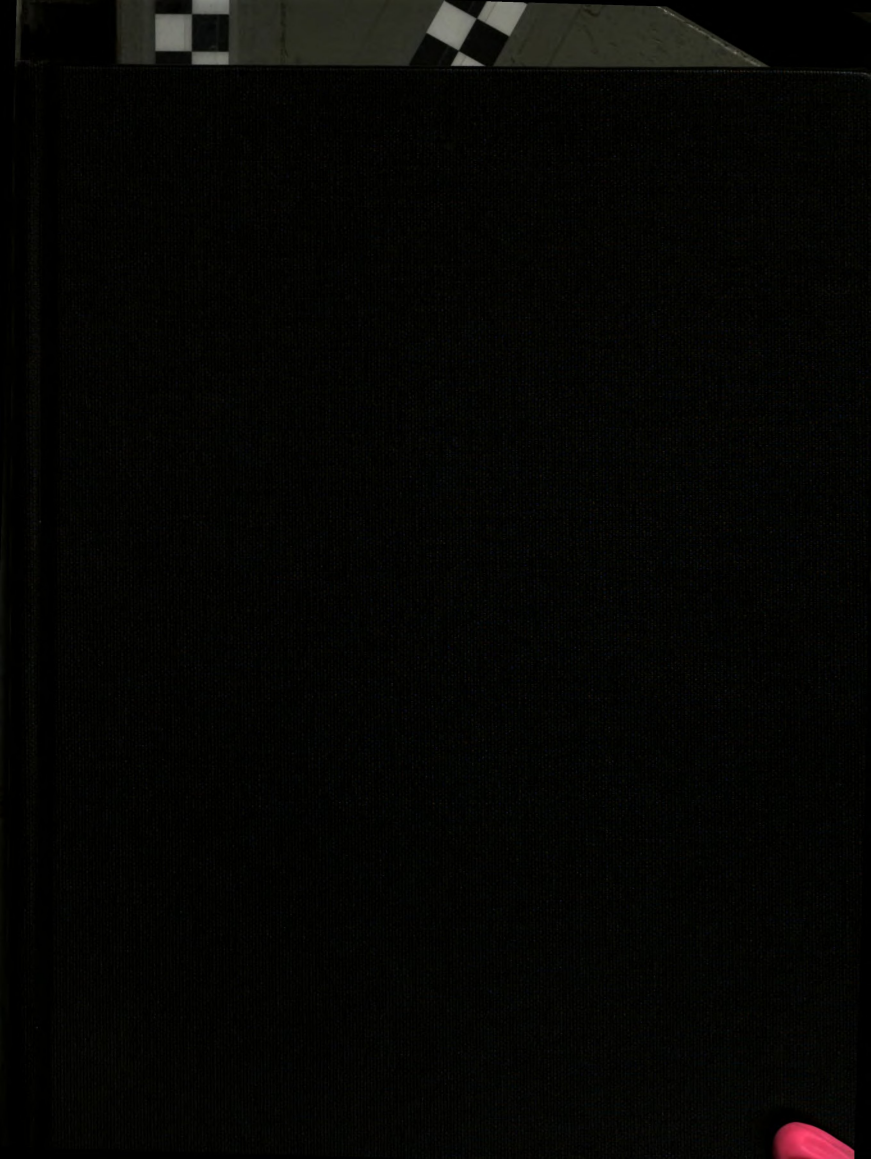


This is to certify that the

dissertation entitled

GALECTIN-1 IN THE CELL NUCLEUS:

EVIDENCE FOR A ROLE IN PRE-mRNA SPLICING

presented by

Anandita Vyakarnam

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry

Major professor

Date 12/20/96

GALECTIN-1 IN THE CELL NUCLEUS: EVIDENCE FOR A ROLE IN PRE-mRNA SPLICING

By

Anandita Vyakarnam

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1996

ABSTRACT

GALECTIN-1 IN THE CELL NUCLEUS: EVIDENCE FOR A ROLE IN PRE-mRNA SPLICING

By

Anandita Vyakarnam

Galectins are a family of β-galactoside specific lectins that share conserved amino acid residues in their carbohydrate recognition domains. Two members of this family, galectin-1 and galectin-3 are widely distributed proteins and are present mainly inside cells. Previous studies have shown that galectin-3 is localized to the cell nucleus and is a component of ribonucleoprotein complexes.

Using a cell free assay, galectin-3 has been shown to be a required factor in pre-mRNA splicing. HeLa cell nuclear extracts that are splicing competent become splicing deficient when depleted of galectin-3 by adsorption on a lactose affinity column. The splicing activity can be restored by addition of recombinant galectin-3. However, depletion of galectin-3 from nuclear extracts by using a monoclonal antibody specific for galectin-3 did not inhibit splicing activity. Furthermore, the splicing activity of this galectin depleted extract was still sensitive to saccharide inhibition. This provided a hint that there may be other galectins present in HeLa nuclear extract that are also involved in pre-mRNA splicing.

The goal of this dissertation was to define the number and identity of other galectins in HeLa nuclear extracts that may be involved in pre-mRNA splicing. We show that galectin-1 is present in the nuclear extract and participates in pre-mRNA splicing in a cell free system. Depletion of galectin-1 or galectin-3 alone by specific antibody adsorption failed to remove all splicing activity and the residual activity was still sensitive to inhibition by sugars. Nuclear extracts depleted of both galectin-1 and galectin-3, either by lactose affinity adsorption or by double antibody adsorption, are no longer splicing competent. Either galectin-1 or galectin-3 alone is sufficient to reconstitute splicing activity of the depleted extracts.

Galectin-1 has been shown to be present intracellularly. However, the issue of its nuclear localization is controversial. We have analyzed the intracellular distribution of galectin-1 using laser confocal microscopy. We find that galectin-1 is present both in the nucleus and cytoplasm of HeLa cells. In addition, the type of permeabilization reagent used has a profound effect on the observation of galectin-1 in the nucleus. We have also compared the distribution of galectin-1 and galectin-3 in the nucleus with the distribution of the splicing factor SC35, which is localized in a speckled staining pattern in the nucleus. HeLa cells that were permeabilized and then fixed, were stained simultaneously with antibodies specific for galectin-1 or galectin-3 and a monoclonal antibody against SC35. Our results indicate that galectin staining in the nucleus colocalizes with the speckled staining pattern of SC35 although galectins also show a more diffused staining.

Dedicated to my parents

ACKNOWLEDGEMENTS

I would like to thank Dr. John L. Wang for his guidance and support through these years. He has provided me many opportunities that helped me in becoming a better researcher and achieve the goals of my education. I would also like to thank the members of my guidance committee: Drs. Laurie Kaguni, William Deal, Justin McCormick and Rich Schwartz for their valuable advice and input from time to time during my research.

I would like to express my gratitude to Dr. Ronald Patterson who has been very helpful to me in setting up splicing assays. His many suggestions and long standing interest in my thesis project were much appreciated. I am grateful to Dr. Sue Dagher who was always available for advice and help in troubleshooting. I would also like to thank Dr. Melvin Schindler and Sharon Grabski for their valuable suggestions and help especially with confocal microscopy. I would like to thank Karen Lakkides for helping me with immunofluorescence studies and Lynn Johnson for helping me in the initial stages of splicing project.

My stay at Michigan State would not have been the same without the friends I have made here. I would like to thank the members of the Wang and Schindler laboratories, past and present, for all their help and support. I would like to express my appreciation to Patty Voss, Mark Kadrofske, Yeou-Guang Tsay and Eric Arnoys for their wonderful friendship and many stimulating discussions. They have made my time in graduate school a very memorable experience. I would also like to thank a number of

friends outside of the lab, particularly Linda, Carol, Beta and Marty for their support and constant enthusiasm.

I especially want to thank my parents for their unwavering support and encouragement throughout the years. My mother's love and sacrifice has always been an inspiration for me. I would like to thank my sister for her love, friendship and understanding over the years. Finally, I want to thank my husband, Murty for all his love, friendship, encouragement and patience over the years. His support and understanding made it easier to meet the challenges of graduate school.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I.

LIST OF FIGURES

CHAPTER I.

CHAPTER II.

CHAPTER III.

CHAPTER I

LITERATURE REVIEW

LITERATURE REVIEW ON GALECTINS

Galectins- A Family of Animal Lectins

Lectins are defined as non-enzyme, non-immunoglobulin proteins that have at least one carbohydrate binding domain (1). Lectins have been identified in various species ranging from viruses and bacteria to plants and vertebrates. Since glycoconjugates have recently been found in the cytoplasm and the nucleus (2), there has been a great interest in the intracellular localization of lectins. This dissertation deals with the intracellular localization of galectin-1 and -3 and a role for these lectins in pre-mRNA splicing.

Galectins are a family of animal lectins characterized by: (a) affinity for β-galactoside-containing glycoconjugates and (b) conserved amino acid residues in their carbohydrate recognition domain (CRD) (3). They are soluble proteins that lack a signal peptide and are found mainly intracellularly. However, there is evidence that some members are also externalized to the cell surface and in the extracellular matrix. Galectins do not require metal ions for their carbohydrate binding activity, which distinguishes them from the C-type lectins that require calcium ions for sugar binding (4). Members of the galectin family are present in a vast variety of tissues and various species ranging from lower invertebrates like sponges and nematodes to higher vertebrates including humans. They have been called by different names by various research groups (soluble lectins, S-type lectins, galaptins, β-galactoside binding lectins etc.) but recently a

consensus on nomenclature has been reached by investigators in the field to use the term 'galectin' (3).

Structural Classification of Galectins

Based on their molecular architecture, members of the galectin family are sub-divided into 3 categories (Figure 1) (5).

(i) *Proto Type galectins*: These have a polypeptide (Mr ~14,000) that contains a consensus CRD.

(ii) *Chimera Type galectins*: These are multi-domain proteins consisting of a N-terminal domain which is proline and glycine rich and a C-terminal domain which is homologous to the Proto Type galectin CRD.

(iii) *Tandem Repeat Type galectins*: These are composed of two tandemly repeated lectin domains each homologous to the Proto Type galectin CRD.

At present, eight mammalian galectins have been reported (Table 1). Galectin-1, -2, -5 and -7 are Proto Type galectins. Of these, galectin-1 and -2 exist as noncovalent dimers under physiological conditions whereas galectin-5 and -7 are monomeric. Galectin-1 is isolated as a homodimer with subunit polypeptide molecular weight of ~14 kDa and is not subjected to any post-translational modifications (6). It is abundant in smooth and skeletal muscle, but is also found in many other cell types. Galectin-2, originally found in human hepatoma, is isolated as a homodimer and has 43% amino acid identity to human galectin-1 (7). Galectin-5 was isolated from rat lung as a monomer of ~18 kDa and has now been localized to rat erythrocytes (8). Galectin-7 is a 15 kDa monomeric protein found in

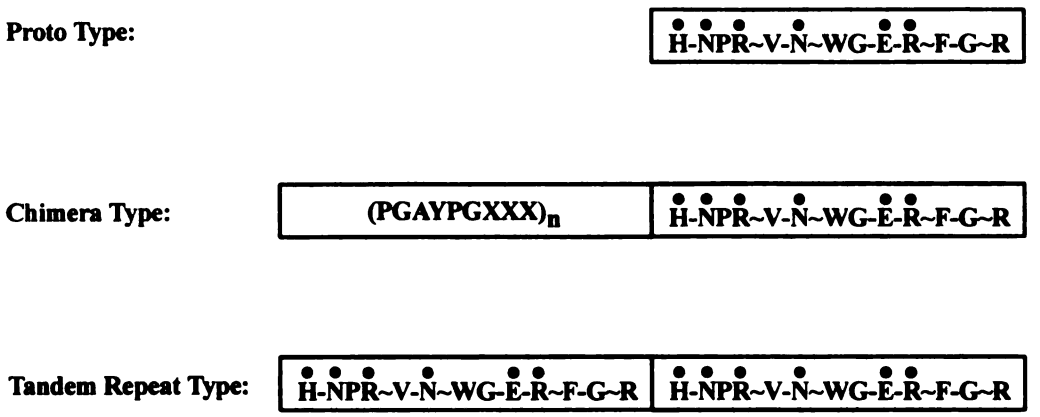


Figure 1. Schematic diagram illustrating the polypeptide architecture of galectins. The Proto Type is composed of a single lectin domain that contains the CRD. The Chimera Type galectin has two parts, a C-terminal half containing the galectin CRD and an N-terminal half rich in proline and glycine residues. The Tandem Repeat Type has two homologous CRD domains. Conserved amino acid residues in galectin CRD are indicated. Dark circles denote residues that directly interact with the carbohydrate by hydrogen bonding. In the N-terminal domain of Chimera Type, nine amino acid sequence motif that is tandemly repeated 8-12 (n) times is shown. The single letter amino acid code is used. X, any amino acid.

Table 1. Galectin members in mammalian species.

Galectin	Structural Type	Subunit Mr	Subunit Structure
Galectin-1	Proto	14,500	Homodimer
Galectin-2	Proto	14,600	Homodimer
Galectin-3	Chimera	35,000	Monomer
Galectin-4	Tandem Repeat	36,000	Monomer
Galectin-5	Proto	16,000	Monomer
Galectin-6	Tandem Repeat	34,000	Monomer
Galectin-7	Proto	15,000	Monomer
Galectin-8	Tandem Repeat	35,000	Monomer

human keratinocytes (9). Galectin-3, previously known as CBP35, was first isolated from mouse 3T3 fibroblasts (10). It consists of a single polypeptide of ~35 kDa and is at present, the only representative of the Chimera Type of galectins. The carboxy terminal half of galectin-3 contains the CRD and the amino terminal half is proline and glycine rich because the sequence PGAYPGXXX is repeated 8-12 times, depending on the species. This domain shows sequence similarity to heterogeneous ribonucleoproteins (hnRNP) (11). Galectins-4, -6 and -8 are Tandem Repeat Type galectins. Galectin-4 was found as an abundant rat intestinal 36 kDa lectin (12). Galectin-6 has been mentioned (3) but its characterization is yet to be published. Galectin-8 was originally cloned from rat liver with a molecular weight of 34 kDa. (13). Unlike galectin-4, which is abundant in the intestine, galectin-8 is expressed in liver, kidney, cardiac muscle, lung and brain. Overall, galectin-8 has 34% identity to galectin-4.

Galectins were initially investigated mainly as vertebrate lectins from various species like electric eel, chick, bovine, rat, mouse and human. However, recently galectins have also been found in invertebrates (Table 2). A Tandem Repeat Type 32 kDa galectin was identified in the nematode *Caenorhabditis elegans* as the first invertebrate galectin (14, 15). The internal sequence homology between the two tandemly repeated domains is 32% amino acid identity between the first and second domains. Both the first and second parts have 25-30% amino acid identity to vertebrate lectins. A 16 kDa Proto Type galectin has also been identified in *Caenorhabditis elegans* (16). A Tandem Repeat Type galectin has also been found in another member of the nematode family, *Onchocerca volvulus* and this galectin has 71% amino acid identity to

Table 2. Galectins classified according to polypepetide architecture.

Structural Type	Animal class	Species	Example
Proto	Mammals	Human, mouse, rat cow, etc.	Galectins-1, -2, -5, -7
	Birds	Chicken, quail	14, 16 kDa galectins
	Amphibia	Frog *(Xenopus, Rana)*	14, 16 kDa galectins
	Fish	Electric eel, conger eel	14, 16 kDa galectins
	Nematodes	*C. elegans*	16 kDa galectin
		O. volvulus	16 kDa galectin
	Sponges	*Geodia cydonium*	15.1, 15.3 kDa galectins
Chimera	Mammals	Human, mouse, rat	Galectin-3
Tandem-repeat	Mammals	Human, rat, mouse, pig	Galectins-4, -6, -8
	Nematodes	*C. elegans*	32 kDa galectin

C. elegans 32 kDa galectin (17). Furthermore, two galectin cDNAs were cloned in the marine sponge *Geodia cydonium* which code for Proto Type galectins (18). Therefore, the galectin family is now known to be widely distributed in the animal kingdom.

Carbohydrate Recognition Domain of Galectins

The sequence of each CRD has been shown to be mainly encoded by 3 exons (19, 20). Most of the residues that are conserved among galectins are found in the sequence encoded by the middle one of these 3 exons. In the crystallographic structure of galectin-1 and -2 (21-23), this sequence includes four contiguous β-strands and intervening loops and contains all residues that interact directly with the carbohydrate ligand. However, deletion of sequences encoded by the other two exons that encode the CRD also impairs activity.

The key structure recognized by galectin CRD is lactose and other related β-galactosides (24-28). Mechanisms underlying the specific interaction have been elucidated by binding experiments, site-directed mutagenesis and x-ray crystallography. X-ray crystal structures of galectin-1 (22, 23) and galectin-2 (21) confirm that the major interaction is with the galactose residue in lactose. Some of the highly conserved residues of the CRD were found to form hydrogen bonds with hydroxyl groups of lactose; 4-OH of Gal with His45, Asn47 and Arg49, 6-OH of Gal with Asn58 and Glu68 (residue numbers are those of human galectin-2). However, interaction with glucose residue in lactose is also significant as reflected in 100-fold higher affinity for lactose compared with galactose (24). The 3-OH of glucose hydrogen bonds with conserved residues Arg49,

Glu68 and Arg70. Trp67 seems to have been conserved mainly for its structural importance or stabilization of the binding site by close contact with C-6 of galactose. Substitution of tyrosine by site-directed mutagenesis slightly reduced lactose binding affinity but the mutant was still adsorbed strongly on lactose affinity column. Substitutions of a series of the conserved hydrophilic residues (His45, As47, Arg49, Asn58, Glu68 and Arg70) led to an almost complete loss of the binding ability (29, 5). However, both N-terminal and C-terminal regions of the 14 kDa polypeptide also seem to be important for carbohydrate binding probably by maintaining structure of CRD, because mutant proteins of the bovine 14-kDa lectin lacking N-terminal or C-terminal 10 residues were found to be inactive (30).

Physical Properties and Subcellular Localization of Galectin-3

Galectin-3 was initially isolated from extracts of mouse 3T3 fibroblasts on the basis of its galactose/lactose specific carbohydrate binding activity(10). Its homologs have since been found in a variety of tissues and species . Sequence and hydropathy analysis of the amino acid sequence for galectin-3 indicates that it has two distinct structural domains (for review, see reference (31)). The C-terminal half, containing the CRD, exhibits both hydrophilic and hydrophobic regions, as is characteristic of many globular proteins . In contrast, the N-terminal half consists, in large part, of a repeated 9-residue motif, PGAYPGXXX; the sequence is repeated eight times in human and hamster, ten times in mouse and rat and twelve times in canine homolog . This proline- and glycine-rich sequence is sensitive to collagenases whereas the CRD is very resistant to proteolysis (32). Collagenase digestion has been used to generate the C-domain and

this domain has been shown to be sufficient for carbohydrate binding activity . Differential scanning calorimetry of galectin-3 and its individual domains yielded transition temperatures of 39 °C and 56 °C for N- and C-terminal domains respectively (33). Lactose binding by the C-domain shifted the transition temperature to 65 °C, whereas sucrose failed to yield the same effect. These results suggest that the individual domains of galectin-3 polypeptide are folded independently and that ligand binding by galectin-3 is accompanied by a conformational change that significantly stabilizes the polypeptide against thermal denaturation.

Mouse 3T3 fibroblasts, fixed with paraformaldehyde and permeabilized with Triton X-100, yielded prominent nuclear and variable cytoplasmic staining with anti-galectin-3 antibodies (34). The nuclear staining appeared to be diffuse, covering the entire nucleus with the exception of about five black holes that correspond to nucleoli. At the ultrastructural level, electron microscopic analysis using anti-galectin-3 antibodies yielded immunogold labeling of interchromatic spaces and at the borders of condensed chromatin (perichromatic fibrils) (35). This localization is similar to the reported distribution of the non-snRNP splicing factors which will be discussed in detail later in this review. The nuclear staining of galectin-3 in immunofluorescence and ultrastructural experiments was sensitive to treatment of the permeabilized cells with ribonuclease A, but not to parallel treatment with deoxyribonuclease I (34, 35). The nuclear and cytoplasmic localization of galectin-3 has also been documented by quantitative immunoblotting analysis of subcellular fractions, relative to enzyme markers characteristic of the various fractions (36).

The pI of murine galectin-3 is 8.7, as determined both by calculation from the deduced amino acid sequence and experimentally by isoelectric focusing of recombinant protein obtained by expression in *E. coli* (37). When extracts of mouse 3T3 cells were subjected to two-dimensional gel electrophoresis and immunoblotting, two spots were observed corresponding to pI values of 8.7 and 8.2. The pI 8.2 species represents a post-translational modification of the pI 8.7 form by the addition of single phosphate group. It has been shown that canine homolog of galectin-3 is phosphorylated mainly at Ser6 with Ser12 carrying a small percent of total radiolabel (38). In addition, Casein Kinase I is shown to phosphorylate recombinant human galectin-3 in vitro at Ser6. The phosphorylated (pI 8.2) form of galectin-3 is found both in the cytosol and nucleus, whereas the unmodified (pI 8.7) species is found exclusively in the nucleus. In quiescent populations of cell cultures, galectin-3 expression level is low and it is present primarily in the phosphorylated form, is located predominantly in the cytoplasm (37). Serum stimulated cells have an increased level of galectin-3. The phosphorylated form is present both in the cytosol and the nucleus but the amount of unmodified form increases dramatically and is present only in the nucleus.

Studies on human colonic epithelial tissue specimens have shown that galectin-3 expression and its nuclear versus cytoplasmic localization vary along the crypt to surface axis (39). Galectin-3 is concentrated in the nuclei of differentiated epithelial cells. During the progression from normal mucosa to adenoma to carcinoma, the expression and nuclear localization of galectin-3 changed significantly. In the carcinoma cells, galectin-3 is absent from the nuclei and is localized only to the cytoplasm. These

observations suggest that the exclusion of galectin-3 from the nucleus may be related to the neoplastic progression of colon cancer.

Although the majority of galectin-3 is intracellular, there is also evidence that galectin-3 can be externalized. From analyzing the cDNAs corresponding to galectin-3 from a number of species and sources, there does not appear to be a typical signal sequence for sequestration into the endoplasmic reticulum and the endomembrane pathway for secretion. Thus it is most probably externalized by a non-classical mechanism. Externalization has been documented in macrophages (40, 41) and in polarized kidney and intestinal epithelial cells, where the secretion is specifically toward the apical surface (42, 43). Galectin-3 has also been found at the cell surface of dorsal root ganglion neurons (44) and in murine fibrosarcoma cells (45). Thus, galectin-3 joins a growing list of polypeptides that exhibit both intracellular (and, in particular, nuclear) and extracellular localization, despite the lack of an obvious signal sequence. The best examples of this list include (a) proteins with nuclear functions, such as the simian virus 40 (SV40) large T antigen (46); (b) members of the fibroblast growth factor (FGF) family (47); (c) interleukins IL-α and 1β (48); and (d) the annexins (49). More recently, there has been much interest on reports that the BReast CAncer 1 (BRCA1) gene product can be found in the nucleus (50), cytoplasm (51), as well as outside of cells (52).

Subcellular Localization of Galectin-1

The subcellular localization of galectin-1 shares features observed for galectin-3: predominantly intracellular but with some externalization. Since this issue, particularly

with respect to the nuclear versus cytoplasmic localization, is one focus of my dissertation, it will be discussed in Chapter 3, in the context of the data obtained in my studies.

LITERATURE REVIEW ON PRE-mRNA SPLICING

Pre-mRNA Splicing and SnRNPs

Pre-mRNA splicing takes place via a two-step mechanism involving sequential transesterification reactions (Figure 2) (for reviews, see references (53, 54)). The first step involves cleavage of pre-mRNA at the 5' exon-intron boundary generating a free exon 1 with a 3'-OH terminus and an intron lariat-exon 2. The lariat structure results from the formation of a 2'-5' phosphodiester bond linking 5' terminus of the intron to the ribose 2'-OH group of an adenosine residue in the intron, usually located 20-40 nucleotides upstream of the 3' splice site. The second catalytic step is cleavage of 3' splice site, ligation of the two exons and excision of the intron lariat.

The splicing reaction takes place within a multi-component complex called the spliceosome. The spliceosome is comprised of RNA-protein subunits called small nuclear ribonucleoprotein particles (snRNPs) (Tables 3, 4). Each of the snRNP subunits contain one or more small nuclear RNAs (snRNAs) and a set of snRNP proteins. In mammalian species, there are six major snRNAs named U1 through U6 ranging in size from 107-216 nucleotides. The snRNAs U1-U5 are transcribed by RNA polymerase II and snRNA U6 by RNA polymerase III. With the exception of U6, newly transcribed monomethyl G-capped UsnRNA is transported to the cytoplasm where the cap is modified to a trimethyl G structure and the snRNA assembles with proteins to form the snRNP (for a review see references (53, 54). The major snRNPs are located in the nucleoplasm except for U3snRNP which is localized in the nucleolus. Each snRNP

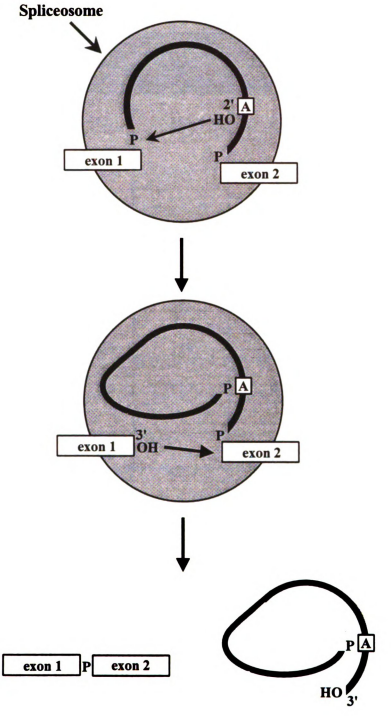


Figure 2. Schematic diagram of spliceosome-catalyzed splicing.

Table 3. Major mammalian SnRNPs.

SnRNP	Nuclear Localization	SnRNA			Protein Composition	
		RNA	size (nucleotides)	RNA polymerase	Protein	Mr
U1	Nucleoplasm	U1	165	II	Core proteins	
					U1 70K	70,000
					A	34,000
					C	22,000
U2	Nucleoplasm	U2	189	II	Core proteins	
					A'	31,000
					B"	28,000
U3	Nucleolus	U3	216	II	U3 34K	34,000
U4	Nucleoplasm	U4	139	II	Core proteins	
U5	Nucleoplasm	U5	117	II	Core proteins	
					U5 100K	100,000
U6	Nucleoplasm	U6	107	III		

Table 4. SnRNP core proteins.

Protein Name	Mr
B'	29,000
B	28,000
D'	18,000
D	16,000
E	13,000
F	12,000
G	11,000

particle contains a common set of "core" proteins designated B, B', D, D', E, F and G that range in size from 11-29 kDa. They interact with a conserved sequence present in U1, U2, U4 and U5 snRNAs (55). Binding of the core proteins to the snRNA is directed by a sequence motif of AU_nG (n = 2-6). Insertion of this sequence into heterologous RNA has been shown to induce the assembly of the snRNP core proteins suggesting that this sequence is necessary and sufficient for core protein binding. Patients with autoimmune diseases like systemic lupus erythmatosus (SLE) possess antibodies against nuclear antigens called Sm antigens (56). Anti-Sm antibodies recognize antigenic determinants on B and D polypeptides of the snRNP core proteins.

In addition to the core proteins, each of the spliceosomal snRNPs also contains specific proteins. For example, the A, C and 70 kDa proteins are specific for U1snRNP, while A' and B'' are specific for U2 snRNP. It is not clear what roles these different proteins might play in the splicing mechanism. It is possible that some of them may be required for snRNP assembly or transport.

Spliceosome Assembly Pathway

Spliceosome assembly is a multistep process, the key steps of which are shown in figure 3. Prior to the formation of first splicing complex, newly transcribed pre-mRNA is bound by hnRNP proteins to form the hnRNP complex or the "H complex". Formation of the "H complex" does not require functional 5' or 3' splice sites, therefore, it is not considered as a functional intermediate in the spliceosome formation (57). First step is the formation of a "commitment complex" which is a stable complex containing

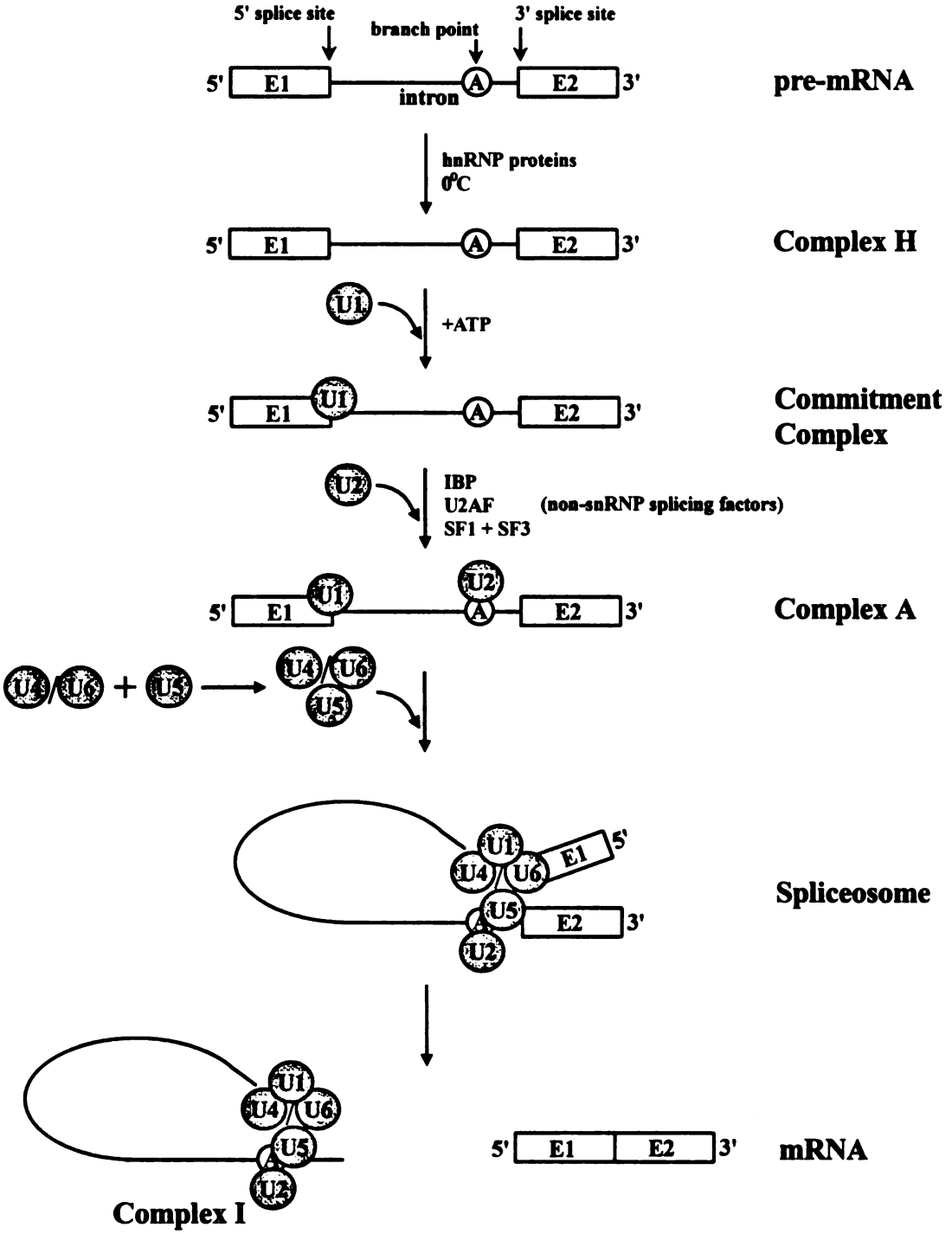


Figure 3. The pre-mRNA splicing pathway in mammalian cells.

U1snRNP. Although the pre-mRNA contained within this complex has not yet been spliced, at this stage it can no longer be displaced by the addition of an excess of competitor pre-mRNA to the *in vitro* splicing reaction. The binding of U1snRNP involves the formation of base pairs between the intron sequence at the 5' splice site and a complementary sequence at the 5' terminus of U1snRNA. U1snRNP is also involved in interactions with the sequences at the 3' end of the intron. Such interactions may play a role in bringing together the separate ends of the intron within the spliceosome. Next a pre-spliceosome complex called the "A complex" is formed by the ATP dependent binding of U2snRNP to the pre-mRNA branch site. Binding of U2snRNP requires U1snRNP (58) and three protein factors SF1, SF3 and U2 auxiliary factor (U2AF) (59-61). The spliceosome or "B complex" is formed following the addition of a preassembled U4/U6.U5 tri-snRNP particle. U4 and U6 snRNAs bind to each other through an extended complementary region to form the U4/U6 snRNP complex (62) and then the U4/U6 snRNP particle associates with U5 snRNP (57). This tri-snRNP is then bound to "A complex" to form the "B complex" or spliceosome. At this point the two transesterification reactions (figure 2) take place within the spliceosome complex. The spliceosome complex dissociates and the newly joined exons are released from the snRNP-intron "I complex". The intron RNA is degraded and snRNPs are recycled for further splicing.

Non- snRNP Splicing Factors

In addition to the snRNP subunits, spliceosomes are also comprised of ~100 non-snRNP protein splicing factors which are assembled onto pre-mRNA in an orderly

fashion to form spliceosomes. These factors function at multiple steps during spliceosome assembly. They are involved both in recruiting spliceosomal snRNPs to the pre-mRNA and in mediating interactions between spliceosomal components to specify splice sites and define exons and introns. Several of these protein factors have been identified including ASF/SF2 (63,64), U2AF (U2snRNP Auxiliary Factor) (61) 65-kDa subunit (U2AF65), U2AF 35-kDa subunit (U2AF35), SC35 (65), U1snRNP 70-kDa protein (66) etc. (Table 5). A sequence comparison of these factors reveal a number of similarities including the presence of RNA recognition motifs (RRM) and SR motifs making them members of the SR protein family (for a review see reference (67).

SR proteins were originally described by Roth and coworkers based on five criteria (68). i) They contain a shared phosphoepitope recognized by the monoclonal antibody mAb104. ii) They copurify in a two-step salt precipitation procedure (soluble in 65% ammonium sulfate and precipitated in 20 mM $MgCl_2$). iii) They share a similar protein structure containing at least one RNA recognition motif (RRM) and an RS domain. iv) Their sizes on SDS-PAGE are conserved from *Drosophila* to man. v) They can complement splicing deficient cytoplasmic S100 extracts, indicating a redundant function in splicing. A second class of proteins called SR protein related polypeptides (SRrp) also contain SR domain but lack RRMs. Together, SR proteins and SR protein related polypeptides constitute the superfamily of RS domain proteins.

The length and composition of the SR domain and its location within the primary sequence of the protein is variable (figure 4). For example, the SR motif is located at the

Table 5. Superfamily of pre-mRNA splicing factors containing an SR domain.

SR proteins		SR protein related polypeptides	
Humans	**Other Species**	**Humans**	**Other Species**
SRp20	x16 (mouse) = SRp20	U170k	Tra (*Dros*.)
SRp30a/SF2/ASF	pr264 (chicken) = SC35	U2AF65	Tra-2 (*Dros*.)
SRp30b/SC35	HRS (rat) = Srp40	U2AF35	SWAP (*Dros*.)
SRp30c	B52 (*Dros*.) = SRp55	HCC1	xU170k (*Xenopus*) = U170k
SRp40	RBP-1 (*Dros*.)	HRH1	dU170k (*Dros*.) = U170k
SRp55	SR-1 (plant)	CLK-1	mU2AF65 (mouse) = U2AF65
SRp75	Npl3 (*S. cerevisiae*)	CLK-2	dU2AF50 (*Dros*.) = U2AF65
9G8		CLK-3	dU2af38 (*Dros*.) = U2AF35

Dros., Drosophila
=, homologous
Table modified from reference (67)

N-terminus of U2AF65, while in ASF/SF2 and SC35 the SR motif occurs at the C-terminus. The SR motif itself can either consist of an uninterrupted stretch of SR dipeptides or of a more dispersed SR-rich region. RRM motifs have been found in diverse number of proteins that bind RNA. RRM includes two conserved sequences, RNP-1 and RNP-2. The RRM domain has been shown to be necessary for high affinity and sequence specific binding of protein to RNA.

In addition to the SR protein family members described above, an RS domain is present in other essential splicing factors and regulators. In fact, the prototype for RS domain proteins are the *Drosophila* splicing regulators SWAP (69), Tra (70), Tra-2 (71). These three proteins are not essential splicing factors, but are specific splicing regulators. Tra and Tra-2 are required for female-specific splicing of the doublesex (Dsx) transcript (for review see reference (72)), and SWAP undergoes autoregulation by suppressing the splicing of its own transcript. These and other SR domain proteins are collectively referred to as SR related polypeptides, or 'SRrps'.

Among the SRrps is the U1 snRNP-associated 70-kDa protein (U1 70k) that contains an RS domain (67). U2snRNP auxiliary factor (U2AF) is a protein splicing factor required for stable binding of U2snRNP to pre-mRNA during spliceosome assembly (62). It binds specifically to polypyrimidine tract/3' splice site region of pre-mRNAs. U2AF is comprised of two associated polypeptides of 35 kDa (U2AF35) and 65 kDa (U2AF65) which are present in equimolar amounts. U2AF65 is an essential splicing factor that has an N-terminal RS domain followed by three RNA-binding domains at the

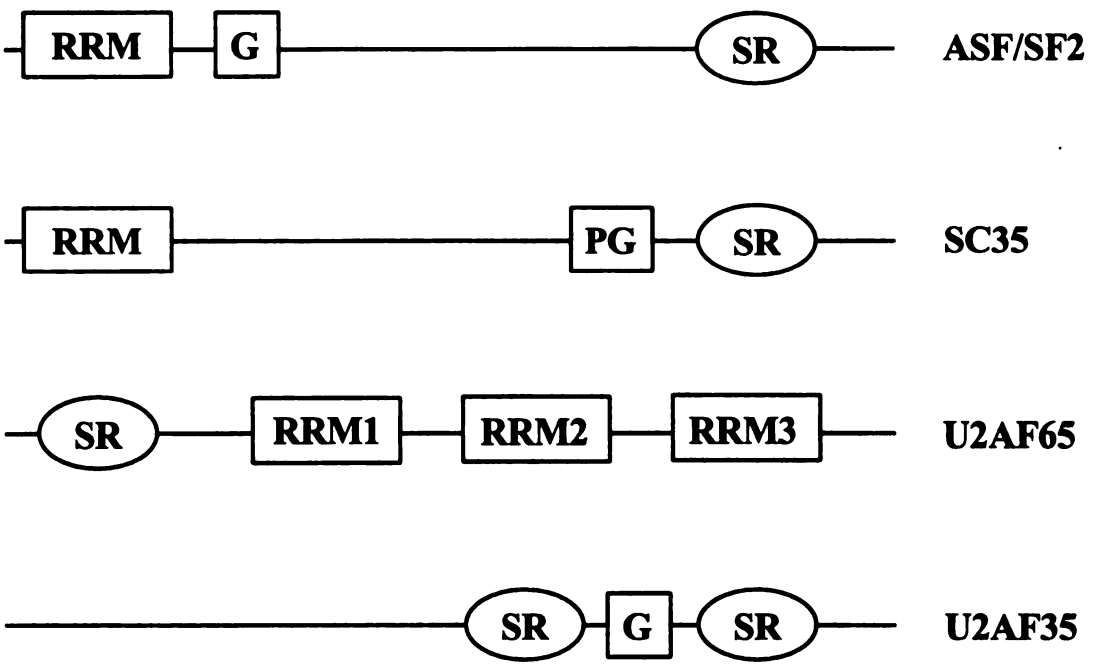


Figure 4. Schematic diagram summarizing the context and organization of specific domains in the SR family of proteins.

The RRM box indicates the RNA recognition domain. The relative location of the serine/arginine rich domain (SR) is shown for each protein. PG and G indicate proline/glycine and glycine rich regions, respectively.

C-terminus (73). However, similar to Tra and SWAP, the small subunit of U2AF (known as U2AF35 in human and U2AF38 in *Drosophila*) contain an SR domain but lacks an RNA-binding motif, and appears unnecessary for constitutive splicing in nuclear extracts from which the U2AF is removed (74). However, recent studies show that U2AF35 is involved in specific protein-protein interactions with SR proteins (75), and both subunits of U2AF are specifically associated with the mammalian commitment complex (76). Finally, the most recent addition to the superfamily of SR domain proteins is a kinase called CLK/STY which can interact with SR proteins in yeast two-hybrid system and phosphorylate these proteins. It is a dual specificity kinase that can autophosphorylate at both serine/threonine and tyrosine residues (77, 78). Human cells express two other related kinases CLK-2 and CLK-3. These kinases contain an N-terminal domain embedded with multiple non-adjacent RS or SR dipeptides and a C-terminal catalytic core present in all serine/threonine kinases. It remains to be tested whether all of these kinases can interact and phosphorylate SR proteins.

In cellular extracts lacking SR proteins, such as S100 (79, 80) no specific splicing complex could be detected, indicating that SR proteins act early in spliceosome assembly. These observations qualify SR proteins as essential splicing factors. Individual SR proteins can each complement this splicing deficient extract suggesting that SR proteins have redundant functions in splicing (81). However, despite their redundant function in complementing S100, individual SR proteins have distinct specificity and efficiency in splicing different pre-mRNA substrates. For example, HIV tat pre-mRNA splicing in HeLa nuclear extracts depends on SF2/ASF but not SC35 (82). In another case, SC35 but

not SF2/ASF can replace the 9G8 antigen to restore splicing in 9G8 depleted nuclear extracts (83). In addition, different SR proteins display distinct activity in 5' splice site selection (84). For example, Srp40 and Srp55 promote the use of a distal 5' splice site for SV40 large T splicing, whereas SC35 stimulates the use of a proximal splice site for the SV40 small t splicing. A specific SR protein effects 5' splice site selection of a given pre-mRNA by either selectively recruiting U1 snRNP to a specific site or by promoting U1 snRNP occupation of all potential binding sites, where the most proximal site is selected by an additional mechanism (85).

U1 snRNP depleted nuclear extracts are inactive for pre-mRNA splicing. Interestingly, addition of excess SR proteins to these extracts rescues both spliceosome assembly, in particular the association of U2 snRNP with the branch site, and splicing activity, demonstrating that SR proteins can bypass the requirement for U1 snRNP (86). U1 snRNP independent specification of the 5' splice site in the reconstituted reaction occurs after the U2 snRNP-branch site interaction and is mediated by U6 snRNA.

Commitment of a pre-mRNA to splicing requires the initial binding of a specific SR protein to the RNA molecule. However, in contrast to the U2AF65, which binds specifically to the polypyrimidine tract at the 3' splice site, the RNA-binding specificity for most SR proteins remains unknown. In fact, purified SC35 alone was shown to bind nonspecifically to RNA, but in splicing complexes SC35 was found to be associated with a complex bound to the 3' splice site (87), indicating that other spliceosomal components

may cooperate with SC35 to acquire an RNA-binding specificity during spliceosomal assembly.

Nuclear Structure and RNA Splicing

The subnuclear locations of a number of proteins and RNAs important for RNA splicing have been studied by light and electron microscopy. The most striking feature is the concentration of snRNPs U1, U2, U4-U6and U5 in a specific distribution within the nucleus that has been termed the speckled pattern (88). The speckled pattern is composed of 20-50 intensely stained and irregularly shaped regions in the nucleus that are connected in places and set against a diffuse nuclear labeling . The diffuse nuclear staining of snRNPs may represent an excess soluble population, snRNPs in transit to or from nascent transcripts, or it may represent snRNPs in transit to speckles from their assembly sites in the cytoplasm. The use of antibodies to trimethylguanosine cap of some of the snRNAs and of antisense probes to the snRNA (89, 90) revealed that the RNA components of the snRNPs are also organized in the nucleus as a speckled pattern. Several non-snRNP splicing factors such as SC35, SF2/ASF have also been localized to speckled nuclear regions (for a review, see reference (88)).

Electron microscopy with antibodies to snRNPs and SC35 showed that the speckled pattern seen by immunofluorescence corresponds to structures termed interchromatin granules and perichromatin fibrils (91, 92). Interchromatin granules corresponds to the larger intensely stained and irregularly shaped speckles seen by immunofluorescence; they have little to no [^{3}H] uridine labeling in their interior and may

represent the sites of splicing factor storage and/or assembly. Perichromatin fibrils are found on the surface of and between interchromatin granule clusters; they are rapidly labeled with [^{3}H] uridine and are thought to represent nascent pre-mRNA transcripts. It was proposed that snRNPs from interchromatin granule clusters move to the sites of active transcription (perichromatin fibrils) to splice nascent pre-mRNA transcripts (93, 94). Three dimensional reconstruction techniques have shown that the snRNPs are not present in isolated islands, instead portions of speckled pattern are connected to form a latticework that extends between the nucleolar surface and the nuclear lamina-envelope. The organization of splicing factors in a latticework is probably dynamic and reflects the physiological state of the cell. Therefore, at any given time the shape, connections between, and organization of, speckles may vary (95).

To analyze the spatial and temporal organization of pre-mRNA splicing in mammalian cells, several individual RNA species transcribed from endogenous templates have been localized and compared to the localization of splicing factors. The induced expression of c-fos transcripts was found to be closely associated with splicing factors (96). In addition it has been shown that the association of nascent RNA transcripts with splicing factors is intron dependent during transient or stable expression suggesting a close link between transcription and splicing (95).

REFERENCES

(1) Barondes, S.H. (1988) Trends Biochem. Sci. **13**, 480-482.

(2) Hart, G.W., Haltiwanger, T.S., Holt, G.D., and Kelly, W. G. (1989) Ann. Rev. Biochem. **58**, 841-874.

(3) Barondes, S.H., Castronovo, V., Cooper, D.N.W., Cummings, R.D., Drickamer, K., Feizi, T., Gitt, M.A., Hirabayashi, J., Hughes, R.C., Kasai, K., Leffler, H., Liu, F.-T., Lotan, R., Mercurio, A.M., Monsigny, M., Pillai, S., Poirer, F., Raz, A., Rigby, P., Rini, J. M., and Wang, J.L. (1994) Cell **76**, 597-598.

(4) Drickamer, K. and Taylor, M.E. (1993) Ann. Rev. Cell Biol. **9**, 237-264.

(5) Hirabayashi, J. and Kasai, K. (1993) Glycobiology **3**, 297-304.

(6) Barondes, S.H. (1984) Science **223**, 1259-1264.

(7) Gitt, M.A., Massa, S.M., Leffler, H., and Barondes, S.H. (1992) J. Biol. Chem. **267**, 10601-10606.

(8) Gitt, M.A., Wiser, M.F., Leffler, H., Herrman, J., Xia, Y.R., Massa, S.M., Cooper, D.N.W., Lusias, A.J., and Barondes, S.H. (1995) J. Biol. Chem. **270**, 5032-5038.

(9) Madsen, P., Rasmussen, H.H., Flint, T., Gromov, P., Kruse, T.A. Honore, B., Vorum, H., and Celis, J.E. (1995) J. Biol. Chem. **270**, 5823-5829.

(10) Roff, C.F. and Wang, J.L. (1983) J. Biol. Chem. **258**, 10657-10663.

(11) Jia, S. and Wang, J.L. (1988) J. Biol. Chem. **263**, 6009-6011.

(12) Oda, Y., Herrman, J., and Gitt, A.M., Turck, C.W., Burlingame, A.M., Barondes, S.H., and Leffler, H. (1993) J. Biol. Chem. **268**, 5929-5939.

(13) Hadari, Y.I., Paz, K., Dekel, R., Mestrovie, T., Accili, D., and Zick, Y. (1995) J. Biol. Chem. **270**, 3447-3453.

(14) Hirabayashi, J., Satoh, M., Ohyama, Y., and Kasai, K. (1992) J. Biochem. **111**, 553- 555.

(15) Hirabayashi, J., Satoh, M., and Kasai, K. (1992) J. Biol. Chem. **267**, 15485-15490.

(16) Hirabayashi, J., Ubukata, T. and Kasai, K. (1996) J. Biol. Chem. **271**, 2497-2505.

(17) Klion, A. D. and Donelson, J.E. (1994) Mol. Biochem. Parasitol. **65**, 305-315.

(18) Pfeifer, K., Haasermann, M., Gamulin, V., Bretting, H., Fahrenholz, F., and Muller, W.E.G. (1993) Glycobiology **3**, 179-184.

(19) Gitt, M.A. and Barondes, S.H. (1991) Biochemistry **30**, 82-89.

(20) Gritzmacher, C.A., Mehl, V.S., and Liu, F.T. (1992) Biochemistry **31**, 9533-9538.

(21) Lobsanov, Y.D., Gitt, M.A., Leffler, H., Barondes, S.H., and Rini, J.M. (1993) J. Biol. Chem. **268**, 27034-27038.

(22) Liao, D.I., Kapadia, G., Ahmed, H., Vasta, G.R., and Herzberg, O. (1994) Proc. Natl. Acad. Sci. USA **91**, 1428-1432.

(23) Bourne, Y., Bolgiano, B., Liao, D.-I., Strecker, G., Cantan, P., Herzberg, O., Feizi, T., and Cambillau, C. (1994) Struct. Biol. **1**, 863-870.

(24) Leffler, H. and Barondes, S.H. (1986) J. Biol. Chem. **261**, 10119-10126.

(25) Abbott, W.M., Hounsell, E.F., and Fiezi, T. (1988) Biochem. J. **252**, 283-287.

(26) Lee, R.T. Ichikawa, Y., Allen, H.J., and Lee, Y.C. (1990) J. Biol. Chem. **265**, 7864-7871.

(27) Sato, S. and Hughes, R.C. (1992) J. Biol. Chem. **267**, 6983-6990.

(28) Knibbs, R.N., Agrwal, N., Wang, J.L., and Goldstein, I.J. (1993) J. Biol. Chem. **268**, 14940-14947.

(29) Hirabayashi, J. and Kasai, K. (1994) Glycoconjugate J. **11**, 437-442.

(30) Abbot, W.M., and Feizi, T. (1991) J. Biol. Chem. **266**, 5552-5557.

(31) Anderson, R.L. and Wang, J.L. (1992) Trends Glycosci. Glycotechnol. **4**, 43-52.

(32) Raz, A., Pazerini, G., and Carmi, P. (1989) Cancer Res. **49**, 3489-3493.

(33) Agrwal, N., Sun, Q, Wang, S.Y., and Wang, J.L. (1993) J. Biol. Chem. **268**, 14932-14939.

(34) Wang, S.Y., Voss, P.G., Patterson, R.J., and Wang, J.L. (1995) Antibody, Immunoconjugate, and Radiopharmaceuticals **8**, 311-324.

(35) Hubert, M., Wang, S.Y., Wang, J.L., Seve, A.P., and Hubert, J. (1995) Exp. Cell Res. **220**, 397-406.

(36) Moutsatsos, I.K., Davis, J.M., and Wang, J.L. (1986) J. Cell Biol. **102**, 477-483.

(37) Cowles, E.A., Agrwal, N., Anderson, R.L., and Wang, J.L. (1990) **265**, 17706-17712.

(38) Huflejt, M.E., Turck, C.W., Lindstedt, R., Barondes, S.H., and Leffler, H. (1993) J. Biol. Chem. **268**, 26712-26718.

(39) Lotz, M.M., Andrews, C.W., Korzelius, C.A., Lee, E.C., Steele, G.D., Clarke, A., and Mercurio, A.M. (1993) Proc. Natl. Acad. Sci. USA **90**, 3466-3470.

(40) Cherayil, B.J., Weiner, S.J., and Pillai, S. (1989) J. Exp. Med. **170**, 1959-1972.

(41) Ho, M.K. and Springer, T.A. (1982) J. Immunol. **128**, 1221-1228.

(42) Lindstedt, R., Apodaca, G., Barondes, S.H., Mostove, K., and Leffler, H. (1993) J. Biol. Chem. **268**, 11750-11757.

(43) Sato, S., Burdett, J., and Hughes, R.C. (1993) Exp. Cell Res. **207**, 8-18.

(44) Hynes, M.A., Gitt, M.A., Barondes, S.H., Jessell, T.M., and Buck, L.B. (1990) J. Neurosci. **10**, 1004-1013.

(45) Raz, A. and Lotan, R. (1987) Cancer Mestastasis Rev. **6**, 433-452.

(46) Butel, J.S. and Jarvis, D.L. (1986) Biochim. Biophys. Acta **865**, 171-195.

(47) Mason, I.J. (1994) Cell **78**, 547-552.

(48) Siders, W.M., Klimovitz, J.C., and Mizel, S.B. (1993) J. Biol. Chem. **268**, 22170-22174.

(49) Kojima, K., Ogawa, H., Seno, N., Yamamoto, K., Irimura, T., Osawa, T., and Matsumoto, I. (1992) J. Biol. Chem. **267**, 20536-20539.

(50) Scully, R., Ganesan, S., Brown, M., DeCaprio, J.A., Cannistra, S.A., Feunteun, J., Schnitt, S., and Livingston, D.M. (1996) Science **272**, 123-125.

(51) Chen, Y., Chen, C.-F., Riley, D.J., Allred, D.C., Chen, P.L., VonHoff, D., Osborne, C.K., and Lee, W.H. (1995) Science **270**, 789-791.

(52) Jensen, R.A., Thompson, M.E., Jetton, T.L., Szabo, C.I., van der Meer, R., Helon, B., Tronick, S.R., Page, D.L., King, M.C., and Holt, J.T. (1996) Nature Genetics **12**, 303-308.

(53) Lamond, A.I. (1993) BioEssays **15**, 595-603.

(54) Sharp, P.A. (1994) Cell **77**, 805-815.

(55) Luhrmann, R., Kastner, B. and Bach, M. (1990) Biochim. Biophys. Acta **1087**, 265-292.

(56) Lerner, M.R. and Steitz, J.A. (1979) Proc. Natl. Acad. Sci. USA **76**, 5495-5499.

(57) Konarska, M.M. and Sharp, P.A. (1987). Cell **49**, 763-774.

(58) Barbino, M.L., Blencowe, B.J., Ryder, U., Sproat, B.S., and Lamond, A.I. (1990) Cell **63**, 293-302.

(59) Kramer, A. and Utans, U. (1991) EMBO J. **10**, 1503-1509.

(60) Ruskin, B., Zamore, P.D. and Green, M.R. (1988) Cell **52**, 207-219.

(61) Zamore, P.D. and Green, M.R. (1989) Proc. Natl. Acad. Sci. USA **86**, 9243-9247.

(62) Hashimoto, C. and Steitz, J.A. (1984) Nuc. Acids. Res. **12**, 3283-3293.

(63) Ge, H. and Manley, J.L. (1990) Cell **62**, 25-34.

(64) Krainer, A., Conway, G.C. and Kozak, D. (1990) Cell **62**, 35-42.

(65) Fu, X.D.,and Maniatis, T. (1990) Nature **343**, 437-441.

(66) Krainer, A., Mayeda, A., Kozak, D., and Binns, G. (1991) Cell **66**, 383-394.

(67) Fu, X.D. (1995) RNA **1**, 663-680.

(68) Zahler, A.M., Lane, W.S., Stolk, J.A., and Roth, M.B. (1992) Genes Dev. **6**, 837-847.

(69) Chou, T.B., Zachar, Z. and Bingham, P.M. (1987) EMBO J. **6**, 4095-4104.

(70) Boggs, R.T., Gregor, P., Idriss, S., Belote, J.M., and McKeown, M. (1987) Cell **50**, 739-747.

(71) Amrein, H., Gorman, M. and Nothiger, R. (1988) Cell **55**, 1025-1035.

(72) Baker, B.S. (1989) Nature **314**, 521-524.

(73) Zamore, P.D., Patton, J.G., and Green, M.R. (1992) Nature **355**, 609-614.

(74) Zhang, M., Zamore, P.D., Carmo-Fonseca, M., Lamond, A.I., and Green, M.R. (1992) Proc. Natl. Acad. Sci. USA **89**, 8769-8773.

(75) Wu, J.Y. and Maniatis, T. (1993) Cell **75**, 1061-1070.

(76) Stakins, D. and Reed, R. (1994) Mol. Cell Biol. **14**, 7670-7682.

(77) Ben-David, Y., Letwin, K., Tannock, L., Bernstein, A., Pawson, T. (1991) EMBO J. **10**, 1273-1278.

(78) Howell, B.W., Afar, D.E., Lew, J., Douville, E.M., Icely, P.L., Gray, D.A., and Bell, J. (1991) Mol. Cell Biol. **11**, 568-572.

(79) Krainer, A., Conway, G.C. and Kozak, D. (1990) Genes Dev. **4**, 1158-1171.

(80) Zahler, A.M., Neugebauer, K., Stolk, J.A., Roth, M.B. (1993) Mol. Cell Biol. **13**, 4023-4028.

(81) Mayeda, A., Zahler, A.M., Krainer, A, Green, M.R. (1992) Proc. Natl. Acad. Sci. USA **89**, 1031-1300.

(82) Fu, X.D. (1993) Nature, **365**, 82-85.

(83) Cavaloc, Y., Popielarz, M., Fuchs, J.P., Gattoni, R., and Stevenin, J. (1994) EMBO J. **13**, 2639-2649.

(84) Zahler, A.M., Neugebauer, K., Stolk, J.A., Roth, M.B. (1993) Science **260**, 219-222.

(85) Eperon, I.C., Ireland, D.C., Smith, R.A., Mayeda, A., and Krainer, A.R. (1993) EMBO J. **12**, 360703617.

(86) Tarn, W.Y. and Steitz, J.A. (1994) Genes Dev. **8**, 2704-2717.

(87) Fu, X.D.,and Maniatis, T. (1992) Proc. Natl. Acad. Sci. USA **89**, 1725-1729.

(88) Spector, D.L. (1993) Annu. Rev. Cell Biol **9**, 265-315.

(89) Carmo-Fonseca, M., Pepperkok, R., Carvalho, M.T., and Lamond, A.I. (1992) J. Cell Biol. **117**, 1-14.

(90) Huang, S. and Spector, D. (1992) Proc. Natl. Acad. Sci. USA **89**, 305-308.

(91) Spector, D., Fu, X.D., and Maniatis, T. (1991) EMBO J. **10**, 3467-3481.

(92) Fakan, S., Leser, G. and Martin, T.E. (1984) J. Cell Biol. **98**, 358-363.

(93) Jimenz-Garcia, L.F. and Spector, D.L. (1993) Cell **73**, 47-57.

(94) Pombo, A., Ferreira, E., Bridge, E., and Carmo-Fonseca, M. (1994) EMBO J. **13**, 5075-5085.

(95) Huang, S. and Spector, D.L. (1996) J. Cell Biol. **133**, 719-732.

(96) Huang, S. and Spector, D.L. (1991) Genes Dev. **5**, 2288-2302.

CHAPTER II

Galectin-1 in the Cell Nucleus. I. Evidence for a Role in Pre-mRNA Splicing

Anandita Vyakarnam[1], Sue F. Dagher[2,3], John L. Wang[1], and Ronald J. Patterson[4]

Departments of [1]Biochemistry and [4]Microbiology and [2]Genetics Program,
Michigan State University, East Lansing, MI 48824

[3]Present address: Department of Biochemistry, Dartmouth Medical School,
Hanover, NH 03755

Acknowledgments: We thank Drs. Sam Barondes and Hakon Leffler for their generous gifts of recombinant rat galectin-1 and rabbit antiserum directed against recombinant rat galectin-1. This work was supported by grants MCB 91-22363 from the National Science Foundation, GM-38740 from the National Institutes of Health, No. 99 from the Mizutani Foundation for Glycoscience, and an Interdisciplinary Research Award from the Cancer Center at Michigan State University.

ABSTRACT

Galectins are a family of β-galactoside-binding proteins that contain characteristic amino acid sequences in the carbohydrate recognition domain of the polypeptide. The polypeptide of galectin-1 contains a single domain, the carbohydrate recognition domain. The polypeptide of galectin-3 contains two domains, a carbohydrate recognition domain fused onto a proline-, and glycine-rich domain. In previous studies, we had shown that galectin-3 is a required factor in the splicing of nuclear pre-mRNA, assayed in a cell-free system. We now document that: (a)nuclear extracts derived from HeLa cells contain both galectins-1 and -3; (b)depletion of both galectins from the nuclear extract, either by lactose affinity adsorption or by double antibody adsorption, resulted in the concomitant loss of splicing activity; (c)depletion of either galectin-1 or galectin-3, by specific antibody adsorption, failed to remove all of the splicing activity and the residual activity was still saccharide inhibitable; (d)either galectin-1 or galectin-3 alone was sufficient to reconstitute, at least partially, the splicing activity of nuclear extracts depleted of both galectins; and (e)although the carbohydrate recognition domain of galectin-3 (or galectin-1) is sufficient to restore splicing activity to galectin-depleted nuclear extract, the concentration required for reconstitution was far greater than that of the full-length galectin-3 polypeptide.

INTRODUCTION

Galectins are a family of widely distributed proteins that: (a)bind to β-galactosides; and (b)contain characteristic amino acid sequences in the carbohydrate recognition domain (CRD) of the polypeptide. At present, eight mammalian galectins have been reported and classified into three subgroups, according to the content and organization of the domains in their respective polypeptides (for reviews, see references 1 and 2). The Prototype subgroup consists of polypeptides (~ 14 kD) with a single domain, the CRD. Galectins-1, -2, -5, and -7 are members of this subgroup. Another subgroup is the Tandem Repeat type, which has three members: galectins-4, -6, and -8. These galectins have two domains, each a CRD, connected by a linker region. Finally, the Chimera subgroup is, at present, represented by a single member, galectin-3. Its polypeptide contains two domains, a CRD fused onto a Pro-, Gly-rich domain.

In previous studies, we had reported the localization of galectin-3 to the cell nucleus, in the form of a ribonucleoprotein (RNP) complex (3, 4). We had also identified it as one of the many proteins required for the splicing of pre-mRNA, assayed in a cell-free system. This conclusion was based on several key findings (5): (a)nuclear extracts (NE) derived from HeLa cells, capable of carrying out splicing, contain galectin-3; (b)NE depleted of galectin-3 by affinity adsorption on lactose (Lac)-agarose become deficient in splicing; (c)the activity of the Lac-agarose depleted extract could be reconstituted by the addition of purified recombinant galectin-3; and (d)saccharides that bind galectin-3 with

high affinity inhibit the cell-free splicing reaction. These results strongly suggested that the lectin is a required factor in cell-free splicing of pre-mRNA.

When we attempted to deplete the splicing activity using a galectin-3 specific monoclonal antibody, the anti-Mac-2 (anti-M2) antibody (6, 7), we were surprised to find that there was some, but not complete, loss of splicing activity in the galectin-3-depleted extract. We thus performed a series of experiments to resolve this apparent dilemma. In the present communication, we report that, in addition to galectin-3, galectin-1 is also a component of NE and it also plays a role in the splicing of pre-mRNA. The accompanying chapter documents the nuclear localization of galectin-1, using laser confocal scanning microscopy.

MATERIALS AND METHODS

Preparation of HeLa Nuclear Extracts

HeLa S3 cells were obtained from American Type Culture Collection (CCL 2.2) and grown in suspension culture in Minimum Essential Medium containing 10% defined / supplemented bovine calf serum (HyClone), 100 U/ml penicillin and 100 U/ml streptomycin at 37°C in humidified atmosphere of 5% CO_2. Nuclear extract (NE) was prepared in buffer C [20 mM Hepes (pH 7.9), 25% v/v glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM phenyl methylsulfonyl fluoride (PMSF), 0.5 mM dithiothreitol (DTT)] as described (8). NE was frozen as aliquots in dry ice/ethanol bath and stored at -80°C. Typically NEs with protein concentrations of 10-13 mg/ml were prepared. Protein concentrations were determined by the method of Bradford using bovine serum albumin as the protein standard (9).

Antibodies and Affinity Columns

Anti-M2 is a rat monoclonal antibody directed against the Mac-2 antigen (6) which has been shown to be galectin-3 (7). The antibody was purified from serum-free cell culture supernatant derived from hybridoma line M3/38.1.2.8.HL.2 obtained from ATCC (TIB 166). Anti-transferrin receptor (TR) antibody is an isotype matched rat monoclonal antibody ($IgG_{2a}\kappa$) that was used as a control for anti-M2. The hybridoma (R17 217.1.3) producing the anti-(TR) antibody was obtained from ATCC (TIB219). The polyclonal rabbit antiserum against recombinant rat galactin-1 (anti-G1) was a gift from

Dr. Hakon Leffler (UCSF, CA). Human autoimmune serum reactive with the Sm antigens of snRNPs (ENA anti-Sm) was purchased from The Binding Site (San Diego, CA).

For saccharide affinity column, α-lactose insolubilized on 6% beaded agarose (Lac-agarose) was purchased from Sigma. Cellobiose-agarose was used as a control for NE depletion. For immunodepletions, antibodies were immobilized on pre-swollen protein G-Sepharose beads (Sigma). Beads were washed with 20 mM Hepes (pH 7.9), 0.5 M NaCl. 150 µl of washed beads were mixed with 150 µl serum or 100 µg monoclonal antibody. The mixture was adjusted to 20 mM Hepes (pH 7.9), 0.5 M NaCl and incubated for 1 h at room temperature with continual rocking. The beads were washed with 1 ml of 0.2 M sodium borate, pH 9.0 and resuspended in the same buffer. Dimethylpimelimidate (Pierce) was added to final concentration of 20 mM to covalently crosslink antibodies to protein G-Sepharose. After 1 hour incubation at room temperature, the beads were washed and incubated in 1ml of 0.2 M ethanolamine (pH 8.0) for one hour. The coupled antibody-Sepharose mixture was washed five times with 1 ml of 20 mM Hepes (pH 7.9), 0.5 M NaCl and used for immunodepletions.

Depletion and Reconstitution of Nuclear Extract

NEs were depleted of galectins by adsorption on either a saccharide affinity column or an antibody affinity column. For saccharide affinity depletions, 150 µl of Lac-agarose beads were washed with 20 mM Hepes (pH 7.9), 0.5 M NaCl and 30 µl NE were added and the mixture incubated on ice for 20 min in disposable spin columns (Millipore). For imunodepletions, 30 µl of NE were incubated with either 60 µl of anti-

Mac-2-Sepharose or 90 µl of anti-galectin-1-Sepharose beads for single antibody depletions and with a mixture of both for double antibody depletions. After incubation on ice for 20 min with appropriate affinity matrices, the unbound fraction was removed. The beads were washed with 12 µl of buffer 1 [60% Dignam buffer D adjusted to 0.42 M NaCl, buffer D: 20 mM Hepes (pH 7.9), 20% v/v glycerol, 0.1M KCl, 0.2 mM EDTA, 0.5 mM PMSF and 0.5 mM DTT] and this wash was added to the unbound fraction. The beads were washed three times with 1 ml of buffer 1 and the bound material eluted by boiling in 100 µl of Laemmli sample buffer (10).

Aliquots of nondepleted NE and unbound fractions of saccharide depletions or immunodepletions were dialyzed in a microdialyzer against 60% buffer D for 40 min at 4°C. The final protein concentration of the depleted and dialyzed extracts was 6-8 mg/ml. The dialyzed fractions were then assayed for splicing activity. In reconstitution experiments, recombinant proteins were added to the unbound fractions prior to dialysis. In experiments directly comparing the efficiency of various recombinant proteins in reconstituting splicing activity, aliquots taken from the same unbound fraction were reconstituted with different amounts of proteins. In experiments testing the effect of exogenously added carbohydrates, the extracts were incubated with thiodigalactoside for 5 min at room temperature after dialysis but prior to the addition of pre-mRNA substrate.

***In Vitro* Splicing Assay**

MINX pre-mRNA, used as a substrate in the splicing assay, was transcribed from the plasmid provided by Susan Berget (Baylor College of Medicine, Houston) using SP6

polymerase (Gibco). The MINX pre-mRNA was labeled with ^{32}P-GTP and the monomethyl cap was added during transcription.

Splicing reaction mixtures (10 µl) contained dialyzed NE (4 µl) or unbound fraction (8 µl), ^{32}P-MINX pre-mRNA, 2.5 mM $MgCl_2$, 1.5 mM ATP, 20 mM creatine phosphate, 0.5 mM DTT and 20 units of RNasin (Promega). Splicing reactions were carried out at 30°C for 45 min. Proteinase K/SDS was added to a final concentration of 4 mg/ml proteinase K and 0.1% SDS and the sample was incubated at 37°C for 15 min. Each splicing sample was diluted to 100 µl with 125 mM Tris (pH 8.0), 1 mM EDTA, 0.3 M sodium acetate. RNA was extracted with 200 µl of phenol/chloroform (50:50 v/v) followed by 200 µl of chloroform. RNAs were precipitated with 400 µl of ethanol at -80°C. The extracted RNAs were subjected to electrophoresis through 13% polyacrylamide (bisacrylamide/acrylamide, 1.9:50 wt/wt), 8.3 M urea gels followed by autoradiography.

The intensities of the bands on the gels were quantitated by direct β-particle counting using an AMBIS Systems system. The percent product formation was calculated by dividing the radioactivity present in the final product (ligated exon 1-exon 2) by the total radioactivity present in the pre-mRNA substrate (exon 1-intron-exon 2), the splicing intermediates (exon 2-lariat, exon 1) and the product at the end of the incubation.

Immunoblot Analysis

For immunoblot analysis, protein samples were resolved on 12.5% SDS-PAGE as described by Laemmli (10) and electrophoretically transferred to Immobilon-P membrane (Millipore) in a transfer buffer containing 25 mM Tris, 193 mM glycine and 20% methanol. The membrane was blocked for several hours with 10% non-fat dry milk in T-TBS (10 mM Tris pH 7.5, 0.5 M NaCl, 0.05% Tween-20). After brief washes with T-TBS twice, the membrane was incubated with primary antibody diluted in 1% non-fat dry milk/T-TBS for two hours at room temperature followed by five washes in T-TBS for 15 min each. The membranes were incubated with secondary antibodies conjugated to horseradish peroxidase for 30 min and washed in T-TBS extensively. Proteins were visualized using Enhanced Chemiluminescence detection system (Amersham). To probe a membrane sequentially with two different antibodies, the manufacturer's instructions were followed.

RESULTS

Saccharide Affinity Depletion versus Immunodepletion

We have optimized conditions for the splicing of the MINX pre-mRNA substrate by NE derived from HeLa cells. In the experiments to be reported below, our typical assay shows that approximately 20-30% of the input substrate is converted to product (Fig. 1A, lane 1). As reported in our previous study (5), splicing activity was depleted from NE when the latter was subjected to adsorption, under conditions of high ionic strength (0.42 M NaCl), on Lac affinity beads. The unbound fraction of the Lac column yielded less than 5% product formation (Fig. 1A, lane 3). Parallel control experiments, in which NE was subjected to adsorption on beads containing cellobiose, which does not bind to galectins, failed to yield the same result (5). Splicing activity in the extract depleted on the Lac matrix was restored by the addition of recombinant galectin-3 (Fig. 1A, lane 4). In both the original NE, as well as in the reconstituted fraction, the splicing activity was sensitive to inhibition by thiodigalactoside, a saccharide that binds to galectin-3 with high affinity (Fig. 1A, lanes 2 and 5). These results formed the basis for our suggestion that galectin-3 is a required factor in the splicing of pre-mRNA.

When we attempted to deplete the splicing activity using anti-M2, a monoclonal antibody specific for galectin-3 (6, 7), there was some, but not complete, loss of splicing activity in the galectin-3-depleted extract. An isotype-matched monoclonal, directed against the transferrin receptor (TR), was used as a control for the anti-M2 antibody. The

Figure 1A: Comparison of the splicing activity of NE, NE depleted by Lac affinity adsorption, and depleted NE reconstituted with recombinant galectin-3.
Lane 1: control splicing reaction of nondepleted NE in the presence of ATP.
Lane 2: components of lane 1 plus 50 mM thiodigalactoside (TDG).
Lane 3: the unbound (UB) fraction when NE is subjected to Lac affinity adsorption.
Lane 4: components of lane 3 plus 12 μM recombinant galectin-3 (rG3).
Lane 5: components of lane 4 plus 150 mM TDG.
Products of the splicing reactions were analyzed by electrophoresis through a 13% polyacrylamide-urea gel and autoradiography. The positions of migration of the pre-mRNA substrate, the splicing intermediates (exon 1 and lariat-exon 2) and mature RNA product are indicated on the right.

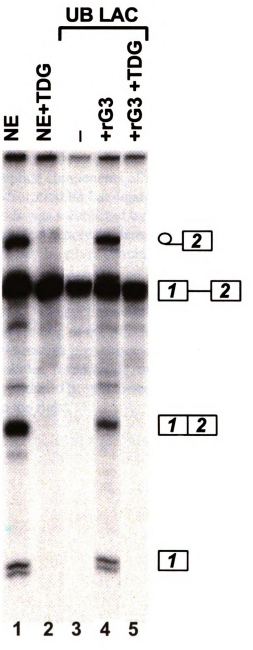
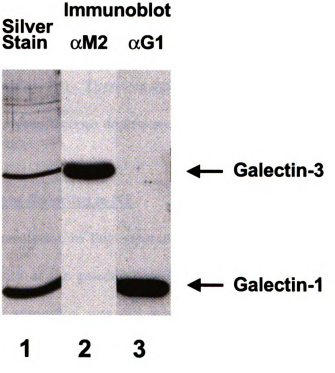

UB LAC
NE
NE+TDG
–
+rG3
+rG3 +TDG
2
1
2
1 2
1
1 2 3 4 5

Figure 1B: Analysis of the polypeptide components of the bound fraction when NE is subjected to affinity adsorption on Lac-agarose columns. The material bound to the Lac-agarose beads was solubilized by SDS-PAGE sample buffer and was electrophoresed on 12.5% polyacrylamide gels. The resolved polypeptides were detected by silver staining or by immunoblotting with specific antibodies.
Lane 1: silver stain. Lane 2: immunoblot with rat monoclonal antibody against galectin-3 (αM2). Lane 3: immunblot with rabbit antibody against rat galectin-1 (αG1). The binding of the primary antibodies in lanes 2 and 3 were revealed with horseradish peroxidase-conjugated goat anti-rat and goat anti-rabbit immunoglobulin, respectively, using the Enhanced Chemiluminescence system. The positions of migration of human galectins-1 and -3 are shown on the right.


Immunoblot
Silver Stain
αM2
αG1
Galectin-3
Galectin-1
1
2
3

unbound fraction of the anti-M2 column retained much of the splicing activity, compared to the corresponding unbound fraction of the anti-TR column (Fig. 2A, lane 4 (15% product) versus lane 1 (21% product)). As a positive control for immunodepletion, the snRNPs were removed from NE by an autoimmune serum reactive against the Sm antigens on the core polypeptides B and D of the snRNPs (designated as anti-Sm) (11). As expected, the Sm-depleted extract (unbound fraction of anti-Sm column) was completely deficient in splicing activity (Fig. 2A, lane 7). Western blot analyses indicated that anti-M2 removed >95% of galectin-3 (galectin-3 band in the bound fraction and undetectable in the unbound fraction), while neither anti-TR nor anti-Sm removed the galectin-3 polypeptide (Fig. 2B). These results indicate that removal of galectin-3 from NE did not result in the simultaneous depletion of splicing activity.

Number and Identity of Galectins in NE

A hint at the resolution of this apparent dilemma was derived from the finding that the splicing activity of the galectin-3-depleted extract, derived from the unbound fraction of the anti-M2 adsorption, was still sensitive to TDG inhibition (Fig. 2A, lane 6). Sucrose, a control disaccharide that does not bind to any of the galectins, failed to yield the same effect (Fig. 2A, lane 5). These results suggest that other Lac-binding proteins, besides galectin-3, may be present in the splicing extract and they are depleted, along with galectin-3, by Lac adsorption but not by anti-M2.

This prompted us to analyze the protein components in the bound fraction of the original Lac affinity adsorption. Using silver staining, we detected two bands whose

Figure 2A: Comparison of the splicing activity of NE after adsorption on protein G-Sepharose beads conjugated with various antibodies. Lanes 1-3: the unbound fraction of rat monoclonal anti-transferrin receptor (αTR) adsorption. Lanes 4-6:the unbound fraction of rat monoclonal anti-galectin-3 (αM2) adsorption. Lanes 7-9: the unbound fraction after adsorption by human autoimmune serum reactive against the Sm epitopes of the snRNP polypeptides (αSm). Splicing reactions with the immunodepleted extracts were assayed in the absence of saccharide (lanes 1, 4, 7), in the presence of 50 mM sucrose (lanes 2, 5, and 8),and in the presence of 50 mM thiodigalactoside (TDG) (lanes 3, 6, and 9). The positions of migration of the pre-mRNA substrate, splicing intermediates, and mature RNA product are highlighted on the right.

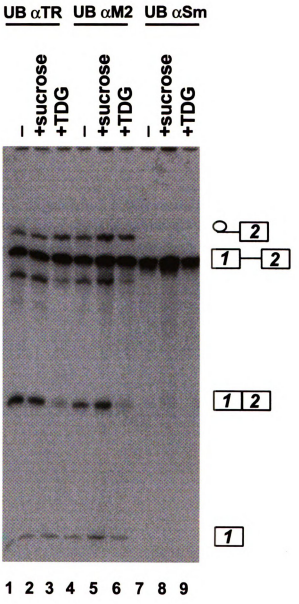
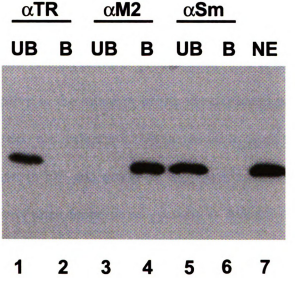

UB αTR
UB αM2
UB αSm
–
+sucrose
+TDG
–
+sucrose
+TDG
–
+sucrose
+TDG
2
1 2
1 2
1
1 2 3 4 5 6 7 8 9

Figure 2B: Comparison of the levels of galectin-3 in NE versus the unbound (UB) and B (bound) fractions of the various immunodepletions. Samples were subjected to SDS-PAGE on 12% polyacrylamide gels and immunoblotting with rat monoclonal anti-galectin-3.


αTR
αM2
αSm
UB B UB B UB B NE
1 2 3 4 5 6 7

mobilities corresponded to polypeptides with M_r ~30,000 and ~15,000 (Fig. 1B). The identity of the ~30 kD polypeptide was ascertained to be galectin-3 by immunoblotting with the anti-M2 antibody. On the basis of the apparent molecular weight of the second silver-stained band, we suspected that it might correspond to galectin-1. This notion was confirmed by immunoblotting with a rabbit antiserum raised against recombinant rat galectin-1 (anti-G1) (Fig. 1B). Because this antibody does cross-react weakly with galectin-2, we carried out direct comparisons between galectins-1 and -2 in terms of: (a)the intensities of their immunoblotted bands when probed with the anti-G1 antibody; and (b)the intensities of the silver-stained bands when known amounts of the respective recombinant proteins were electrophoresed in SDS gels. The level of immunoreactivity in the Western blot, relative to the intensity of the silver-stained gel (Fig. 1B), indicated that the ~15 kD polypeptide was galectin-1. These results suggest that galectin-1 was the other Lac-binding protein in NE and could be responsible for the residual saccharide-inhibitable splicing activity after depletion of galectin-3. Indeed, we were able to detect, by Western blotting analysis, galectin-1 in NE, as well as in the unbound fraction of the anti-M2 column (Fig. 3A, lanes 1 and 2).

Double Antibody Depletion of Galectins-1 and -3

The above results implicate that when NE is subjected to Lac affinity adsorption, both galectins-1 and -3 are bound on the Lac matrix, with concomitant depletion of splicing activity, whereas an antibody directed against any single galectin would fail to completely deplete the activity. Thus, a comparison was made of subjecting NE to incubation with beads derivatized with: (a)anti-M2; (b)anti-G1; (c)anti-M2 plus anti-G1;

Figure 3A: Comparison of the levels of galectins-1 and -3 in NE versus the unbound (UB) and bound (B) fractions of immunodepletion using antibodies directed against galectin-1 or galectin-3. Lane 1: NE. Lanes 2 and 3: rat monoclonal antibody against galectin-3 (αM2). Lanes 4 and 5: rabbit antibodies against rat galectin-1 (αG1). Lanes 6 and 7: αM2 and αG1 used in combination. Lanes 8 and 9: control antibodies, rat antibody against the transferrin receptor (αTR) and rabbit preimmune serum (PI), used in combination. Samples were subjected to SDS-PAGE on 12.5% polyacrylamide gels and immunoblotting with αM2 or αG1. Separate panels are shown for the αG1 and αM2 immunoblots, corresponding to the regions of migration for galectin-1 and galectin-3, respectively.

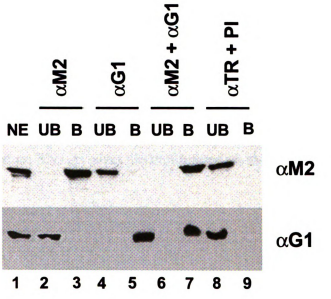
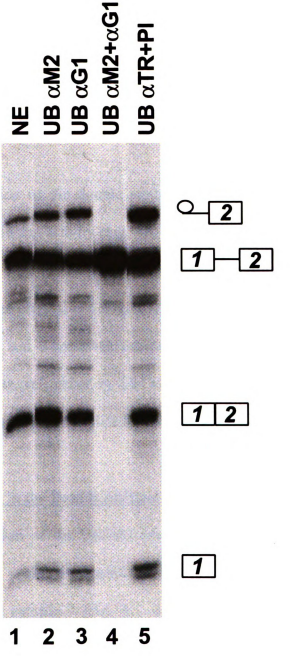

αM2
αG1
αM2 + αG1
αTR + PI
NE UB B UB B UB B UB B
αM2
αG1
1 2 3 4 5 6 7 8 9

Figure 3B: Comparison of the splicing activity of NE versus the unbound (UB) fractions of immunodepletion using antibodies directed against galectin-1 or galectin-3. Lane 1: NE. Lane 2: unbound fraction of the αM2 depletion. Lane 3: unbound fraction of the αG1 depletion. Lane 4: unbound fraction of immunodepletion using both antibodies, αM2 and αG1. Lane 5: unbound fraction of immunodepletion using controls antibodies, αTR and PI. The positions of migration of the pre-mRNA substrate, splicing intermediates, and mature RNA product are highlighted on the right.


NE
UB αM2
UB αG1
UB αM2+αG1
UB αTR+PI
2
1 2
1 2
1
1 2 3 4 5

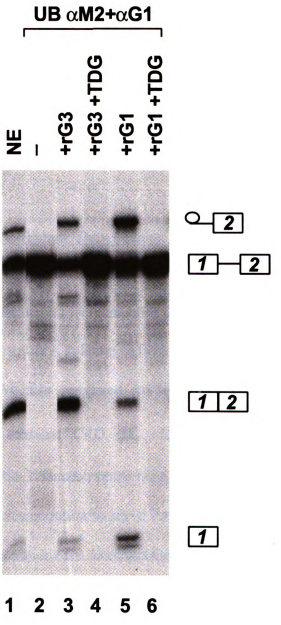
and (d) anti-TR plus rabbit preimmune serum. Western blotting analyses yielded the expected presence of galectins-1 and -3 in the various unbound and bound fractions (Fig. 3A). For example, galectin-3 was quantitatively adsorbed in the bound fractions of matrices containing anti-M2. Similarly, no galectin-1 could be detected in the unbound fractions of matrices containing anti-galectin-1. Finally, both galectins-1 and -3 were found exclusively in the unbound fractions of the beads containing the control antibodies, anti-TR plus preimmune rabbit serum.

The unbound fraction of the double antibody column, anti-M2 plus anti-G1, showed complete loss of splicing activity (Fig. 3B, lane 4). This should be compared to the level of product formation in the unbound fraction of the control adsorption (Fig. 3B, lane 5), as well as that of the original NE (Fig. 3B, lane 1). Together, all of these results suggest that both galectin-1 and -3 contribute independently to the splicing activity of NE and that their complete removal, either by Lac affinity chromatography or by double antibody adsorption, was necessary to deplete the splicing activity.

Reconstitution of Splicing in the Double Antibody-depleted Extracts

The unbound fraction of the anti-M2 plus anti-G1 double antibody column, devoid of splicing activity, could be reconstituted by the addition of either recombinant galectin-1, or recombinant galectin-3 (Fig. 4). Recombinant galectins-1 and -3 were derived from *E. coli* expression systems for the rat and mouse cDNAs, respectively (12, 13). Addition of either recombinant protein to the double antibody-depleted extract restored the splicing activity (Fig. 4, lanes 3 and 5). Moreover, the splicing activity of the

Figure 4: Comparison of the splicing activity of NE, NE immunodepleted by antibodies against galectins-1 and -3, and depleted NE reconstituted with recombinant galectin-1 or recombinant galectin-3.
Lane 1: control splicing reaction of nondepleted NE.
Lane 2: the unbound (UB) fraction when NE is subjected to immunodepletion with rat monoclonal anti-galectin-3 (αM2) and rabbit antibodies against rat galectin-1 (αG1).
Lane 3: components of lane 2 plus 12 μM recombinant galectin-3 (rG3).
Lane 4: components of lane 3 plus 150 mM thiodigalactosdie (TDG).
Lane 5: components of lane 2 plus 26 μM recombinant galectin-1 (rG1).
Lane 6: components of lane 5 plus 150 mM TDG.
The positions of migration of the pre-mRNA substrate, the splicing intermediates, and the mature RNA product are highlighted on the right.


UB αM2+αG1
NE
–
+rG3
+rG3 +TDG
+rG1
+rG1 +TDG
2
1
2
1
2
1
1 2 3 4 5 6

double antibody-depleted extract, reconstituted with either galectin, was again sensitive to TDG inhibiton (Fig. 4, lanes 4 and 6). The effects of galectin-1 or galectin-3 addition were concentration dependent. Compared on an equal molar concentration basis, recombinant galectin-3 was more potent than the corresponding galectin-1 in reconstitution of splicing activity (see below). Finally, neither recombinant galectin-1 nor recombinant galectin-3 had an effect on the splicing assay when added to a nondepleted NE. All of these results indicate that galectin-1 and galectin-3 may serve redundant functions in the splicing activities of the NE.

Comparison of the Concentration Dependence of Galectin-1 and Galectin-3 in Reconstituting Splicing Activity

In this series of experiments, we compared the concentrations of galectin-1 and galectin-3 required to reconstitute a splicing deficient NE, depleted of the galectins by Lac affinity adsorption. We also compared the intact polypeptide of galectin-3 (M_r ~30,000) with its COOH-terminal CRD (M_r ~15,000) (13) in terms of splicing reconstitution. Using galectin-3 concentrations ranging from 2 μM to 18 μM, we found that the minimal threshold concentration required for reconstituting splicing activity was ~3 μM. The optimal concentrations for reconstitution were 6 - 12 μM (Fig. 5A, lanes 3 and 4 and Fig. 5B). Product formation was inhibited at 18 μM, consistent with our previous observations that high concentrations of galectin-3 inhibit the cell-free splicing assay (5).

Figure 5A: Comparison of the splicing activity of NE, NE depleted by Lac affinity adsorption, and depleted NE reconstituted with recombinant galectin-1, recombinant galectin-3, or the COOH-terminal domain of galectin-3.
Lane 1: control splicing reaction of nondepleted NE.
Lane 2: the unbound (UB) fraction when NE is subjected to Lac affinity adsorption.
Lanes 3-5: components of lane 2 plus 6, 12, and 18 μM recombinant galectin-3 (rG3), respectively.
Lane 6-8: components of lane 2 plus 13, 26, and 52 μM COOH-terminal domain derived from recombinant galectin-3 (C-domain), respectively.
Lane 9-11: components of lane 2 plus 13, 26, and 52 μM recombinant galectin-1 (rG1), respectively.
The positions of migration of pre-mRNA substrate, splicing intermediates, and mature RNA product are highlighted on the right.

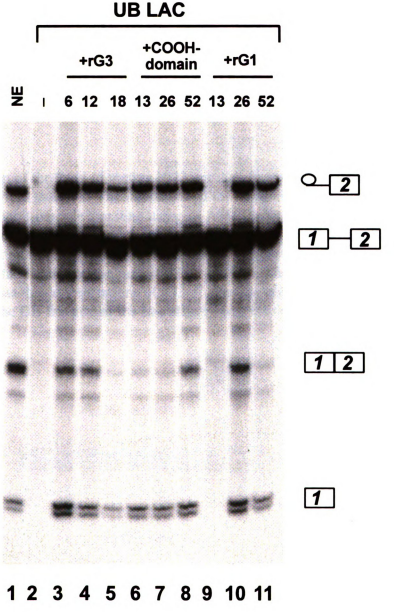
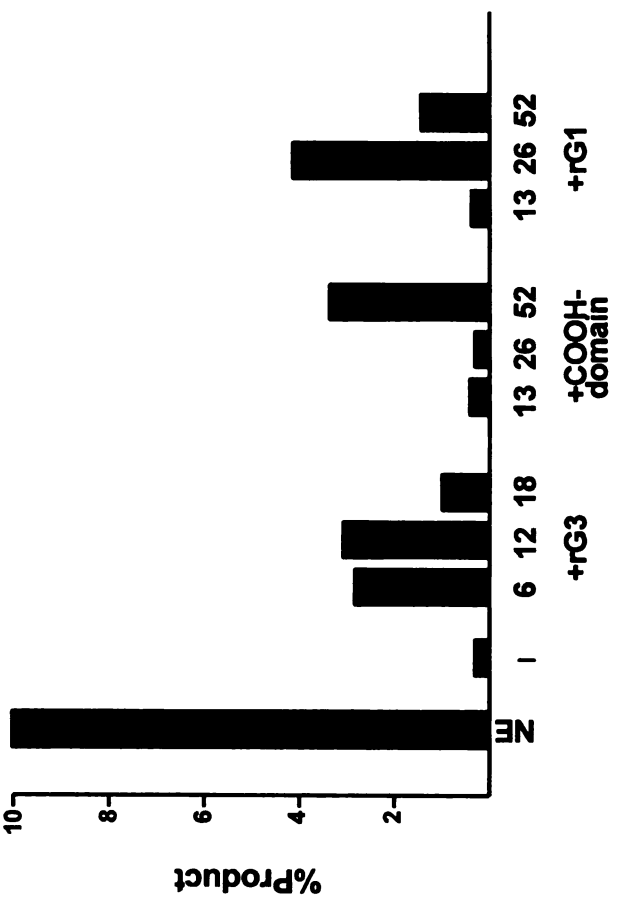

UB LAC
+rG3
+COOH-domain
+rG1
NE
-
6 12 18 13 26 52 13 26 52
2
1 2
1 2
1
1 2 3 4 5 6 7 8 9 10 11

Figure 5B: Quantitation of the level of mature RNA product formation derived from the data shown in panel A. The intensities of the bands on the gels were determined by direct β-particle counting using an AMBIS system. The percent product formed was calculated by dividing the radioactivity present in the RNA product by the total radioactivity present in the pre-mRNA substrate, the splicing intermediates, and the mature RNA product. The numbers at the bottom indicate the concentrations of rG3, C-domain, and rG1 used to reconstitute the UB fraction of the Lac affinity adsorption.


%Product
10
8
6
4
2
NE
I
6
12
18
13
26
52
13
26
52
+rG3
+COOH-domain
+rG1

On the basis of our observations that a higher concentration of recombinant galectin-1 than recombinant galectin-3 was required to reconstitute the splicing deficient NE derived from double antibody depletion experiments (Fig. 4), the concentration range tested for galectin-1 was 13 μM to 52 μM. In contrast to the results obtained with galectin-3, 13 μM was insufficient to reconstitute splicing activity (Fig. 5A, lane 9 and Fig. 5B). Splicing activity was observed with 26 μM galectin-1 (Fig. 5A, lane 10 and Fig. 5B), comparable or even better than that observed with the optimal concentration of galectin-3. Finally, 52 μM showed inhibition of splicing, as was observed with galectin-3 at the high concentration end.

Although the COOH-terminal CRD of galectin-3 can reconstitute the splicing activity of a deficient NE, it required a much higher concentration of protein than was required of the intact galectin-3 polypeptide. Thus, we observed reconstitution of splicing only at the highest concentration tested (52 μM) (Fig. 5A, lane 8 and Fig. 5B). These results suggest that the galectin-3 polypeptide, with a Gly-, Pro-rich domain fused onto the CRD (corresponding to galectin-1 or the COOH-terminal domain of galectin-3), was most efficient in the reconstitution of splicing activity.

DISCUSSION

The key findings documented in the present study include: (a) NE capable of carrying out pre-mRNA splicing, assayed in an *in vitro* system, contain both galectins-1 and -3; (b) depletion of both galectins from NE, either by Lac affinity adsorption or by double antibody adsorption, resulted in the concomitant loss of splicing activity; (c) depletion of either galectin-1 or galectin-3, by specific antibody adsorption, failed to remove all of the splicing activity and the residual activity was still sensitive to saccharide-specific inhibition; (d) either galectin-1 or galectin-3 alone is sufficient to reconstitute, at least partially, the splicing activity of NE depleted of both galectins; and (e) although the CRD is sufficient to restore the splicing activity to galectin-depleted NE, the concentration required for reconstitution was far greater than that of the full-length galectin-3 polypeptide.

One important consideration in the interpretation of these results is that galectin-1 is indeed in the nucleus of a cell. Although it is generally accepted that galectin-1 is predominantly an intracellular protein (1), observations and interpretations pertaining to its nuclear localization are somewhat more difficult to establish (see discussion in chapter III). There are studies that specifically show the localization of galectin-1 in the nuclei, as well as the cytoplasm. There are other studies, however, at the immunofluorescence and ultrastructural levels, which have led to explicit statements that antibodies against galectin-1 failed to label the cell nucleus. To obviate these difficulties in interpretation,

we carried out a study on the intracellular localization of galectin-1 on the very same HeLa cells from which the NE for splicing were derived. Using laser scanning confocal microscopy, we have documented that galectin-1 is indeed in the nucleus of the HeLa cells and that this observation of nuclear localization for galectin-1 is dependent on the detergents used for the permeabilization procedures to allow accessibility of the probing antibody reagent (Chapter III).

The nuclear localization of galectin-1 and the results documented in the present study, demonstrating that depletion of both galectins-1 and -3 are required to remove splicing activity from NE and that either galectin-1 or galectin-3 can restore the splicing activity to a galectin-depleted extract, indicate the two galectins are redundant, at least in terms of this function. The notion of functional redundancy between galectins-1 and -3 is consistent with the results of experiments using transgenic mice in which a null mutation in the gene encoding galectin-1 has been introduced by homologous recombination in embryonic stem cells (14). Homozygous animals carrying the mutant allele lack galectin-1 but there was no apparent damage in terms of development. The mice were viable and fertile. Thus, the function(s) assigned to galectin-1 seems to have been taken over by a redundant relative, possibly galectin-3.

In this connection, it is important to consider the patterns of tissue specific expression of the two galectins. This has been studied carefully during mouse embryogenesis by Poirier and her colleagues (15-17). Both galectins-1 and -3 are first detected on day 4 of mouse development and their expression appear to be limited to the

trophectoderm of the hatched blastocyst. Thus, galectins-1 and -3 overlap in terms of their expression during early embryogenesis. Following gastrulation, however, their patterns of expression no longer parallel each other. Galectin-1 is found in muscle cell precursors of somites while galectin-3 is restricted to the notochord. During the later parts of mouse embryogenesis, galectin-1 expression can be detected in many tissues of the kidney, gut, lung, liver and muscle but not in the chondrocytes of cartilage. In contrast, galectin-3 could be found in the cartilage of the vertebrae, with the hypertrophic chondrocytes exhibiting higher levels of expression than differentiated chondrocytes. Galectin-3 is also found in the suprabasal layer of the epidermis while no transcripts for galectin-1 could be detected. Finally, while galectin-1 is found in the motorneurons and in the sensory neurons of the dorsal root ganglia, galectin-3 could not be observed in the central nervous system.

On the basis of detailed comparisons of the carbohydrate-binding specificities and certain "conserved" versus "variable" amino acid residues within the CRD, Ahmed and Vasta (18) have proposed that the family of galectins actually exhibit two types of CRDs. Galectin-1 contains Type I CRD, with strict conservation of amino acids corresponding to His-52, Asp-54, and Arg-73 and with strict requirements for equitorial -OH groups at C-3 or C-4 of the reducing end of a disaccharide (e.g. low affinity for galactoseβ1,3-N-acetylgalactosamine). On the other hand, galectin-3 contains Type II CRD, in which residues corresponding to His-52, Asp-54, and Arg-73 can be substituted with a different amino acid and with considerable affinity for galactoseβ1,3-N-acetylgalactosamine. Our present results suggest both Type I and Type II CRDs can function in the splicing assay.

This conclusion is based on the facts that: (a) both galectins-1 and -3 must be simultaneously depleted from NE in order to remove splicing activity; (b)either lectin alone can reconstitute the splicing activity of a galectin-depleted NE; and (c)the COOH-terminal CRD of galectin-3 can also restore splicing activity of a galectin-depleted NE.

Although polypeptides corresponding to the CRD are sufficient to reconstitute splicing activity, the minimum concentrations required are 4-8 times higher than that of the intact galectin-3 polypeptide, which contains a Gly-, Pro-rich domain fused onto the CRD. Several possibilities need to be considered to account for this apparent difference in efficiency. First, it is known that endogenous cellular galectin-3 exists as two isoelectric variants: a pI 8.7 species corresponding to the unmodified polypeptide and a pI 8.2 species representing a singly phosphorylated derivative (19). Chemical studies with the canine homolog have identified a major site of phosphorylation to be Ser-6 (~90%) and a minor site at Ser-12 (~10%) (20). In the recombinant murine galectin-3 used in the present studies, residue 12 is Ala so a reasonable possibility is that the polypeptide can undergo phosphorylation at Ser-6 due to kinase(s) present in the galectin-depleted NE, even during the steps of reconstitution, including dialysis and assay. A polypeptide devoid of the Ser-6 phosphorylation site in NH_2-terminal domain (galectin-1 or the COOH-terminal CRD of galectin-3) would thus lack this phosphate group.

Although immublotting of subnuclear fractions and immunofluorescence of permeabilized cells have suggested that galectin-3 is associated with RNP complexes (21, 22) and although there is a recent report that galectin-3 interacts directly with RNA and

single-stranded DNA (23), the specific spliceosomal component that interacts with galectin-3 has not yet been identified. Seve and co-workers have shown that galectin-3 is associated with a glucose-binding protein (24, 25). Athough this interaction can be disrupted with Lac but not by glucose, it is thought the association is via protein-protein interactions rather than protein-carbohydrate recognition. It should be noted, however, that in our previous experiments, the control affinity matrix, composed of cellobiose (glucoseβ1,4glucose)-agarose, did not remove galectin-3 from NE and the unbound fraction of the cellobiose-adsorbed NE was splicing competent (5). In any case, the role of phosphorylation, in terms of enhancing galectin-3's interaction with protein and/or RNA components of spliceosome, is not known. Therefore, we can only speculate that if phosphorylation indeed plays a role, then galectin-1 and COOH-terminal CRD of galectin-3 must overcome their lack of a phosphorylation site on the basis of mass action.

Second, similar arguments apply to the other structural feature of the Pro-, Gly-rich NH_2-terminal domain of galectin-3. The Pro-Gly-Ala-Tyr-Pro-Gly-Xxx-Xxx repeats in this domain (1) may play a role in the interaction between galectin-3 and components of the splicing machinery. Again, galectin-1 and the COOH-terminal CRD of galectin-3, both of which lack this structural feature, must compensate by simple mass action.

The third consideration is that, besides the interaction between the galectin and other spliceosome components, these structural features (phosphorylation and/or Pro-, Gly-rich repeats) may play a role in self-association of the galectin polypeptide. Self-association of the rat and mouse homologs of galectin-3 has been inferred from the

concentration-dependent hemagglutination activity (26, 27) and positive cooperativity in the binding of the lectins to multivalent glycoproteins (26, 28). Both the cooperative binding experiments (26, 28) and cross-linking studies carried out with hamster galectin-3 (29) implicate the NH_2-terminal domain being responsible for the oligomerization of the lectin. Thus, polypeptides lacking the structural features of the NH_2-terminal domain (galectin-1 or the COOH-terminal CRD of galectin-3) would require other mechanisms, including self-association via sequences contained in the CRD, to achieve the same degree of multivalency.

Several studies have indicated that galectin-1 from various species formed homodimers (30-33) and recent crystallographic structures of bovine galectin-1 are consistent with this conclusion (34, 35). In a detailed analysis, Cho and Cummings (36) showed that both recombinant galectin-1 produced in an *E. coli* expression system, as well as galectin-1 endogenous to the cytosol of Chinese hamster ovary cells, existed in a reversible and active monomer-dimer equilibrium. The equilibrium dissociation constant and equilibration time were estimated to be about 7 μM and 10 hours, respectively. Although the concentration dependence was studied in less detail, it has also been shown that the COOH-terminal CRD of galectin-3 can undergo monomer-dimer association (37). Inasmuch as our present data showed that 26 μM recombinant rat galectin-1 was required to reconstitute the splicing activity of a galectin-depleted NE, it appears that protein was functional in the dimeric state. We should hasten to note, however, that the concentration of galectin-1 in NE is estimated to be in the nanomolar range and much of

the requirement for high concentration of the reconstituting protein may have to do with the inefficient steps of dialysis and reassembly of the spliceosome, as discussed previously (14).

REFERENCES

(1) Barondes, S.H., Cooper, D.N.W., Gitt. M.A. and Leffler, H. (1994) J.Biol. Chem. **269**, 20807-20810.

(2) Kasai, K. and Hirabayashi, J. (1996) J. Biochem. **119**, 1-8.

(3) Moutsatsos, I.K., Davis, J.M., and Wang, J.L. (1986) J.Cell Biol. **102**, 477-487.

(4) Laing, J.G. and Wang, J.L. (1988) Biochemistry **27**, 5329-5334.

(5) Dagher, S.F., Wang, J.L. and Patterson, R.J. (1995) Proc. Natl. Acad. Sci. USA **92**, 1213-1217.

(6) Ho, M.K. and Springer, T.A. (1982) J. Immunol. **128**, 1221-1228.

(7) Cherayil, B.J., Weiner, S.J. and Pillai, S. (1989) J. Exp. Med. **170**, 1959-1972.

(8) Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res. **11**, 1475-1489.

(9) Bradford, M.M. (1976) Anal. Biochem. **72**, 248-254.

(10) Laemmli, U.K. (1970) Nature. **227**, 680-685.

(11) Lerner, M.R. and Steitz, J.A. (1979) Proc. Natl. Acad. Sci. USA **76**, 5495-5499.

(12) Cooper, D.N.W., Massa, S.M. and Barondes, S.H. (1991) J. Cell Biol. **115**, 1437-1448.

(13) Agrwal, N., Sun, Q., Wang, S.Y. and Wang, J.L. (1993) J. Biol. Chem. **268**, 14932-14939.

(14) Poirier, F. and Robertson, E.J. (1993) Development **119**, 1229-1239.

(15) Poirier, F., Timmons, P.M., Chen, C.T.J., Guenet, J.L., and Rigby, P.W.J. (1992) Development **115**, 143-155.

(16) Fowlis, D., Colnot, C., Ripoche, M.A., and Poirier, F. (1995) Dev. Dyn. **203**, 241-251.

(17) Colnot, C., Ripoche, M.A., Scaerou, F., Fowlis, D., and Poirier, F. (1996) Trans. Biochem. Soc. **24**, 141-146.

(18) Ahmed, H. and Vasta, G.R. (1994) Glycobiology **4**, 545-549.

(19) Cowles, E.A., Agrwal, N., Anderson, R.L., and Wang, J.L. (1990) J. Biol. Chem. **265**, 17706-17712.

(20) Huflejt, M.E., Turck, C.W., Lindstedt, R., Barondes, S.H., and Leffler, H. (1993) J. Biol. Chem. **268**, 26712-26718.

(21) Laing, J.G. and Wang, J.L. (1988) Biochemistry **27**, 5329-5334.

(22) Wang, S.Y., Voss, P.G., Patterson, R.J., and Wang, J.L. (1995) Antibody, Immunoconjugates, and Radiopharmaceuticals **8**, 311-324.

(23) Wang, L., Inohara, H., Pienta, K.J., and Raz, A. (1995) Biochem. Biophys. Res. Commun. **217**, 292-303.

(24) Seve, A.P., Felin, M., Doyenette-Moyne, M.A., Sahraoui, T., Aubery, M., and Hubert, J. (1993) Glycobiology **3**, 23-30.

(25) Seve, A.P., Hadj-Sahraoui, Y., Felin, M., Doyennette-Moyne, M.A., Aubery, M., and Hubert, J. (1994) Exp. Cell Res. **213**, 191-197.

(26) Hsu, D.K., Zuberi, R.I., and Liu, F.T. (1992) J. Biol. Chem. **267**, 14167-14174.

(27) Ochieng, J., Platt, D., Tait, L., Hogan, V., Raz, T., Carmi, P., and Raz, A. (1993) Biochemistry **32**, 4455-4460.

(28) Massa, S.M., Cooper, D.N.W., Leffler, H., and Barondes, S.H. (1993) Biochemistry **32**, 260-267.

(29) Mehul, B., Bawumia, S., Martin, S.R., and Hughes, R.C. (1994) J. Biol. Chem. **269**, 18250-18258.

(30) Briles, E.B., Li, E., and Kornfeld, S. (1977) J. Biol. Chem. **252**, 1107-1116.

(31) Beyer, E.C., Zweig, S.E., and Barondes, S.H. (1980) J. Biol. Chem. **255**, 4236-4239.

(32) Teichberg, V.I. and Levi, G. (1981) J. Biol. Chem. **256**, 5735-5740.

(33) Roff, C.F. and Wang, J.L. (1983) J. Biol. Chem. **258**, 10657-10663.

(34) Liao, D.I, Kapadia, G., Ahmed, H., Vasta, G.R., and Herzberg, O. (1994) Proc. Natl. Acad. Sci. USA **91**, 1428-1432.

(35) Bourrne, Y., Bolgiano, B., Liao, D.I., Strecker, G., Cantau, P., Herzberg, O., Feizi, T., and Cambillau, C. (1994) Struct. Biol. **1**, 863-870.

(36) Cho, M. and Cummings, R.D. (1995) J. Biol. Chem. **270**, 5198-5206.

(37) Wang, J.L., Anandita (Vyakarnam), Wang, S.Y., and Agrwal, N. (1993) **in** Lectins: Biology, Biochemistry, Clinical Biochemistry, Vol. 9 (P. Chakrabarti, J. Basu, and M. Kundu, eds.) M/S Wiley Eastern Ltd. New Delhi, pp. 89-99.

CHAPTER III

Galectin-1 in the Cell Nucleus. II. A Comparative Localization Study with Other Splicing Components

Anandita Vyakarnam[1], Karen M. Lakkides[1], Ronald J. Patterson[2], and John L. Wang[1]

Departments of [1]Biochemistry and [2]Microbiology, Michigan State University
East Lansing, MI 48824

Acknowledgments: We thank Drs. Sam Barondes and Hakon Leffler for their generous gifts of recombinant rat galectin-1 and rabbit antiserum directed against recombinant rat galectin-1. This work was supported by grants MCB 91-22363 from the National Science Foundation, GM-38740 from the National Institutes of Health, No. 99 from the Mizutani Foundation for Glycoscience, and an Interdisciplinary Research Award from the Cancer Center at Michigan State University.

ABSTRACT

Using both conventional and laser confocal fluorescence microscopy, the intracellular distribution of galectin-1 in HeLa cells was analyzed and compared with the localization of previously documented markers of the nucleus and cytoplasm. The Sm epitopes of the small nuclear ribonucleoprotein complexes (snRNPs) and the non-snRNP splicing factor SC35 yielded only nuclear staining. On the other hand, the enzyme lactate dehydrogenase was cytoplasmic. In contrast to these patterns in which nuclear versus cytoplasmic localizations appeared to be mutually exclusive, galectin-1, as well as galectin-3, yielded both nuclear and cytoplasmic staining simultaneously. Whereas galectin-3 exhibited prominent labeling of the nucleus and weak, diffuse staining of the cytoplasm, the nuclear versus cytoplasmic distribution of galectin-1 was just the opposite, with predominant staining in the cytosol. Confocal microscopy showed galectin-1 fluorescence throughout most of the sections from the top of the cell to the bottom. Through the middle sections, as the plane of focus cuts through the nucleus, there was definite fluorescence staining in the nuclear compartment. Double immunofluorescence analysis showed that, within the nucleoplasm, galectin-3 can be found coincident with the speckled structures observed with SC35. Similar results were also obtained with galectin-1, although in this case, there were areas of galectin-1 devoid of SC35 and vice versa. These results establish the presence of galectin-1 in the nuclei of HeLa cells, a

conclusion consistent with the identification of the protein in nuclear extracts of the same cells and with its documentation as a factor in pre-mRNA splicing.

INTRODUCTION

In the accompanying manuscript (1), we documented that: (a) nuclear extracts (NE) derived from HeLa cells contain both galectins-1 and -3; (b) depletion of both galectins from NE, either by lactose (Lac) affinity adsorption or by double antibody adsorption, resulted in the concomitant loss of splicing activity; (c) depletion of either galectin-1 or galectin-3, by specific antibody adsorption, failed to remove all of the splicing activity and the residual activity was still saccharide inhibitable; and (d) either galectin-1 or galectin-3 alone is sufficient to reconstitute, at least partially, the splicing activity of NE depleted of both galectins. All of the results suggest that the activities of galectin-1 and galectin-3 in the nucleus may be redundant.

A number of lines of evidence have been accumulated to indicate that galectin-3 is found in both the cytoplasm and nucleus of cells. Immunofluorescence staining, using monoclonal, as well as polyclonal, antibodies specifically directed against galectin-3, was carried out on formaldehyde fixed, Triton X-100 permeabilized 3T3 fibroblasts. There was prominent labeling of the nucleus and variable staining of the cytoplasm (2, 3). This dual localization of galectin-3, in the nucleus and cytoplasm, has been confirmed by immunoelectron microscopy (4). Moreover, the nuclear staining of galectin-3 in immunofluorescence and ultrastructural experiments was sensitive to treatment of the permeabilized cells with ribonuclease A (RNase), but not to parallel treatment with

deoxyribonuclease I (DNase) (3, 4). Finally, the nuclear and cytoplasmic localization of galectin-3 has been documented by quantitative immunoblotting analysis of subcellular fractions, relative to enzyme markers characteristic of the various fractions (2).

On the other hand, however, the notion that galectin-1 can be found in the nucleus is controversial. For example, in the original publication reporting the purification and characterization of the chicken homolog, there was an explicit statement that antibodies to the chicken heart protein corresponding to galectin-1 failed to label the cell nucleus of chicken embryo fibroblasts under immunofluorescence (5). More recently, immunofluorescence analysis of hamster galectin-1 showed that it was exclusively cytoplasmic (6). Indeed, the main difficulty in the analysis of galectin-1 is that, unlike galectin-3, the predominant portion of the former protein is found in the cytoplasm. As a result, a strong fluorescent antibody probe staining the cytoplasm surrounding the nucleus sometimes results in an image containing fluorescent flare that exceeds the resolution of the light microscope. To circumvent this difficulty, we took advantage of the availability of a laser confocal microscope to analyze the intracellular distribution of galectin-1. In the course of these investigations, we also found that the type of detergent used to permeabilize the cell had a profound effect on the observation of galectin-1 in the cytoplasm versus the nucleus. The results of these studies are documented in the present communication.

MATERIALS AND METHODS

Immunofluorescence Microscopy

HeLa cells were grown as monolayers in Minimum Essential Media containing 10% defined/supplemented bovine calf serum (HyClone), 100 U/ml penicillin and 100 U/ml streptomycin at 37°C in humidified atmosphere of 5% CO_2. For immunofluorescence microscopy, cells were seeded onto glass coverslips which were placed in 6-well (8 cm^2/well) cluster dishes. Cells grown to 50-70% confluency were washed twice with phosphate buffered saline (PBS, 140 mM NaCl, 2.68 mM KCl, 10 mM Na_2HPO_4, 1.47 mM KH_2PO_4, pH 7.4). They were fixed for 20 min at room temperature with 4% paraformaldehyde in PBS. Cells were washed twice in PBS and permeabilized with 0.2% Triton X-100 in PBS for 5 min. After permeabilization, cells were again washed twice with PBS and incubated with 0.2% gelatin in PBS for at least 1 hour at room temperature. After washing in T-TBS (10 mM Tris pH 7.5, 500 mM NaCl, 0.05% Tween-20), cells were incubated with primary antibody at an appropriate dilution in 0.2% gelatin/T-TBS for 1 hour. Cells were washed three times (15 min each) with T-TBS and incubated for 1 hour with the secondary antibody conjugated with fluorescein isothiocyanate (FITC) at appropriate dilution in 0.2% gelatin/T-TBS. Cells were washed three times in T-TBS for 15 min each and mounted in Perma Fluor (Immunon) on glass slides. In some experiments, cells were permeabilized with 0.05% saponin/PBS for 1 hour at 37°C (6) or 20 μg/ml (0.002%) digitonin/PBS for 10 min at room temperature (7).

In experiments requiring sequential permeabilization with two reagents, fixed cells were first permeabilized with 0.05% saponin for 1 hour at 37°C. Cells were washed three times with PBS and then incubated with 0.2% Triton X-100/PBS for 5 min at room temperature.

For double immunofluorescence analysis, cells were first permeabilized with 0.05% Triton X-100 in permeabilization buffer (300 mM sucrose, 100mM NaCl, 3 mM $MgCl_2$ and 10 mM PIPES pH 7.2) for 5 min at 4°C (8). Cells were rinsed twice with cold permeabilization buffer and then twice with cold PBS. Cells were fixed with 4% paraformaldehyde for 20 min at room temperature, washed extensively with T-TBS and then treated for immunofluorescence staining as described above.. Samples were analyzed using a Meridian Instruments (Okemos, MI) Insight confocal laser scanning microscope.

Antibodies

Anti-galectin-1 is a polyclonal rabbit antiserum raised against recombinant rat galectin-1 (anti-G1) (9) and was a gift from Dr. Hakon Leffler (UCSF, CA). Human autoimmune serum reactive with the Sm antigens of snRNPs (ENA anti-Sm) was purchased from The Binding Site (San Diego). Anti-SC35 (Sigma) is a mouse monoclonal antibody against the non-snRNP splicing factor SC35 (10). Anti-Mac-2 (anti-M2) is a rat monoclonal antibody directed against galectin-3 (11, 12). The antibody was purified from cell culture supernatant derived from hybridoma line M3/38.1.2.8.HL.2 obtained from ATCC (TIB 166). Anti-lactate dehydrogenase (anti-LDH) is a polyclonal rabbit antiserum against pig muscle LDH and was a gift from Dr. John Wilson (Michigan

State University). Secondary antibodies conjugated with fluorescein isothiocyanate (FITC) were used at following dilutions: FITC-goat anti-human immunoglobulin (E.Y. Labs) at 1:500, FITC-goat anti-rabbit immunoglobulin preadsorbed with human serum proteins (Sigma) at 1:500, FITC-goat anti-mouse immunoglobulin at 1:500 and FITC-goat anti-rat immunoglobulin (Sigma) at 1:250. Cy3 conjugated sheep-anti-mouse immunoglobulin (Sigma) was used at 1:250 dilution.

Preparation of HeLa Cell Extracts and Immunoblot Analysis

For preparation of cell extract, HeLa cells were pelleted at 5000g, resuspended in RSB (RSB: 10 mM Tris, pH 8.0, 3 mM $MgCl_2$,10 mM NaCl) and incubated on ice for 10 min. Cells were again pelleted at 2000 rpm for 5 min and resuspended in lysis buffer (25 mM Tris pH 7.4, 50 mM NaCl, 1% NP-40, 0.2% SDS and 200 U/ml aprotinin). Protein concentration of extracts was determined using DC Protein Assay Reagent (Bio-Rad) using BSA as the protein standard (13).

For immunoblot analysis, 50 μg total protein was loaded in each lane. Protein samples were resolved on 12.5% SDS-PAGE as described by Laemmli (14) and electrophoretically transferred to Immobilon-P membrane (Millipore) in a transfer buffer containing 25 mM Tris, 193 mM glycine and 20% methanol. The membrane was blocked for several hours with 10% non-fat dry milk in T-TBS (10 mM Tris pH 7.5, 0.5 M NaCl, 0.05% Tween-20). After brief washes with T-TBS twice, the membrane was incubated with primary antibody diluted in 1% non-fat dry milk/T-TBS for two hours at room temperature followed by five washes in T-TBS for 15 min each. The membrane was then

incubated with secondary antibodies conjugated to horseradish peroxidase for 30 min and washed in T-TBS extensively. Proteins were visualized using Enhanced Chemiluminescence detection system (Amersham).

RESULTS

Nuclear and Cytoplasmic Localization of Galectin-1: Evidence from Laser Scanning Confocal Microscopy

Because the interpretation of the immunofluorscence experiments to be documented below depends on the specificity of the antibody reagents used, the number of components in HeLa cells that react with each antibody reagent was determined by immunoblotting. Anti-G1, a polyclonal rabbit antiserum raised against recombinant rat galectin-1, immunoblotted one predominant polypeptide (M_r ~14,000) whose mobility corresponded to that of authentic galectin-1 (Fig. 1A, lane 1). Much weaker reactivity was observed with a few polypeptides at ~50 kD, which were also observed with rabbit preimmune serum (Fig. 1A, lane 2). Anti-M2, a rat monoclonal antibody directed against galectin-3, yielded a single band corresponding to the human homolog in HeLa cells (Fig. 1A, lane 3). Antibody reagents directed against components that serve as markers for the cytoplasmic and nuclear compartments were also tested. Rabbit anti-pig muscle LDH immunoblotted a single polypeptide (M_r ~37,000) (Fig. 1A, lane 4), consistent with the molecular weight of the enzyme (15). Human autoimmune serum anti-Sm immunoblotted the Sm-B doublets (M_r ~29,000) (Fig. 1A, lane 5), known polypeptide components of the snRNP complexes (16). Finally, mouse monoclonal anti-human SC35 yielded a single band (M_r ~35,000) (Fig. 1A, lane 6), corresponding to this member of the SR family of splicing factors (10).

Figure 1A: Western blot analysis of the the specificity of the antibody reagents. Extracts of HeLa cells (50 μg protein) were subjected to SDS-PAGE and immunoblotting with: (1)rabbit antiserum against recombinant rat galectin-1 (αG1) at 1:1000 dilution; (2)rabbit preimmune serum (PI) at 1:1000 dilution; (3)rat monoclonal antibody against galectin-3 (αM2) at 12.5 μg/ml; (4)rabbit antiserum against pig muscle LDH (αLDH) at 1:1000 dilution; (5)human auto-immune serum reactive against the Sm epitopes of the core polypeptides of snRNPs (αSm) at 1:10,000 dilution; and (6)mouse monoclonal antibody against human SC35 (αSC35) at 1:500 dilution. The binding of the primary antibody was revealed with horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin, goat anti-rat immunoglobulin, goat anti-human immuno-globulin, or goat anti-mouse immunoglobulin, using the Enhanced Chemiluminescence system. The positions of migration of molecular weight markers are indicated on the right.

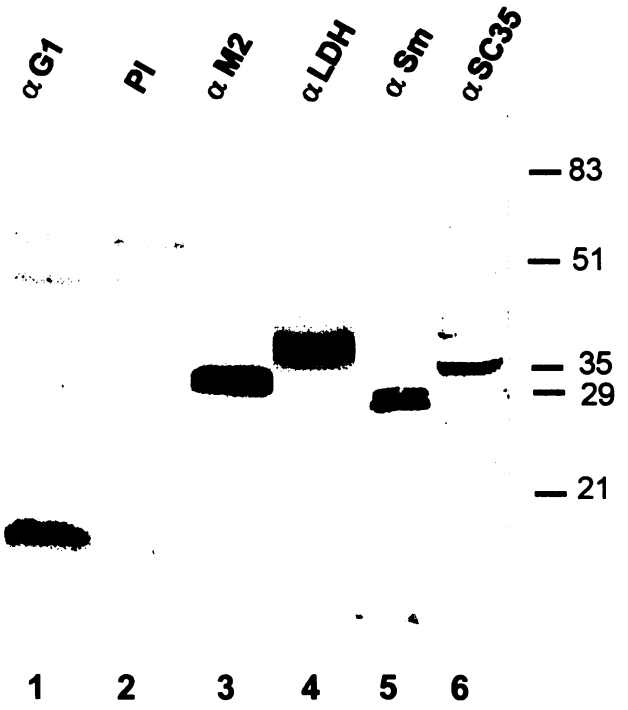
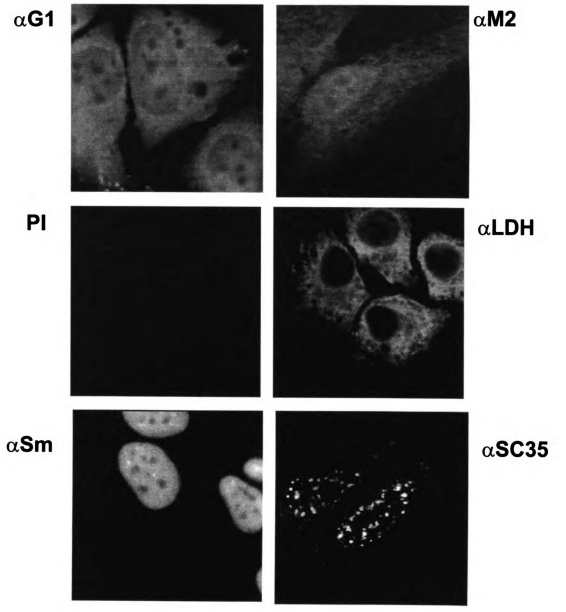

αG1
PI
αM2
αLDH
αSm
αSC35
83
51
35
29
21
1
2
3
4
5
6

Figure 1B: Immunofluorescence staining of galectins-1 and -3 and other nuclear and cytoplasmic markers in human HeLa cells. Cells were fixed with paraformaldehyde (4%) and permeabilized with Triton X-100 (0.2%) and then stained with antibodies: αG1 at 1:50 dilution; PI at 1:50 dilution; αM2 at 12.5 μg/ml; αLDH at 1:50 dilution; αSm at 1:750 dilution; and αSC35 at 1:250 dilution. The binding of the primary antibody was detected by the appropriate fluorescein-conjugated secondary antibody.


αG1
αM2
PI
αLDH
αSm
αSC35

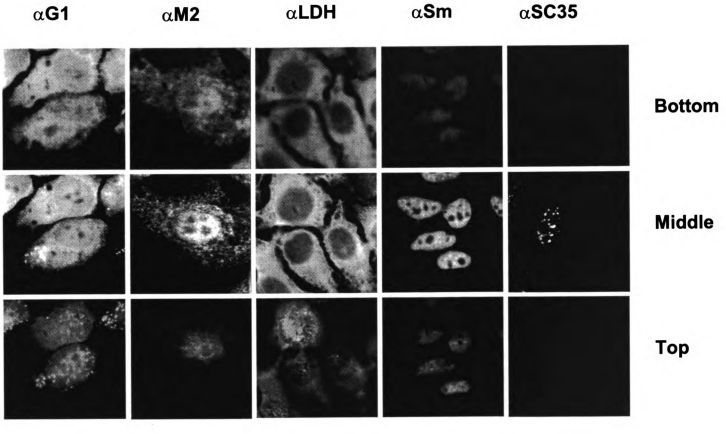
HeLa cells cultured on coverslips were fixed with paraformaldehyde, permeabilized with Triton X-100 and stained with various antibodies. Both antibodies directed against *bona fide* splicing components, anti-Sm and anti-SC35, yielded nuclear staining. The anti-Sm staining exhibited diffuse fluorescence throughout the nucleus, with the exception of ~5 circles devoid of fluorescence (Fig. 1B). These "black holes" correspond to nucleoli (17, 18). Anti-SC35 yielded the characteristic punctate or speckled pattern reported previously (10, 17, 18). Anti-LDH, on the other hand, yielded diffuse distribution of fluorescence in the cytoplasm. The staining obtained with anti-Sm (or anti-SC35) and anti-LDH provided reference patterns expected for a nuclear and cytoplasmic protein, respectively. Finally, anti-M2 yielded labeling of the nucleus, as well as staining of the cytoplasm, in accord with our previous reports (2, 3). As with anti-Sm, the staining of galectin-3 in the nucleus was diffuse throughout the nucleoplasm, with the exception of ~5 "black holes."

The staining patterns obtained with anti-G1 resembled neither those obtained with anti-Sm or anti-M2 nor those seen with anti-LDH. There was extensive staining throughout the cytoplasm, with slightly weaker staining of the nucleus (Fig. 1B). In contrast, the rabbit preimmune serum yielded little or no staining. The galectin-1 staining data appear to be similar to those presented in the original localization of the chicken homolog to galectin-1 in chicken embryo fibroblasts, although the authors explicitly stated that there was no labeling of the cell nucleus (5). Because of the complications arising from cytoplasm overlying and underlying the nucleus, the

fluorescent cells were visualized by laser confocal microscopy, collecting images through consecutive focal planes.

Anti-Sm was used as a positive control for nuclear staining. In a plane above the cell nucleus, there was little fluorescence. Through the middle sections, the plane of focus cuts through the nucleus and yielded intense staining (Fig. 2). As was observed under the conventional fluorescence microscope (Fig. 1B), the Sm antigens showed diffuse distribution throughout the nucleus, except for the 'black holes." Below the nucleus, the various planes of focus no longer showed fluorescence staining. Similar results were also obtained with anti-SC35 (Fig. 2). Anti-M2 also yielded, through the middle sections, bright nuclear staining with the exception of the "black holes" (Fig. 2). Thus, the nuclear distribution of galectin-3 and the Sm epitopes of snRNPs are very similar, a conclusion documented previously (3). Staining for LDH, serving as a negative control, showed that the enzyme was cytoplasmic. Through the middle sections, the anti-LDH yielded a distinct pattern, with inside of the nucleus devoid of fluorescence (Fig. 2). Finally, the staining for galectin-1 showed fluorescence throughout most of the sections (Fig. 2). The middle sections showed fluorescence in both the nucleus and the cytoplasm. All of these results suggest that, although the predominant portion of galectin-1 is found in the cytoplasm, there was definitely nuclear fluorescence due to the same protein.

Figure 2: Subcellular localization of galectin-1 as revealed by laser scanning confocal fluorescence microscopy. Human HeLa cells were fixed with paraform-aldehyde (4%) and permeabilized with Triton X-100 (0.2%) and then stained with rabbit anti- galectin-1 (αG1). The staining of galectin-3 (αM2), snRNP Sm polypeptides (αSm), non-snRNP splicing factor SC35 (αSC35), and LDH (αLDH) provided reference patterns for nuclear and cytoplasmic localizations. The antibodies were used at the same dilution as in Figure 1B. For each staining, images were collected from 9 consecutive focal planes, with the increment of 0.5 μm for each step in the z-direction. Three sections representing the bottom, middle, and top are displayed


αG1
αM2
αLDH
αSm
αSC35
Bottom
Middle
Top

Detection of Nuclear Galectin-1 Dependent on Permeabilization Detergent

The conventional immunofluorescence results we obtained on galectin-1 are, in fact, similar if not identical to the very early study on the immunofluorescence localization of the homolog in chicken embryo fibroblasts (5). The authors of that study apparently discounted the nuclear localization in their statement that "Fluorescence was uniformly distributed throughout the cytoplasm...(but not in the nucleus)." The apparent discrepancy between our results and the most recent localization of galectin-1 in Chinese hamster ovary cells (6) was, however, harder to dismiss. A hint to the resolution of this dilemma was derived from the recent development of assays for studying the import of proteins bearing nuclear localization signals. These assays used digitonin to permeabilize the plasma membrane so that the fluorescently labeled import substrates can gain access to the nulcear translocation machinery at the nuclear pore complex (7, 19). However, the nuclear membrane was not permeabilized so that the import substrates can be transported only in a signal dependent fashion. On the basis of the fact that saponin, the detergent used by Cho and Cummings (6) to permeabilize cells, belonged to the same family of cholesterol-like drugs as digitonin, the possibility was raised that antibodies against nuclear components can gain access in Triton X-100 permeabilized cells of this study while the same reagents were excluded from the nuclear antigens in digitonin/saponin-permeabilized cells.

Thus, parallel cultures of HeLa cells were fixed with paraformaldehyde and then permeabilized with: (a) Triton X-100 (0.2%); (b) digitonin (0.002%; 20 μg/ml); and (c) saponin (0.05%). When subjected to staining with anti-G1, there was a clear difference in the localization observed. Triton-permeabilized cells showed nuclear and cytoplasmic staining (Fig. 3), as was observed in Figures 1B and 2. On the other hand, both digitonin- and saponin-permeabilized cells showed exclusively cytoplasmic staining (Fig. 3).

To confirm the notion that, in saponin-permeabilized cells, galectin-1 was indeed in the nucleus and simply not accessible to the antibody staining reagents, HeLa cells were fixed with paraformaldehyde and then permeabilized with saponin. These saponin-permeabilized cells were then incubated in the presence or absence of Triton X-100. In those cells that were permeabilized only with saponin (no Triton X-100), anti-G1 showed exclusively cytoplasmic staining (Fig. 4). In contrast, saponin permeabilization followed by Triton X-100 yielded both nuclear and cytoplasmic staining. More striking results were obtained with the Sm polypeptides, our reference for nuclear antigens. In the absence of Triton X-100 treatment, the saponin-permeabilized cells showed no anti-Sm staining; these same cells revealed the Sm antigens upon Triton X-100 addition (Fig. 4). Finally, Triton X-100 had no effect on the staining for a cytosolic marker enzyme, anti-LDH. All of these results strongly implicate that galectin-1 is indeed in the nucleus but that it could be detected by immunofluorescence only if the permeabilization procedure allows access of the antibody staining reagents.

Figure 3: Comparison of the immunofluorescence staining pattern of galectin-1 in HeLa cells permeabilized with Triton X-100, digitonin, or saponin. The cells were fixed with paraformaldehyde (4%) and the permeabilized with Triton X-100 (0.2%), digitonin (0.002%; 20 μg/ml), or saponin (0.05%) and then stained with rabbit antiserum against rat galectin-1 (1:20 dilution). The binding of the primary antibody was revealed with fluorescein-conjugated goat anti-rabbit immunoglobulin.

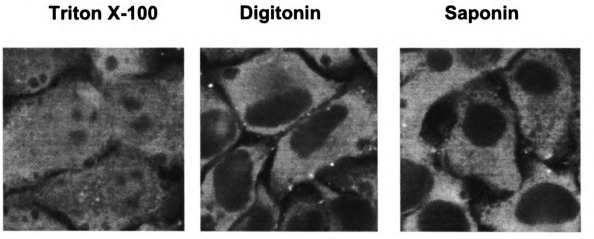
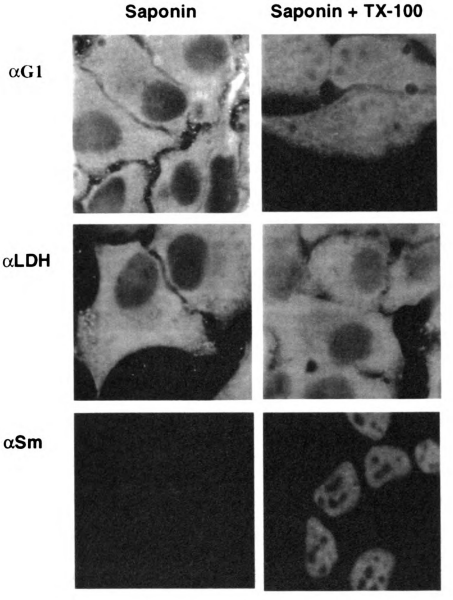

Triton X-100
Digitonin
Saponin

Figure 4: The effect of Triton X-100 on the immunofluorescence staining pattern of galectin-1 in saponin-permeabilized HeLa cells. The cells were first fixed with paraformaldehyde (4%) and then permeabilized with saponin (0.05%). Half of the samples were then treated with Triton X-100 (0.2%). All samples were then subjected to staining with rabbit anti-galectin-1 (αG1), rabbit anti-LDH (αLDH), or human autoimmune serum reactive with Sm (αSm), using the same conditions as that in Figure 1B


Saponin
Saponin + TX-100
αG1
αLDH
αSm

Double Immunofluorescence Analyses: Galectins versus Splicing Factor SC35

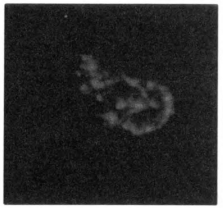
Although the immunofluorescence patterns obtained with fixed and permeabilized cells suggest that the nuclear galectins were diffusely distributed within the nucleoplasm (Fig. 1B and Fig. 2), previous studies have shown that conditions can be found such that galectin-3 and the Sm antigens of snRNPs (3), as well as non-snRNP splicing factors such as SC35 (17, 18), exhibited "speckled" staining patterns. The differences between the diffuse staining pattern and the speckled pattern reflect, at least in part, quantitative differences in either the antigen, the antibody, or both. With high titer antiserum and high levels of antigen (e.g. in fixed and permeabilized cells), the staining intensity could be so strong that it covers the entire nucleus, obscuring distinct subnuclear localization. With loss of antigen from the nuclei during permeabilization and extraction procedures, the staining of discrete structures/regions becomes more distinct, giving rise to a speckled pattern.

HeLa cells were permeabilized with Triton X-100 (without fixation), then fixed and stained simultaneously with anti-M2 and anti-SC35. The binding of anti-M2 was revealed with FITC-conjugated goat anti-rat immunoglobulin, yielding the "green" fluorescence pattern in Figure 5. The binding of anti-SC35 was revealed with cy3-conjugated sheep anti-mouse immunoglobulin, yielding the "red" fluorescence pattern in the same figure. In the composite of the double immunofluorescence patterns, the "yellow" represents regions of coincidence of the "green" and "red" stains. Careful

Figure 5: Double immunofluorescence analysis of galectin-3 and SC35. HeLa cells were permeabilized with 0.05% Triton X-100 (without fixation) and then fixed with paraformaldehyde (4%) and stained simultaneously with: αM2, rat monoclonal antibody against galectin-3 at 12.5 μg/ml; and αSC35, mouse monoclonal antibody against SC35 at 1:250 dilution. The binding of αM2 was revealed by FITC-conjugated goat anti-rat immunoglobulin, yielding the panel with the "green" fluorescence; the binding of αSC35 was revealed by cy3-conjugated sheep anti-mouse immunoglobulin, yielding the panel with the "red" fluorescence. A composite of the the two panels is also shown.

αM2/αSC35

αM2

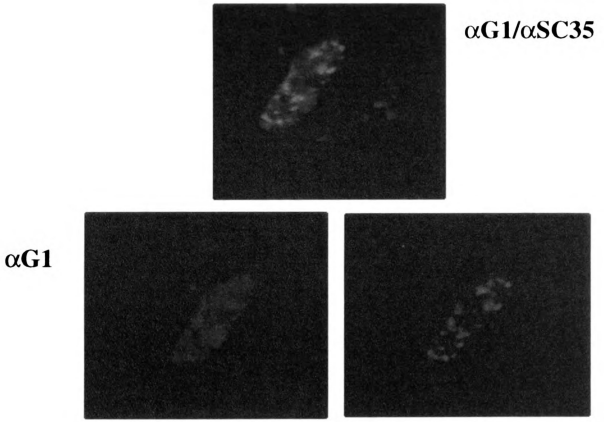


αSC35

analysis of such doubly stained photomicrographs indicates: (a)the "red" anti-SC35 staining appears speckled; (b)there is coincidence between the "red" anti-SC35 spots and the "green" anti-M2 spots, leading to a speckled "yellowish" composite; and (c)the anti-M2 staining for galectin-3 appears to be more diffuse, giving rise to "green" areas in the composite containing no "red" staining. The latter point is important, for it suggests that the sheep anti-mouse immunoglobulin reagent does not cross-react with the rat monoclonal anti-M2 antibody. This conclusion is supported by control experiments, checking for cross-reactivity between the sheep anti-mouse and goat anti-rat with the rat and mouse immunoglobulins, respectively.

A similar double immunofluorescence analysis was carried out with anti-G1 and anti-SC35. The binding of anti-G1 was revealed with FITC-conjugated goat anti-rabbit immunoglobulin ("green" fluorescence); the binding of anti-SC35 was revealed with cy3-conjugated sheep anti-mouse immunoglobulin ("red" fluorescence) (Fig. 6). The speckled SC35 pattern contains regions of coincidence with the anti-G1 staining, giving rise to "yellowish" spots on the composite. However, there are also spots where SC35 staining is not matched by the presence of galectin-1. These appear as "reddish orange" spots on the composite. Finally, there are "green" areas that appear to be devoid of any "red" SC35 staining.

Figure 6: Double immunofluorescence analysis of galectin-1 and SC35. The experimental conditions of Figure 5 were used. αG1, rabbit antiserum directed against recombinant rat galectin-1, was used at 1:5 dilution and was revealed by FITC-conjugated goat anti-rabbit immunoglobulin ("green" panel). αSC35, mouse monoclonal antibody against SC35 was used at 1:250 dilution and was revealed by cy3-conjugated sheep anti-mouse immunoglobulin ("red" panel). A composite of the two panels is also shown.


αG1/αSC35
αG1
αSC35

DISCUSSION

Comparisons of the immunofluorescence staining of live cells versus cells fixed with formaldehyde, followed by permeabilization, indicate that the prototype subgroup of galectins (galectins-1 and -2) are found predominantly in the intracellular compartment (see, for example, reference (5)). Within the cells, the staining is mostly cytoplasmic. Observations and interpretations pertaining to the nuclear localization of galectin-1 are somewhat more difficult to establish.

First, there are studies that specifically show the localization of the prototype galectins in the nuclei as well as the cytoplasm. Cryostat sections of tissues subjected to immunofluorescence showed labeling of both nuclei and cytoplasm in the following cases: anti-CLL-I staining of adult chicken kidney (20) and anti-bovine heart galectin-1 staining of calf pancreas (21). In a series of studies on chick embryonic skin, Akimoto *et al.* (22, 23) reported the chick homolog of galectin-1 in the cytosol and nuclei of cells in the intermediate layer of the epidermis and in some dermal fibroblasts. This nuclear localization, observed at the light microscope level, was confirmed at the ultrastructural level. Similar results were also reported for normal human skin (24).

Second, in a number of studies originally performed for other objectives (e.g. to show overlap in cells expressing lectins and lactoseries glycoconjugates), the investigators provide no specific conclusion regarding the intracellular distribution of

galectin-1. Dorsal root ganglion neurons were subjected to immunofluorescence with anti-rat lung galectin-1 and with a monoclonal antibody directed against a lactoseries glycoconjugate (25). Although the intracellular distribution of the lectin was not discussed in that particular report, the same investigators have more recently concluded, on the basis of the previously published data, that galectin-1 could be detected in both the nucleus and cytoplasm of the neurons (26). Similarly, reassessment of previously published immunocytochemical studies in non-neuronal cells has suggested the presence of galectin-1 in nuclei as well as the cytoplasm (27).

Finally, immunofluorescence and ultrastructural studies have also led to explicit statements that anti-galectin-1 antibodies failed to label the cell nucleus. Two separate studies on mouse C2 myoblasts showed that galectin-1 was found in a diffuse distribution in the cytoplasm but no nuclear localization could be detected (28, 29). Moreover, upon fusion to form myotubes, the intracellular cytoplasmic staining decreases with the appearance of the lectin at the cell surface. In chicken embryo fibroblasts, anti-chicken heart galectin-1 staining was observed only in cells fixed and permeabilized, but not in unfixed cells, suggesting that most of the lectin was intracellular (5). The intracellular staining was ascribed only to the cytoplasm, although the immunofluorescence micrographs are very similar, if not identical, to the present documentation that galectin-1 can be found in nucleus and cytoplasm of HeLa cells. Most recently, immunofluorescence analysis of Chinese hamster ovary cells showed that galectin-1 was exclusively cytoplasmic (6).

On the basis of our present data, it seems likely that at least one parameter that might account for differences in observing nuclear plus cytoplasmic localization of galectin-1, versus exclusive cytoplasmic localization, is the detergent used to permeabilize the cells prior to antibody staining. The use of Triton X-100 permeabilizes both the plasma membrane as well as the nuclear compartment, allowing antibodies reactive against nuclear components to gain access to the antigenic epitopes. On the other hand, saponin (and digitonin) fails to permeabilize the nuclear membrane to allow antibody accessibility. This notion is consistent with the body of literature on the development of a permeabilized cell assay for studying the properties and requirements of protein import through the nuclear pore complex (7, 19). This assay uses digitonin to permeabilize the plasma membrane so that the fluorescently labeled import substrates can bind to the nuclear translocation machinery. At the same time, however, the nuclear membrane is not permeabilized such that the import substrates can be transported only if they carry a nuclear localization signal.

The present microscopic localization of galectin-1 in the HeLa cell nucleus is consistent with the biochemical data indicating that NE contain galectin-1, as well as galectin-3 (1). Both lines of evidence provide the basis for the observations that depletion of both galectins was necessary to achieve a concomitant loss of splicing activity and that either galectin-1 or galectin-3 alone is sufficient to reconstitute the splicing activity of NE depleted of both galectins. The identification of galectins-1 and -3 as splicing factors is also supported by the colocalization of each galectin with speckled patterns observed in immunofluorescence analyses of snRNP antigens and non-snRNP

splicing factors such as SC35 (17, 18). These speckled structures have been found to correspond, at the ultrastructural level, to interchromatin granule clusters and perichromatin fibrils (30). Perichromatin fibrils are readily labeled with short pulses of [^{3}H]uridine (31) and anti-RNA polymerase II antibodies (32), suggesting that they represent nascent transcripts at the sites of mRNA synthesis and early events of pre-mRNA processing. Indeed, immunogold labeling at the ultrastructural level has found galectin-3 in perichromatin fibrils (33). It would, therefore, be of obvious interest to determine if a similar localization can be achieved for galectin-1.

REFERENCES

(1) Chapter 2

(2) Moutsatsos, I. K., Davis, J.M., and Wang, J.L. (1986) J. Cell Biol. **102,** 477-483.

(3) Wang, S.Y., Voss, P.G., Patterson, R.J., and Wang, J.L. (1995) Antibody, Immuno-conjugates, and Radiopharmaceuticals **8**, 311-324.

(4) Hubert, M., Wang, S.Y., Wang, J.L. Seve, A.P., and Hubert, J. (1995) Exp. Cell Res. **220**, 397-406.

(5) Barak-Briles, E., Gregory, W., Fletcher, P., and Kornfeld, S. (1979) J. Cell Biol. **81**, 528-537.

(6) Cho, M. and Cummings, R.D. (1995) J. Biol. Chem. **270**, 5207-5212.

(7) Adam, S.A., Marr, R.S., and Gerace, L. (1990) J. Cell Biol. **111**, 807-816.

(8) Xing,Y. and Lawrence, J.B. (1991) J. Cell Biol. **112**, 1055-1063.

(9) Cooper, D.N.W., Massa, S.M., and Barondes, S.H. (1991) J. Cell Biol. **115**, 1437-1448.

(10) Fu, X.D. and Maniatis,T. (1990) Nature **343**, 437-441.

(11) Ho, M.K. and Springer, T.A. (1982) J. Immunol. **128**, 1221-1228.

(12) Cherayil, B.J., Weiner, S.J., and Pillai, S. (1989) J. Exp. Med. **170**, 1959-1972.

(13) Lowry, O.H., Rosenbough, N.J., Farr, A.L., and Randall, R.J. (1951) J. Biol. Chem. **193**, 265-275.

(14) Laemmli, U.K. (1970) Nature **277,** 680-685.

(15) Weber, K. and Osborn, M. (1969) J. Biol. Chem. **244**, 4406-4412.

(16) Lerner, M.R. and Steitz, J.A. (1979) Proc. Natl. Acad. Sci. USA **76**, 5495-5499.

(17) Spector, D.L., Fu, X.D., and Maniatis, T. (1991) EMBO J. **10**, 3467-3481.

(18) Spector, D.L. (1990) Proc. Natl. Acad. Sci. USA **87**, 147-151.

(19) Moore, M.S. and Blobel, G. (1992) Cell **69**, 939-950.

(20) Beyer, E.C. and Barondes, S.H. (1980) J. Supramol. Struct. **13**, 219-227.

(21) Childs, R.A. and Feizi, T. (1980) Cell Biol. Int. Rep. **4**, 775.

(22) Akimoto, Y., Kawakami, H., Oda, Y., Obinata, A., Endo, H., Kasai, K. and Hirano, H. (1992) Exp. Cell Res. **199**, 297-304.

(23) Akimoto, Y., Obinata, A., Hirabayashi, J., Sakakura, Y., Endo, H., Kasai, K., and Hirano, H. (1995) Cell & Tissue Res. **279**, 3-12.

(24) Akimoto, Y., Hirabayashi, J., Kasai, K., and Hirano, H. (1995) Cell & Tissue Res. **280**, 1-10.

(25) Regan, L. J., Dodd, J., Barondes, S.H., and Jessell, T.M. (1986) Proc. Natl. Acad. Sci. USA **83**, 2248-2252.

(26) Hynes, M.A., Gitt, M.A., Barondes, S.H., Jessell, T.M., and Buck, L.B. (1990) J. Neurosci. **10**, 1004-1013.

(27) Barondes, S.H. (1988) Trends Biochem. Sci. **13**, 480-482.

(28) Cooper, D.N.W. and Barondes, S.H. (1990) J. Cell Biol. **110**, 1681-1691.

(29) Harrison, F.L. and Wilson, T.J.G. (1992) J. Cell Sci. **101**, 635-646.

(30) Fakan, S., Leser, G., and Martin, T.E. (1984) J. Cell Biol. **98**, 358-363.

(31) Fakan, S. and Puvion, E. (1980) Int. Rev. Cytol. **65**, 255-299.

(32) Spector, D.L., O'Keefe, R.T., and Jimenez-Garcia, L.F. (1993) Cold Spring Harbor Symp. Quant. Biol. **58**, 799-805.

(33) Hubert, M., Wang, S.Y., Wang, J.L., Seve, A.P., and Hubert, J. (1995) Exp. Cell Res. **220**, 397-406.

CHAPTER IV

Concluding Statement

CONCLUDING STATEMENT

The identification of nuclear galectin-1 and galectin-3 as novel splicing factors raises many exciting and provocative questions. For example, what is the identity of potential nuclear ligand/partner(s) for galectins-1 and -3? Three classes of molecules could interact with nuclear galectins:

i) RNA (either the pre-mRNA or the snRNA of the snRNPs). Several proteins involved in splicing bind directly to the pre-mRNA or snRNAs required for splicing. RNA-binding motifs have been identified in many of these proteins. Galectins lack identifiable RNA-binding motifs. Moreover, direct interaction of galectins with RNA species has not been demonstrated.

ii) A nuclear glycoconjugate. The obvious candidate ligand for nuclear galectins is a glycoconjugate containing galactose. To date, no such nuclear glycoconjugate has been identified. Recently, glycoproteins with O-linked N-acetylglucosamine (GlcNAc) have been shown to localize to nuclei. Although galectins have no affinity for GlcNAc, it has been shown that GlcNAc residues of nucleoporins are acceptors for galactose addition by galactosyltransferases in a cell-free system. If similar galactose transfer can occur in vivo, the resulting N-acetyllactosamine would exhibit high affinity for galectins.

iii) The nuclear partner for the galectins may be a protein that interacts with the CRD or some other region of the protein. Precedence for the former has been described

for the plant lectin concanavalin A. Hexapeptides from a phage display library have been identified that have an affinity for concanavalin A similar to its normal saccharide ligand. Alternatively, another nuclear splicing component may bind to the galectins through protein-protein interactions. In this case, lactose inhibition of splicing would be an indirect effect, inducing a conformational change resulting in the dissociation of the galectin from its partner. Identification of the nuclear splicing ligand for galectin-1 and -3 is crucial to understand the role of these galectins in splicing.

Another question that remains to be answered: Is the carbohydrate-binding activity of the nuclear galectins required for splicing activity? The ability to bind β-galactoside containing glycoconjugates is the distinguishing feature of galectins. However, since no such conjugates have been identified in mammalian nuclei, it is premature to invoke a specific role for carbohydrate recognition in pre-mRNA processing. Tools are now available to directly test this question. Site-directed mutagenesis of the CRD for galectin-1 has been used to produce mutant forms unable to bind lactose. It needs to be determined whether these mutant galectins restore splicing activity to a galectin-depleted nuclear extract. If carbohydrate binding is not required for splicing function, one possible explanation for inhibition of splicing by galectin-specific saccharides is that binding of lactose to the CRD induces a conformational change in the galectin. This may causes a dissociation of the galectin from its splicing partner, thus inhibiting splicing. Alternatively, specific saccharides could displace a nuclear molecule that may bind to the CRD via molecular mimicry.

Our studies show that one member of the Proto Type and the single member of Chimera Type galectins are involved in pre-mRNA splicing. It would be interesting to determine if other members of the galectin family particularly, Tandem Repeat type galectins, also exhibit splicing activity.

This study demonstrates nuclear localization of galectin-1. Localization of galectin-3 in the nucleus is well documented in literature. However, further studies are required to determine the mechanism of nuclear localization of galectins-1 and -3. Three general modes of nuclear protein import have been described: i) presence of a nuclear localization sequences (NLS) on the polypeptide, ii) diffusion and iii) piggy-back entry. Both galectins-1 and-3 lack NLS. The polypeptide of galectin-1 can associate to form 30 kDa dimers. Galectin-3 remains as monomeric species. At these sizes, both galectins could enter the nucleus by diffusion. Once within the nuclear interior, binding to splicing factors or spliceosomes could result in nuclear retention. Alternatively, these galectins could associate in the cytoplasm with a nuclear-destined protein containing an NLS and enter as a complex. The challenge now is to distinguish between these possibilities, delineate their relative contribution to the nuclear import process and to analyze their regulation.

It is clear that these and many other aspects of galectin function and localization will become apparent as the details of the role of galectins in splicing become unraveled.