and the rabbit TrueBlot reagent (eBioscience) was applied to minimize interference by visualizing heavy and light chains of the rabbit antibody.

Loading control was routinely run by visualizing $\beta$-actin. Immunohistochemical processing of fixed and paraffin-embedded sections of organs of adult (6-month-old) chicken followed an optimized protocol with stringent specificity controls established for respective analysis of the proto-type CGs (*15*), using an AxioImager.M1 microscope (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) equipped with an AxioCam MRc3 and MRc digital camera and Axiovision version 4.6. Photodocumentation of fluorescence emission was recorded after application of the Vector Red Alkaline Phosphatase Substrate Kit I reagents (Axxora, Grünberg, Germany) using the 560 and 630 nm filter combination.

*Sequence Processing*. Location of gene sequence(s) for CG-8 was independently performed using either the known EST (please see above) or a TBLAST search through the most recent version of annotation of the chicken genome available from the NCBI (http://www.ncbi.ulm.nik.gov/Genbank/index.html) guided by the previously defined consensus sequence (*15*). The chromosomal environment was comparatively analyzed using the Ensembl Genome Browser (http://www.ensembl.org/index.html) and/or the NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/). To identify putative binding sites for transcription factors in the proximal promoter region of the CG-8 gene and the genes for the proto-type CGs, the sequence stretching from 2000 bp upstream to 150 bp downstream of the starting point for translation (CG-8 and CG-1A) and transcription (CG-1B and CG-2) was 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 (*15*, *34*).

## RESULTS

*Gene Structure of CG-8*. The entry into the EST database facilitated a systematic search for the respective gene(s) on the genomic level. Independently, a previously defined consensus sequence for the CRD of galectins (*15*) enabled us to scour database information for matching sequences. Both approaches led to one hit on chromosome 3 (minus strand). Figure 1 illustrates the assignment of genes flanking the detected sequence. Upon comparative analysis among species, this pattern was also found in the cases of *Homo sapiens* (on chromosome 1), *Pan troglodytes* (on chromosome 1), *Mus musculus* (on chromosome 13), *Canis familiaris* (on chromosome 4), *Cavia porcellus* (not yet assigned to chromosomes), and *Felis catus* (not yet assigned to chromosomes). Insertion of multiple kinase-like genes between the gene for the ectodysplasin A1 receptor-associated death domain protein and the galectin gene constitutes a deviation from this pattern in the genome of *Rattus norvegicus*, resulting in a 1.5 Mb separation relative to 10 kb in the chicken genome or 35 kb in the human genome. Besides mammals, marsupials (here *Monodelphis domestica*) present this pattern. It is then gradually lost when the mapping is extended to amphibians (i.e., *Xenopus tropicalis*) and several fish species. Noteworthy is the fact that the same pattern of identity and its gradual loss are encountered in the corresponding analysis of the chromosomal environment of galectin-2 genes from these species (not shown), a strong case for homology.

The identified gene consists of sections for two different canonical lectin sites. Thus, the chicken genome contains a single gene with characteristics of a tandem-repeat-type galectin. To solidify its preliminary classification as CG-8 based on the partial

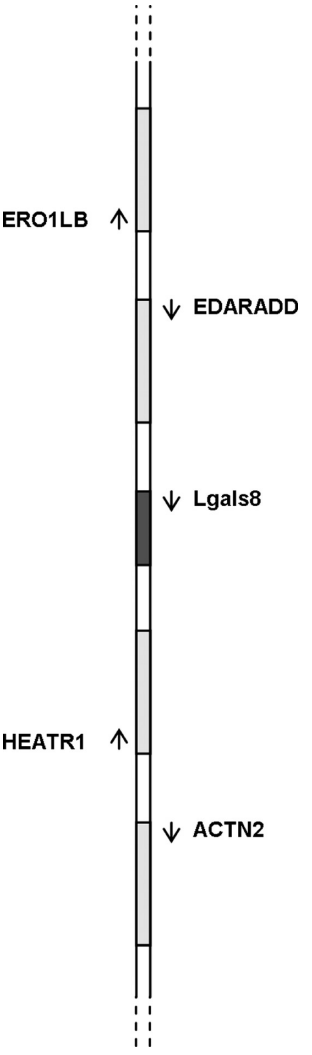


FIGURE 1: Chromosomal environment of the CG-8 gene (dark box) on chromosome 3. Gene orientation is indicated by arrows; box lengths and spacers are not drawn to scale. Abbreviations: Lgals8, CG-8/galectin-8; ACTN2, actinin $\alpha$2; EDARADD, ectodysplasin A1 receptor-associated death domain protein; ERO1LB, endoplasmic reticulum oxidoreductin 1-like $\beta$ (*Saccharomyces cerevisiae*); HEATR1, HEAT repeat containing 1.

EST information, the organization of the gene structure was defined first (Figure 2, top panel). The two CRDs are established by four (N-domain) and three (C-domain) exons. The two remaining exons can be assigned to the linker region, only one being represented in the EST information. Following cloning of complete cDNA from jejunum after rapid amplification of cDNA ends to move beyond the EST-defined stretch, it became evident that exon V can be subject to alternative splicing, therefore giving rise to two forms of mature mRNA of 891 and 948 bp, which differ in the length of the region encoding the linker (Figure 2, bottom panel). Their sequencing solidly verified the occurrence of alternative splicing of exon V. The connecting region between the CRDs appears very short in CG-8I with nine amino acids, extended to 28 amino acids by an addition of 19 amino acids in CG-8II. Central for definitive classification of this lectin are homology evaluations, which could now be based on the sequences for both CRDs. The alignment scores based on respective GenBank entries (accession numbers DQ443719 and AY840215, later NM_001010843) were markedly higher in the cases of sequences for mammalian galectin-8 forms than those of

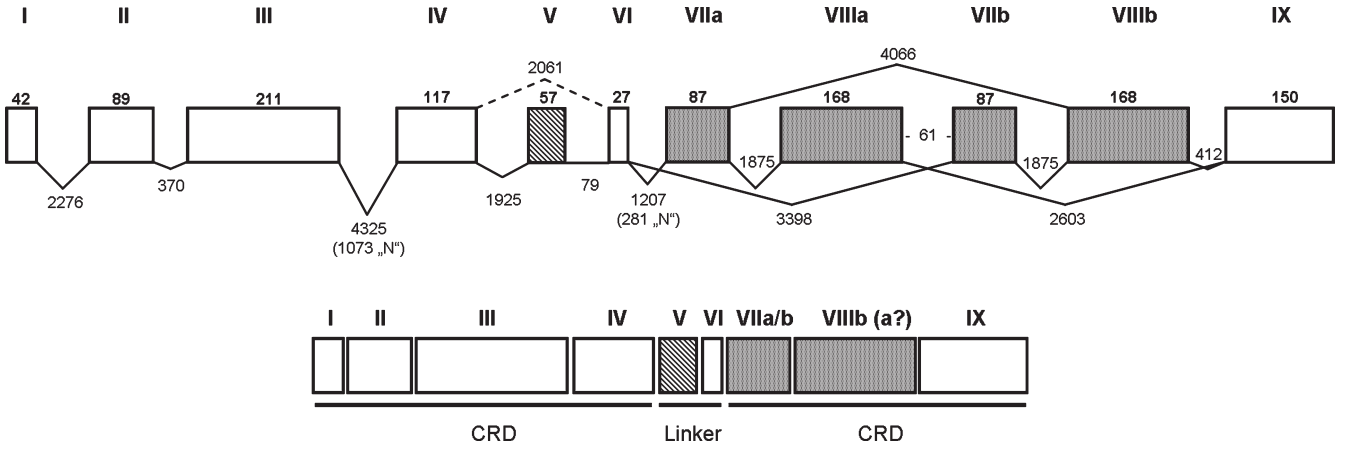


FIGURE 2: Organization of the CG-8 gene and its mRNAs. Exons are given as boxes (size proportional to exon length) and introns as lines (length not proportional) (top panel). Roman numbers reflect the order of the exons; bold Arabic numbers define the length of each exon, and numbers at the center of lines denote the lengths of introns (in base pairs). N indicates stretches of sequence not available in the database (due to incomplete annotation). Exon V (hatched) represents the sequence for linker extension. The genomic region between exons VI and IX is constituted by a tandem-repeat display established by almost perfect duplication (98.8% identical) of a sequence stretch containing both exons VII and VIII and the connecting intron. These two exon pairs are therefore shown as shaded boxes and numbered VII and VIII with a or b, respectively. When the sequences in the exon pairs are compared, the only difference is found at position 33 in exon VIII, with G in VIIIa and A in VIIIb. This deviation does not lead to a change in amino acid sequence (Ser encoded by UCG in VIIIa and UCA in VIIIb). Comparison with the obtained cDNA sequence reveals the presence of exon VIIIb in cDNA from chicken jejunum. Lines connecting exons depict actual and potential splicing events, generating the mature mRNAs with eight or nine exons: exons V and VI establish the linker, and removal of exon V from the mRNA underlies the separation between CG-8I and CG-8II (bottom panel).

the other mammalian tandem-repeat-type proteins or the other CGs when respective comparisons were computed. This result justifies this lectin's designation as CG-8. As typically seen for tandem-repeat-type galectins, the sequences of the two CRDs diverge markedly. To address the question of whether both mRNA species will lead to functional protein, recombinant production was established.

*Mass Profiling of CG-8*. Yields of 25−30 mg/L of culture were obtained. The gel electrophoretic mobilities of the two proteins reflect the calculated disparity in molecular mass between the 296-amino acid protein with the nine-amino acid linker (termed CG-8I, 33703 Da) and the 315-amino acid form with a linker length extended to 28 amino acids (termed CG-8II, 35567 Da) (Figure 3). Of note is the fact that a deviation from the genomic sequence in triplet 45 from UUC to CUC in the cDNA of the latter form entails an F to L substitution. To answer the question of whether the measured mobility difference between CG-8I and CG-8II is exclusively due to the alteration in linker length, the proteins were subjected to mass spectrometric analysis and fingerprinting after tryptic digestion.

Experiments with the unprocessed proteins revealed molecular mass values of 33704 ± 9 Da for CG-8I and 35565 ± 9 Da for CG-8II. These values exclude the presence of any posttranslational modifications. To confirm the variation in linker length as the sole source of the mass difference, trypsin treatment and peptide analyses were performed after full reduction and alkylation. Table 1 shows the list of peptides detected in the fingerprint analysis. The assignment was directed by the cDNA sequences but also had to consider modifications of certain amino acids during processing. In detail, the presence of carboxyamidomethyl cysteines, the formation of an internal disulfide bond, and oxidative modification of some methionines were observed. The sequence coverages by the detected peptides, highlighted in the amino acid sequences of CG-8I and CG-8II given as a footnote to Table 1, were 78 and 54%, respectively. Of particular note is the fact that the region characteristic of the extended form of the linker, namely, amino acids 154−172

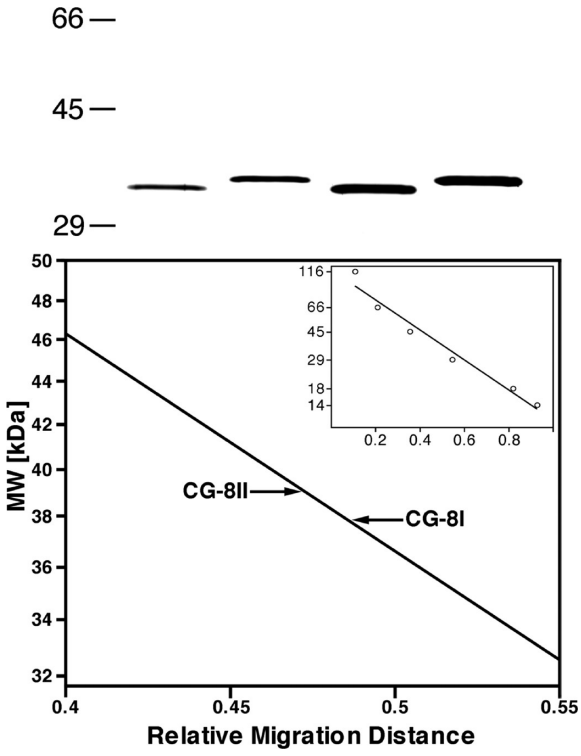


FIGURE 3: Gel electrophoretic mobility of the two CG-8 forms. Purified proteins from recombinant production were separated by polyacrylamide gel electrophoresis under denaturing conditions in the presence of $\beta$-mercaptoethanol in a 12.5% running gel. The top panel shows the bands for CG-8 with standard (CG-8I, 40 and 80 ng per lane) and extended (CG-8II, 30 and 60 ng per lane) lengths of the linker at two given quantities of each form in pairwise arrangement. 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 used standard proteins [$\beta$-galactosidase (116 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa), $\beta$-lactoglobulin (18.4 kDa), and lysozyme (14.2 kDa)].

Table 1: Tryptic Peptides of CG-8I and CG-8II Detected by Fingerprint Analysis[a]

| *m*/*z* | matching mass (Da) | MC | modification | position | peptide |
|---|---|---|---|---|---|
| | | | CG-8I | | |
| 1048.58 | 1048.59 | 0 | | 204−212 | DIALHLNPR |
| 1065.59 | 1065.59 | 0 | | 121−128 | HLLLYNHR |
| 1173.58 | 1173.58 | 0 | | 59−68 | ADVAFHFNPR |
| 1364.60 | 1364.62 | 0 | disulfide, 74.77 | 71−82 | WSGCVVCNTLER |
| 1415.83 | 1415.84 | 0 | | 99−110 | GRPFEIVIMILK |
| 1431.83 | 1431.84 | 0 | MSO, 107 | 99−110 | GRPFEIVIMILK |
| 1465.80 | 1465.82 | 0 | Cys_CAM, 178 | 170−183 | LVSALHPGCTVAIK |
| 1480.64 | 1480.67 | 0 | Cys_CAM, 74.77 | 71−82 | WSGCVVCNTLER |
| 1552.80 | 1552.81 | 1 | | 199−212 | SSDSKDIALHLNPR |
| 1593.66 | 1593.68 | 0 | | 221−233 | NSYLHDSWGEEEK |
| 1634.82 | 1634.83 | 0 | Cys_CAM, 51 | 45−58 | FQVDLQCGSSIKPR |
| 1755.81 | 1755.83 | 1 | Cys_CAM, 74.77 | 69−82 | FKWSGCVVCNTLER |
| 1797.01 | 1797.01 | 0 | | 279−294 | INLLEVTGDVQLLDVR |
| 2100.93 | 2100.97 | 1 | | 83−98 | EKWGWEEITYEMPFQK |
| 2116.95 | 2116.96 | 1 | MSO, 94 | 83−98 | EKWGWEEITYEMPFQK |
| 2470.11 | 2470.20 | 0 | | 147−169 | SIEFVSNGEMPDGLQFGVPYVGK |
| 2486.12 | 2486.19 | 0 | MSO, 156 | 147−169 | SIEFVSNGEMPDGLQFGVPYVGK |
| 2492.17 | 2470.20 + 22 | 0 | $Na^+$ adduct | 147−169 | SIEFVSNGEMPDGLQFGVPYVGK |
| 2508.11 | 2486.19 + 22 | 0 | MSO, 156 $Na^+$ adduct | 147−169 | SIEFVSNGEMPDGLQFGVPYVGK |
| 3510.78 | 3510.92 | 0 | | 11−44 | ISNPIIPYVGTILGGLVPGELIVLHGSIPDDADR |
| | | | CG-8II | | |
| 1048.59 | 1048.59 | 0 | | 223−231 | DIALHLNPR |
| 1065.60 | 1065.59 | 0 | | 121−128 | HLLLYNHR |
| 1173.58 | 1173.58 | 0 | | 59−68 | ADVAFHFNPR |
| 1329.71 | 1329.72 | 0 | | 278−289 | VAVNGVHTLEYK |
| 1364.62 | 1364.62 | | disulfide 74−77 | 71−82 | WSGCVVCNTLER |
| 1423.66 | 1423.62 | | 1 Cys_CAM | 71−82 | WSGCVVCNTLER |
| 1465.81 | 1465.82 | 0 | Cys_CAM, 197 | 189−202 | LVSALHPGCTVAIK |
| 1480.67 | 1480.67 | 0 | Cys_CAM, 74.77 | 71−82 | WSGCVVCNTLER |
| 1552.81 | 1552.81 | 1 | | 218−231 | SSDSKDIALHLNPR |
| 1593.68 | 1593.68 | 0 | | 240−252 | NSYLHDSWGEEEK |
| 1600.86 | 1600.85 | 0 | Cys_CAM, 51 | 45−58 | LQVDLQCGSSIKPR |
| 1755.83 | 1755.83 | 1 | Cys_CAM, 74.77 | 69−82 | FKWSGCVVCNTLERR |
| 1797.01 | 1797.01 | 0 | | 298−313 | INLLEVTGDVQLLDVR |
| 1843.84 | 1843.84 | 0 | | 85−98 | WGWEEITYEMPFQK |
| 2121.09 | 2121.08 | 0 | | 147−167 | SIEFVSNSVQGAQPSSVGVTK |
| 2143.07 | 2121.08 + 22 | | $Na^+$ adduct | 147−167 | SIEFVSNSVQGAQPSSVGVTK |
| 2265.08 | 2265.09 | 0 | | 168−188 | INTENGEMPDGLQFGVPYVGK |
| 2281.05 | 2281.08 | 0 | MSO, 175 | 168−188 | INTENGEMPDGLQFGVPYVGK |

**CG-8I** [b]

```
     1        11        21        31        41        51
     |        |         |         |         |         |
  1  MMSLDGPQKK ISNPIIPYVG TILGGLVPGE LIVLHGSIPD DADRFQVDLQ CGSSIKPRAD
 61  VAFHFNPRFK WSGCVVCNTL EREKWGWEEI TYEMPFQKGR PFEIVIMILK DKFQVSVNKK
121  HLLLYNHRIS LERIDTLGIY GKVQIKSIEF VSNGEMPDGL QFGVPYVGKL VSALHPGCTV
181  AIKGEVNKNP KSFTINLKSS DSKDIALHLN PRLKNKVFVR NSYLHDSWGE EEKEVTNFPF
241  SPGMYFELII FCDAHQFKVA VNGVHTLEYK HRFKQLEKIN LLEVTGDVQL LDVRSW
```

**CG-8II** [b]

```
     1        11        21        31        41        51
     |        |         |         |         |         |
  1  MMSLDGPQKK ISNPIIPYVG TILGGLVPGE LIVLHGSIPD DADRLQVDLQ CGSSIKPRAD
 61  VAFHFNPRFK WSGCVVCNTL EREKWGWEEI TYEMPFQKGR PFEIVIMILK DKFQVSVNKK
121  HLLLYNHRIS LERIDTLGIY GKVQIKSIEF VSNSVQGAQP SSVGVTKINT ENGEMPDGLQ
181  FGVPYVGKLV SALHPGCTVA IKGEVNKNPK SFTINLKSSD SKDIALHLNP RLKNKVFVRN
241  SYLHDSWGEE EKEVTNFPFS PGMYFELIIF CDAHQFKVAV NGVHTLEYKH RFKQLEKINL
301  LEVTGDVQLL DVRSW
```

[a] MC, missed cleavages; MSO, methionine sulfoxide; Cys_CAM, carboxyamidomethyl cysteine. [b] The complete amino acid sequences of recombinant CG-8I and CG-8II, showing amino acid numbering and highlighting sequence coverage by peptide mass fingerprinting, are listed. Theoretical masses were calculated from the sequences using the PeptideMass tool available at the ExPASy Proteomics Server (www.expasy.ch/tools/peptide-mass.html).

(SVQGAQPSSVGVTKINTEN), was covered by peptides 147−167 and 168−188, appearing at *m*/*z* 2121.09/2143.07 and 2265.08/2281.05, respectively. None of these ions were present in tryptic digests of CG-8I. Fittingly, different ions from peptide 147−169 of CG-8I were observed at *m*/*z* 2470.11−2508.11. In turn, none of them was detected in the digests of CG-8II. Turning to the F to L substitution at position 45 noted above, we reason the fingerprinting should uncover a decrease of 33.98 Da in the

mass of peptide 45−58 from CG-8II (ion appearing at $m/z$ 1600.86) compared to the same peptide from CG-8I (ion at $m/z$ 1634.82). This predicted difference was ascertained on the level of the protein. It was the only disparity found between CG-8I and CG-8II apart from linker extension. With the two proteins of stringently controlled sequence at hand, the questions about aggregation in solution, hydrodynamic and cross-linking properties, and the influence of linker length on these features could be addressed, starting experimentally with gel filtration.

*Quaternary Structure and Hydrodynamic Properties*. Gel filtration chromatography was performed in the presence of 0.1 M lactose, which will block any carbohydrate-inhibitable interaction with the resin and has a positive effect on protein solubility. CG-8I and CG-8II eluted predominantly as a single sharp peak centered at 28.3 ± 0.3 and 27.9 ± 0.2 min, respectively. These elution times correspond to apparent molecular masses for globular proteins of 24 ± 2 and 27 ± 1.5 kDa, respectively. For the sake of comparison to homodimeric prototype family members, CG-1A and CG-1B eluted at earlier times (27.4 ± 0.1 and 27.6 ± 0.2 min, respectively), as did human galectin-1 at 27.5 ± 0.2 min (*15*, *16*). These parameters result in an apparent molecular mass of ~30 kDa. Thus, the two CG-8 proteins did not aggregate under these conditions. They exhibited a slightly different chromatographic behavior, attributable to linker-dependent size differences, relative to each other, and also relative to CG-1A/B. The aggregation status was independently studied by ultracentrifugation.

Ultracentrifugation analysis was also performed in the presence of 0.1 M lactose. Sedimentation equilibrium data obtained for CG-8I and CG-8II, at a loading concentration of 0.51 mg/mL, could be fitted to a single ideal component with weight-average molecular masses of 33.1 ± 2.2 and 36.0 ± 1.1 kDa, respectively. The values were slightly higher than the results for CG-1A/B at ~30 kDa. The observed masses were independent of rotor speed, indicating that the samples were homogeneous. Consistent with gel filtration, the sedimentation equilibrium analysis revealed that the two CG-8 proteins behaved as monomers in solution. The hydrodynamic shape of the molecules was examined next by sedimentation velocity experiments.

In the presence of 0.1 M lactose, both CG-8 proteins sedimented as a main peak with sedimentation coefficients of 2.2 ± 0.1 and 2.25 ± 0.15 S, respectively (2.65 and 2.7 for $s^0_{20,w}$, respectively, expressed in terms of the standard solvent of water at 20 °C). These values are similar to those found for homodimeric CG-1A and CG-1B as well as human galectin-1 (*15*, *16*) analyzed under the same conditions. An estimation of the hydrodynamic size of hydration of the proteins was obtained using Sednterp. The partial specific volume and degree of hydration were calculated from the amino acid composition given by the determined sequences (Table 1). Table 2 compiles the ratio between the experimental frictional coefficient ($f$) and the minimum frictional coefficient of a compact anhydrous sphere with the mass and volume of the proteins ($f_0$), and the approximation of the shape of the molecules from this frictional ratio to prolate or oblate ellipsoids of revolution. For the sake of comparison, results of a similar analysis of CG-1B studied both in the presence and in the absence of lactose are included. Noticeably, using either the prolate or oblate models, the shorter $b$-axis was very similar for the three proteins, whereas differences were significant for the longer $a$-axis. Therefore, with the caution imposed by the limitations of the approach, the sedimentation velocity analysis does not provide evidence of major differences between the overall shape of the tandem-repeat-type proteins CG-8I and II and the proto-type galectin CG-1B other than by mass-proportional alteration of the length. Although the analysis of CG-1B in the presence and absence of lactose yielded similar results once the effect of the density and viscosity of the solvent on the sedimentation coefficient was corrected, we cannot exclude an influence of the presence of ligand on CG-8's hydrodynamic properties on the grounds of the marked effect of the presence of lactose on CG-8 solubility. Overall, these results suggest that the studied proto- and tandem-repeat-type CGs share the topological arrangement of the two CRDs.

Table 2: Results of the Sedimentation Velocity Analysis and Hydrodynamic Modeling[a]

| | CG-8I | CG-8II | CG-1B | |
|---|---|---|---|---|
| buffer system and additive | PBS with lactose (0.1 M) | PBS with lactose (0.1 M) | PBS with lactose (0.1 M) | PBS |
| hydrodynamic parameter | | | | |
| $S$ (S) | 2.2 | 2.25 | 2.25 | 2.55 |
| $s^0_{20,w}$ (S) | 2.64 | 2.7 | 2.68 | 2.67 |
| $f/f_0$[b] | 1.32 | 1.35 | 1.26 | 1.26 |
| estimated axial dimensions (nm) | | | | |
| prolate model | | | | |
| a/b | 3.6 | 3.9 | 2.85 | 2.9 |
| 2a | 11.7 | 12.55 | 9.55 | 9.7 |
| 2b | 3.25 | 3.2 | 3.35 | 3.3 |
| oblate model | | | | |
| a/b | 3.8 | 4.15 | 2.95 | 3.0 |
| 2a | 7.8 | 8.15 | 6.8 | 6.9 |
| 2b | 2.05 | 1.95 | 2.3 | 2.25 |

[a] The experiments were conducted in phosphate-buffered saline (PBS) containing 2 mM dithiothreitol, in the presence or absence of 0.1 M lactose. [b] The frictional ratio ($f/f_0$) and the axial dimensions of prolate and oblate models were computed using the Teller method.

*Cross-Linking and Cell Binding*. The indicated shape similarity intimates that the CG-8 proteins will then likely also share the cross-linking capacity of proto-type galectins. Indeed, hemagglutination was observed at minimal concentrations of ~400−600 ng/assay, CG-8I consistently being slightly more active than CG-8II. Cross-linking of erythrocytes by the monomeric but bivalent lectins was impaired by the presence of 3 mM lactose (inhibitory concentration for 50% inhibition), as was the case for CG-1B. Osmolarity controls with mannose were negative. Sugar recognition was also seen in solid-phase assays, which gave saturable and nearly completely inhibitable binding to a lactose-bearing matrix of neoglycoprotein with $K_D$ values in the range of 40 nM. The pH profile of binding activity yielded a broad peak between pH 5 and 8, which is typical for the proto-type CGs tested in parallel. This assay and the evidence from blocking hemagglutination and affinity chromatographic purification on immobilized ligand converge to document reactivity of the CG-8 proteins with lactose, the hallmark of galectin specificity. Evidently, variation in linker length did not significantly affect binding of ligand to the CRDs.

In addition to a further confirmation of specificity, cell binding assays were performed to probe the potential of the linker length difference to affect the galectins' reactivity with surface glycans.

After having ascertained the concentration dependence for biotinylated lectin and inhibition by lactose, excluding a sizable contribution of protein−protein interaction to fluorescent staining (Figure 4, top panel), we determined binding parameters for both CG-8 forms with CHO cells under identical conditions. No notable difference in fluorescent staining by CG-8I/II was revealed. At the same time, the essential nature of galactosylation with a rather small influence of α2,3-sialylation to enhance cell reactivity and of α2,6-sialylation to reduce this parameter was documented (Figure 4, middle and bottom panels). For this purpose, CHO mutants with a reduced level of sialylation (Lec2) or galactosylation (Lec8) and a transfectant with an increased level of α2,6-sialylation were used. Fittingly, a respective decrease in binding parameters was registered after enzymatic α2,3-desialylation of cell surfaces of the transfectant, with fluorescence intensity reduced from 29.9 to 23−25 in representative experiments. Since spatial properties of glycan presentation can differ among cell types, further assays were warranted. To track down a respective influence, binding parameters were also determined for the Capan-1 pancreatic carcinoma cell line. In this cell type, we measured a higher reactivity for CG-8I (64.3 and 84.2% positive cells, intensity levels of 18.1 and 42.4 at 1 and 2 μg of lectin/mL, respectively) than for CG-8II (37.4 and 60.3% positive cells, intensity levels of 16.3 and 24.6 at 1 and 2 μg/mL, respectively) in representative experiments. Consistent with the noted tendency in hemagglutination assays, the extension of linker length, if of influence, can thus account for a signal decrease in cell binding. Whether such results can actually matter for the physiological context depends entirely on the synthesis of the isoforms in vivo. In principle, sequencing after cDNA cloning from jejunum had revealed the occurrence of alternative splicing, and the activity assays reveal both proteins to be active as lectins. The given indication for natural variability due to alternative splicing prompts us to study the occurrence of these two CG-8 forms on the level of mRNA and protein for different tissues. These data, of course, will define the expression profile of the unique tandem-repeat-type galectin in chicken.

*Expression Profiling of CG-8*. We started the respective analysis at the highest level of sensitivity with RT-PCR to exclude false-negative results. The presence of signals was detected rather uniformly, with quantitative differences between the tissue types and the two mRNA species (Figure 5). Because of the small length difference, which makes it difficult to discern the presence of the upper band in gels unambiguously, this mRNA species was amplified separately by a primer set targeting the sequence in exon V, the distinctive property of CG-8II. The obtained results validate the broad-spectrum presence of both mRNAs (Figure 5). On the level of the protein, the processing of the antiserum by stringent negative selection steps against the other four CGs followed by affinity chromatography on resin presenting a CG-8I/CG-8II mixture ensured the absence of cross-reactivity against other CGs and specificity to both forms of CG-8 (Figure 6). Western blotting after immunoprecipitation resulted in visualization of both proteins with no evidence of proteolytic processing such as linker cleavage. Of note is the fact that extracts of several organs contained either exclusively CG-8II or a mixture of both proteins, as also seen in the tested preparations of mononuclear and bone marrow cells (Figure 6). In other words, no tested extract was confined to the presence of CG-8I. Having herewith characterized the profile of protein expression in extracts, we next addressed the issue of CG-8 localization on the cellular level by immunohistochemistry.

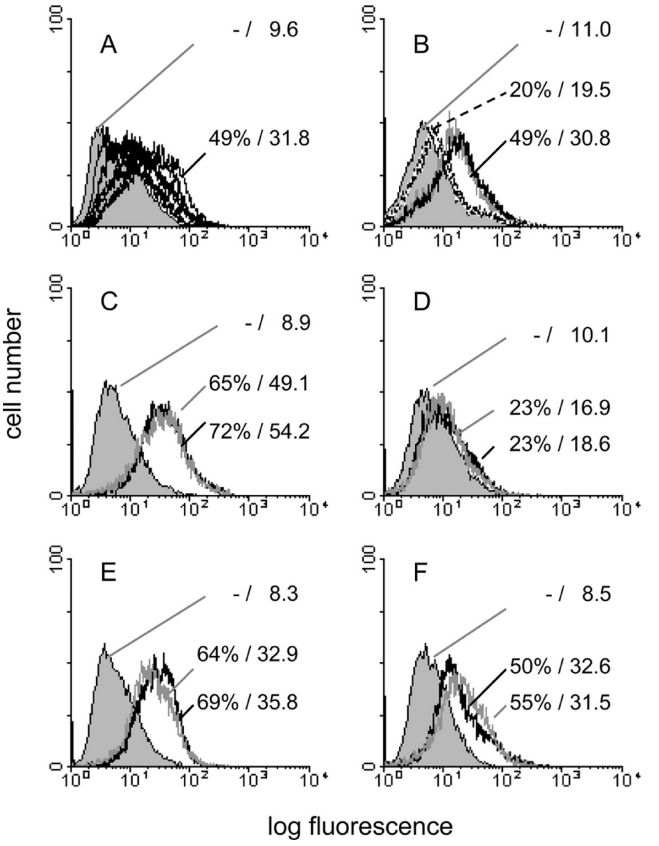


FIGURE 4: Cell surface staining by the two forms of CG-8. Semilogarithmic representation of fluorescent surface staining of the parental CHO line $Pro^{-5}$ (lacking expression of β1, 4-galactosyltransferase VI) and the Lec2/Lec8 mutants by biotinylated CG-8I and II. The concentration dependence (A) (1, 2, 5, 10, and 20 μg of CG-8II/mL), inhibition by lactose (B) (10 and 20 μg of CG-8I/mL, in the presence of 25 mM lactose at 20 μg/mL), and comparative analysis (C) [20 μg/mL CG-8I (gray) and CG-8II (black)] are given for the parental cells. Further comparative scans on staining of Lec8 (D) (decreased level of galactosylation) and Lec2 mutant cells (E) (decreased level of sialylation) as well as on transfectant cells stably overexpressing α2,6-sialyltransferase I (F) are presented. CG-8I (gray) and CG-8II (black) were used with aliquots of the same cell batch at 20 μg/mL. The characteristics of background staining when using the fluorescent indicator without prior incubation of cells with biotinylated galectin are illustrated as a reference in each panel by the shaded area. Quantitative data on the percentage of positive cells (%) and mean fluorescence intensity are inserted for the experiments with a lectin concentration of 20 μg/mL.

Distinct cell positivity was obtained as summarized in Table 3. Neither epithelia nor smooth muscle cell layers were positive. Parenchymatous organs such as lung and kidney presented positive cells in the stroma of parabronchal walls together with alveolar macrophages and embedded in connective tissue, especially between collecting ducts, which surrounds renal corpuscles (Figure 7A−C). Macrophages, together with thymocytes, were stained in the thymus, reflecting the results in Western blotting with extracts obtained either from the organ or from isolated mononuclear cells (Table 3 and Figure 6). A prominent site of positivity is the *l. propria mucosae*. Clustering of CG-8-expressing cells is typical for esophagus, larynx, or trachea; few cells are reactive in the respective interstitial compartment between glands in proventriculus or gizzard, and a regular pattern of intensely stained cells was seen in *l. propria mucosae* surrounding glands or forming cores of villi in the gut (Table 3 and Figure 7D−I). Follicular atresia in the ovary is associated with a pronounced

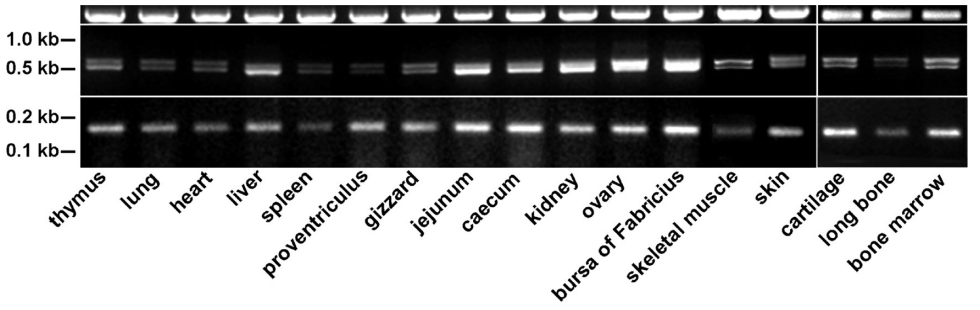


FIGURE 5: Expression profiling of the two CG-8 forms by RT-PCR. The presence of CG-8-specific mRNAs was detected in samples from a panel of tissues. Running the loading control for $\beta$-actin (468 bp; top panel) routinely in parallel, we amplified both types of CG-8-specific mRNA (474 bp for CG-8I and 531 bp for CG-8II) and visualized them using either the same primer set (middle panel) or a primer set specific for the extension in linker length due to exon V in CG-8II (123 bp; bottom panel). Positions of molecular weight markers are indicated (left).

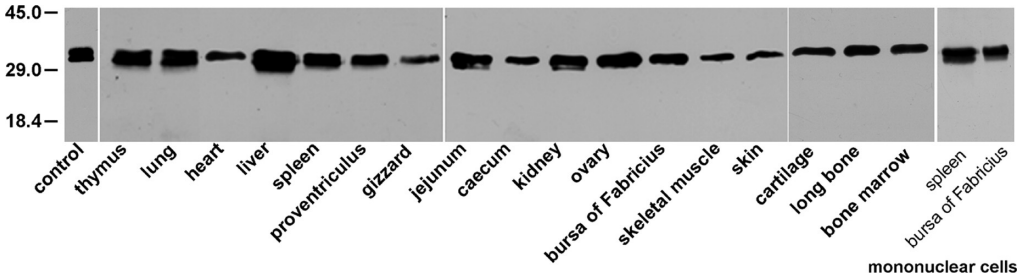


FIGURE 6: Expression profiling of the two CG-8 forms by Western blotting. The presence of lectin in tissue extracts (2.5 mg of total protein processed by immunoprecipitation) and in mononuclear cells (50 $\mu$g) was detected with affinity-purified anti-CG-8 IgG fractions free of cross-reactivity to the other four CGs (0.5 $\mu$g/mL). Positions of molecular weight markers are indicated, and a control with purified CG-8I and II (25 ng each) is given (left).

positivity in cells which are arranged as a thin layer between theca cells and the partially invaginated granulosa cells (Table 3). A distinct cell positivity was also noted in cerebrum and cerebellum (Table 3 and the figure in the Supporting Information). In general, staining was mostly cytoplasmic, with scattered extracellular signals observed in dermis and subcutis of the skin (Table 3). No nuclear staining was detected. A control section documents the lack of antigen-independent staining (Figure 7J). The qualitative differences with respect to tissue parts with strong positivity for proto-type CGs, e.g., hepatocytes and CG-1A, respiratory epithelium and CG-1B, and epithelial lining of gut villi and CG-2, serve as inherent quality controls. As a consequence, they signal pronounced disparities in the profiles of gene regulation, pointing to an underlying sequence diversity in the promoter regions. Systematic computer searches put this hypothesis to the test.

As a first step to give respective research on gene regulation direction, putative target sequences for transcription factors were identified. The resulting list has a series of characteristic features (please see the table in the Supporting Information). A relative abundance of sequence motifs which can bind homeobox proteins such as Cdx1, HOX A3, and Pax2, a distinctive profile for rather ubiquitous activators (positivity for C/EBP, E47, or RFX1, whereas no Sp1-specific sites were found), the occurrence of three sites with potential reactivity for HNF-4$\alpha$2/CouP-TF, and the presence of putative motifs for NF$\kappa$B and the Ahr:Arnt heterodimer (the so-called xenobiotic response element) may underlie the tight regulation and indicate regulatory responses to inflammation and infection. Table 4 summarizes the qualitative differences in this respective region for the genes encoding CG-8 and the three proto-type CGs.

## DISCUSSION

The tandem-repeat-type organization of galectin structure is characterized by the covalent connection of two CRDs through a linker peptide. The sequence divergence of the two carbohydrate-binding modules facilitates cross-linking of ligands by two separate lectin sites, in contrast to homodimeric proto-type proteins with identical domains. Detected first in the nematode *Caenorhabditis elegans*, where five proteins from the set of 11 galectin candidates fall into this category (*35*, *36*), the occurrence of this structural display was thereafter found to be widely present in animals (*4*). Apparently, tandem-repeat-type galectins have developed independently after the phylogenetic separation of vertebrates from nematodes in the Precambrian era more than $6 \times 10^8$ years ago (*35*). Mammals show remarkable diversity in this respect with at least four different galectins of this group. In comparison, the existence of such proteins in chicken was an open question except for partial EST information. We herein report on the presence of one respective gene in chicken, which meets the general sequence criteria for a functional galectin, and its full-length cDNA sequence. The nature of the chromosomal environment in the vicinity of this gene is shared by mammals and marsupials, constituting strong evidence for homology. Systematic alignment of the full-length sequence reaches especially high similarity scores in comparison to proteins from mammals already classified as galectin-8. Gene expression in chicken leads to two proteins by alternative splicing. It pertains to the absence or presence of exon V encoding a 19-amino acid stretch in the linker region. Both proteins were produced in bacteria at high yields, shown to bind $\beta$-galactosides, and shown to act as hemagglutinin, although they are monomers in solution under the conditions of gel filtration and ultracentrifugation. The

Table 3: Immunohistochemical Profiling[a] of the Presence of CG-8 in Various Organs of Adult Chicken

| type of organ | positivity | type of organ | positivity |
|---|---|---|---|
| cerebrum | | gizzard | |
| arachnoid | ++[b,c] | pellicula | − |
| cerebellum | | *glld. ventriculares* | |
| arachnoid | +[b,c] | epithelium | − |
| thymus | | *l. propria mucosae* | +[b] |
| thymocytes, macrophages | +++[b] | stratum compactum | − |
| larynx | | gut[f] | |
| respiratory epithelium | − | epithelial lining of villi and intestinal glands | − |
| *l. propria mucosae* | +[b,d] | *l. propria mucosae* | +++[b] |
| trachea | | kidney | |
| respiratory epithelium | − | connective tissue surrounding glomeruli | ++[b] |
| *l. propria mucosae* | ++[b,d] | tubules and collecting ducts | |
| lung | | epithelium | − |
| parabronchal wall | ++[b] | connective tissue | ++[b] |
| atria | | ovary | |
| respiratory epithelium | − | prehierarchical follicles | |
| interatrial septa | +[b] | healthy | − |
| heart | | early atretic | |
| epicardium | +[b] | granulosa cell layer | − |
| myocardium | +[b] | theca cell layer | ++[b] |
| endocardium | − | late atretic | |
| liver | | granulosa cell layer | +[b] |
| hepatocytes (parenchyma) | − | intercalating cells | +++[b] |
| connective tissue surrounding blood vessels | ++[b] | theca cell layer | ++[b] |
| esophagus | | hierarchical follicles | − |
| epithelium | − | skin | |
| *l. propria mucosae* | +++[b,e] | epidermis | |
| proventriculus | | *stratum corneum* | − |
| *glld. proventriculares superficiales* | | *stratum intermedium* | − |
| epithelium | − | *stratum basalis* | − |
| *l. propria mucosae* | +[b] | dermis | ++[b] |
| *glld. proventriculares profundae* | | subcutis | +[b] |
| epithelium | − | | |
| *l. propria mucosae* | +[b] | | |

[a] Number of positive cells grouped into categories: −, no cells stained; +, single cells; ++, clusters of cells; +++, > 50% of the cell population (intensity rather strong and homogeneous). [b] Only cytoplasmic. [c] Flat layer of closely apposed cells covering the subjacent layer of loosely associated *arachnoid trabecular* cells. [d] Subepithelial. [e] Subepithelial surrounding glands localized in the *l. propria mucosae*. [f] Gut is the duodenum, jejunum, ileum, cecum, and rectum. Abbreviations: *glld.*, *glandulae*; *l.*, *lamina*.

availability of these two forms of CG-8 with variation in linker length enabled us to address the hitherto unresolved questions about the shape of a tandem-repeat-type galectin and the effect of linker elongation. Notably, we can even set these results in relation to the data sets on proto-type CGs, which we had analyzed previously (*15*, *16*).

The determination of the hydrodynamic behavior revealed a rather similar overall shape between CG-8 and CG-1A/B with a mass-proportional increase of length. Obviously, the addition of 19 amino acids to the linker peptide in CG-8II can somewhat increase axial dimensions. In a more general context, the connecting peptide of CG-8II resembles respective portions of mammalian and frog galectin-8 without an extension in length; thus, the reported data on how such a linker region contributes to the overall shape of this assembly of two CRDs in solution are likely to have relevance beyond the avian protein. It will be of interest to examine the respective behavior of mammalian galectin-8 variants with further increases in length (*19*, *20*). In a survey of chicken galectins, the three proto-type and the described tandem-repeat-type proteins equip the cells with a versatile set of lectin tools. When the hydrodynamic shape was characterized, this parameter was delineated to be subject to differential control: marked differences were brought about by redox reactions (CG-1B) (*16*) or revealed by sedimentation velocity (CG-2) (*15*). Alternative splicing acting on linker length is a further means of control (CG-8I and II). Whether these shape differences can modulate cellular functions is an issue that commands attention. Interestingly, the total loss of the hinge in rat galectin-8 was reported to impair the lectin's capacity to mediate CHO cell adhesion and ensuing signaling via extracellular signal-regulated kinase and phosphatidylinositol 3-kinase (*37*). Our respective data on cell binding reveal activity for both forms, as also seen for the isoforms of murine galectin-8 on thymocytes and the induction of apoptosis in the $CD4^{high}$ $CD8^{high}$ subpopulation (*38*). When two different cell types were tested, it turned out that quantitative aspects of binding can differ, in the tested case resulting in a relatively increased level of binding of CG-8I. This result signifies the potential of linker length variation to affect cell binding and gives a clear direction to further work with the labeled CGs as sensors in delineating reactivity patterns. That said, it was essential to document the actual presence of the two isoforms in vivo.

The analysis by Western blotting provided this evidence. It clearly revealed the occurrence of alternative splicing with organ-type differences. In summary, CG-8II was detected predominantly, and the blots ruled out production of any truncated forms. The availability of the other CGs facilitated the insurance of monospecificity of the antibody preparation in the course of

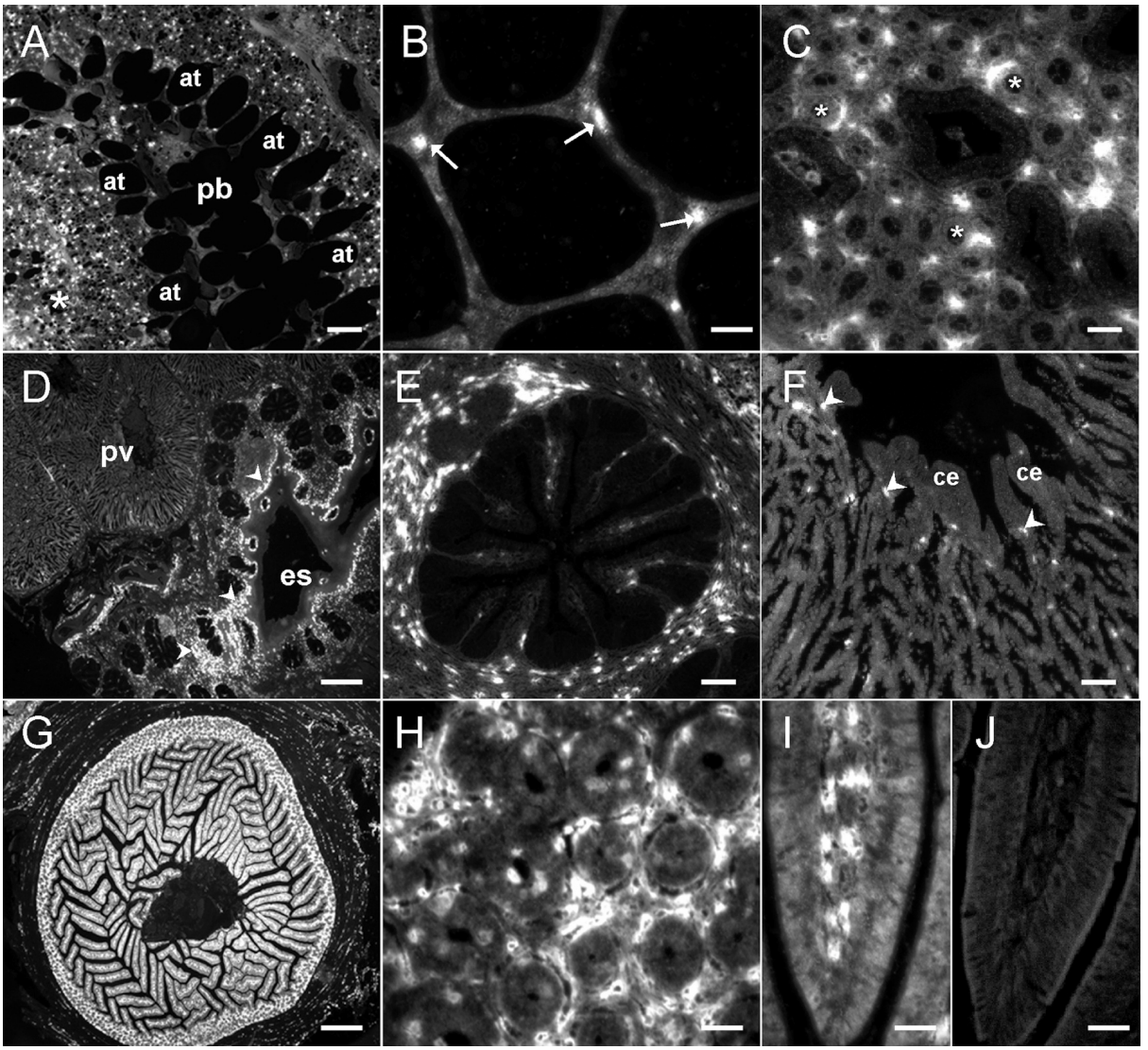


FIGURE 7: Localization of CG-8 in tissue sections by immunohistochemistry. Microphotographs of cross sections of the lumen and the wall of a parabronchus (pb) and connected atria (at) at two levels of magnification (A and B). The presence of strongly CG-8-positive cells was visualized between blood and air capillaries in the parabronchal wall (asterisk) (A); interatrial septa presented alveolar macrophages (arrows) (B). Connective tissue between collecting ducts (asterisk) in the kidney was strongly reactive with the CG-8-specific IgG preparation (C). A cross section through the transition zone between esophagus (es) and proventriculus (pv) is shown at two levels of magnification (D−F). In contrast to CG-8 localization in esophagus, where CG-8-positive cell clusters are abundant (arrowheads) (D, es), which surround glands in the *l. propria mucosae* (E), proventriculus presents with staining confined to single cells in the inconspicuous *l. propria mucosae* between tubules of deep glands (D, pv) and adjacent to the columnar epithelium (arrowheads) (F, ce). In jejunum (G−J), strong signals forming a regular staining pattern of the *l. propria mucosae* (G) originated from the region surrounding intestinal glands (H) and from the cores of villi (I). A control section after processing without the incubation step with the CG-8-specific antibodies ascertained a lack of antigen-independent staining (J). The scale bars are 20 $\mu$m (B and H−J), 50 $\mu$m (C, E, and F), 100 $\mu$m (A), and 500 $\mu$m (D and G).

Table 4: Unique Features of Proximal Promoter Regions of the CG-8 Gene and the Three Proto-Type CG Genes Defined by the Listed Transcription Factor[a,b]

| galectin | presence of target sequence | absence of target sequence |
|---|---|---|
| CG-8 | Ahr:Arnt; NF-$\kappa$B; Nkx2−5; RFX-1 | − |
| CG-1A | c-Myc:Max; FoxD3; NRF-2; Pax-5 (BSAP); STAT1$\alpha/\beta$, 2−6 | − |
| CG-1B | Barbie-Box; CP2; Elk-1 (TCF-A); Pax-6; v-Maf; YY1 | GATA-3; NF-E2 |
| CG-2 | AP-2; cap; HNF-1$\alpha$; MyoD; p300; p53 | CUTL-1; POU2F1/2 (Oct-1) |

[a] In each case, the region from 2000 bp upstream to 150 bp downstream relative to the gene's transcription (CG-1B and CG-2) or translation (CG-1A and CG-8) start defines the proximal promoter region. [b] Putative binding motifs are listed when they are uniquely either present or absent in comparison to the proximal promoter regions of the other three galectin genes. For an explanation of factor names, please see the table in the Supporting Information; additional factors are AP-2 (activator protein 2), Barbie-Box (barbiturate-inducible element), c-Myc:Max [c-Myc (cellular counterpart of viral myelocytomatosis oncogene):Max heterodimer], cap (capping signal), CP2 (CCAAT-binding protein 2), Elk-1 (TCF-A) [*ets*-like factor 1 (transcription factor A)], FoxD3 (fork head box D3), MyoD (myoblast determination gene product), NRF-2 (nuclear respiratory factor 2), p300 (E1A-associated 300 kDa protein), p53 (53 kDa tumor suppressor protein), BSAP (B-cell-specific activator protein), STAT (signal transducer and activator of transcription), v-Maf (avian musculoaponeurotic fibrosarcoma virus AS42 nuclear oncoprotein), and YY1 [Yin and Yang 1 (nuclear factor E1, delta factor)].

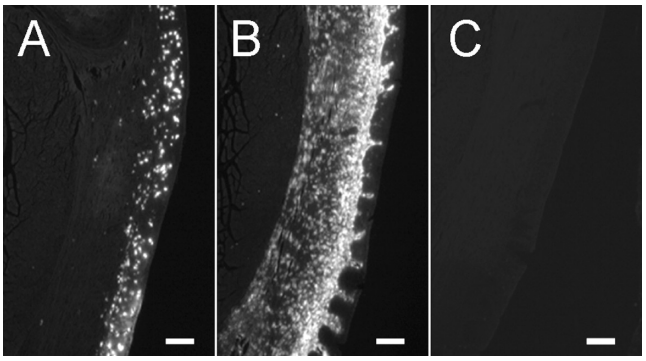


FIGURE 8: Localization of CG-8 and CG-1B and -2 in serial sections of larynx by immunohistochemistry. Microphotographs of serial cross sections from adult larynx after processing with antibody preparations specific for CG-8 (A), CG-1B (B), and CG-2 (C). The presence of CG-8-positive cells is confined to a region adjacent to pseudostratified epithelium (A), whereas CG-1B positivity is seen between and below tubuloacinar mucus-secreting glands, in the extracellular matrix, and in the embedded fibroblasts (B). No signal for CG-2 was detected after identical processing, thus excluding any antigen-independent staining (C). The bar is 100 $\mu$m.

answering the question about the localization of CG-8 by immunohistochemistry. In comparison, human galectin-8 has a rather broad expression profile in terms of organs and cells. Various types of cells, e.g., in lung basal cells, ciliated bronchial cells, chondrocytes, serous cells of the bronchial gland, smooth muscle cells, and endothelial cells, were positive; malignant transformation led to significant changes with tumor-type-dependent up- or downregulation, and immune/parenchymal cells in the vicinity of tumors were apparently stimulated to express galectin-8 (*39*−*44*). The profile of CG-8 localization is less variegated. Moreover, it has rather distinct characteristics in chicken, setting it apart especially from CG-1A and CG-2. The prominent aspect of immunohistochemical CG-8 detection is positivity of immune cells in the *l. propria mucosae*, a salient cellular line of defense, and of macrophages and thymocytes in the thymus. At first sight, a colocalization could be suspected with CG-1B, because its presence in the *l. propria mucosae* had been reported previously (*15*). However, CG-1B resides in the extracellular matrix and fibroblasts, clearly separated from the zone of CG-8-positive immune cells close to the epithelial lining (Figure 8). CG-8 was also, unique for CGs, detected during progression of follicular atresia in the ovary. Here, its pattern of presence and the upregulation were evocative of the behavior of galectin-3 during atresia/luteolysis in murine ovary (*33*). As a first step in delineating the signals underlying CG-8 expression, the initial mapping of the promoter region documents a complex pattern of putative regulatory sequences and deviations from the respective region in genes of proto-type CGs. The unique presence of sites for cardiogenic homeodomain factor Nkx-2.5 initimates the study of CG-8 in the context of early cardiac gene expression (*45*). In view of the disclosed expression profile and the accumulating evidence of a role of mammalian galectin-8 in modulation of neutrophil adhesion and superoxide production via the Mac-1 antigen (CD11b, αM-integrin), in partaking in inducing apoptosis in thymocyte subpopulations and in inflammatory cells of the synovial fluid in rheumatoid arthritis, and in inducing autoantibodies in autoimmune diseases (*38*, *46*−*48*), monitoring CG-8 expression in inflammation stands out as a further attractive study topic inspired by these results. When considering the network aspect, this work will have an additional dimension, that is, to start closing the remaining gap in knowledge about CG localization by defining the respective profile of CG-3, a multifunctional immune regulator in mammals (*33*), in parallel. This chimera-type CG is known to be produced by chicken embryonic fibroblasts in vitro, by hypertrophic chondrocytes, and by osteoclasts (*49*−*51*). Our stepwise program will thus focus next on CG-3 and the relation of its features to those of the other four CGs.

## ACKNOWLEDGMENT

We are grateful to Drs. S. Friday and W. Notelecs for inspiring discussions, to the reviewers for their valuable advice, and to A. Helfrich, B. Hofer, L. López, and L. Mantel for skillful technical assistance.

## SUPPORTING INFORMATION AVAILABLE

One table summarizing the putative transcription factor-binding sequence motifs in the proximal promoter region of the CG-8 gene and one figure illustrating CG-8 localization by immunohistochemistry in cerebrum and cerebellum. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Gabius, H.-J., Ed. (2009) *The Sugar Code. Fundamentals of Glycosciences*, Wiley-VCH, Weinheim, Germany.
2. Gabius, H.-J. (2008) Glycans: Bioactive signals decoded by lectins. *Biochem. Soc. Trans. 36*, 1491–1496.
3. Hirabayashi, J., and Kasai, K.-i. (1993) The family of metazoan metal-independent $\beta$-galactoside-binding lectins: Structure, function and molecular evolution. *Glycobiology 3*, 297–304.
4. Cooper, D. N. W. (2002) Galectinomics: Finding themes in complexity. *Biochim. Biophys. Acta 1572*, 209–231.
5. Gabius, H.-J., Siebert, H.-C., André, S., Jiménez-Barbero, J., and Rüdiger, H. (2004) Chemical biology of the sugar code. *ChemBioChem 5*, 740–764.
6. Houzelstein, D., Gonçalves, I. R., Fadden, A. J., Sidhu, S. S., Cooper, D. N. W., Drickamer, K., Leffler, H., and Poirier, F. (2004) Phylogenetic analysis of the vertebrate galectin family. *Mol. Biol. Evol. 21*, 1177–1187.
7. Den, H., and Malinzak, D. A. (1977) Isolation and properties of $\beta$-D-galactoside-specific lectin from chick embryo thigh muscle. *J. Biol. Chem. 252*, 5444–5448.
8. Nowak, T. P., Kobiler, D., Roel, L. E., and Barondes, S. H. (1977) Developmentally regulated lectin from embryonic chick pectoral muscle. Purification by affinity chromatography. *J. Biol. Chem. 252*, 6026–6030.
9. Beyer, E. C., Zweig, S. E., and Barondes, S. H. (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.
10. Oda, Y., and Kasai, K.-i. (1983) Purification and characterization of $\beta$-galactoside-binding lectin from chick embryonic skin. *Biochim. Biophys. Acta 761*, 237–245.
11. Solís, D., Romero, A., Kaltner, H., Gabius, H.-J., and 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.
12. Varela, P. F., Solís, D., Díaz-Mauriño, T., Kaltner, H., Gabius, H.-J., and Romero, A. (1999) The 2.15 Å crystal structure of CG-16, the developmentally regulated homodimeric chicken galectin. *J. Mol. Biol. 294*, 537–549.
13. Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Müller, W. E. G., Yagi, F., and Kasai, K.-i. (2002) Oligosaccharide specificity of galectins: A search by frontal affinity chromatography. *Biochim. Biophys. Acta 1572*, 232–254.
14. Wu, A. M., Singh, T., Liu, J.-H., Krzeminski, M., Russwurm, R., Siebert, H.-C., Bonvin, A. M. J. J., André, S., and 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.

15. Kaltner, H., Solís, D., Kopitz, J., Lensch, M., Lohr, M., Manning J. C., Mürnseer, M., Schnölzer, M., André, S., Sáiz, J. L., and Gabius, H.-J. (2008) Proto-type 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.
16. López-Lucendo, M. F., Solís, D., Sáiz, J. L., Kaltner, H., Russwurm, R., André, S., Gabius, H.-J., and Romero, A. (2009) Homodimeric chicken galectin CG-1B(C-14): Crystal structure and detection of unique redox-dependent shape changes involving intersubunit and intrasubunit disulfide bridges by gel filtration, ultracentrifugation, site-directed mutagenesis, and peptide mass fingerprinting. *J. Mol. Biol. 386*, 366–378.
17. Shoji, H., Nishi, N., Hirashima, M., and Nakamura, T. (2003) Characterization of the *Xenopus* galectin family. *J. Biol. Chem. 278*, 12285–12293.
18. Hadari, Y. R., Paz, K., Dekel, R., Mestrovic, T., Accili, D., and Zick, Y. (1995) Galectin-8. A new rat lectin, related to galectin-4. *J. Biol. Chem. 270*, 3447–3453.
19. Gopalkrishnan, R. V., Roberts, T., Tuli, S., Kang, D., Christiansen, K. A., and Fisher, P. B. (2000) Molecular characterization of prostate carcinoma tumor antigen-1, PCTA-1, a human galectin-8 related gene. *Oncogene 19*, 4405–4416.
20. Bidon, N., Brichory, F., Hanash, S., Bourguet, P., Dazord, L., and Le Pennec, J.-P. (2001) Two messenger RNAs and five isoforms for Po66-CBP, a galectin-8 homolog in a human lung carcinoma cell line. *Gene 274*, 253–262.
21. Lu, H., Knutson, K. L., Gad, E., and Disis, M. L. (2006) The tumor antigen repertoire identified in tumor-bearing *neu* transgenic mice predicts human tumor antigens. *Cancer Res. 66*, 9754–9761.
22. Gabius, H.-J. (1990) Influence of type of linkage and spacer on the interaction of $\beta$-galactoside-binding proteins with immobilized affinity ligands. *Anal. Biochem. 189*, 91–94.
23. Gabius, H.-J., Wosgien, B., Hendrys, M., and Bardosi, A. (1991) Lectin localization in human nerve by biochemically defined lectin-binding glycoproteins, neoglycoprotein and lectin-specific antibody. *Histochemistry 95*, 269–277.
24. Jiménez, M., Sáiz, J. L., André, S., Gabius, H.-J., and Solís, D. (2005) Monomer/dimer equilibrium of the AB-type lectin from mistletoe enables combination of toxin/agglutinin activities in one protein: Analysis of native and citraconylated proteins by ultracentrifugation/gel filtration and cell biological consequences of dimer destabilization. *Glycobiology 15*, 1386–1395.
25. Gabius, H.-J., Engelhardt, R., Cramer, F., Bätge, R., and Nagel, G. A. (1985) Pattern of endogenous lectins in a human epithelial tumor. *Cancer Res. 45*, 253–257.
26. Mann, K., Weiss, I. M., André, S., Gabius, H.-J., and Fritz, M. (2000) The amino-acid sequence of the abalone (*Haliotis laevigata*) nacre protein perlucin. Detection of a functional C-type lectin domain with galactose/mannose specificity. *Eur. J. Biochem. 267*, 5257–5264.
27. André, S., Unverzagt, C., Kojima, S., Frank, M., Seifert, J., Fink, C., Kayser, K., von der Lieth, C.-W., and Gabius, H.-J. (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.
28. Patnaik, S. K., and Stanley, P. (2006) Lectin-resistant CHO glycosylation mutants. *Methods Enzymol. 416*, 159–182.
29. André, S., Sanchez-Ruderisch, H., Nakagawa, H., Buchholz, M., Kopitz, J., Forberich, P., Kemmner, W., Böck, C., Deguchi, K., Detjen, K. M., Wiedenmann, B., von Knebel-Döberitz, M., Gress, T. M., Nishimura, S.-I., Rosewicz, S., and Gabius, H.-J. (2007) Tumor suppressor p16INK4a: 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.
30. André, S., Pei, Z., Siebert, H.-C., Ramström, O., and Gabius, H.-J. (2006) 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.
31. Saal, I., Nagy, N., Lensch, M., Lohr, M., Manning, J. C., Decaestecker, C., André, S., Kiss, R., Salmon, I., and 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.
32. Lensch, M., Lohr, M., Russwurm, R., Vidal, M., Kaltner, H., André, S., and 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.
33. Lohr, M., Kaltner, H., Lensch, M., André, S., Sinowatz, F., and Gabius, H.-J. (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.
34. Lohr, M., Lensch, M., André, S., Kaltner, H., Siebert, H.-C., Smetana, K.Jr., Sinowatz, F., and Gabius, H.-J. (2007) 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. *Folia Biol. (Prague, Czech Repub.) 53*, 109–128.
35. Hirabayashi, J., Satoh, M., and Kasai, K.-i. (1992) Evidence that *Caenorhabditis elegans* 32-kDa $\beta$-galactoside-binding protein is homologous to vertebrate $\beta$-galactoside-binding lectins. cDNA cloning and deduced amino acid sequence. *J. Biol. Chem. 267*, 15485–15490.
36. Nemoto-Sasaki, Y., Hayama, K., Ohya, H., Arata, Y., Kaneko, M. K., Saitou, N., Hirabayashi, J., and Kasai, K.-i. (2008) *Caenorhabditis elegans* galectins LEC-1-LEC-11: Structural features and sugar-binding properties. *Biochim. Biophys. Acta 1780*, 1131–1142.
37. Levy, Y., Auslender, S., Eisenstein, M., Vidavski, R. R., Ronen, D., Bershadsky, A. D., and Zick, Y. (2006) It depends on the hinge: A structure-functional analysis of galectin-8, a tandem-repeat type lectin. *Glycobiology 16*, 463–476.
38. Tribulatti, M. V., Mucci, J., Cattaneo, V., Agüero, F., Gilmartin, T., Head, S. R., and Campetella, O. (2007) Galectin-8 induces apoptosis in the $CD4^{high}CD8^{high}$ thymocyte subpopulation. *Glycobiology 17*, 1404–1412.
39. Caulet-Maugendre, S., Birolleau, S., Corbineau, H., Bassen, R., Desrues, B., Bidon, N., Delaval, P., Ramée, M.-P., Brichory, F., and Dazord, L. (2001) Immunohistochemical expression of the intracellular component of galectin-8 in squamous cell metaplasia of the bronchial epithelium in neoplastic and benign processes. *Pathol. Res. Pract. 197*, 797–801.
40. Danguy, A., Rorive, S., Decaestecker, C., Bronckart, Y., Kaltner, H., Hadari, Y. R., Goren, R., Zick, Y., Petein, M., Salmon, I., Gabius, H.-J., and Kiss, R. (2001) Immunohistochemical profile of galectin-8 expression in benign and malignant tumors of epithelial, mesenchymatous and adipous origins, and of the nervous system. *Histol. Histopathol. 16*, 861–868.
41. Henno, S., Brichory, F., Langanay, T., Desrues, B., Bidon, N., Delaval, P., Ramée, M.-P., Dazord, L., and Caulet-Maugendre, S. (2002) Expression of Po66-CBP, a galectin-8, in different types of primary and secondary broncho-pulmonary tumors. *Oncol. Rep. 9*, 177–180.
42. Nagy, N., Bronckart, Y., Camby, I., Legendre, H., Lahm, H., Kaltner, H., Hadari, Y., Van Ham, P., Yeaton, P., Pector, J.-C., Zick, Y., Salmon, I., Danguy, A., Kiss, R., and Gabius, H.-J. (2002) Galectin-8 expression decreases in cancer compared with normal and dysplastic human colon tissue and acts significantly on human colon cancer cell migration as a suppressor. *Gut 50* 392–401.
43. Bidon-Wagner, N., and Le Pennec, J.-P. (2004) Human galectin-8 isoforms and cancer. *Glycoconjugate J. 19*, 557–563.
44. Langbein, S., Brade, J., Badawi, J. K., Hatzinger, M., Kaltner, H., Lensch, M., Specht, K., André, S., Brinck, U., Alken, P., and Gabius, H.-J. (2007) Gene-expression signature of adhesion/growth-regulatory tissue lectins (galectins) in transitional cell cancer and its prognostic relevance. *Histopathology 51*, 681–690.
45. Sepulveda, J. L., Belaguli, N., Nigam, V., Chen, C.-Y., Nemer, M., and Schwartz, R. J. (1998) GATA-4 and Nkx-2.5 coactivate Nkx-2 DNA binding targets: Role for regulating early cardiac gene expression. *Mol. Cell. Biol. 18*, 3405–3415.
46. Nishi, N., Shoji, H., Seki, M., Itoh, A., Miyanaka, H., Yuube, K., Hirashima, M., and Nakamura, T. (2003) Galectin-8 modulates neutrophil function via interaction with integrin $\alpha$M. *Glycobiology 13*, 755–763.
47. Sebban, L. E., Ronen, D., Levartovsky, D., Elkayam, O., Caspi, D., Aamar, S., Amital, H., Rubinow, A., Golan, I., Naor, D., Zick, Y., and Golan, I. (2007) The involvement of CD44 and its novel ligand galectin-8 in apoptotic regulation of autoimmune inflammation. *J. Immunol. 179*, 1225–1235.

48. Massardo, L., Metz, C., Pardo, E., Mezzano, V., Babul, M., Jarpa, E., Guzmán, A. M., André, S., Kaltner, H., Gabius, H.-J., Jacobelli, S., González, A., and Soza, A. (2009) Autoantibodies against galectin-8: Their specificity, association with lymphopenia in systemic lupus erythematosus and detection in rheumatoid arthritis and acute inflammation. *Lupus 18*, 539–546.
49. Crittenden, S. L., Roff, C. F., and Wang, J. L. (1984) Carbohydrate-binding protein 35: Identification of the galactose-specific lectin in various tissues of mice. *Mol. Cell. Biol. 4*, 1252–1259.
50. Nurminskaya, M., and Linsenmayer, T. F. (1996) Identification and characterization of up-regulated genes during chondrocyte hypertrophy. *Dev. Dyn. 206*, 260–271.
51. Gorski, J. P., Liu, F.-T., Artigues, A., Castagna, L. F., and Osdoby, P. (2002) New alternatively spliced form of galectin-3, a member of the $\beta$-galactoside-binding animal lectin family, contains a predicted transmembrane-spanning domain and a leucine zipper motif. *J. Biol. Chem. 277*, 18840–18848.