Identification of an Internal Gene to the Human Galectin-3 Gene with Two Different Overlapping Reading Frames That Do Not Encode Galectin-3*

Michael Guittaut‡, Stéphane Charpentier, Thierry Normand, Martine Dubois, Jacques Raimond, and Alain Legrand§

From the Centre de Biophysique Moléculaire (affiliated with the University of Orleans), CNRS UPR4301, Rue Charles Sadron, 45071 Orleans Cedex 02, France

We previously reported that alternative transcripts were initiated within the second intron of the human Galectin-3 gene (LGALS3). We now demonstrate that these transcripts arise from an internal gene embedded within LGALS3 and named galig (Galectin-3 internal gene). Tissue-specific expression of galig was assayed by screening of several human tissues. Contrary to LGALS3, galig appears to be tightly regulated and principally activated in leukocytes from peripheral blood. Cloning and characterization of galig transcripts revealed that they contain two out-of-frame overlapping open-reading frames (ORFs). Transfection of expression vectors encoding enhanced green fluorescent protein (EGFP) chimeras indicated that both ORFs could be translated in proteins unrelated to Galectin-3. The ORF1 polypeptide targets EGFP to cytosol and nucleus whereas ORF2 targets EGFP to mitochondria. These results revealed the exceptional genetic organization of the LGALS3 locus.

Galectin-3 is a β-galactoside-binding lectin involved in a variety of biological processes (1). The expression pattern of Galectin-3 changes during development (2) and production of the protein modulates cell-cell or cell-matrix interactions (3–4). A large number of studies show that expression is dependent on cellular growth properties, correlates with neoplastic transformation (5), and confers resistance to apoptosis (6).

Regulation of Galectin-3 gene (LGALS3) expression is largely unknown. Recently, the complete genomic sequence of human LGALS3 has been reported, and the functional promoter activity of the 5′-flanking region of the gene has been established (7). Previously, we have reported that the second intron of LGALS3 contains an internal promoter, which drives production of alternative transcripts (8).

In this report, we have further investigated the structure of these alternative transcripts. We demonstrate that they are preferentially expressed in peripheral blood leukocytes. They cannot be used for production of Galectin-3 or a modified Galectin-3 because they contain two overlapping open-reading frames (ORF) out-of-frame with the lectin coding sequence. Upon transfection, translation of these two overlapping ORFs target reporter proteins to distinct subcellular compartments. These results indicate that the LGALS3 locus exhibits an unusual genetic structure, which involves the presence of an internal gene embedded within the Galectin-3 gene.

EXPERIMENTAL PROCEDURES

Primer Extension Analysis—RNA was extracted from human osteosarcoma HOS cells (ATCC, CRL-1543) using the guaninium/phenol/chloroform procedure (9). One microgram of RNA was incubated in a final volume of 20 μl containing 30 units of avian myeloblastosis virus reverse transcriptase (Promega, Madison, WI) and 10 picomoles of 32P-end-labeled primer using T4 polynucleotide kinase (New England BioLabs, Beverly, MA) at 42 °C for 2 h. The primer PE (5′-ACACTCA-CAACATCAGAAT-3′) extends from nucleotides 334 to 315 in the second intron of LGALS3 (see Fig. 1). The reaction products were heat-denatured and loaded onto a 6% polyacrylamide sequencing gel.

Detection of Alternative Transcripts by RT-PCR—Two micrograms of total cellular RNA were extracted from SVH-1, a human smooth muscle cell line (10), and from two tumor samples from human colon carcinoma and were reverse transcribed using 10 pmol of random nanomer primers. Thirty PCR cycles were performed using a 3′-primer EX6-R (5′-TTCGCCCTTTTCAGATTATATA-3′), located in the sixth exon at the end of Galectin-3 cDNA (positions 801–780) on the human cDNA (GenBankTM/EBI no. M36682), and 12-F (5′-TCTTGATGTGTTGGATTGTTTTTC-3′), a 5′-primer located from nucleotides 316 to 339 in the second intron of LGALS3 (GenBankTM/EBI no. U10300). A second round of 30 cycles of PCR was initiated using the same 5′-primer and an internal 3′-primer, EX4-R (5′-TGTGGTTGACATTGGGCCTACC-3′) located 336 bp upstream of the primer EX6-R in the fourth exon of the gene. All PCR was carried out under standard conditions with an annealing step of 55 °C.

Tissue-specificity of the alternative transcripts was also performed by reverse transcriptase-PCR on various human tissues using Human Rapid-Scan Gene Expression Panel (Origene Technologies, Inc., Rockville, MD) containing 0.25 ng or 2.5 ng cDNA. The primers used were TSF-1 (5′-TCTGAGTACGGAGGAGTA-3′), corresponding to nucleotides 282–299 in the second intron, and TSR-1 (5′-GGGAAAACCGACTGTCTT-3′), corresponding to nucleotides 588–605 in the fifth exon on the human cDNA (GenBankTM/EBI no. M36682). PCR (35 cycles) was carried out under standard conditions with an annealing step of 57 °C. Galectin-3 transcripts were also detected using the same method with primers EX4-R and GAL25 (5′-ATGGCAGACAATTTTCCAGCTC-3′), located at nucleotides 34–55 on the second exon of the gene. Expression of β-actin was examined as a control with primers supplied by the manufacturer.

Constructions of Reporter Plasmids Encoding EGFP Chimeras—Three ORFs, designated as ORF1, ORF2, and ORF3, were detected in the alternative transcripts. Expression vectors were constructed by

* This work was supported by a grant from the Ligue Nationale Contre le Cancer. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) AF266280.

† Supported by fellowships from the Ministère de l’Enseignement Supérieur et de la Recherche and the Fondation pour la Recherche Médicale.

‡ To whom correspondence should be addressed. Tel.: 33 2 38 25 55 36; Fax: 33 2 38 25 78 07; E-mail: legrand@cnrs-orleans.fr.

1 The abbreviations used are: ORF, open-reading frame; galig, Galectin-3 internal gene; RT-PCR, reverse transcriptase-polymerase chain reaction; bp, base pair(s); GFP, green fluorescent protein; EGFP, enhanced GFP; CMV, cytomegalovirus; SV40, simian virus 40.
insertion of the cDNA encoding EGFP at the 3’-end of each ORF (see Fig. 3A).

EGFP cDNA was isolated from pEGFP-N1 (Invitrogen, Groningen, The Netherlands) and modified to remove the ATG codon to avoid initiation of potential translation because of potential leaky scanning process. In these plasmids, designated as pORF1-EGFP, pORF2-EGFP, and pORF3-EGFP, the 5’-end of the chimeric cDNA is located at position 242 in the second intron of LGALS3. Transcription is driven by the CMV promoter. Similar vectors were constructed in which EGFP cDNA was replaced by the luciferase gene in pGL3basic (Promega).

**Cell Transfection—**Plasmids were transfected into HO5 cells (seeded on glass coverslips at 8 × 10⁴ cells per 12-well plate) or HEK 293 cells (ATCC, CRL-1573, seeded at 2 × 10⁵ cells per well) using DNA-polyethyleneimine complexes (11). Briefly, 5 μg of DNA were mixed in 7.5 μl of a 10 mM polyethylenimine solution (Sigma Aldrich, France) in a total volume of 1 ml of serum-free medium (Dulbecco’s minimum Eagle’s medium, Life Technologies, Inc.). Cells were incubated for 2 h with this solution and maintained for 24 h in fresh Eagle’s minimum essential medium (Life Technologies, Inc.) (HO5 cells) or Dulbecco’s minimum Eagle’s medium (HEK 293 cells) supplemented with 10% fetal calf serum (Life Technologies, Inc.). For EGFP detection, HO5 cells were analyzed by fluorescence microscopy using standard fluorescein filters (excitation at 488 nm and detection at 520 nm). For mitochondriostaining, cells were incubated for 15 min with MitoTracker Red (50 nm, Molecular Probes, Inc, Eugene, OR). MitoTracker Red was excited at 568 nm and detected through a > 600 nm filter. For luciferase assays, 400 ng of pRL-SV40, which encodes Renilla luciferase (Promega), were added and used as an internal standard of transfection efficiency and reproducibility. Luciferase activities were assayed using Dual-Luciferase Reporter Assay System (Promega) and an automated luminometer (Lumat LB9501, EG&G Berthold, Bad Wildbad, Germany).

**Preparation of Cellular Lysates—**HEK 293 cells were trypsinized 24 h after transfection, centrifuged for 5 min at 250 × g and washed twice with phosphate-buffered saline. The pellet was suspended in 1.8 ml of lysis buffer (137 mM NaCl, 30 mM KCl, 15 mM Tris, 2 mM EDTA, 1 mM CaCl₂, 0.5% Triton X-100, 1 mM phenylmethylsulfonide fluoride) containing 200 μl of protease inhibitor mixture (Sigma Aldrich, France) and incubated 30 min on ice. The lysates were centrifuged at 15,000 × g for 5 min at 4 °C. Pelleted nuclei and membranes were suspended in 1 ml of protease inhibitor mixture (Sigma Aldrich, France) and maintained for 24 h in fresh Eagle’s minimum essential medium (Life Technologies, Inc.) (HO5 cells) or Dulbecco’s minimum Eagle’s medium (HEK 293 cells) supplemented with 10% fetal calf serum (Life Technologies, Inc.). For EGFP detection, HO5 cells were analyzed by fluorescence microscopy using standard fluorescein filters (excitation at 488 nm and detection at 520 nm). For mitochondrial staining, cells were incubated for 15 min with MitoTracker Red (50 nm, Molecular Probes, Inc, Eugene, OR). MitoTracker Red was excited at 568 nm and detected through a > 600 nm filter. For luciferase assays, 400 ng of pRL-SV40, which encodes Renilla luciferase (Promega), were added and used as an internal standard of transfection efficiency and reproducibility. Luciferase activities were assayed using Dual-Luciferase Reporter Assay System (Promega) and an automated luminometer (Lumat LB9501, EG&G Berthold, Bad Wildbad, Germany).

**RESULTS**

**Multiple Initiation Sites of Transcription in Alternative Transcripts of LGALS3—**Transcription initiation sites of the alternative transcripts resulting from the activity of the internal promoter of LGALS3 were identified by primer extension analysis (Fig. 1). Reverse transcription was initiated from a primer located within the second intron of the gene to avoid contamination signals resulting from extension of Galectin-3 mRNA. This was particularly beneficial considering the low amount of these transcripts when compared with those produced by the proximal promoter (8). Several signals of low intensities were detected indicating the presence of different initiation sites (Fig. 1B). Among them, five major sites were evident. They were all located in the second intron. Two of them were at positions 284 and 287 from the 5’-end of the second intron sequence and three were at positions 262, 263, and 266 (Fig. 3B).

**Characterization of Alternative Transcripts—**Transcripts originating from the internal promoter were detected by RT-PCR using forward (12-F) and reverse (EX6-R) primers located in the second intron and the sixth exon of LGALS3, respectively (Fig. 1A). One human smooth muscle cell line (10) and two human colon carcinomas tumor samples were tested. Positive signals, identified by a 752-bp fragment were detected after two rounds of PCR amplification confirming the low abundance of these transcripts (Fig. 2). In one tumor sample, a second band of lower size (around 450 bp) was detected. Both bands were cloned and sequenced. Sequencing indicated that the larger transcript was composed of the second intron, from the initiation transcription sites, and the third to sixth exons of LGALS3 (Fig. 3). The smaller band presented an internal deletion between nucleotides 95 and 389 of the large transcript. These sites correspond to consensus donor and acceptor sites for splicing events, indicating the presence of internal splicing within these transcripts.

To avoid potential problems inherent to the RT-PCR technique and to the low amount of transcripts detected in the cells analyzed, we have verified that the alternative transcripts can be cloned from a cDNA library constructed without use of PCR amplification. Because the possible cell specificity of the internal promoter was not known, we had screened at first the dbEST database using the sequence of the second intron. Two clones from two different cDNA libraries were revealed to be positive and contained sequences identical to the second intron. The first clone (GenBankTM/Eli, AA094057) was isolated from a human fetal heart library. The second one (GenBankTM/Eli, R23407 and R44192) was isolated from a human infant brain library and was completely resequenced. This clone is another splicing variant with a donor site located at position 49 instead of 95 in the larger transcript, the acceptor site being located at position 389. This acceptor site, used in both variants, are also used in the mature Galectin-3 mRNA and corresponds to the junction between the second intron and third exon (Fig. 3). These results confirmed that the structure of the alternative transcripts is not resulting from cloning artifacts because of PCR.

![Fig. 1. Determination of transcription initiation sites in mRNA issued from the internal promoter in LGALS3. A, schematic representation of the human LGALS3 gene. The proximal promoter is located upstream of exon 1 (ex.1, Ref. 7), and the internal promoter is located in the 5’-region of the second intron (int.2, Ref. 8). The ATG translation initiation codon used for Galectin-3 production is located in the second exon (ex.2). Introns, represented by thick lines, are not drawn to scale. PE is a primer used for primer extension analysis and is located at positions 334–315 in the second intron of LGALS3 (GenBankTM/Eli, U10300). 12-F, EX6-R, and EX4-R are primers used for the detection of transcripts by RT-PCR. 12-F is a forward primer located in the second intron and EX6-R, a reverse primer located in the sixth exon. EX4-R is a lower primer used for internal PCR (see “Experimental Procedures” for precise location). B, Primer extension analysis was carried out on mRNA from human HOS cells using primer PE. The ladder was determined using a sequencing reaction performed on a control plasmid (not shown). Numbers represent the size of the extended products in base pairs.](http://www.jbc.org/content/2653/2/2653.full.html)
Identification of Putative Overlapping ORFs—The alternative transcripts contain three potential ORFs (Fig. 3). Two of these ORFs, designated as ORF1 and ORF2, have potential translation initiation sites located in the third exon and spread over 318 and 291 nucleotides, respectively. ORF1 and ORF2 are out-of-frame with the Galectin-3 coding sequence. The translation initiation site of the third ORF, assigned as ORF3, is located in the fourth exon. If used, this last ORF should produce a truncated form of Galectin-3 corresponding to the C-terminal half of Galectin-3, which contains the complete carbohydrate binding domain of the lectin (12–13).

Translation of Out-of-Frame Overlapping ORFs—Reporter vectors were constructed to determine potentially translated ORFs. EGFP cDNA was inserted in-frame at the 3'–end of each ORF (Fig. 3A). HOS cells were transfected and analyzed by fluorescence microscopy. No fluorescence could be detected in cells transfected with pORF3-EGFP in which the EGFP is in-frame with the putative truncated Galectin-3, thus demonstrating that ORF3 is not translated (Fig. 4D). However, both other vectors (pORF1-EGFP and pORF2-EGFP), produced a clear fluorescence signal in transfected cells indicating translation of these ORFs. Distribution of fluorescence is strikingly different for these two vectors. The plasmid pORF1-EGFP, with an ATG located at position 394, induced a diffuse signal localized in cytosol and nucleus (Fig. 4B). This localization is similar to that of the nonchimeric EGFP (Fig. 4A). The plasmid pORF2-EGFP, with an ATG codon at position 434, produced a strong fluorescence associated with filamentous organelles and...
Characterization of an Internal Gene to Human LGALS3

**Fig. 4.** Transfection of cells with vectors encoding EGFP fusion proteins. HOS cells were transfected with pORF1-Luc and pORF2-Luc, two plasmids in which ORF1 and ORF2 are fused to luciferase cDNA, respectively. Cells were cotransfected with pRL-SV40, which encodes *Renilla* luciferase. This last plasmid is used as an internal standard. Luciferase activity (Luc) was assayed 2 days after transfection and was expressed relative to *Renilla* luciferase activity (RLuc).

**Fig. 5.** Expression of luciferase-based fusion proteins in transfected cells. HOS cells were transfected with pEGFPN1, a vector encoding EGFP (A), or fusion vectors encoding ORF1-EGFP (B), ORF2-EGFP (C), or ORF3-EGFP (D). Fluorescence microscopy was performed 2-days post-transfection.

**Fig. 6.** Immunodetection of chimeric proteins in pORF1-EGFP and pORF2-EGFP transfected cells. Twenty-four hours post-transfection, cells were fractionated. Membranes and nuclei-enriched fractions (lane M) and cytosolic fractions (lane C) were run on a polyacrylamide gel and transferred and incubated with a murine monoclonal anti-EGFP antibody.

**Fig. 7.** Mitochondrial localization of ORF2-EGFP. Left, confocal fluorescence image of HOS cells transfected with a vector encoding ORF2-EGFP. Right, same cells were costained with MitoTracker Red. Mitochondria in non-EGFP-transfected cells are also stained with MitoTracker Red.

is completely excluded from the nucleus (Fig. 4C).

Plasmids pORF1-EGFP and pORF2-EGFP were modified by replacing EGFP cDNA with the luciferase gene. Transient transfection of HOS cells with these vectors produced similar luciferase activity indicating that both ORFs are translated with the same efficiency (Fig. 5).

Chimeric proteins from HEK 293 transfected cells were analyzed, after cell fractionation, by Western blotting using an anti-EGFP antibody (Fig. 6). The cytosolic fraction of pORF1-EGFP transfected cells showed a strong signal matching with a 38-kDa chimeric protein, whereas nuclei- and membrane-containing fractions were weakly positive. This confirms the major cytosolic distribution of this fusion protein. Fractions of pORF2-EGFP transfected cells enriched in membranes and nuclei exhibited a signal corresponding to a 38-kDa protein, the expected size if translation of ORF2 occurred. Cytosolic fractions were negative confirming fluorescence microscopy analysis.

**Mitochondrial Localization of Fusion Protein ORF2-EGFP—**

Whereas ORF1-EGFP fusion protein has a cytosolic distribution, the protein produced by pORF2-EGFP was associated with organelles resembling typical mitochondria of rod-like, thread-like, and granular structures. Therefore, cells transfected with the pORF2-EGFP plasmid were incubated with a mitochondrial-specific dye marker (MitoTracker Red). Confocal microscopy demonstrates that MitoTracker Red and EGFP were colocalized in transfected cells confirming that ORF2-EGFP was associated with mitochondria (Fig. 7). This localization was confirmed after subcellular fractionation of cells transfected with chimeric constructs bearing luciferase as another reporting system (data not shown).

**Tissue-specific Expression of Alternative Transcripts—** Tissue-specific transcriptional activity of the internal promoter was analyzed in 24 human tissues by RT-PCR using a single round of 35 cycles. The primers used were designed to amplify a 923-bp fragment. The most abundant transcripts were found in peripheral blood leukocytes (Fig. 8) and to a lesser extent in placenta, heart, muscle, stomach, and testis. They were barely detectable in spleen, liver, adrenal gland, uterus, skin, and bone marrow and were negative in other tissues. Heart and muscle exhibited a high level of spliced transcripts revealed by a 629-bp fragment.

The Gene Expression Panel was also used to make semi-quantitative comparisons of alternative transcripts and *LGALS3* transcripts. As expected, *LGALS3* transcripts were highly abundant in various tissues, which was confirmed by the reference to actin transcripts. In all cases when expressed, the *LGALS3* transcripts were more abundant than the alternative transcripts. However, even if both transcripts were detected at different levels, it should be noted that some tissues expressed only alternative transcripts (muscle, stomach, and uterus), others produced both transcripts (PBL, heart, placenta, testis, skin) and others produced only Galectin-3 transcripts (kidney, lung, colon). These results suggest the independent functioning of the two promoters.
Characterization of an Internal Gene to Human LGALS3

Fig. 8. Detection of alternative transcripts in various human tissues. Human Rapid-Scan™ Gene Expression Panel was used to detect the alternative transcripts in various human tissues using RT-PCR. The primers used were designed to amplify a 923- or 629-bp fragment. LGALS3 transcripts were detected as a 457-bp DNA and actin transcripts as a 640-bp DNA. PCR was performed using 0.25 ng or 2.5 ng (10×) template cDNA.

DISCUSSION

The results presented in this study reveal the exceptional genetic organization of the LGALS3 locus listed as the following. (i) LGALS3 contains an internal gene, which is regulated and expressed mostly in peripheral blood leukocytes. (ii) This internal gene contains two different overlapping ORFs, which are translated upon transfection into proteins distinct from Galectin-3.

LGALS3 Contains an Internal Gene Named galig—We have shown that the mRNA transcribed from the internal promoter of LGALS3, initiated at multiple sites located in the second intron (Fig. 1). One type of mRNA contains the second intron and exons 3–6 of LGALS3. Another type, not detected in all cells analyzed, is produced by internal splicing, which removes most of the second intron sequence except the first 90 nucleotides (Fig. 3). Screening of dbEST database revealed the presence of both transcripts in human heart and infant brain cDNA libraries confirming the endogenous presence of these transcripts as polyadenylated mRNA in human tissues.

These alternative transcripts share with the major mRNA of LGALS3, the sequence spanning the third to sixth exon of the gene. Because the ATG codon used for Galectin-3 translation is located in the second exon, a sequence not present in the alternative transcripts, none of the alternative transcripts could encode a full-length Galectin-3 protein. The three potential ORFs, designated as ORF1, ORF2, and ORF3, are overlapping and are each positioned in a different reading frame. Of these three ORFs, only ORF3 would produce if translated a truncated form of Galectin-3 restricted to the carbohydrate recognition domain of the lectin (Refs. 12 and 13 and Fig. 3). Transfection experiments, using vectors expressing EGFP chimeras indicated that ORF3 was not translated. Surprisingly, fluorescence was detected in transfected cells with the two other vectors indicating that the alternative transcripts can produce both proteins encoded by ORF1 and ORF2 (Fig. 4). Western blotting and immunodetection using anti-EGFP antibodies confirmed that both ORFs were translated (Fig. 6). This implies an alternative initiation of translation at different AUG codons, which can be attributed to a leaky scanning process or to an internal entry of ribosomes (14–15). According to Kozak (16), the leaky scanning process would be favored if the first AUG lies in a weak context. This model is supported by the less favorable translation context of the ORF1 start sites, which exhibit A at position +4 (16). In contrast, the AUG codon for ORF2, despite its downstream position, lies in a canonical Kozak sequence (with A and G at positions −3 and +4, respectively). To assess experimentally the efficiency of initiation at these different AUG, the expression vectors were modified by replacing the EGFP cDNA with the luciferase gene. Upon transfection, the same luciferase activity was detected with both vectors, suggesting that the two ORFs were translated with the same efficiency.

These results justify that the structure of LGALS3 has to be redefined. Actually, this locus contains two overlapping genes, which could produce entirely distinct proteins. We propose to designate this internal gene as galig (standing for Galectin-3 internal gene). Such a genetic organization is extremely rare in mammalian genomes and has been reported so far in only two cases, the p16INK4A/p19ARF genes and the growth hormone/GH/DTA genes (17–18).

galig Can Be Translated into Two Distinct Proteins upon Cell Transfection—In addition to its noticeable genetic organization, the most striking property of galig is the capacity to encode two entirely distinct proteins (ORF1 and ORF2) from a single mRNA. In this regard, galig would appear to be as yet almost a unique example in mammals. Indeed, translation of different proteins from a single mRNA usually occurs from the use of in-frame AUG, resulting in proteins isoforms differing only at their N termini (19–27). To our knowledge, among higher eukaryotes, only the chick a2(I) and type III collagen genes might present a situation similar to galig. These genes contain internal promoters in introns 2 and 23, respectively. The resulting transcripts exhibit several overlapping ORFs, which appeared to be out-of-frame with the collagen coding sequences (28–30). However, in both cases, the question of the potentially translated ORFs has not been addressed.

The few reported cases in mammalian in which two completely distinct proteins are encoded from a single mRNA are the bicistronic mRNA producing genes (31–35). However, contrary to galig, the ORFs are organized in tandem such as a typical prokaryotic polycistronic mRNA.

ORF1 and ORF2 present a repetitive organization. This was expected because these sequences are colinear with the repetitive domain of the lectin mRNA (1). BLAST searches against ORF1 and ORF2 sequences revealed no significant homology with any known protein.

The ORF1-predicted protein would contain 106 amino acids with an apparent molecular mass of 11,253 Da. This sequence is highly rich in leucine (20%), proline (13%), and glycine (12%) residues (Fig. 3). ORF1-EGFP has a cytosolic and nuclear distribution (Figs. 4 and 6).

The ORF2-predicted protein is an 11,168 Da protein of 97 residues. The sequence is highly hydrophobic and positively charged, because of a large number of arginine residues (12% of total residues). A remarkable feature is the high content of tryptophan residues. 12% of the residues are tryptophan, the
average of other human proteins is 10 times lower. This rich content in tryptophans confers hydrophobic properties that may account for the membrane localization of the ORF2-EGFP protein (Fig. 6). Consistent with the mitochondrial localization of the ORF2-EGFP fusion protein, this sequence exhibits the common properties of mitochondrial-imported proteins such as the enrichment of arginine, leucine, and serine residues (36).

Tissue Specificity of galig Expression—Detection of galig transcripts in HOS cells and colon tumor cells revealed a low expression level. Based on this observation, the rationale that the appearance of galig transcripts may have resulted from a leaky transcription of a cryptic promoter rather than from an independently functioning promoter could not be excluded. Screening of several human tissues indicated clearly that galig is a tightly regulated gene whose expression is most efficient in mature forms of leukocytes. Whereas the precise quantification of galig mRNA has not been addressed in these experiments, it is known to be highly expressed when activated (37, 38). Indeed, the amount of LGALS3 transcripts, revealed no galig transcripts, thus indicating an independent functioning of the two promoters. In other tissues expressing both galig and LGALS3 mRNA (heart, testis, placenta, and skin, Fig. 8), cDNA amplification caused by infiltrating leukocytes cannot be ruled out. As Galectin-3 is thought to regulate cell proliferation, further investigation of a putative link between the functions of these overlapping genes would be extremely valuable.

Acknowledgments—We thank Dr. R. C. Hughes (NIMR (National Institute for Medical Research), London, United Kingdom) and Dr. M. Tiberi (LRI (Loeb Health Research Institute), Ottawa, Canada) for critically reviewing the manuscript.

REFERENCES
1. Barondes, S. H., Cooper, D. N., Gitt, M. A., and Leffler, H. (1994) J. Biol. Chem. 269, 20807–20810
2. Fowlis, D., Colnot, C., Ripoche, M. A., and Poirier, F. (1995) Dev. Dyn. 203, 241–251
3. Bao, Q., and Hughes, R. C. (1995) J. Cell Sci. 108, 2791–2800
4. Bao, Q., and Hughes, R. C. (1999) Glycobiology 9, 489–495
5. Perrillo, N. L., Marcus, M. E., and Baum, L. G. (1998) J. Mol. Med. 76, 402–412
6. Yang, R. Y., Hsu, D. K., and Liu, P. T. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 6737–6742
7. Kadrofske, M. M., Openo, K. P., and Wang, J. L. (1998) Arch. Biochem. Biophys. 349, 7–20
8. Raimond, J., Rouleux, F., Monsigny, M., and Legrand, A. (1995) FEBS Lett. 363, 165–169
9. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156–159
10. Legrand, A., Greenpan, P., Naggal, M. L., Nachtigal, S. A., and Nachtigal, M. (1991) Am. J. Pathol. 139, 629–644
11. Bossis, O., Lezoualc’h, F., Zanta, M. A., Mergny, M. D., Scherman, D., Demeneix, B., and Behr, J.-P. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7307–7311
12. Hsu, D. K., Zuberi, R. I., and Liu, P. T. (1992) J. Biol. Chem. 267, 14167–14174
13. Agrwal, N., Sun, Q., Wang, S. Y., and Wang, J. L. (1993) J. Biol. Chem. 268, 14382–14389
14. Koizumi, M. (1992) Annu. Rev. Cell Biol. 8, 197–225
15. Gray, N. K., and Wickens, M. (1998) Annu. Rev. Cell Dev. Biol. 14, 399–458
16. Kozak, M. (1997) EMBO J. 16, 2482–2492
17. Labarriere, N., Selvais, P. L., Lemaigre, F. P., Michel, A., Matter, D. M., and Rousseau, G. G. (1995) J. Biol. Chem. 270, 19295–19298
18. Quelle, D. E., Zindy, F., Ashmun, R. A., and Sherr, C. J. (1995) Cell 83, 993–1000
19. Voss, J. W., Yao, T. P., and Rosenfeld, M. G. (1991) J. Biol. Chem. 266, 12832–12835
20. Askhi, M., Hamada, F., Sugimoto, T., Sumida, S., Akitsuka, T., and Toyoshima, K. (1993) J. Biol. Chem. 268, 22723–22728
21. Callegari, V., Bottardi, S., Coggi, S., Apesteguia, I., and Santoro, C. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 11598–11602
22. Vagner, S., Genasc, M. C., Maret, A., Bayard, F., Almaric, F., Prats, H., and Prats, A. C. (1995) Mol. Cell. Biol. 15, 35–44
23. Packham, G., Brimmell, M., and Cleveland, J. L. (1997) Biochem. J. 326, 807–813
24. Akagi, N., Nakahi, D., Finkelestein, Y., Le, S. Y., Elroy-Stein, O., and Levi, B. Z. (1998) Oncogene 17, 227–236
25. Okazaki, S., Ito, T., Uj, M., Watanabe, T., Yoshimatsu, K., and Iba, H. (1998) Biochem. Biophys. Res. Commun. 250, 347–353
26. Yang, X., Chenkenko, G., Hao, Y., Ding, Z., Pater, M. M., Pater, A., and Tang, S. C. (1998) Oncogene 17, 981–989
27. Ayoubi, T. A., and Van De Ven, W. J. (1996) FASEB J. 10, 453–460
28. Bennett, V. D., and Adams, S. L. (1997) J. Biol. Chem. 265, 2223–2230
29. Nah, H.-D., Niu, Z., and Adams, S. L. (1994) J. Biol. Chem. 269, 16443–16448
30. Zhang, Y., Niu, Z., Cohen, A. J., Nah, H.-D., and Adams, S. L. (1997) Nucleic Acids Res. 25, 2470–2477
31. Lee, S. J. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 4250–4254
32. Sloan, J., Kinghorn, J. R., and Unkles, S. E. (1999) Mol. Cell. Biol. 14, 7635–7545
33. Ritchie, H., and Wang, L. H. (1997) Biochem. Biophys. Res. Commun. 231, 425–438
34. Sloan, J., Kinghorn, J. R., and Unkles, S. E. (1999) Nucleic Acids Res. 27, 854–858
35. Stallmeyer, B., Drugeon, G., Reiss, J., Haenni, A. L., and Mendel R. R. (1999) Am. J. Hum. Genet 64, 698–705
36. Horwich, A. (1990) Annu. Rev. Cell Biol. 6, 625–633
37. Hsu, D. K., Zuberi, R. I., and Liu, P. T. (1992) J. Biol. Chem. 267, 14167–14174
38. Nangia-Makker, P., Ochieng, J., Christman, J. K., and Raz, A. (1993) Cancer Res. 53, 5033–5037
Identification of an Internal Gene to the Human Galectin-3 Gene with Two Different Overlapping Reading Frames That Do Not Encode Galectin-3
Michaël Guittaut, Stéphane Charpentier, Thierry Normand, Martine Dubois, Jacques Raimond and Alain Legrand

J. Biol. Chem. 2001, 276:2652-2657.
doi: 10.1074/jbc.M002523200 originally published online November 3, 2000

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 38 references, 18 of which can be accessed free at http://www.jbc.org/content/276/4/2652.full.html#ref-list-1