This is to certify that the
dissertation entitled

Nuclear galectins and their role in pre-mRNA splicing

presented by

Weizhong Wang

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Microbiology and Molecular Genetics

Ronald Patterson
Major Professor's Signature

12/14/06
Date

Nuclear galectins and their role in pre-mRNA splicing

By

Weizhong Wang

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Microbiology and Molecular Genetics

2006

Abstract

Nuclear galectins and their role in pre-mRNA splicing

By

Weizhong Wang

Galectins are a family of carbohydrate binding proteins that share conserved sequence elements in their carbohydrate-binding sites. In previous studies, we have documented that galectin-1 and galectin-3 are necessary and redundant splicing factors in an *in vitro* splicing system.

We wanted to know whether the splicing activity of galectins depends on the carbohydrate binding activity. To address this question, we obtained mutant galectin-1 construct from Hirabayashi's lab. Residue Asn46 is critical for carbohydrate binding activity of galectin-1. By mutating this residue to Asp, the carbohydrate binding activity of galectin-1 was abolished. This mutant was tested along with the wild type galectin-1 in depletion and reconstitution experiments. The results show that the mutant galectin reconstituted splicing as well as the wild type. This allows us to draw the conclusion that the carbohydrate binding activity of galectin-1 is separate from its splicing activity.

The next goal was to determine how galectins are involved in the splicing process. Our hypothesis was that galectins associate with the spliceosome through protein-protein interactions. Immunoprecipitation was used to test this hypothesis. Pre-mRNA was added to HeLa nuclear extracts to assemble spliceosomes, then they were mixed with antibody against galectin-1 or galectin-3 and the bound fractions were isolated and subjected to RNA or protein analysis. The results show that both antibodies can immunoprecipitate

spliceosomal RNAs and spliceosomal protein components. We conclude that galectins physically associate with spliceosomes throughout the splicing process. To further characterize this association, splicing reactions were sequentially immunoprecipitated with galectin antibodies. We found that galectin-1 containing spliceosomes did not contain galectin-3 and vice versa. These data provided an explanation for the functional redundancy of galectin-1 and galectin-3 in splicing activity. Another finding is that the association of galectins with the spliceosomes can be disrupted easily by ionic conditions. Since most of the proteomic studies on spliceosomes used stringent ionic conditions, this result could explain why galectins haven't been found in these studies. We conclude that galectin-1 and -3 are directly associated with splicing complexes throughout the splicing pathway in a mutually exclusive manner and they bind a common splicing partner through weak protein-protein interactions.

Since galectins associate with spliceosomes, it is intriguing to analyze whether they also associate with endogenous nuclear complexes. Using similar immnoprecipitation techniques, we confirmed our hypothesis. Galectin-1 and -3 are both components of nuclear snRNP containing complexes. This association does share some common features as the galectins' association with spliceosome. They are both sensitive to ionic strength. They only incorporate one type of galectin into the complex and mutually exclude the other. Pre-formed galectin containing complex could not be disturbed by carbohydrate ligands. In the future, galectin containing nuclear complexes might be the tools to reveal the mechanism of galectin incorporating into spliceosomes.

Acknowledgments

First and foremost, I would like to thank my mentor, Dr. Ronald Patterson. His support, instruction, and encouragement are of great value for me both academically and personally. Without him, I could hardly finish my Ph.D research. I will forever remember him as a kind and intellectual adviser.

Next, I'd like to thank the members of my graduate study committee, Dr. John Wang, Dr. Richard Schwartz and Dr. Jerry Dodgson. They gave me precious advice, helped to solve problems encountered and guided me through my research study.

I also want to thank all my friends in the department and collaborating lab, including Jung Park, Kevin Haudek, Patty Voss, Richard Gray, Sara Santos, Shibani Mukherjee, Xiaoyu Wang, Lianjie Li and many others. Their friendship and support will be with me forever.

Last but not least, I want to thank my parents. Their love and encouragement is the most important thing for me in these years.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Literature Review

Part A Galectins

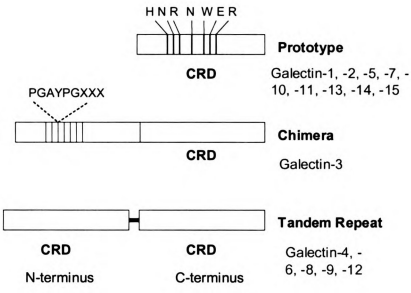
Galectin family

Galectins are a family of carbohydrate-binding proteins. There are two key features to characterize these proteins: (i) binding affinity for β-galactosides and (ii) conserved sequence in their carbohydrate-binding site (1). Galectins have been identified in many species, including mammal, fish, bird, frog, worm, sponges and fungi (2). Each member of the galectin family contains at least one carbohydrate recognition domain (CRD) which has binding affinity for β-galactosides. Several residues in this domain are critical for saccharide ligand binding, including His49, Asn51, Arg53, Asn62, Trp69, Glu72 and Arg74 (using galectin-7 as an example, residues may vary depending on different galectins).

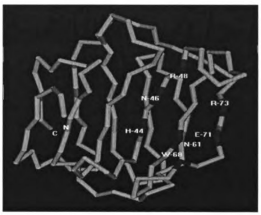
Fifteen galectins have been discovered and numbered sequentially according to the accepted numbers for their genes in the Genome Database. According to their structure, galectins can be separated into three different groups (3) (Figure 1). The first group is the prototype group. It includes galectin-1,-2, -5, -7, -10, -11, -13, -14 and -15. These proteins contain a single CRD of about 130 amino acids. Crystallography has shown that the CRD is folded into two anti-parallel β-sheets forming a sandwich like structure (4, 31). The second group of the galectin family is the chimera type. This group contains only one member, galectin-3. This protein has a unique amino (N) terminal domain linked to the CRD. The three dimensional structure of this N-terminal domain is still unknown. It contains several repeats of Pro-Gly-Ala-Tyr-Pro-Gly-X-X-X, which are

Figure 1: Panel A shows the schematic diagram of the three subfamilies of galectins (proto-, chimera- and tandem repeat). The prototype is composed of a single carbohydrate recognition domain (CRD). The chimera type contains a unique amino terminal domain (key feature is 7-12 repeats of PGAYPGXXX) and a CRD. The tandem repeat type is composed of two similar CRDs. Highly conserved residues of the CRD are shown above the prototype structure. The crystal structure of galectin-1 CRD is shown in panel B (29).

Figure 1


HNR N WER
Prototype
CRD
Galectin-1, -2, -5, -7, -10, -11, -13, -14, -15
PGAYPGXXX
Chimera
CRD
Galectin-3
Tandem Repeat
CRD
CRD
Galectin-4, -6, -8, -9, -12
N-terminus
C-terminus

B


R-48
N-46
R-73
H-44
E-71
N-61
W-68

conserved in all species. The third group of galectins is the tandem repeat group. The members of this group (galectin-4, -6, -8, -9, -12) contain two highly similar repeats of the CRD. These two domains are connected by a short linker peptide (5).

Galectins are believed to be synthesized on free cytoplasmic ribosomes. Although no signal peptide or conventional nuclear localization sequence has been identified in galectins, they distribute into all three compartments of the cell—cytoplasm, nucleus and extracellular space. Some of the galectins bind to membranes and some are in the extracellular matrix (2).

Different galectins exert different functions. Galectin-2 (14 kD), -4 (36 kD), -6 (33 kD), and -11 (15 kD) are mainly expressed in the gastrointestinal tract of different animal species (6). They are usually found in the cytoplasm of the epithelial cells. The functions of these galectins have not been clearly defined. Some evidence suggests that galectin-11 is involved in the immune responses of the digestion tract (7). Galectin-5 (16 kD) was discovered in rat erythrocytes. It can be found in the cytoplasm and the cell surface. Galectin-5 can interact with both free saccharides and mammalian glycoproteins (8). Galectin-7 (15 kD) is a prototype galectin which has a similar folding pattern as galectin-1. Galectin-7 is mainly expressed in stratified epithelia. Interestingly it is found predominantly in the nucleus and can be secreted out of the cell (9). Galectin-7 is also a target for tumor research. It has been suggested to be a pro-apoptotic protein, causing apoptosis after cell damage. Its tumor suppression function is associated with p53 and the JNK pathway (10). Galectin-8 (35 kD) and -9 (35 kD) are tandem repeat galectins and both have several isoforms. Galectin-8 is involved in cell adhesion and spreading (11) while galectin-9 is involved in apoptosis (12). Galectin-9 has also been reported to act as

the uric acid transporter, which is responsible for efflux from urate-producing cells with renal and gastrointestinal excretion (13). Galectin-10 (also known as Charcot-Leyden crystal protein) is expressed uniquely in eosinophilic and basophilic leukocytes. Its sequence only shows moderate similarity to other galectins, while its crystal structure is highly similar to the prototype galectins (14). Galectin-10 is unique in the galectin family as it does not bind to β-galactosides, but it does bind to mannose. Galectin-10 has been found in both the cytoplasm and the nucleus. It has been reported that galectin-10 contains weak lysophospholipase activity (14), but a subsequent study suggests this conclusion is incorrect. When galectin-10 was depleted by its antibody, the lysates derived from blood eosinophils retained their full lysophospholipase activity (15). Galectin-12 (36 kD) is a tandem repeat galectin. It is abundantly expressed in adipose tissue and mainly concentrated in the nucleus. Galectin-12 was suggested to play a role in the apoptosis of adipocytes through suppression of cell growth (16). Galectin-13 (16 kD) was first identified as a soluble human placental tissue protein (17). It has also been found in adult spleen and bladder, fetal kidney, adenocarcinoma of the liver and malignant melanoma. Galectin-14 (17 kD) is found in eosinophils. It has an extended N-terminal domain beyond the normal CRD of galectins (18). Galectin-15 is a newly discovered member of the galectin family. Galectin-15 is expressed by the endometrial luminal epithelium (LE) and superficial ductal glandular epithelium of the ovine uterus. Galectin-15 has been proposed to have an extracellular role in regulating implantation and placentation (by functioning as a heterophilic cell adhesion molecule between the conceptus trophectoderm and endometrial LE). Its intracellular role is to regulate cell survival, differentiation and function (19).

Galectin-1

Galectin-1 is a member of the prototype galectin family. The molecular weight of the galectin-1 polypeptide is about 14 kD. The N-terminus of galectin-1 is acetylated (20). The expression and localization of galectin-1 changes according to the stage of development and differentiation (21, 22). Galectin-1 exhibits dual localization in the cell. Although largely found in the cytoplasm, galectin-1 also exists in the nucleus. In HeLa cells, treatment with detergent that permeabilizes the nuclear membrane reveals obvious nuclear staining of galectin-1 (23). In epidermal Langerhans cells, galectin-1 has been found in the nucleus by electron microscopy (24). In rat osteoblasts, galectin-1 has been determined to be a component of the nuclear matrix (25). Galectin-1 is a potent multifunctional effector that participates in specific protein-carbohydrate and protein-protein interactions. Many important roles of galectin-1 are exerted through its galectin-carbohydrate interactions. For example, by interacting with the ganglioside GM1, galectin-1 is involved in cell growth and neural differentiation (26). Galectin-1 can also bind its protein ligands through protein-protein interactions, as is the case of Gemin-4 (37) and H-Ras (27, 28).

The crystal structure of galectin-1 has been determined by X-ray crystallography (Figure 1). Two identical monomers of galectin-1 assemble into a homodimer. The overall structure of galectin-1 is a β-sandwich consisting of two anti-parallel β-sheets of 5 and 6 strands, respectively (29). The N and C termini of galectin-1 are both located at the dimer interface side and the ligand-binding region is at the far-end (30). This overall structure is shared by most CRDs of the galectin family. A group of highly conserved

amino acid residues interact with the disaccharide ligands, they are His44, Asn46, Arg48, Val59, Asn61, Trp68, Glu71 and Arg73 (31).

Many studies have focused on the extracellular functions of galectin-1. In olfactory neurons, galectin-1 is involved in cell adhesion and establishment and maintenance of synaptic connectivity (32). In T-lymphocytes, galectin-1 can induce apoptosis (33). Galectin-1 also interacts with integrin $\alpha_7\beta_1$ and through integrin binds to laminin (34). Recently, its intracellular functions have attracted attention. In the H-Ras transformed Rat-1 cell, galectin-1 plays an important role in facilitating H-Ras oncogene product anchoring to the membrane, thus helping the cell transformation process (28). Galectin-1 directly binds to H-Ras. This interaction has been shown by cross-linking studies and reciprocal immunoprecipitation using antibodies against both proteins. Several lines of evidence show this interaction is not related to the carbohydrate binding activity of galectin-1. First, this interaction is not sensitive to lactose inhibition, suggesting the interaction is due to protein-protein binding (35). Second, a mutant galectin-1 has been generated by substituting leucine at position 11 by alanine. This mutant galectin-1 retains normal carbohydrate-binding capacity but shows diminished H-Ras GTP-loading and extracellular signal-regulated kinase activation, dislodged H-Ras(G12V) from the cell membrane, and attenuated H-Ras(G12V) fibroblast transformation and PC12-cell neurite outgrowth (36). Another protein binding partner that has been discovered by the yeast two hybrid assay and GST pull down experiment is Gemin-4 (37). Gemin-4 is a component of the survival of motor neuron (SMN) complex. The SMN complex is involved in pre-mRNA splicing. Thus the binding of galectin-1 and Gemin-4 provides evidence to link galectin-1 with splicing.

Galectin-3

Galectin-3 was first discovered as a carbohydrate binding protein, named CBP35 (38). Later, many research groups found this protein independently and named it Mac-2, IgE binding protein, L-29, LR-29, L-34 and so on. These studies implicate galectin-3 in many processes in the cell. Unique in the galectin family, galectin-3 is notable by its unusual N-terminal domain. The N-terminal domain contains multiple internal sequence homologies, each of them consists of nine amino acid repeats of Pro-Gly-Ala-Tyr-Pro-Gly-X-X-X. Differential scanning calorimetry has shown that the N-terminal domain and the C-terminal domain of mouse galectin-3 have transition temperatures of 40°C and 55°C, respectively (39), suggesting these two domains have different folding structures. Galectin-3 can form dimers in the cell. The interaction is through both N-terminal domain and C-terminal domain residues. The C-terminal domain relies on its NWGR motif to form the dimer, this interaction is saccharide ligand inhibitable and abolished in a W181L mutation of the galectin-3 polypeptide (40). In the presence of saccharide ligands, galectin-3 can still form dimers through its N-terminal domain. This has been proved by cross-linking experiments and site-directed mutagenesis (41, 42).

Although galectin-3 lacks a conventional signal sequence for transfer into the endoplasmic reticulum and entry into classical secretory pathway, it does get secreted into the extracellular space (43). Galectin-3 also shuttles between the nucleus and the cytoplasm (44). Its intracellular distribution depends on the proliferative state of the cell, being mainly cytoplasmic in quiescent stage and nuclear in replicating cells (45). Phosphorylation at N-terminal serine 6 and serine 12 by kinases is necessary for galectin-3 shuttling from the nucleus to the cytoplasm, because only the phosphorylated form can

be detected in the cytoplasm (46). Phosphorylation is also important for galectin-3's activity. Phosphorylation of serine residues inhibits galectin-3 adhesion to the nuclear matrix (47). On the other hand, phosphorylation favors the anti-apoptotic function of galectin-3 (48). This has been shown by experiments using site-directed mutagenesis of galectin-3.

Like galectin-1, galectin-3 also exhibits two different modes of interaction with its ligands (49). Using lectin-glycoconjugate interactions, extracellular galectin-3 interacts with the β-galactoside residues of several extracellular matrix (ECM) and cell surface glycoproteins via the CRD. In contrast, using both its N-terminal domain and C-terminal domains, intracellular galectin-3 forms interactions with other proteins via protein-protein associations. One example of such interactions is galectin-3 with TTF-1. GST pull-down assays demonstrated a direct interaction between galectin-3 and the homeodomain of TTF-1 (50). These binding properties could enable galectin-3 to exert multiple functions (51).

Intracellularly, galectin-3 is involved in regulating the cell cycle. It can arrest cells in G1 or at the G2/M phase. The basic regulating steps are downregulating cyclins A and E, upregulating p21 and p27 cyclin inhibitors, and hypophosphorylating Rb protein (52). Galectin-3 is also involved in cell proliferation and cell death. In 3T3 fibroblasts, galectin-3 can stimulate cell proliferation by its upregulation and altered pattern of distribution and phosphorylation (53). The mitogenic potential of galectin-3 has been demonstrated in several experiments. In human lung fibroblasts, cell proliferation is associated with the expression of galectin-3 (54). In human leukemia T-cells, when

transfected with galectin-3 cDNA, the cells show higher growth rates than control transfectants (55). In addition to regulating cell proliferation, galectin-3 favors cell survival by protecting from apoptosis induced by a variety of death signals and anoikis (52, 56). Galectin-3's anti-apoptotic activity seems to be related to the highly conserved NWGR motif sequence. Galectin-3's association with bcl-2 (56) and x/ALG-2-interacting protein 1 (the human homologue of ALG-2 linked protein) (57) is also important to this activity. By its pro-proliferative and anti-apoptotic action, galectin-3 is considered as an immediate early gene possibly implicated in tumor growth. This effect has been shown by the abnormal expression of galectin-3 in several neoplasms (58).

Extracellularly, galectin-3 is involved in regulating cell adhesion. Cell surface galectin-3 binds to a serum glycoprotein, serving as a cross-linking bridge between adjacent cells and promotes homo and heterotypic cell-to-cell interactions (59). On the other hand, galectin-3 associates with the $\alpha 1\beta 1$-integrin receptor, downregulating cell adhesion to the ECM component laminin, thus producing an anti-adhesive effect (60). In the process of tumor invasive and metastasis, these types of adhesion changes play a very important role. Presumably, galectin-3 might increase the interactions among tumor cells and the interaction of tumor cells with the endothelium; it also could alter the adhesion of cells to the basement membrane glycoprotein laminin. Under certain circumstances, galectin-3 may promote cell adhesion to laminin, as shown in neutrophils (61). This effect is believed to play a role in the traversing of neutrophils through the basement membrane at the site of inflammation.

Another important function of extracellular galectin-3 is the modulation of inflammation and allergy. By binding to IgE and the IgE receptor, galectin-3 can induce activation of mast cells and basophils, which is a central event in allergy (62). Galectin-3 also helps eosinophils and neutrophils bind IgE, therefore playing an important role in mediating their IgE-dependent effector function (51). Finally, galectin-3 is reported to be capable of binding, internalizing, and degrading AGE (advanced glycation end products), thus playing a role in diabetes and aging (63).

Part B Splicing and spliceosomes

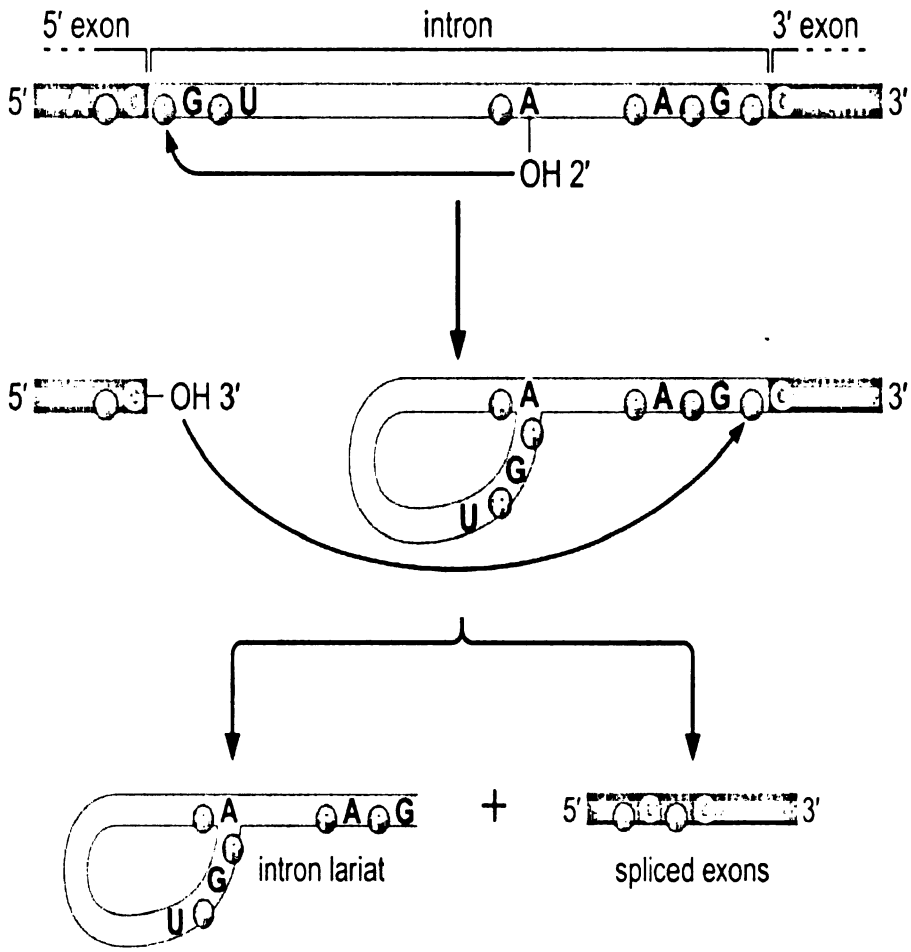
Splicing

Pre-mRNA splicing is a process which removes introns from the pre-mRNA and ligates exons together to form the mature messenger RNA (64). Two transesterification reactions take place in this process. In step 1, the 2'-OH group of the adenosine at the branch point attacks the phosphate group of the 5' splice junction (guanosine) forming exon-1 and lariat-exon-2 intermediates. In step 2, the free 3'-OH of exon-1 attacks the phosphate group of the 3' splice junction (guanosine), resulting in ligation of the exons and release of a free lariat which is degraded later (65) (Figure 2). Several regions on the pre-mRNA are essential for these reactions. The 5' splice junction has the consensus sequence AG:GU. At the 3' splice junction the consensus sequence is AG in a context of semi-conserved CAG:G. A branch point region is also crucial (66). It is usually 30 nucleotides (nt) upstream of the 3' splice site. In yeast, the branch point sequence is an absolutely conserved sequence of UACUA(A)C. In vertebrates, it is variable with PyNPyPu(A)Py. Besides these regions, mammals also contain a polypyrimidine stretch upstream of the 3' splice junction to facilitate recognition.

Pre-mRNA splicing is an important post-transcriptional process in eukaryotic cells. This process is accomplished by a large macromolecular complex called the spliceosome. The spliceosome contains five small nuclear ribonucleoprotein complexes (snRNPs) and a large number of non-snRNP proteins (many are arginine/serine (RS) domain containing proteins). The snRNPs involved in splicing process are U1, U2, U4, U5 and U6 snRNP (67). There are a few exceptions in the splicing process. For example, group 2 introns can splice in the absence of proteins (68). But these occasions are rare.

Figure 2: Schematic diagram depicting the splicing process. In step 1, the 2'-OH group (A) branch point attacks the phosphate group of the 5' splice junction (guanosine) forming exon-1 and lariat-exon-2 intermediates. In step 2, the free 3'-OH of exon-1 attacks the phosphate group of the 3' splice junction (guanosine), resulting in ligation of the exons and a free lariat which is degraded later (105).

Figure 2


5' exon
intron
3' exon
5'
G U
A
A G
3'
OH 2'
5'
OH 3'
A
A G
3'
G
U
A
A G
G
U
intron lariat
+
5'
3'
spliced exons

The snRNPs contain a single low molecular weight RNA and several protein components. The snRNAs are about 200 nucleotides in length. Their 5' end form a 2,2,7-trimethyl-guanosine cap. There are six common sets of snRNP core proteins in this particle in addition to one or more snRNP-specific proteins (67). Besides functioning in pre-mRNA splicing, snRNPs are also involved in 3' end processing of pre-mRNAs (69). The key component in the snRNPs is the snRNA. The snRNAs can base pair with pre-mRNA in the conserved sequence motif and allow the snRNP particles to initiate the assembly of the spliceosome (70).

The cellular amount of snRNPs is variable according to species. In mammalian cells, the snRNAs are rather abundant. There are about $1x10^5$ to $1x10^6$ copies in each human cell. In yeast, there are about 10 times less.

Except for U6, U snRNAs are transcribed by Pol II (DNA dependent RNA polymerase II) in the nucleus. They are then capped with a 7-methyl-guanosine and exported to the cytoplasm. In the cytoplasm, a set of Sm proteins is assembled onto the snRNAs, the 5' end of snRNAs is also hypermethylated, forming the 2,2,7-trimethyl-guanosine cap structure. The cap structure and the Sm protein ring serve as nuclear import signals, bringing the snRNPs back into the nucleus. In the nucleus, snRNPs first appear in the Cajal Body. In this nuclear structure, snRNPs undergo further assembly. Several snRNP-specific proteins are added, bases of the snRNA are modified, and U4/U6.U5 tri-snRNP is assembled (71, 72, 73). U6 goes through a different pathway. It is transcribed by Pol III and does not get exported to the cytoplasm. Thus, U6 only has a mono-methyl cap and remains in the nucleus. Since U6 does not enter the cytoplasm, it

does not acquire the Sm proteins. Instead, eight Sm-like (LSm) proteins replace the Sm proteins and form the core proteins (72, 74).

The snRNPs carry out essential functions during the splicing reaction. Via the interaction of their RNA and protein components with the pre-mRNA, they mediate the recognition and subsequent base pairing of the 5' and 3' splice sites of an intron. In addition, they form the three-dimensional structure which is required for forming the spliceosomes' two active sites. Components of the snRNPs also appear to catalyze the two transesterification reactions that occur during splicing (74).

Spliceosome

Newly transcribed pre-mRNA is covered by hnRNP (heterogeneous nuclear ribonucleoprotein) proteins, forming the initial H complex. This is a non-specific complex and not all H complexes can be chased into later splicing complexes. Spliceosome assembly is a process in which the 5 snRNPs and many non-snRNP proteins orderly assemble onto the pre-mRNA. The conventional model for spliceosome assembly is a step-wise process (75, 76). The first step is the assembly of U1 snRNP onto the 5' splicing site of the pre-mRNA. The consensus sequence at the 5' splice junction is recognized by U1 through an RNA-RNA complementary base pairing (77). The second step is the addition of U2 snRNP. A U2 snRNP specific protein (U2AF) plays an important role in this step. U2AF has two sub-domains. The small subunit is about 35 kD, called U2AF35; the large one is 65 kD, named U2AF65. U2AF35 recognizes the 3' splicing site with the sequence of AG and binds to this region, while U2AF65 binds to the polypyrimidine tract (78). The binding of U2AF recruits the U2 snRNP and does not require ATP hydrolysis. With U1 snRNP and U2AF bound to pre-mRNA, a new splicing

complex (E complex) forms (79). E complex is also called the commitment complex. This is the real starting point of the functional spliceosome. Given the right conditions, all of the E complexes can be chased into higher ordered complexes. There are also other important proteins needed to assemble onto the RNA before U2 snRNP can bind. SF1 (called BBP in yeast) recognizes the branch point A and is critical in facilitating the assembly of U2 snRNP. SF2 and SC35 bring the 5' and 3' splice sites in proximity by interacting with U2AF35 and thus change the spatial structure of the complex (80, 81). Although U2 snRNP has been reported to associate with the E complex, it is widely believed that the binding of U2 snRNP to the branch point forms the next higher ordered complex, the A complex. This is accomplished by the U2 snRNA replacing the SF1 protein and base pairing with pre-mRNA at the branch A region using its 5' end. This switch of protein-RNA interaction to RNA-RNA interaction is the first ATP dependent step in spliceosome assembly and starts the formation of the catalytic complex (80).

After A complex is formed, U4/U6.U5 tri-snRNP comes into play. In this tri-snRNP, U4 and U6 associate through snRNA-snRNA interactions while U5 snRNP binds to them through protein-protein interactions. The binding of this tri-snRNP with the pre-mRNA forms an intermediate complex (B complex) before the spliceosome becomes splicing active (76, 82). Before U6 snRNA can bind to the complementary region on the pre-mRNA, the base pair between U4 and U6 needs to be disrupted. This process is accomplished by several DExD/H box proteins (84). Prp44, also called Brr2, has been implicated in unwinding the U4/U6 helices and likely confers the ATPase activity (85). U4 snRNP protein Prp28 and U6 snRNP protein Prp24 have also been proposed to directly or indirectly influence U4/U6 unwinding (86). A U5 snRNP protein Prp8 has

been suggested to play an important role in regulating the activity of the DExD/H-box proteins (87). Overall this unwinding is an ATP-dependent process and requires multiple protein involvement. After the U6 snRNA dissociates from the U4 base pair, it displaces U1 RNA from the 5' splice site, establishing a new pairing between U6 RNA and the 5' splice site. Following this event, U1 and U4 snRNP are released from the spliceosome.

Accompanying these component changes, the splicing complex also goes through a conformation rearrangement. When U2 and U6 snRNA base pair together, the catalytic core is formed and the active splicing complex (C complex) finally forms (76). These conformational changes appear to require the branchpoint sequence, presumably because the activated U2 snRNP has a role in the process. The role of U5 snRNP is mainly contributed by its protein components. U1-70K interacts with the U5 snRNP proteins Prp44/Brr2 and Prp8 (85, 88) and with protein Exo84. Exo84, in turn, interacts with Prp8 (89). Prp8 can also interacts with U1 snRNP proteins Prp39 and Prp40 (87). In fact, a U1/U5 bi-snRNP has been identified in yeast cell extracts (90), which further strengthens these discoveries.

Recently, a new model for the assembly of spliceosome has been proposed by Stevens et al. They believe that in the nucleus there is a massive ribonucleoprotein complex, containing five U snRNPs (U1, U2, U4 U5 and U6), called penta-snRNP (91). The facts which support that such a complex exists are: a) the snRNPs appear to be in stoichiometic proportions and b) the proteins found in this purification are specific. This complex contains all the snRNP proteins and 13 known splicing associated proteins. In a functional study, Stevens et al. (91) have shown evidence that the penta-snRNP complex is significant for pre-mRNA splicing. Extracts depleted of endogenous snRNAs by

nuclease digestion can be restored of their splicing activity by adding back the isolated penta-snRNP. Moreover, when they used tagged snRNAs for a mixing experiment, they found that during spliceosome assembly the components of the snRNPs do not exchange with each other, which strongly supports the idea that the penta-snRNP works as a pre-existing unit for splicing. Stevens et al. also claimed that the difference between this model and the traditional one may be due to different isolation conditions (91). Although no other significant developments have been reported recently, this model is still considered as an innovative point of view toward spliceosome assembly.

Besides the major splicing complexes, there is a minor class of introns using spliceosomes containing U11, U12 and U5 snRNPs. Unlike the GT-AG rule of the vast majority of introns, this minor class of introns contains AT and AC at their 5' and 3' ends, respectively. In this spliceosome, U11 and U12 resemble the U1 and U2 roles in the major spliceosome, respectively. U5 also functions in this spliceosome, but there is no evidence to suggest the involvement of U4 or U6 snRNPs (92, 93, 94).

As has been stated, besides snRNPs, there are many other components in the spliceosome. Their roles in splicing have not been completely studied. It is generally believed that they facilitate the assembly of the snRNPs and stabilize the spliceosome as an active complex. In recent years, using proteomic studies, many of the spliceosome components have been determined and analyzed.

Proteomic studies of spliceosomes

For a long time, it has been known that spliceosomes are large and complicated complexes. In early studies, researchers used depletion and reconstitution and budding yeast genetic analysis to identify the components of these complexes. But until recent

years, no one has shown a complete picture of the spliceosome. In the past decade, a new method has been used to analyze the spliceosome. It is tandem mass spectrometry. The first splicing complexes analyzed by mass spectrometric methods were the hnRNP complex assembled on mammalian pre-mRNA (95) and subunits of the yeast spliceosome (96, 97). Subsequently, large-scale analysis of human multi-protein complexes on *in vitro* assembled spliceosomes was performed. The first step of such analysis is to isolate the spliceosome from the cell nuclear extract. Different ways of purification have been exploited.

In 1999, Fischer et al. (98) used antibodies against the trimethylguanosine (m3G)-cap to immunoaffinity-purify snRNPs from HeLa nuclear extracts and separated the complexes on 10 to 30% glycerol gradients. Fractions containing different snRNP complexes were pooled and selected with biotinylated 2'-O-methyl oligonucleotides and streptavidin agarose. Thus, these studies provided a thorough analysis of the snRNP complexes.

In 2002, using a similar method, Rappsilber et al. (99) assembled a mixture of spliceosomal complexes on biotinylated, radioactively labeled RNA from adenovirus (AD1) and β-globin (AL4) transcripts. Samples were subjected to gel filtration and affinity selection using streptavidin beads. The peptides bound to the high-performance liquid chromatography (HPLC) column were analyzed.

In another proteomic study, Zhou et al. (100) also used two distinct pre-mRNAs (adenovirus major late, AdML, and Fushi tarazu, Ftz) to form a mixture of all stages of spliceosomes, including complexes that had undergone the first or second catalytic steps of the splicing reaction. In this study, they used a different purification approach. The

substrate pre-mRNA contained three hairpin structures which bind to the MS2 bacteriophage coat protein. Using an MS2-MBP fusion protein (MBP is maltose-binding protein), these complexes could be used for affinity purification.

With the purification of different stage spliceosomes, the next step was to subject them to mass spectrometric analysis. Mass spectrometric methods and human sequence databases have continued to improve in recent years, allowing both increased sensitivity and higher throughput. It is now possible to analyze mixtures of hundreds or even thousands of peptides by liquid chromatography coupled to tandem mass spectrometry (LC MS/MS). Improved software and databases containing most human genes are also now available, allowing automated data processing of the large volume of acquired mass spectra (101).

These proteomic studies have shown a large number of protein components of the spliceosomes. In Reed's study, a total of 145 distinct proteins were identified as shared between the two types of pre-mRNA spliceosomes, including 88 known splicing factors/snRNP proteins/spliceosomal proteins (100). Among these 88 known proteins, 41 are distinct proteins found in U1, U2, U4, U5 and U6 snRNPs. The other 47 proteins include members of the SR family of splicing factors, DExD box splicing factors, proteins that are found in the salt-stable spliceosome 'core', second catalytic step splicing factors and known components of the spliced messenger ribonucleoprotein (mRNP) that promote mRNA transport to the cytoplasm. Interestingly, a number of factors involved in genetic diseases, like retinitis pigmentosa (U4/U6-61K, U4/U6-90K, U5-220K) or cancer (WTAP), are also associated with spliceosomes. Besides these known splicing factors,

they also found 58 proteins that have not been previously identified as spliceosomal proteins in humans. 36 of these proteins have previously been named and the rest are uncharacterized. Some of these 58 proteins have been suggested by others to be involved in splicing. One example is cyclophilin-like proteins. Recent studies revealed a role for one cyclophilin (U4/U6-20K) in splicing, and these data suggest that cyclophilins may have a general function in mediating conformational changes in the spliceosome. The fact that some of these proteins contain domains which resemble RNA recognition motifs or DExD boxes, are known to co-localize with splicing factors in nuclear speckles and/or to localize in the nucleus and many of the proteins have known yeast counterparts that function in mRNA processing strongly surpport they are also splicing related proteins.

Mann's study also revealed the same type of protein catalog of the spliceosomes, only with a larger number of components (99). A total of 311 proteins were identified. More than 40 percent of the identified proteins have a known function related to splicing or RNA processing. The first group is the core snRNP proteins (Sm proteins), all that are present in U1, U2, U4, and U5 snRNPs were identified; six of the seven Lsm proteins that substitute for Sm proteins in the U6 snRNP were also detected. The second group is 39 non-snRNP protein splicing factors, including early-acting factors such as splicing factor 1 (SF1) and late factors such as Slu7(required for the second catalytic step of the splicing reaction) and Aly (for RNA export). The third group is 20 hnRNP proteins which are defined by a common RNA-binding motif and are thought to have diverse functions in RNA protection and processing. Some hnRNPs are known to be splicing factors, such as GRY-RBP/hnRNP Q. The fourth group is 27 other RNA-processing-related proteins. These proteins have functions closely associated with splicing, such as 5' cap binding

proteins (CBP) 20, CBP 80 and the 3' end cleavage and polyadenylation factor (CPSF). This research team also found a large number of unknown novel proteins in their analysis. Compared to their previous study, they found 96 new proteins associated with spliceosomes. Many of these proteins have a domain structure suggesting their role in splicing/RNA processing -- 55 proteins have sequences similar to known splicing factors or domains that implicate them in RNA processing. They also found a number of proteins involved in transcription and cellular regulatory mechanisms, indicating some form of coupling of these processes to splicing.

Although proteomic analysis is a powerful tool to analyze spliceosome, it is important to notice that not all the spliceosome components would be detected. If a protein was small or not abundant, it would generate few detectable tryptic peptides, and thus not be able to be detected by this analysis. Moreover, since most of the proteomic studies use very stringent conditions for complex purification, if a component is not tightly associated with the spliceosome, it is likely this protein will be lost during the purification step.

Part C Galectin involvement in splicing

Galectins have been known for many years. But the intracellular function of galectins has only been recognized in recent years. The splicing activity of galectin-1 and galectin-3 was first hinted by cesium sulfate density gradient centrifugation. In this experiment, galectin-3 was found not only as a free protein, but also co-sedimenting with hnRNP and snRNP complexes, which are related to pre-mRNA splicing. In another experiment, when treated with ribonuclease, the immunoflurescence staining pattern of galectin-3 associated with the nuclear structures was abolished, indicating galectin-3 is associated with ribonucleoprotein complexes. The third line of evidence is that confocal immunofluorescence microscopy shows galectin-3 and galectin-1 co-localize with Sm core proteins and SC35, a non-snRNP splicing factor, in a subnuclear region called speckles. Speckles are the storage sites for splicing factors (23, 103). All of these results point to the possibility that galectin-3 is a splicing factor.

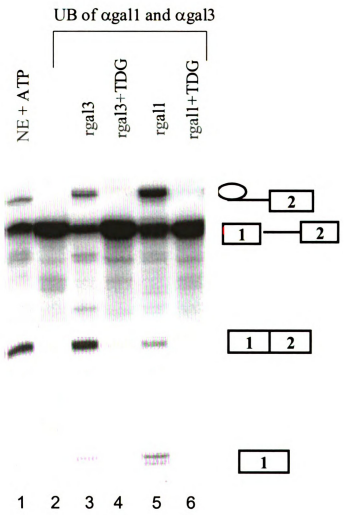
To confirm galectins as splicing factors, functional assays have been performed. Two types of assays were employed. The first assay is using denaturing gel electrophoresis to analyze spliceosomal substrate RNA species. During the splicing reaction, pre-mRNA will be converted into mature mRNA. Together with intermediate RNA species, they will be resolved by this gel system (102). The second way of analysis is to form different splicing complexes and resolve these complexes by native gel electrophoresis. An ideal assay will show H/E, A, B and C complexes (102). The role of galectins in splicing can thus be illustrated by saccharide inhibition and depletion-reconstitution experiments.

HeLa cell nuclear extract (NE) is capable of carrying out *in vitro* splicing reactions. When saccharides with high galectin binding affinity (lactose or TDG) were added, the *in vitro* splicing reaction was inhibited. This phenomenon was not observed when saccharides without affinity for galectins were added. This suggests a role of galectins in splicing. In the process of depletion and reconstitution, removal of galectins from NE, either by lactose affinity chromatography or by antibody-mediated depletion, abolished the splicing activity of the nuclear extract. When galectin-1 or galectin-3 was added to the depleted extract, the splicing activity was restored (Figure 3), confirming galectin-1 and -3 are bona fide splicing factors and their roles in splicing are redundant (103).

Recently, the discovery of a galectin-1 protein ligand gives support for the above conclusion. Using a yeast two hybrid analysis, a binding partner of galectin-1, Gemin-4, was isolated from a HeLa cell cDNA library. The direct interaction between galectin-1 and Gemin-4 was also confirmed by a GST pull down experiment (37). Gemin-4 is a component of the SMN complex. The SMN complex is responsible for the assembly of snRNPs in the cytoplasm and the recycling of snRNPs in the nucleus after one round of splicing (104). Thus, a link between galectin-1/-3 and snRNP complexes was established. The most recent discovery in this area is the association of nuclear galectins with a transcription factor TF II-I. Since pre-mRNA splicing appears to occur both co-transcriptionally and post-transcriptionally, this observation may suggest a linkage role of galectins in mRNA processing.

Figure 3: Depletion and reconstitution of splicing activity of nuclear extract with galectin-1 and -3. Nuclear extract lost its splicing activity through antibody-mediated depletion (lane 2). When recombinant galectin-1 or galectin-3 were added back (lanes 3 and 5), the splicing activity of the depleted nuclear extract was restored. This reconstitution was inhibitable by thiodigalactoside (TDG) (lanes 4 and 6) (103).

Figure 3


UB of αgal1 and αgal3
NE + ATP
rgal3
rgal3+TDG
rgal1
rgal1+TDG
2
1
2
1
2
1
1 2 3 4 5 6

Although these results have suggested that galectins are splicing factors, it must be acknowleged that mutant mice with galectin-1 and galectin-3 genes knocked out are viable and fertile (though some physiological changes could be observed). This indicates that galectins' function could be replaced by other proteins in a living system.

References

1. Barondes, S. H., V. Castronovo, D. N. W. Cooper, R. D. Cummings, K. Drickamer, T. Feizi, M. A. Gitt, J. Hirabayashi, C. Hughes, K. Kasai. 1994. Galectins: a family of animal beta-galactoside-binding lectins. *Cell* 76: 597-598.

2. Hirabayashi, J. 1997. Recent Topics on Galectins. Trends In Glycosci. *Glycotechnol.* 9: 1-184.

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

4. Barondes, S.H., D.N.W. Cooper, M.A. Gitt, and H. Leffler. 1994. Galectins: Structure and function of a large family of animal lectins. *J. Biol. Chem.* 269: 20807-20810.

5. Gitt, M.A., C. Colnot, F. Poirier, K.J. Nani, S.H. Barondes, H. Leffler. 1998. Galectin-4 and galectin-6 are two closely related lectins expressed in mouse gastrointestinal tract. *J. Biol. Chem.* 273: 2954-2960.

6. Wang, J.L., R. Gray, K. Haudek, R.J. Patterson. 2004. Nucleocytoplasmic lectins. *Biochim. Biophys. Acta.* 1673: 75-93.

7. Dunphy, J.L., A. Balic, G.J. Barcham, A.J. Horvath, A.D. Nash, E.N.T. Meeusen. 2000. Isolation and Characterization of a Novel Inducible Mammalian Galectin, *J. Biol. Chem.* 275: 32106-32113.

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

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

10. Kopitz, J., S. Andre, and C. Von Reitzenstein, et al. 2003. Homodimeric galectin-7 (p53-induced gene 1) is a negative growth regulator for human neuroblastoma cells. *Oncogene* 22: 6277-6288.

11. Levy, Y., D. Ronen, A.D. Bershadsky, Y. Zick. 2003. Sustained induction of ERK, protein kinase B, and p70 S6 kinase regulates cell spreading and formation of F-actin microspikes upon ligation of integrins by galectin-8, a mammalian lectin. *J. Biol. Chem.* 278: 14533-14542.

12. Wada, J., K. Ota, A. Kumar, E.I. Wallner, Y.S. Kanwar. 1997. Developmental regulation, expression, and apoptotic potential of galectin-9, a beta-galactoside binding lectin. *J. Clin. Invest.* 99: 2452-2461.

13. Leal-Pinto, E., W. Tao, J. Rappaport, M. Richardson, B.A. Knorr, R.G. Abramson. 1997. Molecular cloning and functional reconstitution of a urate transporter/channel. *J. Biol. Chem.* 272: 617-625.

14. Leonidas, D.D., B.L. Elbert, Z. Zhou, H. Leffler, S.J. Ackerman, K.R. Acharya. 1995. Crystal structure of human Charcot-Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins. *Structure* 3: 1379-1393.

15. Ackerman, S.J., L. Liu, M.A. Kwatia, M.P. Savage, D.D. Leonidas, G.J. Swaminathan, K.R. Acharya. 2002. Charcot-Leyden crystal protein (Galectin-10) is not a dual function galectin with lysophospholipase activity but binds a lysophospholipase inhibitor in a novel structural fashion. *J. Biol. Chem.* 277: 14859-14868.

16. Hotta, K., T. Funahashi, Y. Matsukawa, M. Takahashi, H. Nishizawa, K. Kishida, M. Matsuda, H. Kuriyama, S. Kihara, T. Nakamura, Y. Tochino, N.L. Bodkin, B.C. Hansen, Y. Matsuzawa. 2001. Galectin-12, an adipose-expressed galectin-like molecule possessing apoptosis-inducing activity. *J. Biol. Chem.* 276: 34089-34097.

17. Than, N.G., B. Sumegi, G.N. Than, Z. Berente, H. Bohn. 1999. Isolation and sequence analysis of a cDNA encoding human placental tissue protein 13 (PP13), a new lysophospholipase, homologue of human eosinophil Charcot-Leyden Crystal protein. *Placenta* 20: 703-710.

18. Dunphy, J.L., G.J. Barcham, R.J. Bischof, A.R. Young, A. Nash, E.N.T. Meeusen. 2002. Isolation and Characterization of a Novel Eosinophil-specific Galectin Released into the Lungs in Response to Allergen Challenge. *J. Biol. Chem.* 277: 14916-14924.

19. Gray, C. A., K. A. Dunlap, R.C. Burghardt, T. E. Spencer. 2005. Galectin-15 in ovine uteroplacental tissues. *Reproduction* 130: 231-240.

20. Clerch, L.B., P. Whitney, M. Hass, K. Brew, T. Miller, R. Werner, D. Massaro. 1988. Sequence of a full-length cDNA for rat lung beta-galactoside-binding protein: primary and secondary structure of the lectin. *Biochemistry* 27: 692-699.

21. Cooper, D.N.W., S.H. Barondes. 1990. Evidence for export of a muscle lectin from cytosol to extracellular matrix and for a novel secretory mechanism. *J. Cell Biol.* 110: 1681-1691.

22. Harrison, F.L., T.J. Wilson. 1992. The 14 kDa beta-galactoside binding lectin in myoblast and myotube cultures: localization by confocal microscopy. *J. Cell Sci.* 101: 635-646.

23. Vyakarnam, A., A.J. Lenneman, K.M. Lakkides, R.J. Patterson, J.L. Wang. 1998. A comparative nuclear localization study of galectin-1 with other splicing components. *Exp. Cell Res.* 242: 419-428.

24. Akimoto, Y., J. Hirabayashi, K. Kasai, H. Hirano. 1995. Expression of the endogenous 14-kDa beta-galactoside-binding lectin galectin in normal human skin. *Cell Tissue Res.* 280: 1-10.

25. Choi, J.Y., A.J. van Wijnen, F. Aslam, J.D. Leszyk, J.L. Stein, G.S. Stein, J.B. Lian, S. Penman. 1998. Developmental association of the beta-galactoside-binding protein galectin-1 with the nuclear matrix of rat calvarial osteoblasts. *J. Cell Sci.* 111: 3035-3043.

26. Siebert, H.C., S. Andre, S.Y. Lu, M. Frank, H. Kaltner, J.A. van Kuik, E.Y. Korchagina, N. Bovin, E. Tajkhorshid, R. Kaptein, J.F. Vliegenthart, C.W. von der Lieth, J. Jimenez-Barbero, J. Kopitz, H.J. Gabius . 2003. Unique conformer selection of human growth-regulatory lectin galectin-1 for ganglioside GM1 versus bacterial toxins. *Biochemistry* 42: 14762-14773.

27. Paz, A., R. Haklai, G. Elad-Sfadia, E. Ballan, Y. Kloog. 2001. Galectin-1 binds oncogenic H-Ras to mediate Ras membrane anchorage and cell transformation. *Oncogene* 20: 7486-7493.

28. Prior, I.A., C. Muncke, R.G. Parton, J.F. Hancock. 2003. Direct visualization of Ras proteins in spatially distinct cell surface microdomains. *J. Cell Biol.* 160:165-170.

29. Rini, J.M., Y.D. Lobsanov. 1999. New animal lectin structures. *Curr. Opin. Struct. Biol.* 9: 578-584.

30. Lopez-Lucendo, M.F., D. Solis, S. Andre, J. Hirabayashi, K. Kasai, H. Kaltner, H.J. Gabius, A. Romero. 2004. Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *J. Mol. Biol.* 343: 957-970.

31. Liao, D. I., G. Kapadia, , H. Ahmed, , G. R. Vasta, , O. Herzberg. 1994. Structure of S-lectin, a developmentally regulated vertebrate β-galactoside-binding protein. *Proc. Natl. Acad. Sci.* 91: 1428-1432.

32. Puche, A.C., B. Key. 1995. Identification of cells expressing galectin-1, a galactose-binding receptor, in the rat olfactory system. *J. Comp. Neurol.* 357: 513-523.

33. Gonzalez, M.I., N. Rubinstein, J.M. Ilarregui, M.A. Toscano, N.A. Sanjuan, G.A. Rabinovich. 2005. Regulated expression of galectin-1 after in vitro productive infection with herpes simplex virus type 1: implications for T cell apoptosis. *Int. J. Immunopathol Pharmacol.* 18: 615-623.

34. Gu, M., W. Wang, W.K. Song, D.N. Cooper, S.J. Kaufman. 1994. Selective modulation of the interaction of alpha 7 beta 1 integrin with fibronectin and laminin by L-14 lectin during skeletal muscle differentiation. *J. Cell Sci.* 107: 175-181.

35. Elad-Sfadia, G., R. Haklai, E. Ballan, H.J. Gabius, Y. Kloog. 2002. Galectin-1 augments Ras activation and diverts Ras signals to Raf-1 at the expense of phosphoinositide 3-kinase. *J. Biol. Chem.* 277: 37169-37175.

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

37. Park, J.W., P.G. Voss, S. Grabski, J.L. Wang, R.J. Patterson. 2001. Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. *Nucleic Acids Res.* 29: 3595-3602.

38. Roff, C.F., P.R. Rosevear, J.L. Wang, R. Barker. 1983. Identification of carbohydrate-binding proteins from mouse and human fibroblasts. *Biochem. J.* 211: 625–629.

39. Agrwal, N., Q. Sun, S.Y. Wang, J.L. Wang. 1993. Carbohydrate-binding protein 35. I. Properties of the recombinant polypeptide and the individuality of the domains. *J. Biol. Chem.* 268: 14932-14939.

40. Yang, R.Y., P.N. Hill, D.K. Hsu, F.T. Liu. 1998. Role of the carboxyl-terminal lectin domain in self-association of galectin-3. *Biochemistry* 37: 4086-4092.

41. Hsu, D.K., R.I. Zuberi, F.T. Liu. 1992. Biochemical and biophysical characterization of human recombinant IgE- binding protein, an S-type animal lectin. *J. Biol. Chem.* 267: 14167-14174.

42. Mehul, B., S. Bawumia, S.R. Martin, R.C. Hughes. 1994. Structure of baby hamster kidney carbohydrate-binding protein CBP30, an S-type animal lectin. *J. Biol. Chem.* 269: 18250-18258.

43. R. C. Hughes. 1999. Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim. Biophys. Acta.* 1473: 172-185.

44. Davidson, P.J., M.J. Davis, R.J. Patterson, M.A. Ripoche, F. Poirier, J.L. Wang. 2002. Shuttling of galectin-3 between the nucleus and cytoplasm. *Glycobiology* 12: 329-337.

45. Moutsatsos, I.K., M. Wade, M. Schindler, J.L. Wang. 1987. Endogenous lectins from cultured cells: nuclear localization of carbohydrate-binding protein 35 in proliferating 3T3 fibroblasts. *Proc. Natl. Acad. Sci. U.S.A.* 84: 6452-6456.

46. Takenaka, Y., T. Fukumori, T. Yoshii, N. Oka, H. Inohara, H.R. Kim, R.S. Bresalier, A. Raz. 2004. Nuclear export of phosphorylated galectin-3 regulates its antiapoptotic activity in response to chemotherapeutic drugs. *Mol. Cell. Biol.* 24: 4395-4406.

47. Mazurek, N., J. Conklin, J.C. Byrd, A. Raz, R.S. Bresalier. 2000. Phosphorylation of the ß-galactoside-binding protein galectin-3 modulates binding to its ligands. *J. Biol. Chem.* 275: 36311–36315.

48. Yoshii, T., T. Fukumori, Y. Honjo, H. Inohara, H.R. Kim, A. Raz. 2002. Galectin-3 phosphorylation is required for its anti-apoptotic function and cell cycle arrest. *J. Biol. Chem.* 277: 6852–6857.

49. Liu, F.T., R.J. Patterson, J.L. Wang. 2002. Intracellular functions of galectins. *Biochim. Biophys. Acta.* 1572: 263–273.

50. Paron I., A. Scaloni, A. Pines, A. Bachi, F.T. Liu, C. Puppin, M. Pandolfi, L. Ledda, C. Di Loreto, G. Damante, G. Tell. 2003. Nuclear localization of Galectin-3 in transformed thyroid cells: a role in transcriptional regulation, *Biochem. Biophys. Res. Cummun.* 302: 545-553.

51. Pricci, F., G. Leto, L. Amadio, C. Iacobini, G. Romeo, S. Cordone, R. Gradini, P. Barsotti, F.T. Liu, U. Di Mario, G. Pugliese. Role of galectin-3 as a receptor for advanced glycosylation end products. *Kidney Int.* 77.

52. Kim, H.R., H.M. Lin, H. Biliran, A. Raz. 1999. Cell cycle arrest and inhibition of anoikis by galectin-3 in human breast epithelial cells. *Cancer Res.* 59: 4148–4154.

53. Inohara, H., S. Akahani, A. Raz. 1998. Galectin-3 stimulates cell proliferation. *Exp. Cell Res.* 245: 294–302.

54. Hamann, K.K., E.A. Cowles, J.L. Wang, R.L. Anderson. 1991. Expression of carbohydrate binding protein 35 in human fibroblasts: Variations in the levels of mRNA, protein, and isoelectric species as a function of replicative competence. *Exp. Cell Res.* 196: 82–91.

55. Yang, R.Y., D.K. Hsu, F.T. Liu. 1996. Expression of galectin-3 modulates T-cell growth and apoptosis. *Proc. Natl. Acad. Sci. U S A* 93: 6737–6742.

56. Akahani, S., P. Nangia-Makker, H. Inohara, H.R. Kim, A. Raz. 1997. Galectin-3: A novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family. *Cancer Res.* 57: 5272–5276.

57. Thery, C., M. Boussac, P. Veron, P. Ricciardi-Castagnoli, G. Raposo, J. Garin, S. Amigorena. 2001. Proteomic analysis of dendritic cell-derived exosomes: A secreted subcellular compartment distinct from apoptotic vesicles. *J. Immunol.* 166: 7309–7318.

58. Perillo, N.L., M.E. Marcus, L.G. Baum. 1998. Galectins: Versatile modulators of cell adhesion, cell proliferation, and cell death. *J. Mol. Med.* 76: 402–412.

59. Inohara, H., A. Raz. 1995. Functional evidence that cell surface galectin-3 mediates homotypic cell adhesion. *Cancer Res.* 55: 3267–3271.

60. Ochieng, J., M.L. Leite-Browning, P. Warfield. 1998. Regulation of cellular adhesion to extracellular matrix proteins by galectin-3. *Biochem. Biophys. Res. Commun.* 246: 788–791.

61. Kuwabara, I., F.T. Liu. 1996. Galectin-3 promotes adhesion of human neutrophils to laminin. *J. Immunol.* 156: 3939–3944.

62. Frigeri, L.G., R.I. Zuberi, F.T. Liu. 1993. Epsilon BP, a beta-galactoside-binding animal lectin, recognizes IgE receptor (Fc epsilon RI) and activates mast cells. *Biochemistry* 32: 7644–7649.

63. Vlassara, H., Y.M. Li, Y. Imani, D. Wojciechowicz, A. Yang, F.T. Liu, A. Cerami. 1995. Identification of galectin-3 as a high-affinity binding protein for advanced glycation end products (AGE): A new member of the AGE-receptor complex. *Mol. Med.* 1: 634–646.

64. Sharp, P. A. 1994. Split genes and RNA splicing (Nobel Lecture). *Cell* 77, 805-815.

65. Moore, J.M., C.C. Query, P.A. Sharp. 1993. Splicing of precursors to mRNA by the spliceosome. *The RNA World.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.

66. Query, C. C., S. A. Strobel, P. A. Sharp. 1996. Three recognition events at the branch site adenine. *EMBO J.* 15: 1392-1402.

67. Dreyfuss, G., L. Philipson, I. Mattaj. 1988. Ribonucleoprotein particles in cellular processes. *J. Cell Biol.*, 106:1419-1425.

68. Podar, M., P.S. Perlman, R.A. Padgett. 1998. The two steps of group II intron self-splicing are mechanistically distinguishable. *RNA.* 4: 890–900.

69. Cotton, M., O. Gick, A. Vasserot, G. Schaffner, M.L. Birnstiel. 1988. Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the in vitro 3' RNA processing reaction. *EMBO J.* 7: 801–808.

70. Bindereif, A., M.R. Green. 1987. An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. *EMBO J.* 6: 2415–2424.

71. Mattaj, I. W. 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell* 46: 905–911.

72. Bertrand, E., R. Bordonné. 2004. Assembly and traffic of small nuclear RNPs. *Prog. Mol. Subcell. Biol.* 35: 79–97.

73. Darzacq, X., B.E. Jády, C. Verheggen, A. M. Kiss, E. Bertrand, T. Kiss. 2002. Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* 21: 2746–2756.

74. Will, C.L., R. Lührmann. 2001. Spliceosomal UsnRNP biogenesis, structure and function. *Current Opinion in Cell Biology* 13:290-301.

75. Reed, R., L. Palandjian. 1997. Spliceosome assembly. In: A.R. Krainer, Editor, *Eukaryotic mRNA processing*, IRL Press, Oxford, pp. 103–129.

76. Brow, D. A. 2002. Allosteric cascade of spliceosome activation. *Ann. Rev. Genet.* 36: 333-360.

77. Ruby, S.W., J. Abelson. 1988. An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* 242: 1028-1035.

78. Guth, S., C. Martinez, R. Gaur, J. Valcarcel. 1999. Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). *Mol. Cell. Biol.* 19: 8263–8271.

79. Das, R., Z. Zhou, R. Reed. 2000. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol. Cell* 5: 779–787.

80. Rutz, B., B. Seraphin. 1999. Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *RNA* 5:819–831.

81. Graveley, B.R. 2000. Sorting out the complexity of SR protein functions. *RNA* 6: 1197–1211.

82. Umen, J.G., C. Guthrie. 1995. The second catalytic step of pre-mRNA splicing. *RNA* 1: 869–885.

83. Maroney, P.A., C.M. Romfo, T.W. Nilsen. 2000. Functional recognition of the 5′ splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Mol. Cell* 6: 317–328.

84. Tanner, N.K., P. Linder. 2001. DExD/H box RNA helicases: from generic motors to specific dissociation functions *Mol. Cell* 8: 251–262.

85. van Nues, R.W., J.D. Beggs. 2001. Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of Saccharomyces cerevisiae. *Genetics* 157:1451-1467.

86. Kwan, S.S., D.A. Brow. 2005. The N- and C-terminal RNA recognition motifs of splicing factor Prp24 have distinct functions in U6 RNA binding. *RNA* 11: 808-820.

87. Kuhn, A.N., D.A. Brow. 2000. Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. *Genetics* 155: 1667-1682.

88. Bottner, C.A., H. Schmidt, S. Vogel, M. Michele, N.F. Kaufer. 2005. Multiple genetic and biochemical interactions of Brr2, Prp8, Prp31, Prp1 and Prp4 kinase suggest a function in the control of the activation of spliceosomes in Schizosaccharomyces pombe. *Curr. Genet.* 48: 151-161.

89. Awashti, S., R. Palmer, M. Castro, C.D. Mobarak, S.W. Ruby. 2001. New roles for the Snp1 and Exo84 proteins in yeast pre-mRNA splicing. *J. Biol. Chem.* 276: 31004-31015.

90. Ruby S.W. 1997. Dynamics of the U1 small nuclear ribonucleoprotein during yeast spliceosome assembly. *J. Biol. Chem.* 272: 17333-17341.

91. Stevens, S.W., D.E. Ryan, H.Y. Ge, R.E. Moore, M.K. Young, T.D. Lee, J. Abelson. 2002. Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol. Cell* 9: 31-44.

92. Frilander, M.J., J.A. Steitz. 2001. Dynamic exchanges of RNA interactions leading to catalytic core formation in the U12-dependent spliceosome. *Mol. Cell* 7: 217-226.

93. Will, C.L., R. Luhrmann. 2005. Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol. Chem.* 386: 713-724.

94. Tarn, W.Y., J. Steitz. 1996. A novel splicesome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* 84: 801-811.

95. Calvio, C., G. Neubauer, M. Mann, A.I. Lamond. 1995. Identification of hnRNP P2 as TLS/FUS using electrospray mass spectrometry. *RNA* 1: 724-733.

96. Neubauer, G., A. Gottschalk, P. Fabrizio, B. Séraphin, R. Lührmann, M. Mann. 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc. Natl. Acad. Sci.* 94: 385-390.

97. Gottschalk, A., J. Tang, O. Puig, J. Salgado, G. Neubauer, H.V. Colot, M. Mann, B. Seraphin, M. Rosbash, R. Luhrmann. 1998. A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA* 4: 374-393.

98. Fischer, U., R. Luhrmann. 1999. An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. *Science* 249: 786-790.

99. Rappsilber, J., U. Ryder, A.I. Lamond, M. Mann. 2002. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* 12: 1231-1245.

100. Zhou, Z., L.J. Licklider, S.P. Gygi, R. Reed. 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature* 419: 182-185.

101. Griffin, T.J., R. Aebersold. 2001. Advances in proteome analysis by mass spectrometry. *J. Biol. Chem.* 276: 45497-45500.

102. Dagher, S.F., J. L. Wang, R.J. Patterson. 1995. Identification of galectin-3 as a factor in pre-mRNA splicing. *Proc. Natl. Acad. Sci. USA*, 92: 1213-1217.

103. Vyakarnam, A., S. F. Dagher, J. L. Wang, R. J. Patterson. 1997. Evidence of a role for galectin-1 in pre-mRNA splicing. *Mol. Cell. Biol.* 17: 4730-4737.

104. Pellizzoni, L., N. Kataoka, B. Charroux, G. Dreyfuss. 1998. A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* 95: 615-624.

105. Watson J.D., N.H. Hopkins, J.W. Roberts, J.A. Steitz, A.M. Weiner. 2004. Molecular Biology of the Gene, Fifth Edition.

Chapter 2

Immunoprecipitation of spliceosomal RNAs by antisera to galectin-1 and galectin-3

Nucleic Acids Res. 2006 Oct 6; Epub

Abstract

We have shown that galectin-1 and galectin-3 are functionally redundant splicing factors. Now we provide evidence that both galectins are directly associated with spliceosomes by analyzing RNAs and proteins of complexes immunoprecipitated by galectin-specific antisera. Both galectin antisera co-precipitated splicing substrate, splicing intermediates and products in active spliceosomes. Protein factors co-precipitated by the galectin antisera included the Sm core polypeptides of snRNPs, hnRNP C1/C2 and Slu7. Early spliceosomal complexes were also immunoprecipitated by these antisera. When splicing reactions were sequentially immunoprecipitated with galectin antisera, we found that galectin-1-containing spliceosomes did not contain galectin-3 and vice versa, providing an explanation for the functional redundancy of nuclear galectins in splicing. The association of galectins with spliceosomes was (i) not due to a direct interaction of galectins with the splicing substrate and (ii) easily disrupted by ionic conditions that had only a minimal effect on snRNP association. Finally, addition of excess amino terminal domain of galectin-3 inhibited incorporation of galectin-1 into splicing complexes, explaining the dominant negative effect of the amino domain on splicing activity. We

conclude that galectins are directly associated with splicing complexes throughout the splicing pathway in a mutually exclusive manner and they bind a common splicing partner through weak protein-protein interactions.

Introduction

Before transport to the cytoplasm for translation, RNA transcripts in the form of pre-messenger RNA (pre-mRNA) assemble into a macromolecular structure termed the spliceosome. During subsequent remodeling events of spliceosomes, introns are removed and exons are ligated to form mature mRNA. *In vitro* cell-free assays using simple splicing substrates and nuclear extracts (NE) have established the basic sequence of biochemical events and ordered series of complexes in the pathway (1-4). In this dynamic, multi-step process, over 300 proteins and five ribonucleoprotein particles (snRNPs) orchestrate two trans-esterification reactions that result in the formation of mRNA (5-9). We previously identified two polypeptides in NE from HeLa cells (galectin-1 and galectin-3) required for splicing (10, 11). Depletion of both galectins from NE, either by lactose-agarose affinity or double antibody adsorption chromatography, abolished splicing activity and halted spliceosome assembly at an early step. Addition of either galectin restored both splicing activity and spliceosome formation. These data suggested that galectins are, indeed, splicing factors and that they are functionally redundant.

Galectins are a family of carbohydrate-binding proteins with specificity for galactose or galactose-containing glycoconjugates that localize both intracellularly and extracellularly (12). Two of the galectins (galectins-1 and –3) are ubiquitous in mammalian tissues and have been documented in the nucleus and cytosol. Galectin-1 (gal-1) is composed of a single carbohydrate recognition domain (CRD) of ~130 amino

acids. Galectin-3 (gal-3) has an amino terminal domain (ND) of ~130 amino acids containing multiple repeats of the amino acid sequence PGAYPGXXX of unknown function fused to a CRD whose sequence is conserved when compared to the CRD of gal-1 and the other galectins. As we investigated the role of gal-1 and gal-3 in splicing, we noted several key differences. First, recombinant gal-3 is ~5-10 times more efficient at reconstitution of splicing than gal-1 or the isolated CRD of gal-3 (11). Second, although both galectins co-localize with known splicing factors (i.e., the Sm core polypeptides of snRNPs and SC35) and each other in nuclear speckles, there are regions of non-overlap (13). Lastly, the isolated ND of gal-3 inhibits splicing chemistry and spliceosome formation in a dominant negative manner when added to splicing competent NE (14).

Although these data are consistent with galectins being splicing factors, they do not address whether their role in splicing is direct or indirect. In this study, we provide evidence for a direct role of galectins in the splicing pathway by showing that galectin-specific antibodies co-immunoprecipitate pre-mRNA, splicing intermediates and products throughout the entire splicing pathway. We have investigated various characteristics of galectin incorporation into spliceosomes and conclude they interact via weak protein-protein interactions with another spliceosomal component.

Materials and Methods

Antibodies

Polyclonal rabbit antiserum against recombinant rat gal-1 was a gift from Doug Cooper (University of California, San Francisco) (14). Additionally, we raised a second

polyclonal rabbit antiserum against recombinant human gal-1 expressed as a fusion protein with glutathione S-transferase (GST). A human gal-1 cDNA clone was kindly provided by Jun Hirabayashi (15). The gal-1 cDNA insert was isolated after BamH1 digestion and inserted into pGEX-2T. Following expression in DH5α, the fusion protein was purified by glutathione-agarose affinity chromatography and used to immunize rabbits. Several polyclonal rabbit antisera to murine gal-3 were produced as described in Agrwal et al. (16). When used in immunoprecipitation experiments (Figures 4 and 6), each galectin antiserum was mono-specific for a galectin (i.e., anti-gal-1 precipitated only gal-1 from splicing reactions and anti-gal-3 precipitated only gal-3). Human autoimmune serum (ENA anti-Sm) was purchased from The Binding Site. Anti-Slu7 antibodies were purchased from Santa Cruz Biotechnology and anti-Ran antibodies from BD Biosciences. A monoclonal antibody against hnRNP C1/C2 (4F4) was provided by Gideon Dreyfuss (University of Pennsylvania). For immunoprecipitation experiments, all antibodies were covalently cross-linked to protein G-Sepharose fast flow 4B beads (Sigma) as previously described (11, 14) generally using a 2:1 ratio of antiserum to protein G beads.

Co-immunoprecipitation of ^{32}P-labeled-MINX RNAs

Splicing reactions containing 60% by vol. HeLa NE were assembled with ^{32}P-MINX RNA without or with ATP (1.5 mM) and creatine phosphate (20 mM) and incubated at 30°C for the times indicated as previously described (10, 11, 14). HeLa cells were obtained from the National Cell Culture Center. Typically 60–100 μl of the splicing reaction mixture was diluted to 0.5 ml with 60% buffer D (D60) and incubated with 30-50 μl antibody cross-linked protein G-Sepharose beads at 4°C for 1 - 2 hr. After washing with D60 containing 0.05% Triton X-100 (three washes, each with 1 ml buffer), the

bound material was eluted with 20 μl of SDS-PAGE sample buffer. The eluted sample was divided into two aliquots. RNA was purified from one by incubating at 37°C for 20 min with proteinase K (4 mg/ml final concentration) and diluting to 100 μl with 125 mM Tris (pH 8), 1 mM EDTA, 300 mM sodium acetate. RNA was extracted with 200 μl of phenol-chloroform (50:50 [vol/vol]), followed by 200 μl of chloroform. RNA was precipitated with 300 μl of ethanol at –80°C. The extracted RNA was dissolved in 10 μl of sample buffer (9:1 / formamide: bromophenol blue) and subjected to electrophoresis through 13% polyacrylamide (bisacrylamide-acrylamide 1.9:50 [wt/wt])-8.3 M urea gels. The radioactive RNA species were revealed by autoradiography and quantitated by phosphorimage analysis (Molecular Dynamics). The other aliquot was subjected to electrophoresis in 12.5% or 15% SDS-PAGE (bisacrylamide-acrylamide 0.9:30 [wt/wt]) and analyzed by western blotting.

The salt sensitivity of the association of galectins and snRNPs with spliceosomes was determined as follows. Active splicing complexes (incubated for 60 min with ATP at 30°C) were incubated with antibody-coupled beads. Following the initial binding, the beads were washed either with 60, 130 or 250 mM KCl-containing buffer and the bound material was eluted with SDS-PAGE sample buffer. The RNA was extracted and analyzed by denaturing gel electrophoresis as described above.

Native gel electrophoresis

To evaluate binding of the galectins to the splicing substrate, recombinant gal-1 or recombinant gal-3 (expressed in DH5α and purified by lactose-agarose affinity chromatography), or NE was incubated with ^{32}P-MINX under splicing conditions.

Following incubation, the reactions were subjected to native gel electrophoresis (4% polyacrylamide gels [bisacrylamide:acrylamide 1:80 wt/wt]) (17) and complex formation revealed by autoradiography. Agarose native gel electrophoresis to identify H or E complexes was performed as described by Das et al. (18) followed by autoradiography.

Results

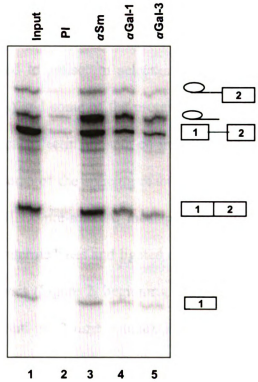
Galectins-1 and –3 are associated with spliceosomes

Previously we showed that depletion of gal-1 and gal-3 from HeLa NE abolished pre-mRNA splicing and arrested spliceosomes before formation of active complexes (i.e., only early complexes formed) (10). Addition of either recombinant gal-1 or gal-3 to the galectin-depleted NE resulted in reconstitution of splicing activity and active spliceosome formation (10, 11). We concluded that these two nuclear galectins are redundant splicing factors. However, there was little information to suggest whether their involvement in the splicing pathway was direct or indirect.

To test for a direct association of galectins with spliceosomes, we determined whether radiolabeled splicing substrate in splicing complexes could be precipitated by galectin-specific antisera. As seen in Figure 1 (lane 1), after incubation in HeLa NE with ATP for 60 min, the ^{32}P-MINX splicing substrate is converted to processing intermediates and products. Typically, 20-35 % of the RNA detected following 60 min is ligated product. In the absence of ATP, no processing of the MINX substrate occurs (see below). When splicing reactions that had been incubated for 60 min are immunoprecipitated by galectin specific antisera, all species of RNA, the pre-mRNA

Figure 1: Analysis of spliceosomal RNA species immunoprecipitated by various antisera. Equal aliquots of splicing reactions incubated for 60 min with ^{32}P-MINX and ATP were subjected to antibody adsorption and the bound RNA analyzed by denaturing gel electrophoresis and autoradiography. Lane 1, RNA in the splicing reaction subjected to immunoprecipitation; lane 2, RNA bound to pre-immune serum; lane 3, RNA bound to anti-Sm; lane 4, RNA bound to anti-gal-1 and lane 5, RNA bound to anti-gal-3. The positions of the various RNA species are diagrammed on the right in this and subsequent figures analyzing spliceosomal RNAs.

Figure 1



substrate, the intermediates and the mature products, are found in the precipitates (lanes 4 and 5). In contrast, pre-immune serum did not precipitate significant quantities of labeled RNA (lane 2). Human autoimmune serum reactive against the Sm polypeptides of the snRNPs precipitated all species of radiolabeled spliceosomal-associated MINX as expected (lane 3). Thus, gal-1 and gal-3, like Sm proteins, are associated with active spliceosomes.

To characterize in greater detail galectin-containing splicing complexes, splicing reactions were subjected to antiserum selection during a time course experiment and RNAs (Figures 2 and 3) and some of the proteins (Figure 4) in the immunoprecipitates were characterized. Each galectin antisera co-immunoprecipitated the splicing substrate throughout the time course of the splicing reaction. At the earliest times sampled, both antisera precipitated MINX pre-mRNA, most probably in the form of H/E complexes (see below). Splicing intermediates and ligated exons were precipitated as they appeared during the kinetic analysis (Figure 2, beginning at 20 min for the anti-gal-1 time course and at 40 min for the anti-gal-3 time course). An internal control for the specificity of RNA precipitation is apparent in these analyses. Degraded RNAs observed at 0, 5 and 10 min (highlighted by arrows to the left of the input lanes in Figure 2) are not detected in the immunoprecipitated complexes at these times. Both galectin antisera appeared to immunoprecipitate the excised lariat RNA species preferentially (compare the ratios of lariat to pre-mRNA species in the precipitates and in the input at the 40 and 60 min time points in Figure 2). Both observations (i.e., no precipitation of degraded RNAs and preferential precipitation of free lariat) argue against non-specific adsorption of radioactive RNA species to beads since the precipitated RNAs do not reflect the same

Figure 2: Time course analysis of spliceosomal RNAs precipitated by anti-gal-1 or anti-gal-3 antiserum. ^{32}P-MINX was incubated with ATP in NE at 30°C. At the times indicated, the reactions were subjected to galectin affinity adsorption and the bound fraction separated into two parts. RNA was extracted from one portion and analyzed by denaturing gel electrophoresis. MINX RNA in splicing reactions at 0, 5, 10, 20, 40 and 60 min (input) and MINX RNA precipitated by anti-gal-1 or anti-gal-3 antiserum. Arrows at left indicate degraded RNA species in the input material.

Figure 2

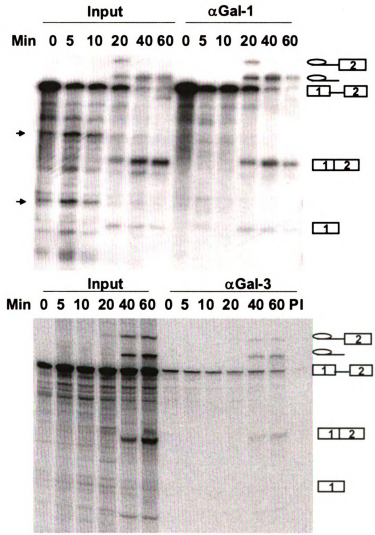

Input
αGal-1
Min 0 5 10 20 40 60 0 5 10 20 40 60
2
1 2
1 2
1
Input
αGal-3
Min 0 5 10 20 40 60 0 5 10 20 40 60 PI
2
1 2
1 2
1

Figure 3: Immunoprecipitation of ^{32}P-labeled RNA assembled into early (H and E) complexes. Panel A: MINX RNA was incubated with NE at 30°C in the absence of ATP for the indicated times. The reaction mixtures were loaded directly onto 1.5% agarose gels. Following electrophoresis, autoradiography was performed to detect RNA-containing complexes. Panel B: Splicing reactions containing MINX RNA in the absence of ATP at 0, 20, 40 and 60 min incubation at 30°C (lanes 1-4) were incubated with anti-gal-1 immobilized and the bound material subjected to analysis by denaturing gel electrophoresis (lanes 5-8).

Figure 3

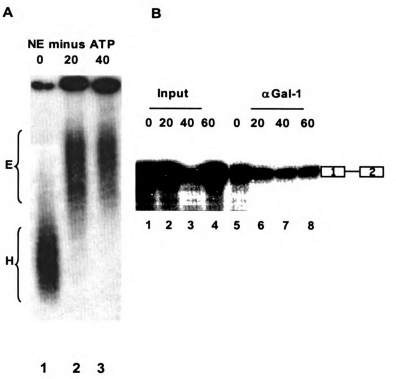
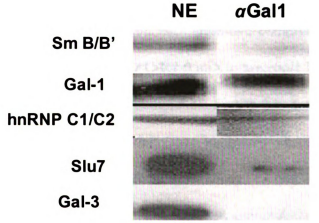

A
NE minus ATP
0 20 40
E
H
1 2 3
B
Input
αGal-1
0 20 40 60 0 20 40 60
1 2 3 4 5 6 7 8

Figure 4: Analysis of proteins co-immunoprecipitated by gal-1 antiserum. HeLa NE (NE) and bound fractions from anti-gal-1 immunoprecipitates (αG1) were blotted for Sm B/B', hnRNP C1/C2, Slu7, gal-1 and gal-3.

Figure 4



relative amounts of the different RNA species in the sample used for immunoselection (input).

Detecting MINX pre-mRNA in galectin immunoprecipitates at early times in the splicing reaction prompted us to investigate the association of galectins with complexes characterized as H or E splicing complexes. H complexes form upon addition of a substrate to a splicing extract even when incubated on ice. Following incubation at elevated temperatures, H complexes are converted to E (early) complexes that are the immediate precursors of active spliceosomes. Neither H nor E complex assembly requires ATP (18). To test for galectin association with these two complexes, MINX pre-mRNA was incubated without ATP in NE at 30°C for 0 min to assemble H complexes or 20 and 40 min to chase H complexes into E complexes. Figure 3A shows MINX pre-mRNA formed H complexes upon addition to NE (lane 1). Nearly all of the MINX RNA was chased into E complexes following incubation for 20 and 40 min at 30°C (lanes 2 and 3). In a time course experiment, anti-gal-1 antiserum immunoprecipitated splicing competent pre-mRNA assembled in H complexes at 0 time (Figure 3B, lane 5) and E complexes after 20 through 60 min of incubation at 30°C (Figure 3B, lanes 6-8). Thus, complexes formed in the absence of ATP with mobilities characteristic of H and E pre-splicing complexes contain gal-1. Similar results for gal-3 were obtained (data not shown).

To confirm further that the immunoprecipitated complexes represented spliceosomes, antisera specific for several splicing factors were used to probe the precipitated fractions (Figure 4). Antiserum specific for gal-1 co-precipitated the Sm B/B' core polypeptides of snRNPs, hnRNP C1/C2 and a factor required for in the second transesterification reaction Slu7. In contrast, neither Ran (a protein does not involved in

splicing, data not shown) nor gal-3 was detected in the immunoprecipitate. As expected, gal-1 antiserum precipitated gal-1. Similar results were obtained when the immunoprecipitates of anti-gal-3 antiserum were analyzed by western analysis (data not shown).

Galectins should be detected in spliceosomes isolated by precipitation with antiserum directed against another splicing factor. A reciprocal co-immunoprecipitation experiment was performed using anti-Sm antisera (Figure 5). As expected, spliceosomes selected by the Sm antiserum contained the splicing substrate, intermediates and products. Immunoprecipitation with human IgG revealed only background levels of splicing RNAs. Most importantly, gal-1 was detected in the Sm selected complexes, but not detected in the material precipitated by the control human IgG. These data strongly support our contention that galectins are splicing factors associated with the splicing machinery.

Galectin-1 and galectin-3 reside on distinct splicing complexes

Spliceosomes selected by galectin-specific antisera are a heterogeneous mixture of complexes (e.g., spliceosomes containing pre-mRNA substrate would be expected to contain different RNA and protein components than spliceosomes containing ligated exon product). While the stoichiometry of most of the proteins in spliceosomes has not been determined, it is assumed that under standard splicing conditions using a single splicing substrate each distinct processing pathway intermediate will contain the same complement of proteins. Thus, it is reasonable to predict that spliceosomes containing gal-1 would also contain gal-3. However, in our previous depletion-reconstitution studies, we showed that either gal-1 or gal-3 could restore splicing activity to a galectin-depleted

Figure 5: Immunoprecipitation of spliceosomes by anti-Sm antiserum. Splicing reactions were incubated with ATP for 60 min and then were immunoprecipitated with human anti-Sm antiserum or human IgG (control). Bound fractions were eluted and RNA analyzed by gel electrophoresis and autoradiography (panel A) and gal-1 detected by western blotting (panel B).

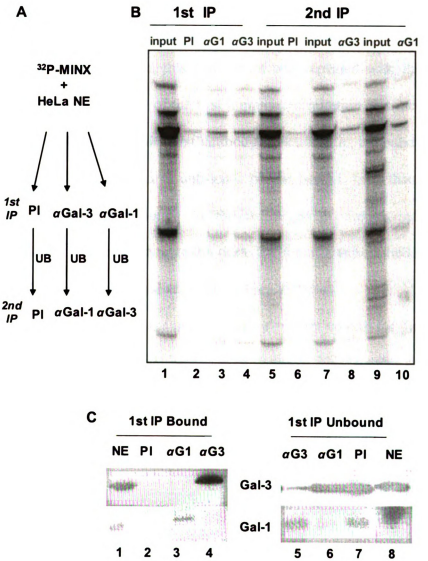
Figure 5

NE. Obviously, under these reconstitution conditions, spliceosomes containing a galectin would contain only that particular member of the galectin family.

To distinguish between these two formal possibilities (spliceosomes precipitated by one galectin antibody contain one or both nuclear galectins), we performed sequential immunoprecipitations as outlined in Figure 6A. Standard splicing reactions were incubated for 60 min and divided into two equal portions. One aliquot was immunoprecipitated with anti-gal-1 and the other with anti-gal-3. The unbound fractions were then subjected to a second immunoprecipitation using the other galectin antiserum. Radiolabeled RNA in the bound fractions from each immunoprecipitation was analyzed (Figure 6B). Roughly the same quantity of spliceosomes was precipitated by the anti-gal-1 antiserum in the two sequential selections (compare lanes 3 and 10). Similar results were obtained following the two anti-gal-3 immunoprecipitations (compares lanes 4 and 8). In order to interpret these results, the efficiency of each galectin antiserum to quantitatively immunoprecipitate its cognate antigen was determined. We analyzed the bound and unbound fractions from the first immunoprecipitation for gal-1 and gal-3 (Figure 6C). The bound fraction from the first anti-gal-1 precipitate showed only gal-1 (Figure 6C, lane 3) with no detectable gal-3. Further, the unbound fraction of this precipitation showed nearly quantitative depletion of gal-1 (lane 6; the amount of gal-1 in this fraction represents less than 10% of the total gal-1 in the reaction used for immunoprecipitation). Similar results were obtained with gal-3. Analysis of the bound fraction of the first anti-gal-3 precipitation showed only gal-3 (lane 4) and nearly all of gal-3 was removed by this immunoprecipitation (lane 5; less than 15% of the total gal-3 in the reaction remained in the unbound fraction of the first precipitation). We interpret

Figure 6: Sequential immunoprecipitation of splicing complexes by gal-1 and gal-3 antisera. Panel A: protocol for sequential immunoprecipitation. Splicing reactions were incubated for 60 min with ^{32}P-MINX and divided into equal aliquots. The aliquots were incubated with the indicated antiserum as described in Materials and Methods. Part of the unbound fraction was then incubated with the other galectin antiserum. After washing, the bound material from the first and second immunoprecipitations was eluted and analyzed for MINX RNA (panel B). Lanes 1, 5, 7 and 9 represent aliquots of the bound fraction used for immunoprecipitation (input). Lanes 2, 3, 4, represent the MINX RNA bound to the indicated antiserum in the first immunoprecipitate and lanes 6, 8 and 10 represent MINX RNA bound to the indicated antiserum in the 2nd immunoprecipitate. Panel C: Western blotting analysis of gal-1 and gal-3 in the bound (lanes 2, 3 and 4) and unbound (lanes 5, 6 and 7) fractions of the first immunoprecipitation.

Figure 6


A
B
1st IP
2nd IP
input PI αG1 αG3 input PI input αG3 input αG1
32P-MINX
+
HeLa NE
1st IP PI αGal-3 αGal-1
UB UB UB
2nd IP PI αGal-1 αGal-3
1 2 3 4 5 6 7 8 9 10
C
1st IP Bound
NE PI αG1 αG3
1 2 3 4
1st IP Unbound
αG3 αG1 PI NE
Gal-3
Gal-1
5 6 7 8

these data to indicate that gal-1 and gal–3 were quantitatively removed during the initial immunoselection and that the two galectins reside on different splicing complexes. Finally, spliceosomal RNAs could be immunoprecipitated by anti-Sm serum from the material remaining after the two sequential galectin adsorptions (data not shown), indicating that some spliceosomal complexes contained neither gal-1 nor gal-3.

As an additional evaluation of the nearly quantitative and specific removal of each galectin by its respective antiserum, this experiment was repeated with the following modification. The unbound fraction of the first immunoprecipitation was incubated with the same antiserum as that used in the first antibody selection (i.e., unbound material of anti-gal-1 precipitation was rebound to anti-gal-1 coated beads). Less than 5% of the initially precipitated RNA was bound to the antiserum in the second round of precipitation. This low level of binding to the same antiserum matched the low levels of galectin found in the unbound material after the first precipitation.

It is important to note that these results provide another control for the specificity of the spliceosomes precipitated by the anti-galectin antisera. The fact that each galectin antiserum precipitates only its respective antigen indicates that these antibodies do not precipitate nuclear proteins/complexes non-specifically and further suggests that galectins do not bind to splicing complexes non-specifically.

Galectin-1 and galectin-3 do not bind MINX RNA directly

The findings described above could be explained by a direct interaction of the galectins with the MINX pre-mRNA. If this interaction was with a unique site in the pre-mRNA, then only one galectin would bind per pre-mRNA and spliceosomes containing

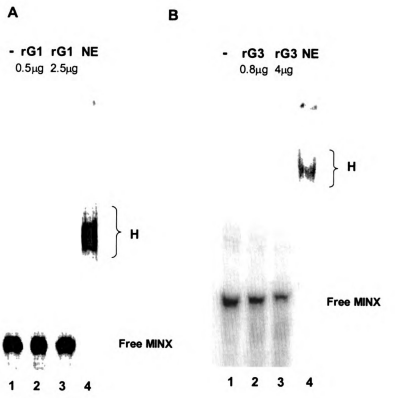
this pre-mRNA would have either gal-1 or gal-3 associated. To test for direct galectin-MINX interactions, we used an electrophoretic mobility shift assay (Figure 7). ^{32}P-MINX RNA migrated in native polyacrylamide gels as shown in lane 1 of Figure 7A and B. Incubation of MINX with recombinant human gal-1 (0.5 or 2.5 μg, Figure 7A, lanes 2 and 3, respectively) or recombinant gal-3 (0.8 or 4 μg, Figure 7B, lanes 2 and 3, respectively) did not alter the mobility of MINX when compared to the mobility of the substrate alone. Various incubation conditions (i.e., changing temperature, time of incubation and incubation with or without ATP) yielded the same results. As a positive control for altered mobility, MINX was incubated with NE for 15 min on ice. As expected, MINX was assembled into an H complex that had a slower mobility in the native gel (lane 4 of Figure 7A and B). We conclude that the association of galectins with spliceosomes occurs through interaction with another splicing component rather than through direct binding to the splicing substrate.

Amino terminal domain of gal-3 blocks gal-1 association with spliceosomes

Splicing activity in a complete HeLa NE is inhibited in a dose dependent manner by the addition of the ND of gal-3. At the highest concentration of ND, neither splicing activity nor spliceosome complex formation could be detected even though gal-1 was available (14). Our finding that each galectin resides on separate splicing complexes suggests a mechanism for this dominant negative effect of the ND. Splicing requires binding of galectins to another splicing factor. Likely this partner is shared by the galectins in a mutually exclusive manner. We suggest that excess ND binds to this partner and blocks the interaction of this partner with both nuclear galectins. To test this

Figure 7: Native gel analysis of splicing substrate incubated with recombinant gal-1, recombinant gal-3 or NE. Panel A: ^{32}P-MINX was incubated on ice for 15 min with 0.5 (lane 2), 2.5 (lane 3) μg recombinant human gal-1, HeLa NE ~25 μg protein (lane 4), or nothing (lane 1). Panel B: ^{32}P-MINX was incubated on ice for 15 min with 0.8 (lane 2), 4.0 (lane 3) μg recombinant gal-3, HeLa NE ~25 μg protein (lane 4) or nothing (lane 1). Following incubation, heparin was added and the samples analyzed on 4% polyacrylamide native gels followed by autoradiography. The migration of free MINX RNA and MINX RNA in H-complexes (indicated between the two panels) was detected by autoradiography.

Figure 7


A
- rG1 rG1 NE
0.5µg 2.5µg
H
Free MINX
1 2 3 4
B
- rG3 rG3 NE
0.8µg 4µg
H
Free MINX
1 2 3 4

prediction, we quantitated spliceosomes immunoprecipitated by anti-gal-1 antiserum in reactions inhibited by the ND of gal-3 (added as a GST fusion protein). The results are shown in Figure 8. As previously reported (14), the GST-ND polypeptide inhibited product formation nearly 100% (lane 5) compared to reactions incubated with (lane 3) or without (lane 1) GST. The addition of GST to the reaction did not inhibit the precipitation of spliceosomes by anti-gal-1 (compare lane 4 to lane 2). In contrast, the addition of GST-ND nearly completely inhibited the association of gal-1 with the splicing machinery (compare lane 6 to lanes 2 and 4). Thus, the ND of gal-3 exerts its dominant negative effect by regulating the incorporation of gal-1 (and gal-3) into splicing complexes.

Galectin association with spliceosomes is salt labile

The strength of the association of galectins with spliceosomes was evaluated in relation to the stable association of the snRNPs with splicing complexes. Splicing complexes formed after a 60 min splicing reaction (Figure 9, lane 1) were incubated with each galectin antiserum or pre-immune serum in 60 mM KCl buffer (the buffer used for optimal splicing efficiency). Aliquots of the antibody-bound spliceosomes were then washed with 60 mM (lanes 2-5), 130 mM (lanes 6-9) or 250 mM (lanes 10-13) KCl buffers. The bound fractions were eluted and analyzed for radiolabeled RNA. Salt concentrations of 130 or 250 mM released most of the splicing substrate from the galectin-selected columns (>90% of the spliceosomes were released as determined by quantitation from phosphorimage analysis) whereas 130 mM KCl had no effect on spliceosomes selected by anti-Sm antiserum. Even when the salt was increased to 250 mM KCl, nearly 20% of the snRNPs remained stably associated with spliceosomes. The

Figure 8: Analysis of spliceosomes immunoprecipitated by anti-gal-1 antiserum in reactions incubated with GST or GST-ND of gal-3. Equivalent splicing reactions were incubated with no addition (lane 1, input), GST (lane 3, +GST: input, 100 μM) or GST-ND (lane 5, +GST-ND: input, 100 μM) for 60 min. Anti-gal-1 beads were added to each reaction and RNA in the bound fractions was analyzed by gel electrophoresis and autoradiography (lanes 2, 4 and 6).

Figure 8

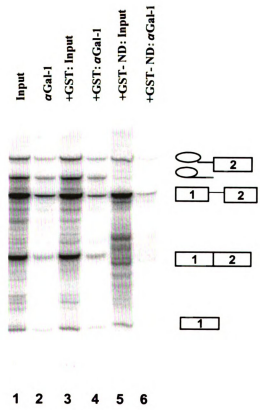
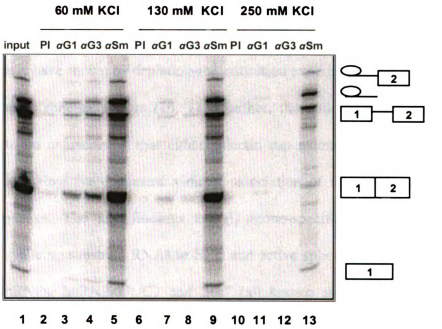

Input
σGal-1
+GST: Input
+GST: σGal-1
+GST- ND: Input
+GST- ND: σGal-1
2
1
2
1
2
1
1 2 3 4 5 6

Figure 9: Analysis of spliceosomes precipitated by various antisera at increasing salt concentrations. HeLa NE was incubated with ^{32}P-MINX for 60 min in the presence of ATP. Immobilized pre-immune serum (lanes 2, 6 and 10); anti-gal-1 (lanes 3, 7 and 11); anti-gal-3 (lanes 4, 8 and 12) or anti-Sm (lanes 5, 9 and 13) was incubated with the reactions in 60 mM KCl-buffer for 60 min. Beads were then washed extensively with 60 mM KCl-buffer (lanes 2-5); 130 mM KCl-buffer (lanes 6-9); or 250 mM KCl-buffer (lanes 10-13). The bound fractions were eluted and RNA analyzed by denaturing gel electrophoresis and quantitated by phosphorimage analysis (Molecular Dynamics) following autoradiography. Lane 1 indicates the RNA subjected to immunoprecipitation.

Figure 9


60 mM KCl
130 mM KCl
250 mM KCl
input
PI
αG1
αG3
αSm
PI
αG1
αG3
αSm
PI
αG1
αG3
αSm
2
1
2
1
2
1
1 2 3 4 5 6 7 8 9 10 11 12 13

loss of spliceosomal RNAs from the antibody columns was due to dissociation of the complexes from each galectin and not due to release of the galectins from their respective antibody. At 130 mM KCl, virtually no gal-1 or gal-3 was released from the immobilized antibodies compared to the galectins bound at 60 mM KCl. At 250 mM KCl, more than 70% of the galectins remained bound to the antibodies (data no shown). We conclude that the association of galectins with the splicing machinery is sensitive to perturbation of ionic strength.

Discussion

We previously have shown by depletion-reconstitution experiments that galectins-1 and -3 are required splicing factors (10, 11). Further, the splicing activity of the galectins appears to be redundant in that either galectin can reconstitute splicing in a galectin-depleted NE. Now we document a direct association of these galectins with spliceosomal complexes. The key findings are (i) mono-specific galectin antisera immunoprecipitate splicing substrate RNAs in H, E and active spliceosomes along with Sm proteins of snRNPs, hnRNP C1/C2 and Slu7 (all known splicing factors), (ii) spliceosomal complexes contain either gal-1, gal-3 or no galectin, (iii) neither galectin interacts directly with the pre-mRNA substrate, (iv) the amino terminus of gal-3 exerts a dominant negative effect on splicing by regulating the entry of gal-1 (and presumably gal-3) into splicing complexes and (v) the association of galectins with spliceosomes is salt sensitive compared to the stable association of the snRNPs.

The data presented are significant in evaluating the role of galectins as splicing factors. First, these data support our hypothesis that nuclear galectins are indeed splicing factors that can be incorporated into the canonical model for pre-mRNA splicing (1). The

fact that both galectin antibodies co-immunoprecipitate pre-mRNA associated with early (i.e., H and E) complexes illustrates the initial entry of galectins during spliceosome assembly. While a precise molecular function of galectins during spliceosome assembly is not known, it appears to be in the recruitment or supplying of snRNPs to pre-mRNA based on two observations: (i) depletion of galectins from splicing extracts inhibits transition of early (e.g., H and/or E) complexes to active spliceosomes (1), and (ii) galectins are associated with gemin4 in SMN complexes (14) which are implicated in recycling snRNPs to pre-mRNA in the early commitment complex (19).

Since the anti-galectin antibodies precipitated not only the pre-mRNA substrate but also the intermediates and products of the splicing reaction generated on the spliceosome, our results suggest that gal-1 and gal-3 are associated with the assembled machinery throughout the reaction cycle. It is clear that these galectin-containing complexes are spliceosomes as the hnRNP C1/C2 polypeptides (known splicing factors [5-9, 20]), the snRNP-specific Sm proteins and Slu7 (2, 5-9) co-precipitate with the substrate intermediates and products. Results of three controls strengthen our interpretation of the data. First, the nuclear shuttling protein Ran, which has not been found associated with spliceosomes, is not co-precipitated by the galectin antisera. Second, gal-1-containing spliceosomes are not precipitated by antibodies specific for gal-3 and vice versa. These results indicate (i) the galectin antisera are mono-specific and (ii) galectins do not adhere non-specifically to splicing complexes. Third, spliceosomes isolated by precipitation with antibodies directed against the Sm polypeptides of snRNPs co-precipitate galectins.

An obvious question of our contention that galectins are splicing factors is that they have not been identified in any of the proteomic analyses of spliceosomes (5-9). Several reasons can be offered to explain this discrepancy. First, most of the spliceosome isolation procedures use 150 – 250 mM salt during binding or washing to select stable complexes. The rationale has been to correctly identify components with a stable association in the spliceosome, rather than catalogue all associations. As we have shown (Figure 9), ionic conditions greater than 60 mM release the galectins from spliceosomal complexes. We contend the buffer conditions we use for immunoprecipitation (which are optimal for *in vitro* splicing activity) allow a more complete cataloguing of spliceosomal proteins. Could the transient/loose association of galectins hint to a regulatory role? Second, the galectins may be in low abundance in spliceosomes. Our observation that not all spliceosomes contain galectins speaks to this point. Third, it is possible that galectins only assist in initiating spliceosome assembly (i.e., only associate with a complex containing the pre-mRNA substrate) and are not stable components of active splicing complexes. Only a thorough and careful evaluation of these early complexes would reveal this association. Finally, the stringency set for the identification of peptides and subsequent database searches results in missing members of a complex. For example, of the four massive proteomic analyses published, none have detected the U6-associated polypeptide LSm5 and there are several instances where one of the four studies detected a core spliceosomal component and the other three did not.

Other significant aspects of our findings include providing an explanation for functional redundancy of the galectins and hinting at the nature of the spliceosome-associated binding partner for the galectins. Reconstitution of a galectin-depleted NE can

be achieved by either gal-1 or gal-3 (10, 11). The sequential immunoprecipitation data provide experimental proof for the exclusive incorporation of only a single galectin into a splicing complex in a complete (i.e., non-depleted) splicing extract. Thus, functional redundancy means spliceosomes contain only one galectin. We show that the ND of gal-3 regulates the entry of gal-1 into spliceosomes. In aggregate, these data suggest the galectins share a common binding partner. This partner is probably a polypeptide splicing factor that weakly interacts with the galectins. In a splicing extract, gal-3 interacts with this partner via its two domains (the ND which contains the PGAYPGXXX repeats of unknown function and the carboxyl terminus which is the CRD) whereas gal-1 only binds via its single CRD. The observation that gal-3 is 8 – 10 times more efficient in reconstituting splicing activity in galectin-depleted extracts compared to gal-1 supports this contention (11). Addition of excess ND binds to this common partner and blocks binding of gal-1 or gal-3. Abrogation of gal-1 and gal-3 binding results in inhibition of splicing. It remains to be determined whether the galectin-binding partner is assembled into splicing complexes when bound to the isolated ND.

Crucial to providing a mechanistic interpretation of these data regarding the association of galectins with spliceosomal complexes is the identification of the splicing partner for the galectins. While we have identified gemin4 as a galectin binding protein (14), neither gemin4 nor other members of the SMN complex (19) have been identified as spliceosomal components in proteomic analyses (5-9). Is this due to the fact that gemin4 and interacting proteins are not spliceosomal proteins or, as with the galectins, that they are weakly associated with spliceosome complexes and released under the conditions used for spliceosome isolation? From a different perspective, gal-1 has been

shown to be a component of the nuclear matrix (21) that has been proposed to serve as a scaffold for the splicing process. The nuclear matrix partner with which gal-1 interacts is unknown. Also, proteomic analysis of interchromatin granule clusters (IGC) has identified a galectin as a member of this nuclear organelle. Over 80% of the proteins identified as IGC components play a role in RNA biogenesis with over 50% having a splicing function (22). Both of these findings lend additional, albeit indirect, support to our assertion that the galectins are splicing factors.

Galectin-containing spliceosomes have been precipitated with several different galectin-specific antisera using three conditions to assemble various splicing complexes. The association of galectins with spliceosomes is through a salt-sensitive protein-protein interaction rather than a galectin-splicing substrate interaction. The commonly shared binding partner/ splicing factor explains the mutually exclusive incorporation of gal-1 or gal-3 into splicing complexes and the dominant negative inhibition of splicing demonstrated by the ND of gal-3.

Acknowledgments

Jung Park contribute more than one third of the work in this paper. We also thank Kevin Haudek, Rich Gray and Patty Voss for their invaluable input and comments on this manuscript. The work has been supported by grants MCB-0092919 (RJP) from the National Science Foundation and GM-38740 (JLW) from the National Institutes of Health.

References

1. Brow, D. A. (2002) Allosteric cascade of spliceosome activation. *Ann. Rev. Genet.*, **36**, 333-360.

2. Chua, K., and Reed, R. (1999) Human step II splicing factor hSlu7 functions in restructuring the spliceosome between the catalytic steps of splicing. *Genes Dev.*, **13**, 841-850.

3. Konarska, M. M., and Query, C. C. (2005) Insights into the mechanisms of splicing: more lessons from the ribosome. *Genes Dev.*, **19**,2255-2260.

4. Hastings, M. L., and Krainer, A. R. (2001) Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol.*, **13,** 302-309.

5. Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A., and Mann, M. (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nature Genet.*, **20**, 46-50.

6. Hartmuth, K., Urlaub, H., Vornlocher, H-P., Gentzel, M., Wilm, M., and Luhrmann, R. (2002) Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc. Natl. Acad. Sci. USA*, **99**, 16719-16724.

7. Rappsilber, J., Ryder,U., Lamond, A., and Mann, M. (2002) Large-scale proteomic analysis of the human spliceosome. *Genome Res.*, **12**, 1-15.

8. Zhou, Z., Licklider, L. J., Gygi, S. P. and Reed, R (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature*, **419**, 182-184.

9. Makarov, E. M., Makarova, O. V., Urlaub, H., Gentzel, M., Will, C. L., Mann, M., and Luhrmann, R. (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science*, **298**, 2205-2208.

10. Dagher, S.F., Wang, J. L., and Patterson, R. J. (1995) Identification of galectin-3 as a factor in pre-mRNA splicing. *Proc. Natl. Acad. Sci. USA*, **92**, 1213-1217.

11. Vyakarnam, A., Dagher, S. F., Wang, J. L., and Patterson, R. J. (1997) Evidence of a role for galectin-1 in pre-mRNA splicing. *Mol. Cell. Biol.*, **17**, 4730-4737.

12. Barondes, S.H., Castronovo, V., Cooper, D. N. W., Cummings, R. D., Drickamer, K., Feizi, T., Gitt, M. A., Hirabayashi, J., Hughes, C., Kasai, K., Leffler, H., Liu, F.-T., Lotan, R., Mercuric, A. M. , Monsigny, M., Pillai, S., Poirier, F., Raz, A., Rigby, P. W., Rini, J. M., and Wang, J. L. (1994) Galectins: A family of animal β-galactoside-binding lectins. *Cell*, **76**, 597-598.

13. Vyakarnam, A., Lenneman, A. J., Lakkides, K. M., Patterson, R. J., and Wang, J. L. (1998) A comparative nuclear localization study of galectin-1 with other splicing components. *Exp. Cell Res.*, **242**, 419-428.

14. Park, J. W., Voss, P. G., Grabski, S., Wang, J. L., and Patterson, R. J. (2001) Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. *Nucleic Acids Res.*, **29**, 3595-3602.

15. Hirabayashi, J., Ayaki, H., Soma, G., and Kasai, K. (1989) Production and purification of a recombinant human 14 kDa beta-galactoside-binding lectin. *FEBS Lett.*, **250**, 161-165.

16. Agrwal N, Sun Q, Wang SY, Wang JL. (1993) Carbohydrate-binding protein 35. I. Properties of the recombinant polypeptide and the individuality of the domains. *J Biol Chem.*, **268**, 14932-14939.

17. Zillman, M., Zapp, M. L., and Berget, S. M. (1988) Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.*, **8**, 814-821.

18. Das, R., Zhou, Z., and Reed, R. (2000) Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol. Cell*, **5**, 779-787.

19. Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998) A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell*, **95**, 615-624.

20. Choi, Y.D., Grabowski, P. J., Sharp, P. A., and Dreyfuss, G. (1986) Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. *Science*, **231**, 1534-1539.

21. Choi, J. Y., van Wijnen, A. J., Aslam, F., Leszyk, J. D., Stein, J. L., Stein, G. S.. Lian, J. B., and Penman, S. (1998) Developmental association of the beta-galactoside-binding protein galectin-1 with the nuclear matrix of rat calvarial osteoblasts. *J. Cell Sci.*, **111**, 3035-3043.

22. Saitoh, N., Spahr, C. S., Patterson, S. D., Bubulya, P., Neuwald, A. F., and Spector, D.L. (2004) Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell*, **15**, 3876-90.

Chapter 3

Dissociation of the splicing and the carbohydrate-binding activities of galectin-1

ABSTRACT

A site-directed mutant of galectin-1 (Gal1), in which an asparagine at residue 46 is replaced by aspartic acid, was expressed and purified as a fusion protein with glutathione S-transferase (GST). This mutant, designated as GST-Gal1(N46D), was compared with the corresponding fusion protein containing wild-type galectin-1, GST-Gal1(WT), in three assays: (a) binding to asialofetuin-Sepharose, as a test of the carbohydrate-binding activity; (b) reconstitution of splicing in a galectin-depleted nuclear extract of HeLa cells, as a test of the *in vitro* splicing activity; and (c) reconstitution of splicing in a nuclear extract of HeLa cells containing 150 mM thiodigalactoside, as a test of sensitivity to saccharide inhibition. The binding of GST-Gal1(N46D) to asialofetuin-Sepharose was less than 10% of that observed for GST-Gal1(WT), indicating that the mutant was deficient in carbohydrate-binding activity. Both GST-Gal1(WT) and GST-Gal1(N46D) were equally efficient, however, in reconstitution of splicing activity in a galectin-depleted nuclear extract, suggesting that carbohydrate-binding, *per se*, was not required for the splicing activity. Previous experiments had shown that the endogenous splicing activity of HeLa cell nuclear extracts can be inhibited by the addition of thiodigalactoside. GST-Gal1(WT) cannot overcome this inhibition. In contrast, the addition of GST-Gal1 (N46D) to a thiodigalactoside-inhibited nuclear extract resulted in

splicing of the pre-mRNA substrate. Additional experiments show that Gal1 is incorporated into H, E and active splicing complexes. Addition of thiodigalactoside released Gal1 from H splicing complexes, but not from committed complexes (i.e., E and active spliceosomes). These results indicate that the carbohydrate-binding and the splicing activities of galectin-1 can be dissociated.

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 of the polypeptide (1, 2). In previous studies, we had reported the localization of galectin-1 (Gal1)[1] and galectin-3 (Gal3) in the cell nucleus (3, 4) and identified them as two of the many proteins required for the splicing of pre-mRNA, assayed in a cell-free system (5, 6). The key findings from these studies included: (a) nuclear extracts (NE) derived from HeLa cells, capable of carrying out splicing of pre-mRNA, contained both Gal1 and Gal3; (b) depletion of both galectins from NE, either by lactose (Lac) affinity chromatography or by antibody adsorption, resulted in the concomitant loss of splicing activity; and (c) either recombinant Gal1 or recombinant Gal3 was able to reconstitute splicing activity in a galectin-depleted extract.

More recently, we have identified Gal1 and Gal3 as novel components of macromolecular complexes containing the survival of motor neuron (SMN) protein, Gemin2 and the core snRNP polypeptides bearing the Sm epitope (7). Both galectins interact directly with at least one component of these complexes, Gemin4. The SMN-

containing complexes are found in the cytoplasm and in discrete nuclear bodies called gems (gemini of coiled bodies) (8-10). In the former, the SMN complex plays a critical role in the biogenesis of snRNPs (11, 12). In the nucleus, SMN is required for pre-mRNA splicing by supplying snRNPs to the H-complex in the assembly of spliceosomes (13, 14). This H-complex juncture is indeed where galectins appear to be required, as demonstrated by the effect of galectin depletion (6) and by the effect of addition of the NH_2-terminal domain of galectin-3, which exerts a dominant negative effect on spliceosome assembly and the splicing reaction (7).

In the course of our studies, we had observed and reported that saccharides which bind to Gal1 and Gal3 with high affinity, such as Lac and thiodigalactoside (TDG), inhibited the cell-free splicing reaction when added to a complete NE (5, 6). In contrast, saccharides that do not bind to Gal1 and Gal3, such as cellobiose, failed to have any effect. Does inhibition of splicing by Lac and TDG indicate that the nuclear ligand for the galectins interacts with the carbohydrate-binding site and that it is competitively displaced upon Lac/TDG addition? Or, does Lac/TDG addition alter the conformation of the galectins, causing the release or altered interaction with a nuclear ligand bound via protein-protein, rather than protein-carbohydrate, interactions? In the present communication, we report the identification of a site-directed mutant of Gal1, deficient in carbohydrate-binding activity, that is still capable of functioning in the splicing assay. We show that TDG can release Gal1 from H splicing complexes, but not from committed complexes. The results suggest that carbohydrate-binding, *per se*, is not required for the splicing activity of galectins.

EXPERIMENTAL PROCEDURES

Glutathione S-transferase (GST) Fusion Proteins

The cDNAs for wild-type (WT) and Asn46Asp (N46D) mutant forms of human Gal1 have been described (15). Each cDNA was subcloned into the BamHI restriction site of the pGEX-2T vector (Pharmacia) to produce a fusion protein between GST and Gal1: GST-Gal1(WT) and GST-Gal1(N46D). Each of the constructs was subjected to DNA sequencing to verify: (a) the juncture of the fusion protein between GST and Gal1; and (b) the wild-type and mutant amino acid at residue 46.

The constructs were transformed into *E. coli* (strain BL21-codon Plus (DE3); Stratagene) and GST-Gal1(WT) and GST-Gal1(N46D) were purified from 1 liter cultures on the basis of GST binding to glutathione-agarose beads (Pierce). The purity of the protein preparations was assessed by SDS-PAGE (16). The polypeptides were revealed by silver staining (17) or by immunoblotting (7), following procedures that have been described previously. Affinity purified polyclonal rabbit anti-Gal1 and anti-GST antibodies were used for the immunoblotting.

To generate the antibodies, the immunogen used was GST-Gal1(WT) purified on the basis of binding to two columns: (a) glutathione-agarose and elution with glutathione; and (b) Lac-agarose (Sigma) and elution with Lac. Approximately 70 ml of antisera, pooled from four bleeds of rabbit #55, were subjected to ammonium sulfate fractionation (50% of saturation). The immunoglobulin-containing precipitated fraction was solubilized in phosphate-buffered saline (PBS) and passed over a 5-ml column of GST-agarose. The unbound (flow-through) fraction was immediately loaded over the same

column (six passes over the same column to insure binding). The bound fraction was eluted with 0.1 M glycine-HCl (pH 2.2) and this was dialyzed immediately against PBS to neutralize the pH. The bound and eluted material from the GST affinity column is designated as affinity purified anti-GST. Purified GST-Gal1(WT) (the immunogen) was bound to glutathione-agarose and covalently cross-linked with dimethylpimelimidate (20 mM; Pierce). The reaction was carried out in 0.2 M sodium borate (pH 9) for 1 hour at room temperature; the cross-linked beads were washed twice with 0.2 M ethanolamine, pH 8, followed by a 2-hour incubation at room temperature in the same buffer to block unreacted groups. The unbound fractions of the antisera, depleted of anti-GST antibodies, were passed over this GST-Gal1(WT) affinity column (six passes to insure binding). The bound and eluted material from this column is designated as affinity purified anti-Gal1. At 1:2000, 1:5000, and 1:10,000 dilutions, this affinity purified anti-Gal1 blots Gal1 in NE of HeLa cells and purified GST-Gal1(WT) but does not blot GST.

Assay of Carbohydrate-binding Activity

The preparation of the saccharide affinity beads, asialofetuin (ASF)-Sepharose 4B, has been described (18). Approximately 180 nmoles of ASF were coupled per ml of Sepharose beads. GST proteins (350 ng each of GST, GST-Gal1(WT), and GST-Gal1(N46D)) were incubated with 35 µl of ASF-Sepharose for 2 hours at 4 °C. The incubations were carried out in 60% buffer D (buffer D is 20 mM Hepes, pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM phenyl-methylsufonyl fluoride, and 0.5 mM dithiothreitol) containing 0.1% NP-40 (Pierce) in the presence and absence of Lac (100 mM). The material not bound to the beads was removed by centrifugation (1000 x g) and resuspension; the beads were washed four times in 60% buffer D containing 0.1%

NP-40. The GST proteins bound to the beads were subjected to SDS-PAGE and immunoblotting with anti-GST and anti-Gal1 antibodies.

The chemiluminescent signal provided by horseradish peroxidase conjugated to the secondary antibody was detected using the Western Lightning reagent (Perkin Elmer Life Sciences). This signal was quantitated with a BioRad model GS505 Molecular imager system and associated software. Known amounts of GST, GST-Gal1(WT), and GST-Gal1(N46D) were used to establish standard curves. The quantitative value of the immunoblotted band derived from the incubation carried out in the presence of Lac represented non-specific binding not inhibitable by Lac; this accounted for about 3% of the total binding for both the wild-type and mutant proteins. Lac-inhibitable specific binding was calculated by subtracting this value from the total binding.

Assay of Splicing Activity

HeLa S3 cells were grown in suspension culture by the National Cell Culture Center (Minneapolis, MN). NE was prepared in buffer C (20 mM Hepes, pH. 7.9, 25% glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM phenylmethylsufonyl fluoride, 0.5 mM dithiothreitol), as described by Dignam *et al.* (19). Protein concentrations were determined by the Bradford assay (20); in this study, the protein concentration of the NE was ~4 mg/ml.

NEs were depleted of galectins by adsorption on ASF-Sepharose (18). The beads (150 µl) were washed with 20 mM Hepes, pH 7.9, 0.5 M NaCl; 30 µl NE were added and incubated on ice for 20 minutes in disposable spin columns (Millipore). The unbound fraction was removed and the beads were washed with 12 µl of 60% buffer D adjusted to 0.42 M NaCl and this wash was combined with the unbound fraction. Aliquots of

nondepleted NE and unbound fractions of the depletion were dialyzed in a microdialyzer against 60% buffer D for 40 minutes at 4 °C using a dialysis membrane with a 10-kDa cutoff. In reconstitution experiments, GST, GST-Gal1(WT), or GST-Gal1(N46D) were added to the unbound fractions prior to dialysis. The dialyzed fractions were then assayed for splicing activity.

The plasmid used to transcribe the MINX pre-mRNA substrate was obtained from Dr. Susan Berget (Baylor College of Medicine, Houston, TX) (21). The MINX pre-mRNA was labeled with [^{32}P]GTP and the monomethyl cap was added during SP6 polymerase (Gibco BRL) transcription. Splicing reaction mixtures contained a total volume of 10 μl: dialyzed NE (4 μl) or unbound fraction (8 μl), [^{32}P]MINX pre-mRNA, 2.5 mM $MgCl_2$, 1 mM ATP, 20 mM creatine phosphate, 0.5 mM dithiothreitol, and 20 U of RNasin (Promega). In experiments testing the effect of exogenously added carbohydrates, the extracts were incubated with 150 mM TDG or 150 mM cellobiose for 5 minutes at room temperature prior to the addition of pre-mRNA substrate. Splicing reactions were incubated at 30 °C for 45 minutes. Proteinase K-SDS solution was added to a final concentration of 4 mg/ml and 0.1%, respectively. The sample was incubated at 37 °C for 15 minutes. Each sample was then diluted to 100 μl with 125 mM Tris, pH 8, 1 mM EDTA, 0.3 M sodium acetate. RNA was extracted with 200 μl of phenol-chloroform (50:50 (vol/vol)), followed by 200 μl 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.

Preparation of H complexes involved addition of pre-mRNA to NE in buffer D in the absence of ATP, creatine phosphate and $MgCl_2$ and incubation on ice for 45 min. E complexes were formed by incubation at 30 °C of a reaction identical to that used for H complex formation. Active spliceosomes were formed as above.

Quantitation of the amount of radioactivity was carried out on a STORM phosphorimager (Molecular Dynamics). 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, the splicing intermediates (exon 2-lariat and exon 1), and the product.

Immunoprecipitation of Splicing Complexes

Polyclonal rabbit antiserum to Gal1 or pre-immune serum were attached to protein G-agarose beads (described in chapter 2). After extensive washing, the beads were incubated with preformed H, E or active spliceosomes in the presence or absence of TDG or cellobiose. After washing, the bound material was eluted with PAGE-SDS sample buffer and aliquots assayed for proteins or splicing substrate RNAs.

RESULTS

GST-Gal1(N46D) is Deficient in Carbohydrate-binding Activity

Hirabayashi and Kasai (15) had demonstrated that substitutions at highly conserved residues of the carbohydrate recognition domain of human Gal1, such as Asn46, Glu71, and Arg73, resulted in loss of saccharide-binding activity. The cDNAs corresponding to the WT and N46D mutant were each subcloned into the pGEX-2T

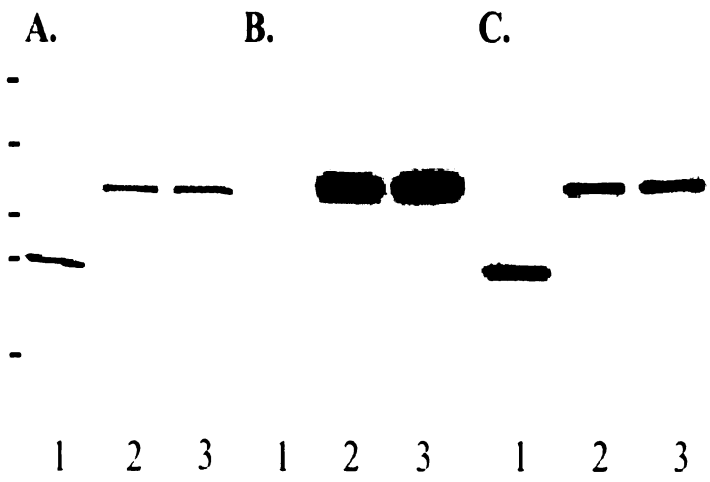
vector to produce a fusion protein between GST and Gal1 so that we can take advantage of the easy purification of the proteins on the basis of GST binding to glutathione beads. Nucleotide sequencing analysis of the constructs verified the sequence of the fusion protein; in particular, residue 46 in the Gal1 sequence was confirmed to be Asn and Asp, respectively for GST-Gal1(WT) and GST-Gal1 (N46D). In parallel, we also expressed GST from the same vector to be carried as a control in the functional tests below.

On SDS-PAGE and silver staining (Figure 1, panel A, lanes 1-3), each of the purified GST proteins yielded a single band with a mobility corresponding to the expected molecular weight: GST, ~27 kDa; GST-Gal1(WT), ~42 kDa; and GST-Gal1(N46D), ~42 kDa. Immunoblotting with affinity purified anti-Gal1 antibodies yielded a single band at the same molecular weight for GST-Gal1(WT) and GST-Gal1(N46D) (Figure 1, panel B, lanes 2-3). No reaction was observed between anti-Gal1 and GST (Figure 1, panel B, lane 1). Finally, immunoblotting of the respective GST protein preparations with affinity purified anti-GST antibodies yielded the same single band patterns as were observed by silver staining (Figure 1, panel C, lanes 1-3). These results establish the purity of the protein reagents to be compared in the functional assays below.

Purified GST-Gal1(WT) and GST-Gal1(N46D) were compared in terms of their binding to ASF-Sepharose beads. After washing, the material bound to the beads was subjected to SDS-PAGE and immunoblotting with anti-GST antibodies. Using known amounts of GST-Gal1 (WT or mutant) to establish standard curves, we quantitated the Lac-inhibitable binding, as well as binding not inhibitable by Lac. The latter accounted

Figure 1: Characterization of the preparations of fusion proteins containing wild-type galectin-1 or its N46D mutant by SDS-PAGE. Lane 1: GST; lane 2: GST-Gal1(WT); and lane 3: GST-Gal1(N46D). The proteins (30 ng in each lane) were electrophoresed through 12.5% acrylamide gels. Panel A: silver staining; Panel B: immunoblotting with affinity-purified anti-Gal1 antibodies; Panel C: immunoblotting with affinity-purified anti-GST antibodies. The binding of the primary antibodies in panels B and C were revealed with horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin and the enhanced chemiluminescent system. The positions of migration of molecular weight standards (80 kDa; 52 kDa; 35 kDa; 30 kDa; and 22 kDa) are indicated on the left.

Figure 1



for about 3% of the total binding observed for both the WT and mutant proteins. In terms of Lac-inhibitable binding, 10-15% of the GST-Gal1(WT) added to the assay was bound specifically; in contrast, less than 1% of the GST-Gal1(N46D) added to the assay was bound in a Lac-inhibitable fashion. More importantly, we found that the binding of the N46D mutant was drastically reduced (to less than 10%) relative to the wild-type protein (Figure 2). The level of binding of GST-Gal1(N46D) was similar to that observed for GST. The same overall conclusion was obtained using either anti-GST antibodies (Figure 2) or anti-Gal1 antibodies for the quantitation (the latter obviously could not detect GST itself, as was shown in Figure 1, panel B, lane 1).

Reconstitution of Splicing in Galectin-depleted NE by GST-Gal1(WT) and GST-Gal1(N46D)

NE was prepared in buffer C, which contained 0.42 M NaCl to dissociate splicing or RNP complexes. This NE was incubated with ASF-Sepharose. Western blotting analysis with antibodies against Gal1 and Gal3 documented that both proteins were present in the NE and in the bound fraction of the affinity beads. Neither Gal1 nor Gal3 was detected in the unbound fraction. Thus, the galectins were quantitatively depleted from the extract.

NE, depleted of the galectins by saccharide affinity adsorption, was also depleted in terms of splicing activity (Figure 3, lanes 1-2). We had previously documented that GST, by itself, had no effect on the cell-free splicing assay (7) so the effects of purified GST-Gal1(WT) and GST-Gal1(N46D) could be directly compared. The same preparations of the WT and mutant proteins that exhibited drastic differences in saccharide-binding activity (Figure 2) both reconstituted splicing in the galectin-depleted

Figure 2: Comparison of the carbohyrate-binding activity of GST-Gal1(WT) and GST-Gal1(N46D). GST proteins (350 ng each of GST, GST-Gal1(WT) and GST-Gal1(N46D)) were incubated with ASF-Sepharose for 2 hours at 4 °C in 60% buffer D containing 0.1% NP-40. Parallel incubations were carried out in the presence and absence of 100 mM Lac. The GST proteins bound to the ASF-beads were quantitated by Western blotting with anti-GST. The values shown represent Lac-inhibitable specific binding.

Figure 2

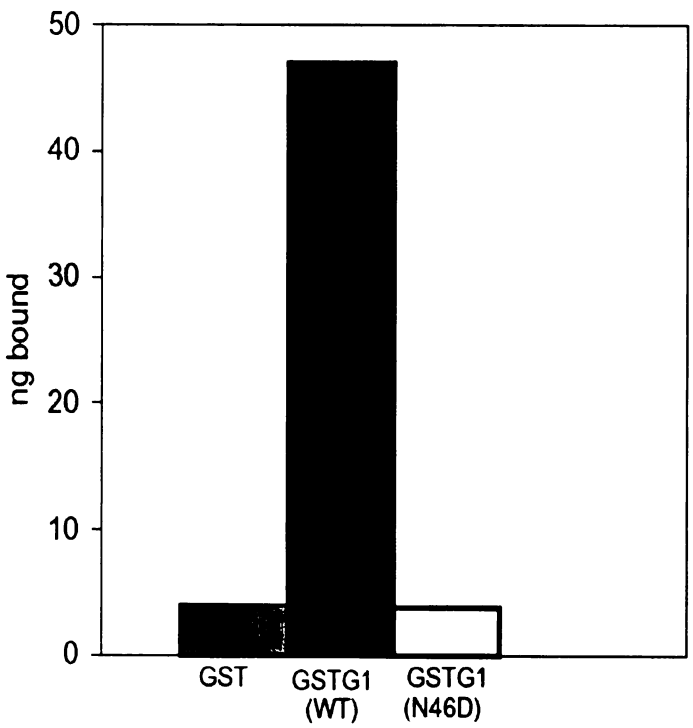
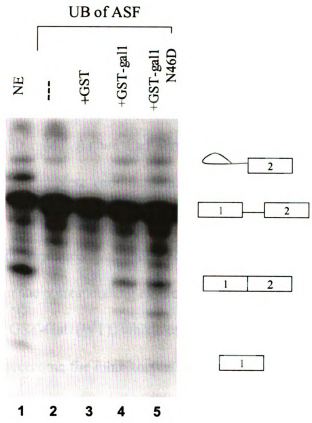


Figure 3: Comparison of the splicing activities of NE, NE after adsorption on ASF-Sepharose, and depleted NE reconstituted with GST, GST-Gal1(WT) and GST-Gal1(N46D). Lane 1: the complete (non-depleted) NE; lanes 2-5: the unbound (UB) fractions from the ASF-Sepharose adsorption, assayed without (lane 2) and with reconstitution with GST (lane 3), GST-Gal1(WT) (lane 4) and GST-Gal1(N46D) (lane 5). The concentration of the GST proteins was 20 μM. 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.

Figure 3



NE (Figure 3, lanes 4-5). In contrast, GST by itself could not reconstitute the splicing activity in the depleted NE (Figure 3, lane 3). The concentration of GST proteins used in the reconstitution was 20 μM. At this concentration, the level of product formation (~7%) in the reconstitution assay was comparable to that achieved previously with recombinant galectin-1 (5).

Effect of GST-Gal1(N46D) on NE Whose Splicing Activity is Inhibited by TDG

The above results suggest that the N46D mutant of Gal1, deficient in carbohydrate-binding activity, was still functional in terms of its splicing activity. A confirmation of this notion was obtained when we carried out the splicing assay in NE containing TDG (150 mM). As we had shown previously (5, 6), TDG inhibited the splicing activity of a complete (non-depleted) NE when added to the cell-free assay whereas cellobiose, which does not bind to the galectins, failed to yield the same effect. The level of product formation was much higher in NE containing cellobiose (~30%; Figure 4, lane 2) than the corresponding extract containing TDG (~1%; Figure 4, lane 1).

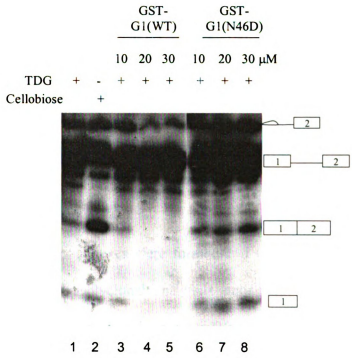
Addition of GST-Gal1(WT), which can bind to oligosaccharide-bearing ligands (Figure 2), did not overcome the inhibitory effect of TDG on splicing (Figure 4, lanes 3-5). In contrast, GST-Gal1(N46D), which was deficient in carbohydrate binding (Figure 2), was not sensitive to inhibition by the saccharide. Addition of the mutant fusion protein resulted in a dose-dependent increase in splicing activity (Figure 4, lanes 6-8).

Effect of TDG on the Association of Gal1 with Splicing Complexes

If the carbohydrate binding function is not required for the splicing function, how do galectin-specific saccharides inhibit the splicing activity in a NE? Since there are no known nuclear glycoconjugates for galectins to bind in their splicing role, we believe that

Figure 4: Comparison of the effects of cellobiose, TDG, GST-Gal1(WT) and GST-Gal1(N46D) on the splicing activity of NE. Lane 1: TDG addition (150 mM); lane 2: cellobiose addition (150 mM); lanes 3-5: TDG (150 mM) and GST-Gal1(WT) (10, 20, and 30 μM); lanes 6-8: TDG (150 mM) and GST-Gal1 (N46D) (10, 20, and 30 μM). 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, and mature RNA product are indicated on the right.

Figure 4


GST-
G1(WT)
GST-
G1(N46D)
10 20 30 10 20 30 μM
TDG + - + + + + + +
Cellobiose +
2
1 2
1 2
1
1 2 3 4 5 6 7 8

they participate by interacting with a splicing partner through protein-protein interactions. Addition of saccharides with affinity for galectins would directly or indirectly interfere with this required interaction.

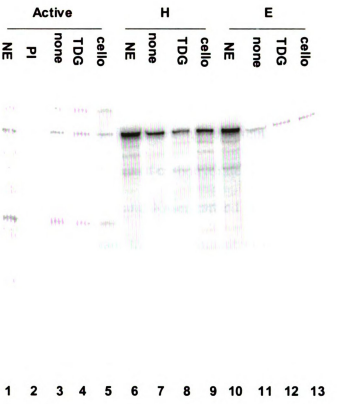
One way to test this suggestion is to determine whether galectin-specific saccharides release galectins that are associated with splicing complexes. As shown previously, galectins are assembled in splicing complexes throughout the splicing pathway (see chapter 2). Three distinguishable splicing complex intermediates can be assembled using NE: these are the H, E and active complexes. H complexes form immediately upon addition of pre-mRNA to a NE in the absence of ATP or elevated temperatures. The H complexes are chased into E (early) complexes following incubation at 30 °C. The E complexes are functionally committed and the addition of ATP converts them to active spliceosomes, another heterogeneous collection of complexes that catalyzes splicing chemistry.

H, E and active complexes were formed and Gal1 incorporation monitored by immunoprecipitation with polyclonal anti-Gal1 antiserum. The results are shown in Figure 5. Gal1 is associated with each complex (lanes 3, 7 and 11). We now determined whether a saccharide with high affinity for galectins (TDG) could dissociate Gal1 from these splicing complexes. As a negative control, the effect of a disaccharide with no detectible galectin binding activity (cellobiose) was compared. Cellobiose had no effect on Gal1 assembled into any of the splicing complexes (lanes 5, 9 and 13). In contrast, TDG can release Gal1 from H complexes (compare lane 8 to lane 7). Interestingly, TDG had no effect on Gal1 in E (compare lane 12 to lane 11) or active complexes (compare

lane 4 to 5). Once Gal1 is incorporated into complexes committed to the splicing pathway, it is inaccessible for interaction with specific saccharide ligands.

Figure 5: Comparison of the effects of cellobiose and TDG on the release of Gal1 from splicing complexes. Lanes 1, 6 and 10: NE used as input material for each class of splicing complex; lanes 2-5, active spliceosomes; lanes 6-9, H complexes; lanes 10-13, E complexes. Lane 2 was precipitated with pre-immune serum; lanes 3-5, 7-9, and 11-13 were precipitated with anti-Gal1 in the presence of cellobiose (150 mM) (lanes 5, 8 and 13); TDG (150 mM) (lanes 4, 7 and 12); or no additions (lanes 3, 7 and 11).

Figure 5


Active
H
E
NE
PI
none
TDG
cello
NE
none
TDG
cello
NE
none
TDG
cello
1 2 3 4 5 6 7 8 9 10 11 12 13

DISCUSSION

The key findings of the present study include: (a) the N46D mutant of Gal1, expressed as a fusion protein with GST, is deficient in carbohydrate-binding activity, compared to the WT protein; (b) GST-Gal1(N46) can reconstitute splicing activity in NE depleted of galectins, just like GST-Gal1(WT); (c) unlike the WT protein, GST-Gal1(N46D) is insensitive to the presence of TDG and therefore, can overcome the inhibition of splicing in NE containing the saccharide; and (d) Gal1 can be released from H splicing complexes by TDG but it cannot be released from committed splicing complexes (E or active spliceosomes). These finding represent the first demonstration of a function for a galectin that is independent of carbohydrate-binding activity.

In previous studies, Hirabayashi and Kasai carried out site-directed mutagenesis on the human Gal1 cDNA to generate the N46 mutant (15). When bacterial lysates expressing this mutant were passed over asialofetuin-Toyopearl columns, Gal1(N46D), as detected by Western blotting, was found exclusively in the flow-through fraction together with the bulk of the *E. coli* proteins. There was no sign of retardation on the affinity columns that might reflect some retention of weak carbohydrate-binding capacity. Thus, it was concluded that substitution of Asn by Asp at residue 46 resulted in complete loss of saccharide-binding activity (15, 22). Like the parent mutant protein, Gal1(N46D), our present fusion construct GST-Gal1(N46D) also resulted in a polypeptide deficient in binding to asialofetuin. The level of binding of GST-Gal1(N46D) was found to be similar to that observed for GST. That residue 46 of Gal1 is critical in saccharide-binding is consistent with the results of X-ray crystallographic analysis of the carbohydrate

recognition domain of the galectins, in which the Asn residue at this position serves as an acceptor of a hydrogen bond from the hydroxy group at C-4 of galactose (23 - 26).

In any site-directed mutagenesis experiment, a critical issue is whether the mutation has caused a disruption of the overall folding of the polypeptide. Thus, structural data will ultimately be required to provide a definitive basis for proper interpretation of the results obtained with the mutant polypeptide. It should be noted, however, that our previous documentation that Gal1 can independently reconstitute splicing activity in NE depleted of the galectins provided an activity of the protein that can be assayed without measuring carbohydrate-binding (5). Our present results showing that GST-Gal1(N46D) can reconstitute splicing despite being deficient in carbohydrate-binding suggest that we have been able to dissociate the two activities in the N46D mutant. This, in turn, suggests that the mutant polypeptide must have retained sufficient structure to preserve the splicing activity while the carbohydrate-binding activity was compromised. We draw analogy to two previously documented examples. Prostaglandin endoperoxide synthase, which catalyzes the committed step in the synthesis of prostaglandins and thromboxanes, exhibits two activities: (a) cyclooxygenase activity; and (b) peroxidase activity. A site-directed mutant (Tyr385Phe) lacked cyclooxygenase activity but retained peroxidase activity, arguing against the notion that the single amino acid substitution resulted in gross misfolding of the polypeptide (27). Similarly, Bicoid, the anterior determinant of *Drosophila*, controls embryonic gene expression by transcriptional activation and translational repression. Replacement of arginine at residue 54 (e.g. R54S) shifts the binding properties of the homeodomain to prefer DNA over RNA recognition. This abolishes mRNA translational repression without affecting

transcriptional activation (28). Thus, we interpret our present data to indicate that the saccharide-binding activity is not required for the splicing activity.

We have taken advantage of this property of the N46D mutant, loss of carbohydrate-binding activity but retention of splicing activity, to overcome the inhibition of splicing observed in complete (non-depleted) NE containing TDG. This raises the question of how to reconcile the two apparently disparate findings: while carbohydrate-binding is not required for splicing activity, saccharides such as Lac and TDG nevertheless exert an inhibitory effect when added to the splicing assay. It is possible that binding of saccharide ligands to the carbohydrate-binding site results in a conformational change that disrupts the interaction of the galectin polypeptide with a component of the splicing machinery. Endogenous Gal1 and Gal3 in the NE would be sensitive to TDG because they can bind to the saccharide. Likewise, exogenously added GST-Gal1(WT) cannot overcome the TDG inhibition because it also binds to the carbohydrate ligand. In contrast, because GST-Gal1(N46D) is deficient in carbohydrate-binding activity, it will not undergo the conformational change and therefore, can retain its functional splicing interactions.

The finding that Gal1 association with H splicing complexes is sensitive to saccharide dissociation is also very interesting. In this instance, the site is accessible to galectin-specific ligands and binding alters the overall conformation of the galectin which prohibits other interactions. The transition from H to E and active spliceosomes causes rearrangements that now cause the carbohydrate-binding site to be inaccessible to TDG and resistant to release from the splicing complexes.

To the best of our knowledge, a conformational change in the Gal1 polypeptide upon saccharide binding has not been reported. Moreover, most of the crystallographic structures of the galectin family of polypeptides have been determined in complexes containing saccharides (23-26). The lone exception is that of galectin-7, in which the X-ray structure was determined in both free and carbohydrate-bound forms (29). A comparison of the three-dimensional structures of the two forms showed that the polypeptide does not undergo any significant conformational change upon carbohydrate-binding. On the other hand, evidence for a conformational change in the COOH-terminal carbohydrate recognition domain of Gal3 upon Lac binding has been suggested by differential scanning calorimetry studies (30). In addition, several studies have documented that Gal3 binds to multivalent ligands with positive cooperativity (31-33) and it was suggested that saccharide-binding might expose hydrophobic surfaces for interactions between the carbohydrate recognition domains of separate Gal3 molecules (34). Thus, a rigorous test of our hypothesis must await physico-chemical studies on the Gal1 polypeptide itself, as well as tests of the effect of Lac/TDG on the interaction between Gal1 (and Gal3) with components of the splicing machinery.

FOOTNOTES

This work was supported by grants MCB 97-23615 from the National Science Foundation (RJP) and GM-38740 from the National Institutes of Health (JLW). Patty Voss contributed part of the work.

Abbreviations: Gal1, galectin-1; Gal3, galectin-3; NE, nuclear extract; Lac, lactose; SMN, survival of motor neuron; snRNP, small nuclear ribonucleoprotein; TDG,

thiodigalactoside; ASF, asialofetuin; GST, glutathione S-transferase; WT, wild-type; PBS, phosphate-buffered saline

REFERENCES

1. Barondes, S.H., Castronovo, V., Cooper, D.N.W., Cummings, R.D., Drickamer, K. *et al.*, (1994) Cell **76**, 597-598.

2. Hirabayashi, J. (ed.) (1997) Trends Glycosci. Glycotechnol. **9**, 1-184.

3. Vyakarnam, A., Lenneman, A.J., Lakkides, K.M., Patterson, R.J., and Wang, J.L. (1998) Exp. Cell Res. **242**, 419-428.

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

5. Vyakarnam, A., Dagher, S.F., Wang, J.L., and Patterson, R.J. (1997) Mol. Cell. Biol. **17**, 4730-4737.

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

7. Park, J.W., Voss, P.G., Grabski, S., Wang, J.L., and Patterson, R.J. (2001) Nucleic Acids Res. **27**, 3595-3602.

8. Liu, Q., and Dreyfuss, G. (1996) EMBO J. **15**, 3555-3565.

9. Dietz, H. (1998) Nature Genet. **20**, 321-322.

10. Matera, A.G., and Frey, M.R. (1998) Am J. Hum. Genet. **63**, 317-321.

11. Fischer, U., Liu, Q., and Dreyfuss, G. (1997) Cell **90**, 1023-1029.

12. Mattaj, J.W. (1998) Curr. Biol. **8**, 93-95.

13. Pellizzoni, L. Kataoka, N., Charroux, B., and Dreyfuss, G. (1998) Cell **95**, 615-624.

14. Meister, G., Buhler, D., Laggerbauer, B., Zobawa, M., Lottspeich, F., and Fischer, U. (2000) Hum. Mol. Genet. **9**, 1977-1986.

15. Hirabayashi, J., and Kasai, K. (1991) J. Biol. Chem. **266**, 23648-23653.

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

17. Blum, H., Berer, H., and Gross, H.J. (1987) Electrophoresis **8**, 93-99

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

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

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

21. Zillman, M., Zapp, M.L., and Berget, S.M. (1988) Mol. Cell. Biol. **8**, 814-821.

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

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

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

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

26. Seetharaman, J., Kanigsberg, A., Slaaby, R., Leffler, H., Barondes, S.H., and Rini, J. (1998) J. Biol. Chem. **273**, 13047-13052.

27. Shimokawa, T., Kulmacz, R.J., DeWitt, D.L., and Smith, W.L. (1990) J. Biol. Chem. **265**, 22073-22076.

28. Niessing, D., Driever, W., Sprenger, F., Taubert, H., Jackle, H., and Rivera-Pomar, R. (2000) Mol. Cell **5**, 395-401.

29. Leonidas, D.D., Vatzaki, E.H., Vorum, H., Celis, J.E., Madsen, P., and Acharya, K.R. (1998) Biochemistry **37**, 13930-13940.

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

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

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

33. Probstmeier, R., Montag, D., and Schachner, M. (1995) J. Neurochem. **64**, 2465-2474.

34. Barboni, E.A.M., Bawumia, S., Henrick, K., and Hughes, R.C. (2000) Glycobiology **10**, 1201-1208.

Chapter 4

Nuclear galectin containing complexes

Introduction

The sedimentation of galectin-3 with hnRNP and snRNP complexes in cesium chloride buoyant density gradients is an interesting finding because it was performed without adding exogenous pre-mRNA. This leads to a hypothesis that the galectins might associate with nuclear complexes other than the spliceosome. To prove this hypothesis, several experiments were carried out and the existence of such complexes was confirmed. Some interesting properties of these complexes have also been observed in our experiments.

Observations

A. Galectin-1 and galectin-3 associate with nuclear complexes which contain snRNPs.

Using antibodies against galectin-1 and -3, snRNAs of U1, U2, U4, U5 and U6 were precipitated from HeLa cell nuclear extract. The procedure is similar to that described in Chapter 2. HeLa nuclear extract was incubated at 30°C in the presence of ATP for 30 minutes to remove any pre-existing splicing complexes. The extract was then incubated at 4°C with antibody coupled protein-A beads for 1 hour, followed with

washing and 5xSDS sample buffer elution. The bound fraction was then subjected to denaturing gel separation and resolved RNAs were transferred to hybond-N nylon membrane using standard northern blotting transfer procedure. To reveal the RNAs after transfer, we use both riboprobes transcribed from snRNA clones and 32p-labeled oligo nucleotide probes as following:

U1: **TCC CCT GCC AGG TAA GTA TC**

U2: **TTA GCC AAA AGG CCG AGA AGC GAT**

U4: **GGG GTA TTG GGA AAA GTT TTC**

U5: **ATC TGA AGA GAA ACC AGA GTA T**

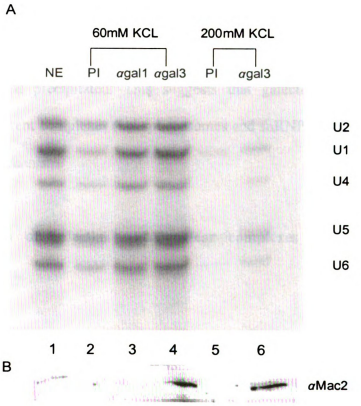
U6: **TTC TCT GTA TCG TTC CAA T**

The hybridization conditions are: using 20 ml hybridization solution [containing 3ml 20xSSC solution, 2 ml 50x Denhart's solution, 1 ml 1M NaH_2PO_4, 0.2 ml 10%SDS, 0.2 ml herring sperm DNA and 3 ml (for oligo probe) or 10 ml (for riboprobe) formamide], hybridize membrane at 42 °C over night and wash with 2xSSC solution for 1 hour.

Figure 1 shows the antibody precipitated snRNAs after northern blotting. As we expected, pre-immune serum only precipitated background levels of snRNAs, while significant amounts of snRNAs were precipitated by anti-galectin-1 and anti-galectin-3. These data strongly indicate the presence of galectin-1 and galectin-3 in snRNP containing complexes.

Figure 1: Galectin-1 and galectin-3 associate with nuclear complexes which contain snRNPs and this association is sensitive to ionic strength. In panel A, using antibodies against galectin-1 and -3, snRNAs of U1, U2, U4, U5 and U6 were precipitated from HeLa cell nuclear extract. Lane 1 shows nuclear snRNAs (U1, U2, U4, U5 and U6) as positive control. In lane 2, preimmune serum only precipitates background level snRNPs. In lanes 3 and 4 anti-galectin-1 and -3 precipitate snRNAs, respectively. Lanes 5 and 6 show similar precipitation under high ionic strength condition (200mM KCl). Note the amount of snRNAs precipitated by galectin antibody was greatly reduced under such condition. Panel B shows the western blotting of galectin-3 in the bound fraction with the same lane order.

Figure 1


A
60mM KCL
200mM KCL
NE
PI
αgal1
αgal3
PI
αgal3
U2
U1
U4
U5
U6
1
2
3
4
5
6
B
αMac2

B. The association of galectins with nuclear complexes is sensitive to ionic strength.

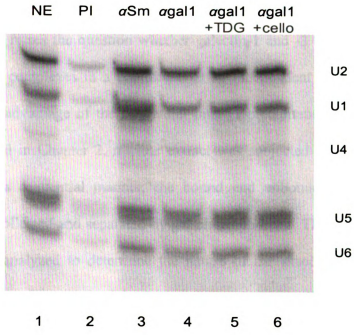
Since ionic strength is a very important factor in galectins' association with spliceosomes, it was interesting to see how ionic conditions affect the association of galectins with nuclear complexes. We used the same method for immunoprecipitation, but increased the salt concentration of the binding and washing solutions from 60mM to 200mM. As shown in Figure 1, increasing the salt concentration significantly reduced the amount of RNA that precipitated. This suggests that galectin-1 and -3 probably incorporate into different complexes (i.e. spliceosomes and snRNP containing complexes) through a similar mechanism.

C. The association of galectins with nuclear complexes is not affected by carbohydrate ligands.

In previous experiments testing carbohydrate ligands effect on galectins' interaction with spliceosomes, TDG inhibited the precipitation of pre-mRNA in H complex, but not in E and active splicing complexes, indicating the association of galectins with the active spliceosomes can not be disrupted by carbohydrate ligands. A similar experiment performed without pre-mRNA will give us a clue of the role of the CRD in galectins' association with nuclear complexes. HeLa nuclear extract was pre-incubated with 200mM TDG or cellobiose. Splicing reaction has been documented to be inhibitable by TDG at this concentration. Then the nuclear extract was subjected to anti-galectin-1 immunoprecipitation. Figure 2 shows the effect of TDG on anti-galectin-1

Figure 2: The association of galectins with nuclear complexes is not affected by carbohydrate ligands. HeLa nuclear extract was pre-incubated with 200mM TDG or cellobiose. Then the nuclear extract was subjected to anti-galectin-1 immunoprecipitation. Comparing to regular precipitation (lane 4), there were no significant changes in the amount of snRNAs precipitated with the addition of TDG (lane 5) or cellobiose control (lane 6). Preimmune (lane 2) and anti-Sm (lane 3) served as negative and positive controls.

Figure 2



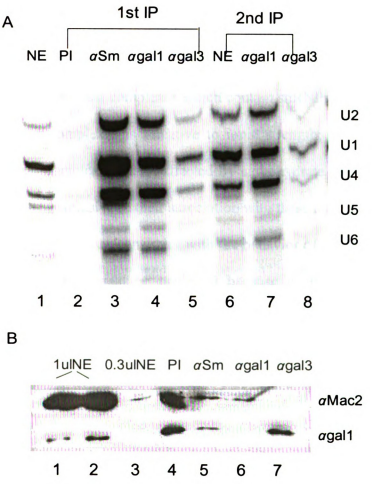
precipitating snRNAs. There is no change in the quantity of snRNAs precipitated with the addition of TDG into the extract when comparing lanes 4, 5 and 6.

D. Galectin-1 and -3 are also mutually exclusive in their association with nuclear complexes.

We previously discovered that the association of galectin-1 and -3 with spliceosomes is mutually exclusive, indicating there are at least two different types of spliceosomes. This raises the question whether galectin-1 and -3 associate with snRNP containing nuclear complexes in a coordinate or an independent way. To address this question, we took advantage of the sequential immunoprecipitation method. Using the procedure described in Chapter 2, nuclear extract was subjected to anti-galectin-1 and anti-galectin-3 in a sequential manner, the bound and unbound fraction from each precipitation was collected and separated by gel electrophoresis. The snRNA and protein components were analyzed to determine the nature of the associations. As shown in Figure 3, in the first immunoprecipitation, preimmune serum did not precipitate any snRNAs. Anti-Sm precipitated snRNAs and served as a positive control. Anti-galectin-1 and anti-galectin-3 both precipitated snRNAs in this round as they did in previous experiments. When the unbound fraction of the first precipitation passed through the opposite antibody beads in the second round of precipitation, significant amounts of snRNAs were still precipitable. The western blot data shows that after extract went through the first antibody column, >85% of the galectin-1 or galectin-3 protein was removed from the extract, but the opposite protein was still present in the first

Figure 3: Galectin-1 and -3 are mutually exclusive in their association with nuclear complexes. HeLa nuclear extract was subjected to anti-galectin-1 and -3 in a sequential manner, the bound and unbound fraction from each precipitation was collected and the snRNA and protein components were analyzed, respectively. Panel A shows the snRNAs precipitation. In the first immunoprecipitation, anti-galectin-1 and anti-galectin-3 both precipitated snRNAs from the nuclear extract (lanes 4 and 5). Preimmune (lane 2) and anti-Sm (lane 3) served as negative and positive controls. When the unbound fraction of the first precipitation passed through the opposite antibody in the second precipitation, significant amounts of snRNAs were still precipitable (lanes 7 and 8). Panel B shows the western blotting (αMac2 was used to detect galectin-3) of the unbound fraction of the first precipitation. Lanes 6 and 7 shows >85% of the galectin-1 or galectin-3 protein was removed from the extract by its own antibody, but the opposite galectin was still present.

Figure 3


A
1st IP
2nd IP
NE PI αSm αgal1 αgal3 NE αgal1 αgal3
U2
U1
U4
U5
U6
1 2 3 4 5 6 7 8
B
1ulNE 0.3ulNE PI αSm αgal1 αgal3
αMac2
αgal1
1 2 3 4 5 6 7

unbound fraction, indicating the removal of one galectin does not affect the other. These data suggest that each galectin associates with a subset of snRNPs independently of one another.

Discussion

In this series of experiments, we got the following major results: 1, Galectin-1 and galectin-3 associate with nuclear complexes which contain snRNPs; 2, The association of galectins with nuclear complexes is sensitive to ionic strength; 3, The association of galectins with nuclear complexes is not affected by carbohydrate ligands; 4, Galectin-1 and -3 are also mutually exclusive in their association with these nuclear complexes.

Most of these findings are consistent with the behavior of galectin-1 and -3 in the spliceosomes. These results could be predicted because it is easy to conceive that galectin-1 and -3 are more likely to be a component of pre-existing splicing related complexes than a free protein in the nuclear extract. Dreyfuss' study has indicated that nuclear snRNP particles go through regenerating cycles. Nuclear galectins are likely to be involved in such process. They probably incorporate into the particles and, as we have suggested, deliver the snRNPs to the spliceosomes. The precipitation of snRNAs from nuclear extract by galectin antibodies strongly supports this idea.

Like the association with spliceosomes, the association of galectins with nuclear complexes is sensitive to ionic strength. This suggests galectins probably loosely bind to some protein components in the nuclear complex. A weak interaction does not mean it is an unimportant one. In fact, many regulatory factors bind to their binding partners with a

rather "weak" interaction. Galectin-3 has been suggested to be a regulatory factor in other research. In thyroid cancer cell, galectin-3 directly interacts with a transcription factor TTF-1, stimulates the DNA binding activity of this protein and thus contributes to the proliferation of the thyroid cell. Galectin-3 has also been reported to interact with Sufu protein, which is a negative regulator of the Hedgehog signal transduction pathway. Galectin-3 can affect the localization of this protein between nucleus and cytoplasm and hence play a role in regulating this specific pathway. Overall, the role of galectins in splicing maybe far more sophisticated than we expected.

The fact that TDG did not affect galectins' association with the snRNPs is not unexpected. In previous studies, we have found out that adding excess TDG can only affect the association of galectins with the H complex, but not that of galectins with the E and active splicing complexes. We have proposed that when TDG binds to galectin-1 or -3, it will induce a conformational change and thus prohibit galectins from binding to their protein partners in the splicing complex. This is probably the case of the H complex. On the other hand, in the nuclear complexes or active splicing complexes, because galectins have incorporated stably into a complex, it will be covered by other proteins and won't be accessible to TDG anymore. Hence TDG won't affect the association of galectins with these complexes.

The sequential precipitation study proved our theory that in the nuclear extract, there are pre-existing galectin containing snRNP complexes and galectin-1 and -3 are mutually exclusive in these complexes (i.e. one complex only containing one type of galectin). When pre-mRNA is added to the extract, galectins might facilitate the complexes they associated with to interact with the pre-mRNA, deliver the necessary

factors to the splicing reaction. This finding is consistent with galectin-spliceosome association pattern. Since galectin-1 and -3 do not coexist in the pre-existing complexes, they would not associate with the spliceosome in a co-existing manner.

Chapter 5

Conclusion and future plans

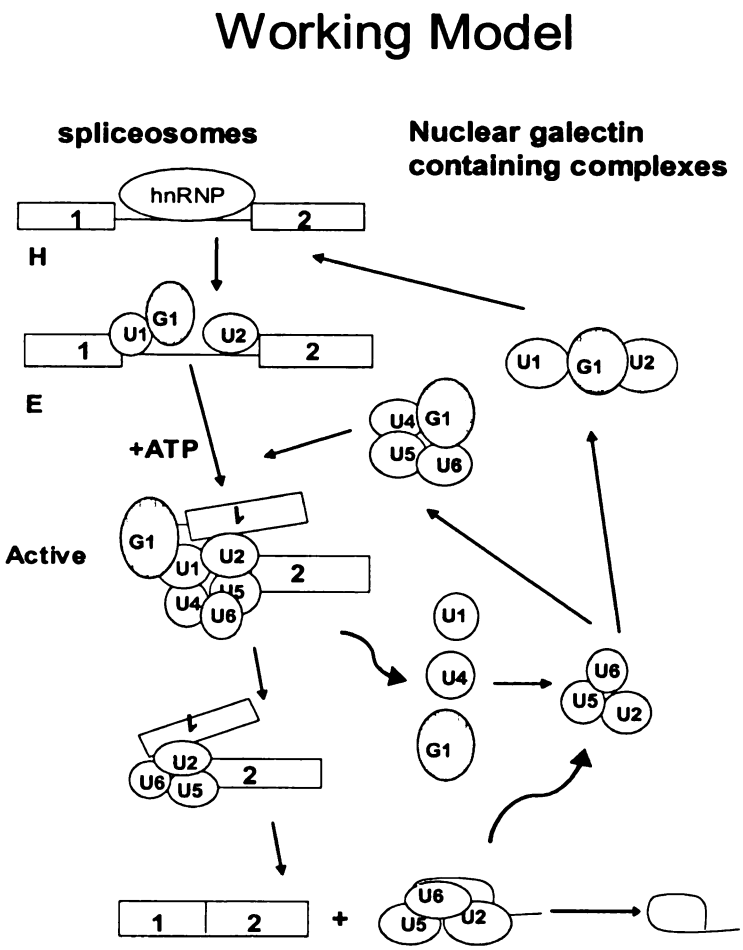
Pre-mRNA splicing is a very important post-transcriptional process mediated by a large complex called the spliceosome. Previous studies have confirmed that nuclear galectin-1 and -3 are essential splicing factors. After years of working on galectins and splicing, we have achieved some important discoveries and elucidated part of the roles of galectins in the splicing process.

My work has partially solved the problem of how galectin-1 and -3 exert their functions in the splicing process. Galectin-1 and -3 are associated with nuclear snRNP containing complexes without pre-mRNA. After pre-mRNA is made, galectins will facilitate the delivery of the snRNP particles onto the pre-mRNA which allowed the sequential forming of the E, A, B and C splicing complexes. The function of galectins starts at the assembly of the E complex and they may slowly dissociate from the spliceosomes during the splicing process (Figure 1). Each of the nuclear galectins functions by itself and excludes the other from the same complex.

There are still many questions need to be addressed concerning galectins' roles in splicing. We need to isolate the nuclear galectin containing complexes and determine the constituents of these complexes. Silver staining of polyacrylamide gels, mass spectrometry and database searches could be used to enumerate and identify the polypeptides of these complexes. This could give us clues to determine the real binding partners of galectins in the splicing process and further elucidate the mechanism of splicing. Whether galectins have other important roles in the spliceosome after they

Figure 1: Schematic diagram illustrating the role of galectin-1 and -3 in the splicing process. Galectin-1 and -3 are associated with nuclear snRNP containing complexes without pre-mRNA. Each of such complexes contains only one type of galectin. After pre-mRNA is made, galectins facilitate the delivery of the snRNP particles onto the pre-mRNA which allow the sequential forming the E, A, B and C splicing complexes. When mature mRNA was released, galectins might be involved in the regeneration process of the snRNPs.

Figure 1


Working Model
spliceosomes
Nuclear galectin containing complexes
hnRNP
1
2
H
G1
U1
U2
E
+ATP
U4
U5
U6
Active
U1
U4
G1
U6
U5
U2
+

delivering the snRNPs is another important question. Many splicing factors are involved in spliceosomal protein and RNA rearrangement and relocation. Do nuclear galectins also have such role in these processes? Lastly, what happens after the mRNA is released and the spliceosome dissociates? We have seen data suggesting that galectins could be involved in the regeneration of snRNP particles in the nucleus. This should be a new area for us to explore.

Another line of my research has discovered that the carbohydrate binding activity of galectins is separated from their splicing activity. The evidence is that mutant galectins without carbohydrate binding activity can still reconstitute splicing reaction in a galectin depleted system. To further confirm this conclusion, we should construct a mutant galectin which abolished its splicing activity and test its ability to bind to lactose or TDG. A positive result of such experiments will strengthen our conclusion drastically. It is also important to find out the entire conformational changes of galectins after they bind to lactose. It will help us to understand why lactose can inhibit splicing and determine the binding surface of galectins with their partners in the spliceosome.

The data contained in this thesis have laid a foundation for these future studies. The entire role of nuclear galectins in splicing will be eventually determined.