This is to certify that the
dissertation entitled

Role of galectin-1 in pre-mRNA Splicing

presented by

Jung W. Park

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Deptment of Microbiology and Molecular Genetics

Ronald J. Patterson
Major Professor's Signature

8/10/03
Date

Role of galectin-1 in pre-mRNA splicing

By

Jung W. Park

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

2003

Abstract

Role of galectins in pre-mRNA splicing

By

Jung W. Park

In previous studies, we had shown that galectin-1 and -3 are required splicing factors using cell-free splicing assays. To elucidate their role(s) in splicing, we performed a yeast two-hybrid screen with galectin-1 as a "bait" to identify interacting partners. Gemin4, a component of the SMN (Survival of Motor Neuron) complex, which is involved in snRNP assembly in the cytoplasm and implicated in snRNP recycling in the nucleus, was identified as a putative ligand. We confirmed the yeast two-hybrid result by in vitro binding assay in which both galectin-1 and -3 interacted with the carboxy terminal 50 amino acids of Gemin4 and immunoprecipitation experiments in which we showed the association of galectin-1 and -3 in a complex containing SMN, Gemin2, and Sm core polypeptides. In addition, we documented that galectin-1 and -3 are associated with early commitment and active spliceosomal complexes by showing coimmunoprecipitation of pre-mRNA

substrates and RNA intermediates contained in spliceosomal complexes by antibodies specific for galectins. However, preferential association of galectin-1 with intron lariat and galectin-3 with ligated exons (mature RNAs), respectively, suggests that galectin-1 and galectin-3 are involved in two distinct steps in splicing pathway. We further demonstrated that SMN is contained in these spliceosomal complexes and this finding was confirmed by data showing immunoprecipitation of splicing substrates by antibodies specific for SMN. Both the amino terminal domain of galectin-3 and the carboxy terminal 50 amino acids of Gemin4 act as dominant negative polypeptides by inhibiting splicing activity and arresting spliceosomal complex formation at the H complex, indicating a direct role of galectin-3 and Gemin4 in splicing. Taken together, we conclude that galectin-1 and galectin-3 are components of the splicing machinery and have role(s) in assembly of the early spliceosome. An additional function of galectin-1 in pre-mRNA splicing may be to facilitate the reassembly or recycling of multi-snRNPs particles through interaction with the SMN complex, which is involved in the assembly of multiple nuclear RNP complexes.

Acknowledgments

First and foremost, I would like to thank my mentor, Dr. Ronald Patterson. He fully supported, helped, and motivated me throughout my graduate school. Without him, I could have not accomplished this. His patience and genuine interest in my growth as a scientist were indispensable for me to stay focused during times I was bewildered and losing directions. He was more than a typical academic advisor. I will cherish the moments we discussed about our research, life, and science.

Next, I would like to gratefully thank my family. There were moments when I thought I could not finish the graduate school. My parents, Young-Im and Min-Hyung Park, continuously supported, encouraged me during hard times. They taught me that you have to work hard to achieve meaningful things in life.

I would like to express my deep gratitude to my friends for the friendships we have developed during graduate school. I will certainly miss Weizhong Wang, Kevin Haudek, Chancey Spooner, Sandi Clemens, Sheldon Leung, Kwisook Kim, Namjoon Kim, Jongkyong Kim for their moral support and friendships.

Finally, I would like thank my wife, Sunghee, Daniel, and Elizabeth for their love and support. Without them, I would not be here.

I dedicate this work to Sunghee, Daniel, and Elizabeth Park for their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Literature Review

Structure of galectins

Galectins are a family of lectins that specifically bind to β-galactoside sugars through a highly conserved carbohydrate recognition domain (CRD) (Barondes et al., 1994a). These lectins, formerly known as S-type or S-lac lectins, are found in many eukaryotic tissues ranging from lower invertebrates such as sponges to mammals including humans (Cooper and Barondes 1999). Members of the galectin family share two distinct properties: similar amino acid sequence and specific affinity for ß-galactoside residues (Gal(ß1 --> 4) GlcNAc). Galectins can be classified into three subgroups according to the content and organization of domains. Among the 14 members of the mammalian galectin family (galectin-1 to galectin-14) discovered so far, the first group is the prototype (includes galectin-1, -2, -5, -7, -10, -11, -13 and -14) which contains one CRD of about 130 amino acids (Figure 1). The three-dimensional structure of galectin-1, -2, -7, and -10, elucidated by X-ray crystallography, revealed that the CRD is folded with two anti-parallel β-sheets forming a jelly roll-like structure (Barondes et al., 1994). Most of the prototype galectins

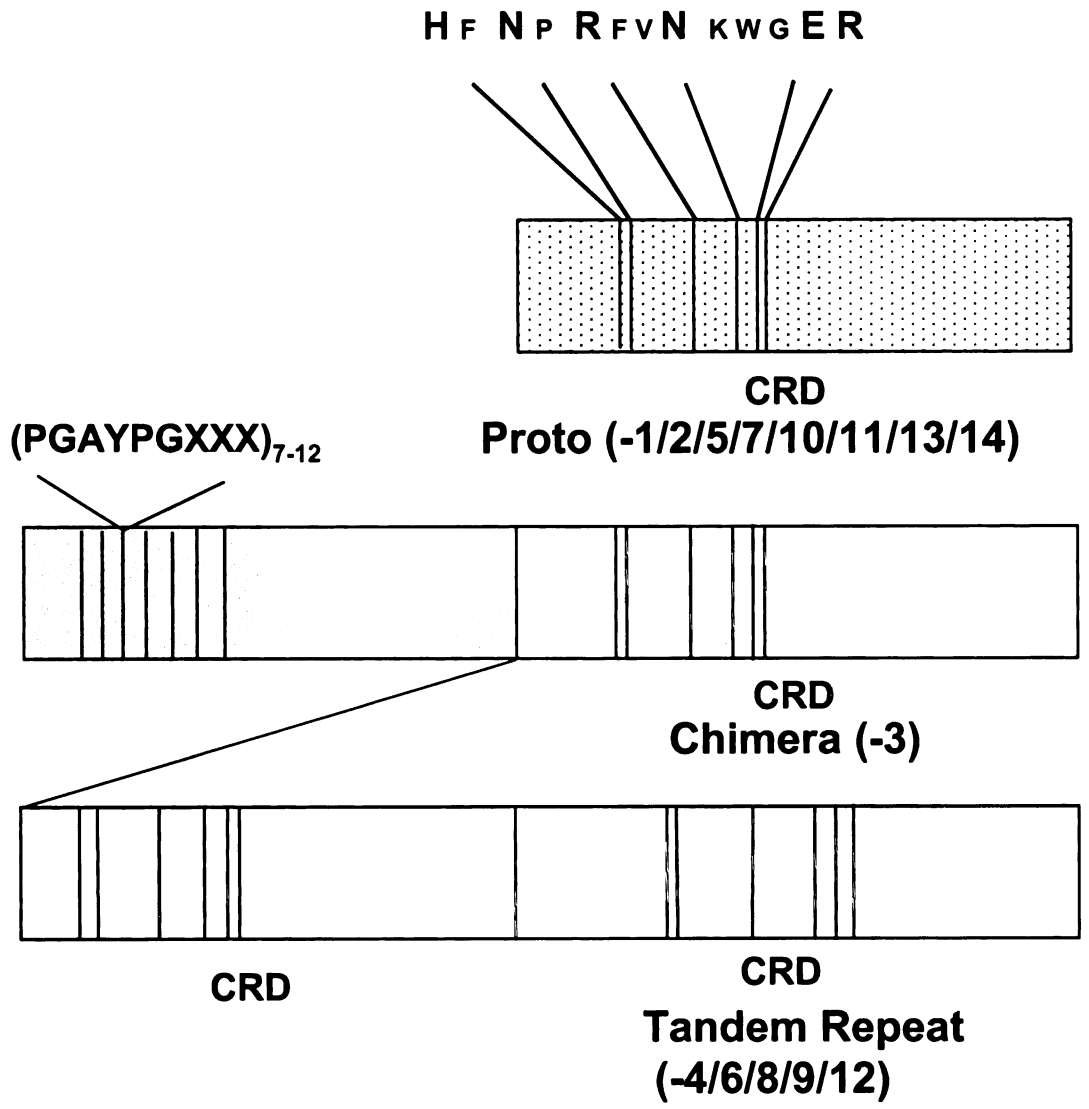

HF NP RFVN KWG E R
CRD
Proto (-1/2/5/7/10/11/13/14)
(PGAYPGXXX)7-12
CRD
Chimera (-3)
CRD
CRD
Tandem Repeat
(-4/6/8/9/12)

Figure 1. Schematic representation of the three types of galectins (proto-, chimera-, and tandem repeat). The proto-type is composed of a single carbohydrate recognition domain (CRD). The chimera-type is composed of two parts; a carboxy terminal domain containing the CRD and an amino terminal domain of unknown function. The amino terminal domain contains 7-12 repeats of (PGAYPGXXX). Tandem repeat-type is composed of two similar CRDs. Highly conserved residues among all galectins are shown in bold letters above the proto-type structure.

form non-covalent homodimers under physiological conditions (Kasai and Hirabayashi 1996). The prototype galectins were the first galectins identified by their hemagglutination activity which is attributable to their ability to form divalent homodimers. Others, such as rat galectin-5, are monomeric (Gitt et al., 1995).

The members of the second group (tandem repeat) include galectin-4, -6, -8, -9, and -12. They are monomers with a molecular mass of ~35 - 36 kDa composed of two highly similar tandem repeats of 130 amino acids each connected by a short linker peptide (Figure 1). No crystal structure of tandem repeat type galectins has been attained yet. Since tandem repeat type galectins are made of two CRDs, they could simultaneously bind to multivalent ligands, possibly two distinct types of ligands (Arata et al., 1997). Comparative sequence studies have shown a large difference between galectin-4 and -6 only in the linker peptide region, implying a biological importance of linker peptide possibly influencing the ligand cross-linking properties of tandem repeat galectins (Gitt et al., 1998).

The third type of galectin family is called chimera type because it is composed of two distinct domains. The only chimera type galectin so far identified is galectin-3. Galectin-3 is made of a CRD in its carboxy domain and a unique amino terminal domain with several repeats of pro-

gly-ala-tyr-pro-gly-X-X-X peptide, characteristic of the collagen gene superfamily. Tertiary structure of the CRD of galectin-3 has been resolved by X-ray crystallography and its structure is similar to that of galectin-1 and -2. However, the CRD of galectin-3 fails to form homodimers by analogous CRD:CRD interactions unlike galectin-1 and -2 (Seetharaman et al., 1998). However, the full length galectin-3 binds to multiglycosylated proteins with positive cooperativity indicating the ability to form dimers. This multivalent activity of galectin-3 is not seen when the CRD fragment lacking the amino terminal domain alone is used (Hsu et al., 1992; Messa et al., 1993). This implies that galectin-3 may be associated into multimeric complexes through amino terminal domains (Barboni et al., 2000). However, the CRD of galectin-3 itself may play a role in self-association (Yang et al., 1998).

Another unique structural feature of galectin-3 is its phosphorylation on serine 6. Although the enzyme catalyzing galectin-3 phosphorylation has not been identified, serine 6 was shown to an effective substrate for casein kinase I *in vitro* (Huflejt et al., 1993). Although the functional significance of the phosphorylation has yet to be established, the phosphorylation of galectin-3 lowers binding to laminin and purified mucin from colon cancer (Mazurek et al., 2000).

Subcellular localization of galectins

Another important biological feature of galectins is their unique cellular localization pattern. Although neither signal peptide nor nuclear localization sequences on galectin molecules has yet been identified, several members of the galectin family localize in all three compartments; nucleus, cytoplasm, and extracellular space (either free or membrane bound (Hirabayashi 1997). It is thought that all galectins are synthesized on free cytoplasmic ribosomes. Furthermore, the amino terminus of both galectin-1 and galectin-3 is acetylated, which is typical of cytosolic proteins (Clerch et al., 1988; Herrmann et al., 1993). These data suggest that galectins' roles at the time of their first appearance could have resided inside cells, but later they have acquired extracellular functions through development (Kasai and Hirabayashi 1996). In any case, galectin-1 has been clearly shown to be externalized to the extracellular matrix in mouse C2 myoblasts through a novel secretory mechanism (Cooper and Barondes 1990). There is also evidence that galectin-3 is externalized to the cell surface and extracellular matrix via "membrane blebbing" (Mehul and Hughes 1997).

Some galectins have been shown to localize in both nuclear and cytoplasmic compartments. Earlier work on

intracellular localization of galectins showed only cytoplasmic staining as demonstrated in CHO cells (Cho and Cummings 1995)and chick embryo fibroblasts (Briles et al., 1979). However, those results turned out to be due to technical problems in that the detergents used in permeabilizing the plasma membrane could not permeabilize the nuclear membrane. Consequently, treating *HeLa* cells with detergents that are capable of permeabilizing the nuclear membrane clearly showed galectin-1 and galectin-3 staining in the nucleus (Vyakarnam et al., 1998). In addition, galectin-1 was identified as a component of the nuclear matrix in rat osteoblasts (Choi et al., 1998)and the localization of galectin-1 in the nucleus was also demonstrated by an electron microscopic study of Langerhans cells in the epidermis (Akimoto et al., 1992). Also, the intracellular localization of galectin-1 is strongly supported by findings in which an oncogene, H-Ras (Paz et al., 2001) and Gemin4 (Park et al., 2001) were identified as interacting ligands for galectin-1.

Galectin-3 was first identified as a major macrophage cell surface antigen, Mac-2 (Cherayil et al., 1989), implicated in protein-carbohydrate interaction with laminin (Runyan et al., 1988). Metastasis of malignant cells was also related to the presence of galectins on the cell surface (Raz and Lotan 1987).

Intracellular localization of galectin-3 was initially observed by an immunofluorescence study of mouse 3T3 fibroblasts, fixed with paraformaldehyde, permeabilized with Triton X-100, and stained with antibodies against galectin-3. There was a strong labeling of both the nucleus and cytoplasm (Moutsatsos et al., 1986). The nuclear versus cytoplasmic distribution of galectin-3 was shown to be dependent on the proliferative state of the cells. Sparse, proliferating cells showed nuclear localization of galectin-3; in contrast, confluent, quiescent cells showed mostly cytoplasmic localization of galectin-3 (Moutsatsos et al., 1987). The phosphorylation state of galectin-3 was shown to be also dependent on the proliferative state of 3T3 fibroblasts: only phosphorylated galectin-3 is found in cytoplasm in quiescent cells and both phosphorylated and unphosphorylated forms of galectin-3 are found in the nucleus in proliferating cells (Cowles, 1990 #238). Treatment of permeabilized mouse 3T3 fibroblasts with ribonuclease A released galectin-3 from the nucleus, whereas treatment with DNase I did not. In addition, mouse 3T3 galectin-3 co-fractionated with hnRNP and snRNP complexes in Cs_2SO_4 gradients at densities between 1.28 and 1.32 g/ml (Laing and Wang 1988). It also has been observed that only phosphorylated galectin-3 is exported out of digitonin-treated 3T3 fibroblasts. More interestingly, the export of

galectin-3 was inhibited by addition of leptomycin B, a drug which specifically disrupt the CRM1-mediated export process. In the same report, galectin-3 was shown to be exported as a large molecular weight complex (~650 kDa)(Tsay et al., 1999). Therefore, the omnipresent feature of galectins may be related to their specific subcellular functions depending on their available ligands in each compartment.

Biological functions of galectins: intracellular and extracellular

The biological significance of galectins is evident based on the following observations. Any given organism expresses at least one galectin and often multiple members of the galectin family. Different tissues or cell types within the organism contain distinct galectins. Also, expression of galectins is regulated in a tissue-specific manner and developmentally (Colnot et al., 1996). Galectins have been initially implicated to play some role in modulating cell-cell and cell-matrix interactions because of their carbohydrate binding properties. In fact, galectin-1 was found in the cytoplasm as well as extracellular matrix and has been shown to interact with laminin through protein-carbohydrate interaction between the carbohydrate recognition domain (CRD) of galectin-1 and the carbohydrate

residues of laminin(Cooper et al., 1991b). Galectin-1 is implicated in cell growth and neural differentiation by interacting with ganglioside GM (Koiptz et al., 1998). Galectin-1 induces apoptosis of T-cells when added exogenously (Perillo et al., 1995). In addition, galectin-1 knockout mice resulted in defects of olfactory neurons (Tenne-Brown et al., 1998). Intracellularly, galectin-1 is involved in pre-mRNA splicing (Vyakarnam et al., 1997).

Although the biological role of tandem repeat galectins is not known at this time, they presumably are involved in similar processes as galectin-1 or galectin-3 considering their similar amino acid sequence, localization, and regulated level of expression. Galectin-9, also known as Ecalectin, initially isolated from serum of patients with Hodgkin's disease, is implicated in eosinophil chemoattraction (Matsushita et al., 2000).

Galectin-1/-3 knockout mice were generated but did not manifest any obvious defect in implantation or early development. However, the same study has found that another galectin, galectin-5, is expressed in an early stage of embryogenesis. Furthermore, the same group reports that galectin-3 may be involved in bone development (Colnot et al., 1998).

Galectin-3, previously known as CBP35, Mac-2, IgE-binding protein, CBP30, L-29, LR-29, and L-34 according to

its localization in different tissue types or its functional properties, is reported to bind laminin (Sato et al., 1992; Woo et al., 1990), be a modulator of cell-cell and cell-matrix interaction (Lamm and Lamond 1993), be a modulator of immune responses (Rabinovich et al., 2002), bind to lipopolysaccharide (LPS) (Mey et al., 1996), inhibit apoptosis (Yang et al., 1996), cause metastasis (Hebert and Monsigny, 1994), and interact with a novel cysteine and histidine-rich cytoplasmic protein in a carbohydrate independent manner (Menon et al., 2000). Galectin-3 localizes in the nucleus and is involved in pre-mRNA splicing (Dagher et al., 1995). Although it is not clear how one protein can be involved in many different cellular functions, it is not an unprecedented idea in which a protein could play multiple roles depending on the localization, interacting partner, and cellular signal (Smalheiser et al., 1996). Evidence for galectin-3 as well as galectin-1 in pre-mRNA splicing will be discussed later. Although galectin-3's precise role(s) in the nucleus has to be established, its involvement in pre-mRNA splicing may be important for the regulation of cell growth.

Mechanism of pre-mRNA splicing

Pre-mRNA splicing, removal of intervening sequences from precursor mRNA, is one of the essential post-

transcriptional processes of most eukaryotic mRNAs (Sharp, 1994). The catalysis of pre-mRNA splicing involves two separate transesterification reactions of the pre-mRNA substrate. Step 1 involves a nucleophilic attack on the phosphodiester bond at the 5′ intron-exon junction from a 2′ ribose hydroxyl group of an adenosine residue at the branch point in the intron. This generates reaction intermediates corresponding to a free 5′ exon and a lariat attached to the 3′ exon. In step 2, a second nucleophilic attack occurs from free 3′ ribose hydroxyl group of the 5′ exon on the phosphodiester bond at the 3′ intron-exon junction. This results in the formation of products corresponding to ligated exons (mRNA) and fully excised intron lariat (Moore et al., 1993) (Figure 2).

In all metazoans, pre-mRNA splicing is mediated by the macromolecular complex called the spliceosome in an energy-dependent manner. The spliceosome is composed of five small nuclear ribonucleoprotein complexes (snRNPs), U1, U2, U4, U5, and U6 and a number of non-snRNP proteins. Discovery of group II introns which self splice *in vitro* by an identical two-step transesterification mechanism, independent of ATP hydrolysis and protein components, opened the question if those two requirements are indispensable (Michel and Ferat 1995). Genetic studies involving *in vitro* mutagenesis have

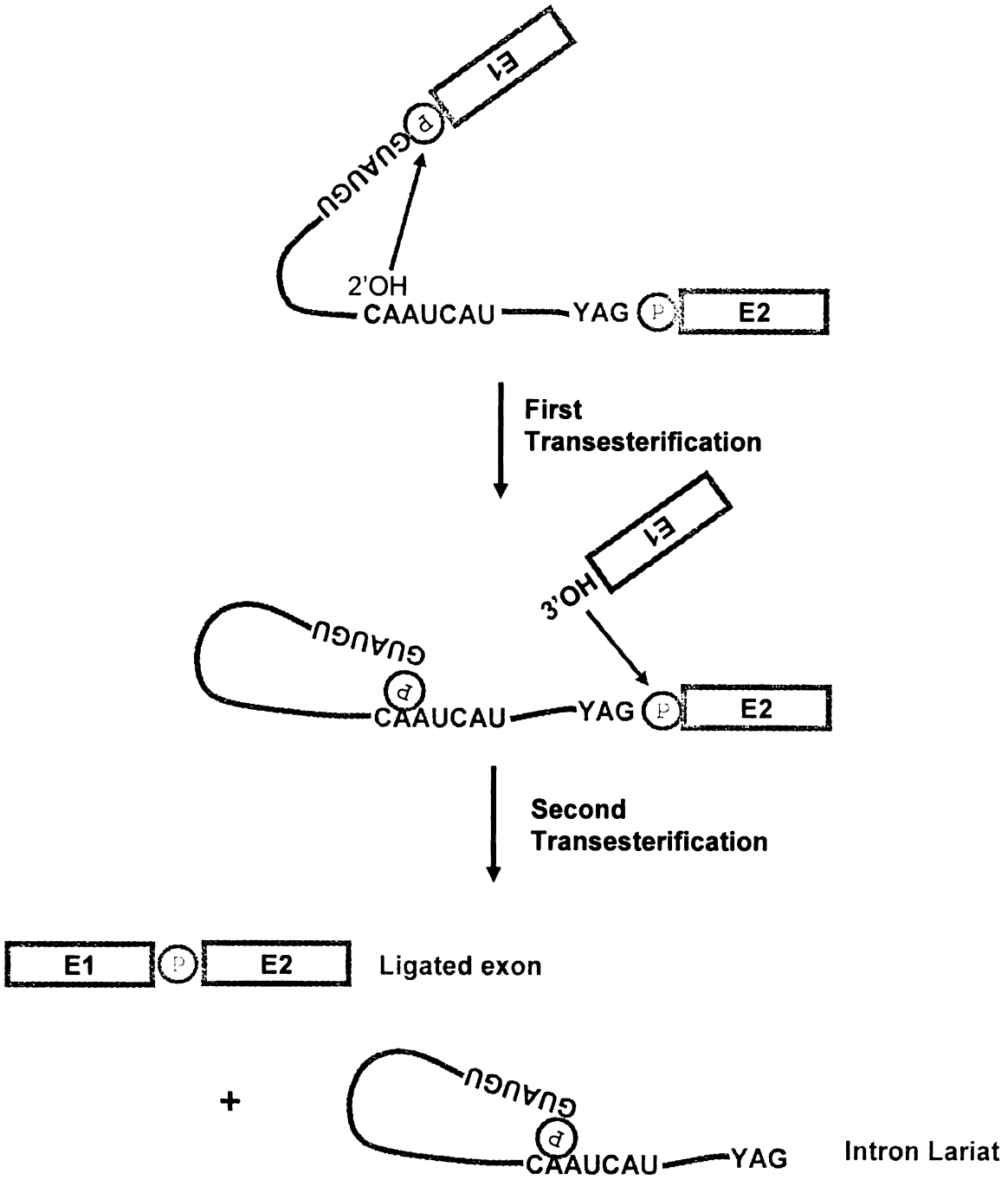

E1
GUAUGU
2'OH
CAAUCAU
YAG
E2
First
Transesterification
3'OH
Second
Transesterification
E1
E2
Ligated exon
+
Intron Lariat

Figure 2: Schematic diagram depicting the mechanism of spliceosome mediated pre-mRNA splicing. Step 1 is a transesterification reaction in which 2′-OH of branch point (adenosine) nucleophilically attacks the phosphate group of 5′ splice junction (guanosine), resulting in a free exon 1 and a lariat-exon 2 intermediates. Step 2 is another transesterification reaction in which free 3′-OH of exon 1 nucleophilically attacks the phosphate group of 3′ splice junction (guanosine), resulting in the ligated exons and lariat.

shown that the base-pairing interactions between snRNA and pre-mRNA are necessary for the assembly of the spliceosome (Wassarman and Steitz 1992). These interactions between snRNA and pre-mRNA have been confirmed by UV cross-linking at specific stages of spliceosome assembly (Staley and Guthrie 1998). The mechanistic similarities between splicing of nuclear pre-mRNA involving the spliceosome and group II introns, which are self-catalytic and ATP-independent, support the idea that RNA-RNA interactions play a central role, and may even catalyze the splicing reactions (Valadkhan and Manley 2001).

However, many biochemical and genetic studies have revealed function(s) of those protein factors comprising the machinery in achieving accurate and precise splicing. Although RNAs are directly involved in the catalysis of pre-mRNA splicing, snRNP proteins and a large number of non-snRNP proteins must play critical roles in positioning, stabilizing, and disrupting RNA-RNA interactions in the spliceosome. The roles of those components will be discussed in the next section in detail.

In mammals, the consensus sequence of the 5′ splice site is AG/GUAUGU (/ is the exon/intron junction). Three distinct sequences are conserved at the 3′ splice site; the branch point (PyNPyPyRAC), a polypyrimidine tract, and the 3′ splice site (PyAG/N). A specific association of each

snRNP with those conserved sequence elements is monitored, mediated, and optimized by many non-snRNP proteins. These proteins, many of them termed SR proteins due to the high content of serine and arginine residues, bind to both intronic and/or exonic enhancer elements to activate the recruitment of snRNPs and subsequent assembly of the spliceosome. There are intronic and/or exonic silencer elements which result in the negative regulation of pre-mRNA splicing and play key roles in splice site selection. A close communication of these enhancers and silencers with many non-snRNP proteins is also an important regulatory mechanism in alternative splicing (Black 2000).

Spliceosome assembly

Considerable progress has been made in understanding the model of spliceosome assembly. It is an ordered process by which five small nuclear ribonucleoprotein particles and many non-snRNP splicing factors are recruited and assembled onto the RNA substrate sequentially (Reed 2000)(figure 3). The assembly begins with the association of the U1 snRNP with pre-mRNA. U1 snRNP recognizes 5′ splice site through RNA-RNA complementary interaction. The base pairing between the 5′ splice consensus sequence and a highly conserved sequence at the 5′ end of U1 snRNA initiates a cascade of

events leading to a large complex in a highly systematic manner.

Subsequently, U2AF and SF1/Branch-point Binding Protein (BBP) bind to the polypyrimidine tract and branch-point respectively (Abovich and Rosbash 1997). U2AF large subunit (65 kDa) and small subunit (35 kDa) serve as a platform for U2 snRNP binding or at least facilitate its binding. The main function of U2AF small subunit of 35 kDa (U2AF35) appears to be in recognizing the 3′ splice site (AG/N) for further recruiting other splicing factors such as SR proteins, which will be described below (Wu et al., 1999). Also, U2AF35 serves as a bridge molecule between U1 snRNP bound to 5′ splice site and 3′ splice site by interacting with two additional SR proteins, SC35 and SF2/ASF (Splicing Factor 2/Alternative Splicing Factor) bringing two splice sites in proximity (Wu and Maniatis 1993). This complex commits the pre-mRNA to the splicing pathway and is thus called commitment complex (CC) or early complex (E complex) (Michaud and Reed 1991). It is known that E complex does not require ATP hydrolysis to form (Michaud and Reed 1991). Although E complex formation is energy-independent, it is more efficient at 30°C indicating a possible role of ATP hydrolysis in E complex formation (Das et al., 2000). Next, U2 snRNP binds to the branch-point to form the A

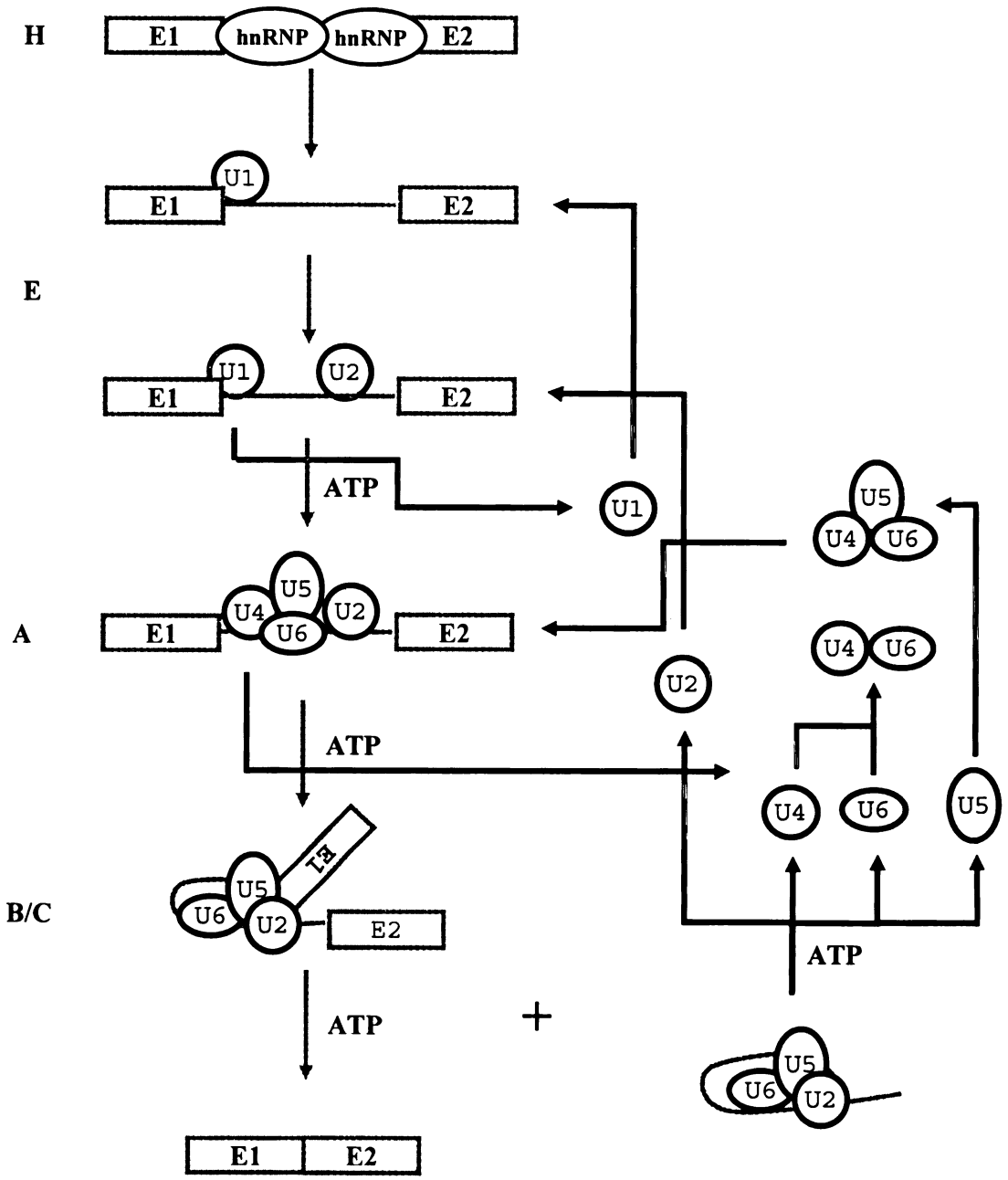

H
E1
hnRNP
hnRNP
E2
U1
E
U2
ATP
A
U5
U4
U6
B/C
E1
E2
+

Figure 3: Schematic diagram delineating the mechanism of spliceosome assembly and disassembly. The association of pre-mRNA with hnRNP results in the formation of ATP-independent H complex. Addition of U1 and U2 snRNPs results in an intermediate complex (early (E) complex or commitment complex). The association of a preassembled U4/U5/U6 tri-snRNPs with E complex results in A complex, which is converted to B and C complexes in ATP-dependent manner. After the ligated exon (mature mRNA) is released, the lariat form of the intron is still associated with snRNPs in a complex. While the snRNPs are recycled, the lariat intron is debranched and degraded.

complex (Bennett et al., 1992). U2 snRNP replaces BBP by switching protein:RNA interaction of BBP and branch-point to RNA:RNA base pairing between branch-point and 5′ end of U2 snRNA and the replacement requires ATP hydrolysis (Nilsen, 1998). This switch is a mutually exclusive event and required for further progression in assembly.

The next important events in the splicing pathway are the association of U4/U5/U6 tri-snRNP particles with the early spliceosome and dynamic rearrangement of snRNPs within the spliceosome. The major remodeling of the spliceosome is considered as a switch from inactive spliceosome to a catalytic machine (Murray and Jarrell 1999). U4/U5/U6 tri-snRNP complexes which associate through U4 and U6 snRNA base pairing and protein-protein interaction of U4/U6 snRNP complex with U5 snRNP are added to the spliceosome. The intermediate complex right before the activation is called B complex. The first, U4:U6 base pairing is disrupted and releases U4 during the activation of spliceosome. Next, U1 base pairing with 5′ splice site is replaced by the base pairing between the 5′ end of U6 snRNA and the 5′ splice site. U1 and U4 snRNPs are no longer needed and are released from the spliceosome prior to catalysis (Kuhn et al., 1999; Staley and Guthrie 1999). These rearrangements and disassembly require ATP and numerous DEXD/H box proteins. A delicate balance of splicing enhancers and inhibitors can

control the fate of the spliceosome. Prp8, a yeast homolog of human U5-220kD, is regarded as an inhibitor by which U4/U6 interaction is stabilized, as demonstrated by the fact that a Prp8 mutation exacerbates the effect of a U4 mutation destabilizing U4/U6 association. In contrast, prp28, a yeast homolog of human U5-100kD in conjunction with Brr2 (U5-200kD), destabilizes the U4/U6 complex and thus considered as a stimulator of splicing activation (Kuhn et al., 1999; Staley and Guthrie 1999). Prp38 is another protein required for the progression to the active spliceosome by destabilizing U4/U6 association (Xie et al., 1998). The active splicing machinery formed through the exchange of factors and rearrangements of snRNPs are conventionally known as C complexes. Although both genetics studies in yeast *Saccharomyces cerevisiae* and parallel biochemical studies have provided important information with regard to spliceosome assembly, there are still many other splicing factors that have not been characterized for a specific function. However, powerful genetic and biochemical tools available today will help elucidate the function(s) of those protein components of the spliceosome in the future. Several attempts have been made to identify the protein composition of the spliceosome. Spliceosomes assembled on biotin-tagged RNA substrate were isolated by size exclusion chromatography, purified through strepavidin

affinity column, and subjected to SDS-PAGE for analysis (Bennett et al., 1992). Identity of proteins was determined by western blotting or peptide sequencing upon excision of a protein band from a gel. However, these processes were not only limited since these techniques heavily depended on the availability of antibodies but also laborious and time-consuming. More recently, a combined use of tandem mass spectrometry with automated database searches has permitted a rapid identification of a large number of peptides assembled on pre-mRNA (Jurica et al., 2002; Rappsilber et al., 2002; Zhou et al., 2002). Recently, pre-assembled penta-snRNP, which contains five snRNPs in stoichiometic ratios and more than 60 splicing factors, has been purified as ~45S complex (Stevens et al., 2002). The functionally active penta-snRNPs were demonstrated to be preformed unit rather assembled stepwise manner by mixing experiments which show no mixing of tagged snRNAs with endogenous snRNPs during spliceosome assembly. A similar preassembled complex was detected in mammalian systems and the complex exists in the absence of pre-mRNA (Konarska et al., 1988). This holospliceosome hypothesis suggests that the detection of complexes sequentially on native gels is caused by the ordered stepwise stabilization rather than stepwise recruitment of components (Brow, 2002; Nilsen, 2002).

snRNP particles: structure, biogenesis, import into the nucleus

The small nuclear ribonucleoproteins U1, U2, U4, U5, and U6 are essential components of the spliceosome. Each snRNP is made of U RNA and a set of core snRNP proteins (Sm proteins) and non-core snRNP-specific proteins. There are eight Sm proteins; B' (29 kDa), B (28 kDa), D1 (16 kDa), D2 (16.5 kDa), D3 (18 kDa), E (12 kDa), F (11 kDa), and G (9 kDa(Luhrmann et al., 1990). The biogenesis of U6 snRNP is quite different from that of the other snRNPs (Friesen and Dreyfuss 2000). U1, U2, U4, and U5 snRNA are transcribed by RNA polymerase II and acquire a monomethyl cap. Subsequently, the RNAs are exported to the cytoplasm and bind to the Sm core proteins to form the Sm core before they are hypermethylated to a trimethyl cap for nuclear import (Zeller et al., 1983; Mattaj 1986). In contrast, U6 snRNA is transcribed by RNA polymerase III (Dahlberg and Lund 1991) and stays in the nucleus (Terns et al., 1993). U6 snRNP has monomethyl cap and is made of U6 snRNA and eight similar but distinct Sm-like(Lsm) proteins (Lsm1 to Lsm8) (Achsel et al., 1999). Besides the core polypeptides, each snRNP also contains a specific set of proteins that have functions in the splicing reaction by coordinating

interactions of snRNP with the spliceosome (Table 1)(Jurica et al., 2002).

SnRNP assembly, which occurs in the cytoplasm, requires the assembly of the Sm proteins on the Sm binding site, which is a uridine-rich sequence present in the U snRNAs, to form the Sm core (Mattaj and De Robertis 1985). A properly assembled Sm core and the hypermethylated 5' cap are both required to recruit the import receptors necessary for snRNP translocation into the nucleus. Once in the nucleus, snRNPs associate with specific proteins that are unique to each snRNA, and function in pre-mRNA splicing.

Galectins as pre-mRNA splicing factors

Previously, galectin-1 and -3 were identified as required splicing factors in HeLa cell nuclear extracts (Dagher et al., 1995; Vyakarnam et al., 1997). An initial hint that galectins were splicing factors was demonstrated by the observation that certain saccharides which bind to galectins with high affinity can inhibit splicing *in vitro*. Lactose, thiodigalactoside (TDG), and melibiose inhibited splicing while others with no affinity for galectins did not inhibit splicing. Moreover, nuclear extracts depleted of galectins by lactose-sepharose affinity chromatography

Table 1.

A summary of snRNP specific proteins

The table was adapted and modified from reference (Jurica, 2002)

snRNP	Human	S. cerevisiae	Complex
U1	U1-70 KDa	Snp1p	E/A/B
	U1-A	Mud1p	E/A/B
	U1-C		E/A/B
U2	U2-A'	Lea1p	E/A/B/C
	U2-B"	Msl1p	E/A/B/C
	SF3a60	Prp9p	E/A/B/C
	SF3a66	Prp11p	E/A/B/C
	SF3a120	Prp21p	E/A/B/C
	SF3b53	Hsh49p	A/B/C
	SF3b150	Cus1p	E/A/B/C
	SF3b130	Rse1p	E/A/B/C
	SF3 160		E/A/B/C
U5	U5-220	Prp8p	C
	U5-200	Brr2p	H/A/B/C
	U5-116	Snu114p	H/A/B/C
	U5-102	Prp6p	A/B/C
	U5-100	Prp28p	A/B/C
	U5-40		A/B/C
	U5-15	Dib1p	-
Tri-snRNP	HPRP3	Prp3p	A/B
	HPRP4	Prp4p	A/B
	SART1	Snu66p	C
	PRP31	Prp31p	-
	SNU13	Snu13p	-
		Snu23p	-
		Prp38p	-
		Spp381p	-

failed to support splicing, but splicing activity was restored upon addition of mouse recombinant galectin-3 purified from an *E. coli* expression system. In parallel studies, the formation of spliceosome complexes was analyzed by native gel electrophoresis which allows identification of native splicing complexes (Konarska and Sharp 1986). This electrophoretic separation of complexes involved in pre-mRNA splicing produces at least 4 bands: H-complex and active spliceosomes (A,B,C). Depletion of galectins by lactose-sepharose affinity chromatography from *HeLa* nuclear extracts arrested spliceosome formation at the H-complex, suggesting that galectin-3 is required for the formation of active spliceosomes. The addition of recombinant galectin-3 to the depleted extract resulted in the formation of active spliceosomes, which is consistent with the reconstitution of splicing activity.

Unexpectedly, the removal of galectin-3 by antibody adsorption from splicing extracts was found not sufficient to disrupt splicing activity (Vyakarnam et al., 1997). The discrepancy in splicing activity between anti-galectin-3 antibody depletion versus saccharide affinity depletion suggested that another galectin in *HeLa* nuclear extracts could substitute galectin-3's role in splicing. Since galectin-1 was known to be present in HeLa cells, galectin-1 was a potential candidate to replace galectin-3's role in

splicing. Thus, passage of HeLa nuclear extracts over lactose-sepharose affinity column or adding galectin-specific saccharide ligands to nuclear extracts abolished the splicing activity by removing, either physically or functionally, both galectin-1 and -3 from the extract. Monoclonal anti-galectin-3 antibody, however, exclusively removed galectin-3 while galectin-1 remains physically and functionally intact in the splicing extract. This idea was tested by double antibody depletion experiments in which both galectin-1 and -3 were removed by anti-galectin-1 and anti-galectin-3 from splicing extracts. Splicing activity was completely inhibited as a result. Either recombinant galectin-1 or recombinant galectin-3 could restore the splicing activity of the galectin-depleted extracts. Moreover, the restoration of splicing activity by either galectin in double-antibody-depleted extracts was again sensitive to TDG inhibition. In a comparison of the efficiency of reconstitution of splicing, recombinant full-length galectin-3 was more efficient than recombinant galectin-1 or the carboxy domain of galectin-3. These data strongly suggest that galectin-1 and galectin-3 share redundant functions in pre-mRNA splicing (Vyakarnam et al., 1997) and that the amino domain facilitates splicing.

In addition, galectin-1 and galectin-3 colocalize with Sm core proteins and SC35, one of the non-snRNP splicing

factors, in the subnuclear region called speckles as demonstrated by confocal immunofluorescence microscopy. Nuclear speckles are considered as storage sites for splicing factors and some transcription factors (Vyakarnam et al., 1998).

How galectins participate in the splicing pathway has not been established. Identifying a splicing factor that interacts with nuclear galectins could provide a major stepping stone toward understanding the mechanism and role of the galectins during pre-mRNA splicing. The identification of *bona fide* splicing factors as ligand(s) for galectins could help delineate the role(s) of galectins in the splicing pathway *in vivo*. The yeast two hybrid system is a sensitive *in vivo* approach used to detect protein-protein interactions even in cases where the interaction is weak and/or transient (Fields and Song 1989). This powerful genetic technique was used to identify spliceosomal components that could serve as ligand(s) for galectins.

In an initial search of ligands for galectin-3, it was discovered that galectin-3 fused to the DNA binding domain (DNA-BD) alone was sufficient for the activation of transcription from a Gal4 promoter (Park 1996)(Unpublished data).

In contrast, using galectin-1 as bait, gemin4 was identified as an interacting protein using the yeast two hybrid screen (Park et al., 2001). Gemin4 is one component of the macromolecular complex called the SMN complex. The SMN complex has been implicated with a role in pre-mRNA splicing possibly in recycling snRNPs or/and supplying snRNPs to early spliceosomes (Pellizzoni et al., 1998a). A more comprehensive review on the SMN complex will be given in the next section. The identification of gemin4 as a putative ligand for galectin-1 is an important advancement in our understanding of the role(s) of galectins in pre-mRNA splicing since interactions between gemin4 and galectins in supplying snRNPs for the assembly of the early spliceosome could serve as a good molecular model to test.

SMN biology

The survival of motor neuron (SMN) protein has been identified as the disease gene in spinal muscular atrophy (SMA), an autosomal recessive neuromuscular disorder. The SMN gene encodes a protein product of ~ 36 kDa. SMA is characterized by degeneration of motor neurons of the spinal cord and progressive muscular atrophy (Czeizeil and Hamula 1989). The human SMN locus is part of a large inverted duplication on chromosome 5, with one copy near the

centromere and the other near the end of chromosome (Coovert et al., 1997). In more than 98% of SMA cases, the disease results from deletion or mutation of the telomeric copy of the SMN gene while the centromeric SMN gene is normal (Bussaglia et al., 1995; Rodrigues et al., 1995). Individuals lacking just the centromeric SMN gene are normal. Interestingly, sequencing of both centromeric and telomeric genes revealed only 5 nucleotide substitutions (Lorson et al., 1999). The same report shows that a single nucleotide substitution resulted in skipping of exon 7. Furthermore, lacking both telomeric and centromeric copies of SMN genes causes embryonic lethality (Talbot et al., 1998; Jablonka et al., 2000). In a mouse model system, transgenic mice were generated from mouse lines deficient for mouse Smn with various numbers of human SMN gene copies (Hsieh-Li et al., 2000). The severity of the pathological changes in these mice inversely correlated with the amount of SMN protein that contained the region encoded by exon 7.

Composition of the SMN complex

The SMN protein forms a macromolecular complex with a large number of proteins. At least, seven other polypeptides have been identified. Monoclonal antibody against SMN protein coimmunoprecipitated several polypeptides including Gemin2 (~32 kDa)(also known as SIP-1)

(Fischer et al., 1997), Gemin3 (~100 kDa) (Charroux et al., 1999), Gemin4 (~97 kDa) (Charroux et al., 2000) , Gemin5 (~170 kDa) (Gubitz et al., 2002), Gemin6 (~16 kDa) (Pellizzoni et al., 2002) , Gemin7 (~15 kDa) (Baccon et al., 2002; Pellizzoni et al., 2002), and several Sm core proteins (Pellizzoni et al., 2001).

Gemin2 was initially identified as an interacting protein for SMN protein. It was shown to have a critical role in the assembly of snRNPs in the cytoplasm by antibody microinjection experiments in *Xenopus* oocytes (Fischer et al., 1997). Gemin3, Gemin4, and Gemin5 were identified by amino acid sequencing of polypeptides that were coimmunoprecipitated by antibody against SMN protein (Charroux et al., 1999; Charroux et al., 2000). Gemin3 and Gemin4 show no significant homology to any protein found in the databases although Gemin3 contains DEAD motif, usually found in RNA helicases, implicating role(s) in rearrangement of RNA-RNA interaction and Gemin5 contains thirteen novel WD domains with no known function. Finally, Gemin6 and Gemin7 were identified by similar approaches as for Gemin3, Gemin4, and Gemin5 except that they were identified by coimmunoprecipitation by anti-FLAG antibody from stable cell lines expressing FLAG-SMN proteins. Again, Gemin6 and Gemin7 show no significant homology to any protein found in the database.

Functions of SMN protein

SMN is composed of several functional regions. SMN interacts with Gemin2, also called SMN Interacting Protein 1 (SIP1) through the N-terminal 40 amino acids (Liu et al., 1997). SMN oligomerizes through the C-terminus (amino acid 242-278)(Lorson et al., 1998; Pellizzoni et al., 1999). The domain (amino acid 28 - 91) encoded by exon 2 is necessary and sufficient for RNA binding (Lorson and Androphy 1998). The tudor domain, the central region of SMN (amino acid 90 - 160), has been shown to facilitate binding of Sm proteins (Buhler et al., 1999). The tudor domain exhibits a conserved negatively charged surface that is shown to interact with the C-terminal Arg and Gly-rich tails of Sm proteins (Friesen and Dreyfuss 2000; Friesen et al., 2001). Other proteins containing RG-rich domains have also been documented to binding to the SMN protein through the tudor domain. Fibrillarin and GAR1, components of Box C/D and Box H/ACA snoRNPs, interact with the SMN protein through their RG-rich domains in the amino terminus and in both amino and carboxy termini, respectively (Jones et al., 2001; Pellizzoni et al., 2001). RNA helicase A was identified as a ligand of the SMN protein by a two hybrid screen and RNA polymerase II was shown to interact with the SMN complex (Pellizzoni et al., 2001).

Elsewhere, the SMN protein was found to interact with ZPR1, a cytoplasmic zinc-finger protein that redistributes to the nucleus in mitogen-treated cells, possibly through their C-terminal domains. Deletion of either the region encoded by exon 7 for the SMN protein or the region at the carboxy terminal half of ZPR1 disrupted the interaction of those two proteins (Gangwani et al., 2001). The nuclear transcriptional activator E2 of papillomavirus was also identified as an interacting ligand for the SMN protein by a yeast two hybrid screen (Strasswimmer et al., 1999). Although interactions of the SMN protein with multiple protein partners are likely to be mutually exclusive considering the fact that interactions occurs through overlapping regions, how each interaction occurs and their relevant biological functions are unclear at this time. However, the overall theme observed in these interactions reflects the role of the SMN protein as a master assembler of RNA-Protein (RNP) complexes including transcriptosome, spliceosome, and ribosome.

The remaining thesis consists of 2 manuscripts, one published in Nucleic Acid Research (volume **27**: 3595-3602) under the title of 'Association of galectin-1 and galectin-3 with gemin4 in complexes containing the SMN protein.', and one in preparation, followed by conclusion and possible avenues for future studies.

Reference

Abovich, N. and M. Rosbash. 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* **89**: 403-412.

Achsel, T., H. Brahms, B. Kastner, A. Bachi, M. Wilm, and R. Luhrmann. 1999. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J* **18**: 5789-5802.

Akimoto, Y., H. Kawakami, Y. Oda, A. Obinata, H. Endo, K. Kasai, and H. Hirano. 1992. Changes in expression of the endogenous beta-galactoside-binding 14-kDa lectin of chick embryonic skin during epidermal differentiation. *Exp Cell Res* **199**: 297-304.

Albrandt, K., N.K. Orida, F.T. Liu. An IgE-binding protein with a distinctive repetitive sequence and homology with an IgG receptor. 1987. *Proc Natl Acad Sci USA* **84**: 6859-6863.

Arata, Y., J. Hirabayashi, and K. Kasai. 1997. The two lectin domains of the tandem-repeat 32-kDa galectin of the nematode *Caenorhabditis elegans* have different binding properties. Studies with recombinant protein. *J Biochem* **121**: 1002-1009.

Baccon, J., L. Pellizzoni, J. Rappsilber, M. Mann, and G. Dreyfuss. 2002. Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J Biol Chem* **277**: 31957-62.

Briles, E.B., W. Gregory, P. Fletcher, S. Kornfeld. 1979. Vertebrate lectins, comparison of properties of beta-galactoside-binding lectins from tissues of calf and chicken. *J Cell Biol* **81**: 528-537.

Barboni, E.A.M., S. Bawumia, K. Hendrick, and R.C. Hughes. 2000. Molecular modeling and mutagenesis studies of the N-terminal domains of galectin-3: evidence for participation with the C-terminal carbohydrate recognition domain in oligosaccharide binding. *Glycobiology* **10**: 1201-1208.

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, H. Leffler, F.-T. Liu, R. Lotan, A.M. Mercurio, M. Monsigny, S. Pillai, F. Poirier, A. Raz, P.W. Rigby, J.M. Rini, and J.L. Wang. 1994a. Galectins: A family of animal β-galactoside-binding lectins. *Cell* **76**: 597-598.

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

Bennett, M., S. Michaud, J. Kingston, and R. Reed. 1992. Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes Dev* **6**.

Black, D.L. 2000. Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell* **103**: 367-370.

Buhler, D., V. Raker, R. Luhrmann, and U. Fischer. 1999. Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. *Hum Mol Gen* **8**: 2351-2357.

Bussaglia, E., O. Clermont, E. Tizzano, S. Lefevbre, L. Burglen, C. Cruaud, and J.A. Urtizberea. 1995. A frame-shift deletion in the survival motor neuron gene in Spanish spinal muscular atrophy patients. *Nature gen* **11**: 335-337.

Charroux, B., L. Pellizzoni, R.A. Perkinson, A. Shevchenko, M. Mann, and G. Dreyfuss. 1999. Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J Cell Biol* **147**: 1181-94.

Charroux, B., L. Pellizzoni, R.A. Perkinson, J. Yong, A. Shevchenko, M. Mann, and G. Dreyfuss. 2000. Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J Cell Biol* **148**: 1177-86.

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

Cho and R.D. Cummings. 1995. Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. II.

Localization and biosynthesis. *The Journal of Biological Chemistry* **270**: 5207-5212.

Choi, J.Y., A.J. van Wijnen, F. Aslam, J.D. Leszyk, J.L. Stein, G.S. Stein, J.B. Lian, and 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.

Clerch, L.B., P. Whitney, M. Hass, K. Brew, T. Miller, R. Werner, and 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.

Colnot, C., D. Fowlis, M.-A. Ripoche, I. Bouchaert, and F. Poirier. 1998. Embryonic implantation in galectin-1/galectin-3 double mutant mice. *Dev Dyn* **211**: 306-313.

Colnot, C., M.-A. Ripoche, F. Scaerou, D. Fowlis, and F. Poirier. 1996. Galectins in mouse embryogenesis. *Biochem. Soc Trans* **24**: 141-146.

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

Cooper, D.N.W. and S.H. Barondes 1999. God must love galectins; He made so many of them. *Glycobiology* **9**: 979-984.

Cooper, D.N.W., S.M. Massa, and S.H. Barondes. 1991. Endogenous muscle lectin inhibits myoblast adhesion to laminin. *J Cell Biol* **115**: 1437-1448.

Coovert, D., T.T. Le, P.E. McAndrew, J. Strasswimmer, T.O. Crawford, J.R. Mendell, S.E. Coulson, E.J. Androphy, T.W. Prior, and A.H.M. Burghes. 1997. The survival motor neuron protein in spinal muscular atrophy. *Hum Mol Gen* **6**: 1205-1214.

Czeizeil, A. and J. Hamula. 1989. A hungarian study on Werdnig-Hoffmann disease. *J Med Gen* **21**: 761-763.

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

Dahlberg, J.E. and E. Lund. 1991. How does III x II make U6? *Science* **254**: 1462-1463.

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

Fields, S. and O. Song. 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340**: 245-247.

Fischer, U., Q. Liu, and G. Dreyfuss. 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**: 1023-9.

Friesen, W.J. and G. Dreyfuss. 2000. Specific sequence of the Sm and Sm-like (Lsm) proteins mediate their interaction with the spinal muscular atrophy disease gene products (SMN). *J Biol Chem* **275**: 26370-26375.

Friesen, W.J., S. Massenet, S. Paushkin, A. Wyce, and G. Dreyfuss. 2001. SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Molecular Cell* **7**: 1111-1117.

Gangwani, L., M. Mikrut, S. Theroux, M. Sharma, and R.J. Davis. 2001. Spinal muscular atrophy disrupts the interaction of ZPR1 with the SMN protein. *Nature cell biol* **3**: 376-383.

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

Gitt, M.A., M.F. Wiser, H. Leffler, J. Herrmann, Y.R. Xia, S.M. Massa, D.N.W. Cooper, A.J. Lusis, and S.H. Barondes. 1995. Sequence and mapping of galectin-5, a beta-galactoside-binding lectin, found in rat erythrocytes. *J Biol Chem* **270**: 5032-5038.

Gubitz, A.K., Z. Mourelatos, L. Abel, J. Rappsilber, M. Mann, and G. Dreyfuss. 2002. Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J Biol Chem* **277**: 5631-6.

Hebert, E. and M. Monsigny. 1994. Galectin-3 mRNA level depends on transformation phenotype in ras-transformed NIH 3T3 cells. *Biol. Cell* **81**: 73-76.

Herrmann, J., C.W. Turck, R.E. Atchison, M.E. Huflejt, L. Poulter, M.A. Gitt, A.L. Burlingame, S.H. Barondes, and H. Leffler. 1993. Primary structure of the soluble lactose binding lectin L-29 from rat and dog and interaction of its non-collagenous proline-, glycine-, tyrosine-rich sequence with bacterial and tissue collagenase. *J Biol Chem* **268**.

Hirabayashi, J. 1997. Reviews in galectins *Trends Glycosci Glycotechnol* **9**: 1-185.

Hsieh-Li, H.M., J.G. Chang, Y.J. Jong, M.H. Wu, N.M. Wang, C. Tsai, and H. Li. 2000. A mouse model for spinal muscular atrophy. *Nature gen* **24**: 64-70.

Hsu, D.K., R.I. Zuberi, and 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.

Huflejt, M.E., C.W. Turck, R. Lindstedt, S.H. Barondes, and H. Leffler. 1993. L-29, a soluble lactose-binding lectin, is phosphorylated on serine 6 and serine 12 in vivo and by casein kinase I. *J Biol Chem* **268**: 26712-26718.

Jablonka, S., W. Rossoll, B. Schrank, and M. Sendtner. 2000. The role of SMN in spinal muscular atrophy. *J Neurol* **247**: 37-42.

Jones, K.W., K. Gorzynski, C.M. Hales, U. Fischer, F. Badbanchi, R.M. Terns, and M.P. Terns. 2001. Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J Biol Chem* **276**: 38645 - 38651

Jurica, M.S., L.J. Licklider, S.P. Gygi, N. Grigorieff, and M.J. Moore. 2002. Purification and characterization of native spliceosome suitable for three-dimensional structural analysis. *RNA* **8**: 426-439.

Kasai, K. and J. Hirabayashi. 1996. Galectins: A family of animal lectins that decipher glycocodes. *J Biochem* **119**: 1-8.

Koiptz, J., C.V. Reitzenstein, M. Burchert, M. Cantz, and H.-J. Gabius. 1998. Galectin-1 is a major receptor for ganglioside GM, a product of the growth-controlling activity of a cell surface ganglioside sialidase, on human neuroblastoma cells in culture. *J Biol Chem* **273**: 11205-11211.

Konarska, M.M. and P.A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNA. *Cell* **46**: 845-855.

Kuhn, A.N., Z. Li, and D.A. Brow. 1999. Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. *Mol Cell* **3**: 65-75.

Laing, J.G. and J.L. Wang. 1988. Identification of carbohydrate binding protein 35 in heterogeneous nuclear ribonucleoprotein complex. *Biochemistry* **27**: 5329-5334.

Lamm, G.M. and A. Lamond. 1993. Non-snRNP protein splicing factors. *Bio Bio Act* **1173**: 247-265.

Liu, Q., U. Fischer, F. Wang, and G. Dreyfuss. 1997. The spinal muscular atrophy disease gene product, SMN and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* **90**: 1013-1021.

Lorson, C. and E.J. Androphy. 1998. The domain encoded by exon 2 of the survival motor neuron protein mediates nucleic acid binding. *Hum Mol Gen* **7**: 1269-1275.

Lorson, C., E. Hahnen, E.J. Androphy, and B. Wirth. 1999. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci USA* **96**: 6307-6311.

Lorson, C., J. Strasswimmer, J.-M. Yao, J.D. Baleja, E. Hahnen, B. Wirth, T.T. Le, A.H.M. Burghes, and E.J. Androphy. 1998. SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nature gen* **19**: 63-66.

Luhrmann, R., B. Kastner, and M. Bach. 1990. Structure of snRNP and their role in pre-mRNA splicing. *Bio Bio Act* **1087**: 265-292.

Matsushita, N., N. Nish, M. Seki, R. Matsumoto, I. Kuwabara, F.-T. Liu, Y. Hata, T. Nakamura, and M. Hirashima. 2000. Requirement of divalent galactoside-binding activity of ecalectin/galectin-9 for eosinophil chemoattraction. *J Biol Chem* **275**: 8355-8360.

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

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

Mehul and R.C. Hughes. 1997. Plasma membrane targeting, vesicular budding and release of galectin 3 from the cytoplasm of mammalian cells during secretion. *J Cell Sci* **110**: 1169-1178.

Menon, R.P., M. Strom, and C. Hughes. 2000. Interaction of a novel cysteine and histidine-rich cytoplasmic protein with galectin-3 in a carbohydrate-independent manner. *FEBS Lett* **470**: 227-231.

Messa, S.M., D.N.W. Cooper, H. Leffler, and S.H. Barondes. 1993. L-29 an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity. *Biochemistry* **32**: 260-267.

Mey, A., H. Leffler, Z. Hmama, G. Normier, and J.P. Revillard. 1996. The animal lectin galectin-3 interacts with bacterial lipopolysaccharides via two independent sites. *J Immu* **156**: 1572-1577.

Michaud, S. and R. Reed. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev* **5**: 2534-2546.

Michel, F. and J.L. Ferat. 1995. Structure and activities of group II introns. *Annu Rev Biochem* **64**: 435-461.

Moore, M.J., C.C. Query, and P.A. Sharp. 1993. Splicing of precursors to mRNA by the spliceosome. *In The RNA World, R.F. Gesteland and J.F. Atkins, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press)*: 303-357.

Moutsatsos, I.K., J.M. Davis, and J.L. Wang. 1986. Endogenous lectins from cultured cells: subcellular localization of carbohydrate-binding protein 35 in 3T3 fibroblasts. *J Cell Biol* **102**: 477-483.

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

Murray, H.L. and K.A. Jarrell. 1999. Flipping the switch to an active spliceosome. *Cell* **96**: 599-602.

Nilsen, T.M. 1998. RNA-RNA interactions in nuclear pre-mRNA splicing. *In RNA structure and function, R. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, NY)*: 279-307.

Park, J.W. 1996. Identification of splicing factors that interact with galectin-3 using yeast two hybrid system. In *Thesis*.

Park, J.W., P.G. Voss, S. Grabski, J.L. Wang, and R.J. Patterson. 2001. Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. *Nucl Acids Res* **27**: 3595-3602.

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

Pellizzoni, L., J. Baccon, B. Charroux, and G. Dreyfuss. 2001a. The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Cur Biol* **11**: 1079-1088.

Pellizzoni, L., J. Baccon, J. Rappsilber, M. Mann, and G. Dreyfuss. 2002. Purification of native survival of motor neurons complexes and identification of Gemin6 as a novel component. *J Biol Chem* **277**: 7540-5.

Pellizzoni, L., B. Charroux, and G. Dreyfuss. 1999. SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. *Proc Natl Acad Sci USA* **96**: 11167-11172.

Pellizzoni, L., B. Charroux, J. Rappsilber, M. Mann, and G. Dreyfuss. 2001c. A functional interaction between the survival motor neuron complex and RNA polymerase II. *J Cell Biol* **152**: 75-85.

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

Perillo, N.L., K.E. Pace, J.J. Seihamer, and L.G. Baum. 1995. Apoptosis of T cells mediated by galectin-1. *Nature* **378**: 736-739.

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

Raz, A. and R. Lotan. 1987. Endogenous galactoside-binding lectins: a new class of functional tumor cell surface molecules related to metastasis. *Cancer Metast Rev* **6**: 433-452.

Reed, R. 2000. Mechanisms of fidelity in pre-mRNA splicing. *Curr Opin Cell Biol* **12**: 340-345.

Rodrigues, N.R., N. Owen, K. Talbot, J. Ignatius, V. Dubowitz, and K.E. Davies. 1995. Deletions in the survival motor neuron gene on 5q13 in autosomal recessive spinal muscular atrophy. *Hum Mol Gen* **4**: 631-634.

Runyan, R.B., J. Versalovic, and B.D. Shur. 1988. Functionally distinct laminin receptors mediate cell adhesion and spreading: the requirement for surface galactosyltransferase in cell spreading. *J Cell Biol* **107**: 1863-1871.

Sato, S. and R.C. Hughes. 1992. Binding specificity of a baby hamster kidney lectin for H type I and II chains, polylactosamine glycans, and appropriately glycosylated forms of laminin and fibronectin. *J Biol Chem* **267**: 6983-6990.

Seetharaman, J., A. Kanigsberg, R. Slaaby, H. Leffler, S.H. Barondes, and J.M. Rini. 1998. X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1 A resolution. *J Biol Chem* **273**: 13047-13052.

Sharp, P.A. 1994. Split genes and RNA splicing. *Cell* **77**: 805-815.

Staley, J.P. and C. Guthrie. 1998. Mechanical devices of the spliceosome motors, clocks, springs, and things. *Cell* **92**: 315-326.

Staley, J.P. and C. Guthrie. 1999. An RNA switch at 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* **3**: 55-64.

Strasswimmer, J., C. Lorson, D.E. Breiding, J.J. Chen, T.T. Le, A.H.M. Burghes, and E.J. Androphy. 1999. Identification of survival motor neuron as a transcriptional activator-binding protein. *Human Molecular Genetics* **8**: 1219-1226.

Talbot, K., I. Miguel-Aliaga, P. Mohaghegh, C.P. Ponting, and K.E. Davies. 1998. Characterization of a gene encoding survival motor neuron (SMN)-related protein, a constituent of the spliceosome complex. *Hum Mol Ge* **7**: 149-156.

Tenne-Brown, J., A.C. Puche, and B. Key. 1998. Expression of galectin-1 in the mouse olfactory system. *Int J Dev Biol* **42**: 791-799.

Terns, M.P., J.E. Dahlberg, and E. Lund. 1993. Multiple cis-acting signals for export of pre-U1 snRNA from the nucleus. *Genes Dev* **7**: 1898-1908.

Tsay, Y.-G., N.Y. Lin, P.G. Voss, R.J. Patterson, and J.L. Wang. 1999. Export of galectin-3 from nuclei of digitonin-permeabilized mouse 3T3 fibroblasts. *Exp Cell Res* **252**: 250-261.

Valadkhan, S. and J.L. Manley. 2001. Splicing-related catalysis by protein free snRNA. *Nature* **413**: 701-707.

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

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

Wassarman, D.A. and J.A. Steitz. 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science* **257**: 1918-1925.

Wu, J.Y. and T. Maniatis. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* **75**: 1061-1070.

Wu, S., C.M. Romfo, T.M. Nilsen, and M.R. Green. 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **402**: 832-835.

Xie, J., K. Beickman, E. Otte, and B.C. Rymond. 1998. Progression through the spliceosome cycle requires prp38p function for U4/U6 snRNA dissociation. *EMBO J* **17**: 2938-2946.

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

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

Zeller, R., T. Nyffenegger, and E.M. DeRobertis. 1983. Nucleocytoplasmic distribution of snRNPs and stockpiled snRNA-binding proteins during oogenesis and early development in *Xenopus laevis*. *Cell* **32**: 425-434.

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

Chapter 2

Association of Galectin-1 and Galectin-3 with Gemin4 in Complexes Containing the SMN Protein

ABSTRACT

In previous studies, we had shown that galectin-1 and galectin-3 are factors required for the splicing of pre-mRNA, assayed in a cell-free system. Using a yeast two-hybrid screen with galectin-1 as "bait," Gemin4 was identified as a putative interacting protein. Gemin4 is one component of a macromolecular complex containing approximately 15 polypeptides, including the SMN (survival of motor neuron) protein. Rabbit anti-galectin-1 co-immunoprecipitated from HeLa cell nuclear extracts, along with galectin-1, polypeptides identified to be in this complex: SMN, Gemin2, and the Sm polypeptides of snRNPs. Direct interaction between Gemin4 and galectin-1 was demonstrated in glutathione S-transferase (GST) pull-down assays. We also found that galectin-3 interacted with Gemin4 and was found as a part of the complex co-immunoprecipitated with galectin-1. Indeed, fragments of either Gemin4 or galectin-3 exhibited a dominant negative effect when added to the cell-free splicing assay. For example, a dose-dependent inhibition of splicing was observed in the presence of exogenously added NH_2-terminal

domain of the galectin-3 polypeptide. In contrast, parallel addition of either the intact galectin-3 polypeptide or the COOH-terminal domain failed to yield the same effect. Using native gel electrophoresis to detect complexes formed by the splicing extract, we found that, with the addition of the NH_2-terminal domain, the predominant portion of the radiolabeled pre-mRNA was arrested at a position corresponding to the H-complex. Inasmuch as SMN-containing complexes have been implicated in the delivery of snRNPs to the H-complex, these results provide strong evidence that galectin-1 and galectin-3, by interacting with Gemin4, play a role in spliceosome assembly *in vivo*.

INTRODUCTION

Pre-mRNA splicing is an essential post-transcriptional process coordinated by a macromolecular complex termed the spliceosome. This assembly is composed of five small nuclear ribonucleoprotein particles (U1, U2, U4, U5, and U6 snRNPs) and non-snRNP proteins, estimated to number between 50 and 100 (Sharp 1994; Fu 1995; Kramer 1996). Many of these splicing factors are involved in one or more of the following steps: snRNP biogenesis, spliceosome assembly and disassembly, and the catalytic events responsible for intron excision and exon ligation (Arenas and Abelson 1997; Raker et al., 1999; Reed 2000).

Galectins-1 and -3 belong to a family of proteins defined by their binding to galactose-containing glycoconjugates (for a recent review, see (Leffler 1997)). The nuclear localization of galectin-3, coupled with its co-sedimentation with nuclear RNPs and snRNPs, provided the initial hint that it might play a role in pre-mRNA splicing (Moutsatsos et al., 1986; Laing and Wang 1988). Using a cell-free splicing assay, subsequent experiments documented: (a) saccharides that bind galectins with high affinity inhibited the splicing reaction; (b) depletion of both galectins from the nuclear extract, either by lactose affinity chromatography or by antibody adsorption, resulted in the concomitant loss of splicing activity; and (c) either recombinant galectin-1 or recombinant galectin-3 was able to reconstitute splicing activity in a galectin-depleted extract (Dagher et al., 1995; Vyakarnam et al., 1997). Finally, double immunofluorescence experiments showed that both galectin-1 and -3 can be colocalized in nuclear speckles with known splicing factors (the core polypeptides of snRNPs bearing the Sm epitope and the non-snRNP splicing factor SC35) (Vyakarnam et al., 1997; Vyakarnam et al., 1998).

In order to define the role played by galectins in the splicing process, it was important to identify and characterize their interacting partner(s) in the nucleus.

To this end, we carried out a yeast two-hybrid screen using galectin-1 as "bait." In the present communication, we identify Gemin4, one component of nuclear complexes containing the survival of motor neuron (SMN) protein, as a partner that directly binds to galectin-1 and galectin-3. We also show that fragments of either of the Gemin4-galectin-3 partners, such as the carboxyl-terminal 50 amino acids of Gemin4 (hereafter designated as Gemin4(C50)) or the ~140-residue NH_2-terminal domain of galectin-3 (designated as ND), exert a dominant negative effect on pre-mRNA splicing. The data suggest that the ND blocks an early step in spliceosome formation, consistent with the notion that SMN-containing complexes are responsible for the delivery of snRNPs to the spliceosome.

MATERIALS AND METHODS

Yeast two-hybrid screen

The cDNA of human galectin-1 (Hirabayashi and Kasai 1991) was kindly provided by Drs. K. Kasai and J. Hirabayashi (Teikyo University, Sagamiko, Kanagawa, Japan). It was subcloned into the BamHI site of the vector pAS2-1 (Clontech) downstream of the yeast GAL4 DNA-binding domain. A HeLa cell cDNA library fused to pGAD-GH vector was purchased from Clontech and served as the "prey." The screen for putative interacting partners was performed following the manufacturer's instructions. The $His^+/LacZ^+$

clones were isolated and their DNAs were amplified by the polymerase chain reaction. Sequence analysis was carried out by the Macromolecular Structure, Sequencing and Synthesis Facility at Michigan State University.

Protein preparations and *in vitro* binding assays

Procedures for the preparation of recombinant murine galectin-3 and its ND and CD (carboxyl-terminal domain, 147 amino acids) have been described (Agrwal et al., 1993). The purity of these protein preparations was characterized by SDS-PAGE and immunoblotting using domain-specific antibodies. In addition, the molecular weights of recombinant galectin-3, ND, and CD were determined by mass spectrometry, using matrix-assisted laser desorption/ionization (PE PerSeptive Biosystems Voyager-DE STR). The same samples were also subjected to amino acid sequence analysis, using a Precise pulsed-liquid protein sequence (PE Applied Biosystems). Both of these procedures, carried out by the Macromolecular Structure, Sequencing and Synthesis Facility at Michigan State University, confirmed the identity and purity of the protein preparations. Bovine galectin-1 (Childs and Feizi 1979) used for the *in vitro* binding assay was kindly provided by Dr. Ten Feizi (Imperial College School of Medicine, Northwick Park, Harrow, Middlesex, UK).

The cDNA of Gemin4(C50), identified through the yeast two-hybrid screen, was subcloned into NcoI and EcoRI restriction sites of pGEX-2T (Pharmacia) to produce the fusion protein, GST-Gemin4(C50). Glutathione-Sepharose beads (Pharmacia) were pretreated with bacterial cell lysate isolated from *E. coli* used for recombinant fusion protein production. Approximately 1 µg GST or GST-Gemin4(C50) was adsorbed to 20 µl of the pretreated beads. Then, 0.5 µg of bovine galectin-1 or recombinant murine galectin-3 was incubated with the GST or GST-Gemin4(C50) beads at 4 °C for 2 hrs in 60% buffer D (buffer D is 20 mM Hepes, pH 7.9, 20% glycerol, 100 mM KCl and 0.2 mM EDTA) containing 0.05% Triton X-100. After extensive washing, the bound material was eluted with SDS-PAGE sample buffer and resolved by electrophoresis.

Immunoprecipitation of complexes from nuclear extracts of HeLa cells

HeLa S3 cells were grown in suspension culture by the National Cell Culture Center (Minneapolis, MN). Nuclear extract (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 phenylmethylsulfonyl fluoride, 0.5 mM dithiothreitol) or buffer D, as described by Dignam *et al* (Dignam et al., 1993). NEs were frozen as aliquots in a dry ice-ethanol

bath and stored at -80 °C. Protein concentrations were determined by the Bradford assay (Bradford 1976). In this study, the protein concentration of NE was approximately 4 mg/ml.

Polyclonal rabbit antiserum against recombinant rat galectin-1 (Cooper et al., 1991) was a gift from Dr. Sam Barondes (University of California, San Francisco, CA). We had previously described the adsorption and covalent coupling of this antibody to protein G-Sepharose beads (Vyakarnam et al., 1997). NE (~80 µg protein) was incubated with beads containing anti-galectin-1 or the corresponding preimmune serum. The incubation was carried out at 4 °C for 2 hrs in 60% buffer D containing 0.05% Triton X-100. After extensive washing with the same buffer, the bound material was eluted with SDS-PAGE sample buffer and resolved by electrophoresis.

SDS gel electrophoresis and immunoblotting

Proteins were resolved on 12% or 15% acrylamide gels electrophoresed in the presence of SDS (SDS-PAGE), as described by Laemmli (Laemmli 1970). The polypeptides were revealed by staining with Coomassie blue, using the Gel Code Blue reagent of Pierce. Polypeptides were also revealed by immunoblotting of Hybond-C membranes (Amersham), following

the general procedures previously described (Dagher et al., 1995).

A rat monoclonal antibody was developed against the Mac-2 antigen (Ho and Springer 1982), which has been shown to be galectin-3 (Cherayil et al., 1989)(αGal-3). The anti-galectin-1 antibodies used for immunoblotting were affinity purified from the serum of rabbits immunized with GST-human galectin-1 (αGal-1). Mouse monoclonal antibodies directed against SMN (αSMN) and Gemin2 (αGemin2) were purchased from Transduction Laboratories. The B, B', and D polypeptides of the snRNPs were detected with human autoimmune serum (ENA anti-Sm) purchased from The Binding Site. HMG 14/17 polypeptides were detected with a rabbit polyclonal antiserum (a gift from Michael Bustin, NIH, Bethesda, MD). The polypeptides were revealed using horseradish peroxidase-conjugated secondary antibodies and the enhanced chemiluminescence system (Amersham Pharmacia).

***In vitro* splicing assay**

The plasmid used to transcribe the MINX pre-mRNA substrate was a kind gift of Dr. Susan Berget (Baylor College of Medicine, Houston, TX). The MINX pre-mRNA was labeled with [^{32}P]-GTP and the monomethyl cap was added during SP6 polymerase (GIBCO BRL) transcription (Zillman et al., 1988).

Samples of NE were dialyzed against 60% Dignam buffer D in the absence or presence of exogenously added proteins: GST, GST-Gemin4(C50), recombinant murine galectin-3, ND, or CD. The dialysis was carried out for 75 min at 4 °C in a microdialyzer with a 6 to 8-kDa cutoff dialysis membrane (Vyakarnam et al., 1997).

Splicing reaction mixtures, in a total volume of 10 µl, contained dialyzed NE (6-8 µl), [^{32}P]MINX pre-mRNA, 2.5 mM $MgCl_2$, 1.5 mM ATP, 20 mM creatine phosphate, 0.5 mM dithiothreitol, and 20 U of RNasin (Promega). Splicing reactions were incubated at 30 °C for 45 - 60 min. The assay was stopped by the addition of proteinase K and SDS to final concentrations of 4 mg/ml and 0.1%, respectively. The samples were incubated at 37 °C for 20 min, 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 of chloroform. The RNAs were precipitated with 300 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. The radioactive RNA bands were revealed by autoradiography.

Gel mobility shift assay for splicing complex formation

Ribonucleoprotein complexes representing intermediates in the splicing pathway were monitored by electrophoresis in nondenaturing gels (Konarska and Sharp 1986; Zillman et al., 1988). Splicing reaction mixtures (10 µl) were incubated at 30 °C for 0 - 20 min. Heparin was then added to a final concentration of 1 mg/ml, along with 1 µl of glycerol containing 0.2% each of bromophenol blue, xylene cyanol, and phenol red. The samples were loaded onto a prerun 4% polyacrylamide gel (bisacrylamide:acrylamide, 1:80 [wt/wt]). Electrophoresis was carried out in 50 mM Tris, 50 mM glycine, pH 8.8, at 4 °C for ~90 min at constant voltage (150 V). The positions of migration of the splicing complexes were revealed by autoradiography of the radioactive pre-mRNA. Quantitation of the amount of radioactivity was carried out on a STORM phosphorimager (Molecular Dynamics, Sunnyvale, CA).

RESULTS

Identification of Gemin4 as an interacting partner of galectin-1 and galectin-3

To identify proteins that interact with the galectins, we performed a yeast two-hybrid screen (Fields and Song 1989), using human galectin-1 as the "bait" protein and a

HeLa cell cDNA library as "prey." We screened 2 x 10^6 independent clones and isolated 12 $His^+/LacZ^+$ clones. Based on sequence analysis, four clones were identified as ribosomal proteins (RPL37, RPS26, RPS3, and 60S acidic ribosomal protein PO), four clones were identified as cytochrome oxidase subunit III, and two clones were identified as ubiquitin-like proteins, human SMT3B/SUMO-3 (Small Ubiquitin-like Modifier). Two clones contained novel sequences at the time of initial isolation. One of the latter, however, was identified as the carboxyl terminal 50 amino acids of Gemin4 (accession number: AF173856) when the sequence was subjected to a BLAST search against the Genbank sequence database. There was a 100% identity between the two sequences (Fig. 1). The sequence identified through the yeast two-hybrid screen is designated as Gemin4(C50).

The cDNA of Gemin4(C50) was subcloned into pGEX-2T and the fusion protein GST-Gemin4(C50) was purified from the bacterial expression system. Direct binding between galectin-1 and Gemin4(C50) was tested in GST pull-down assays. Galectin-1 bound to GST-Gemin4(C50) (Fig. 2A, lane 3) but not to GST alone (Fig. 2A, lane 2). We also tested the binding of galectin-3 to Gemin4(C50) using the same *in vitro* binding assay. Galectin-3 bound directly to GST-Gemin4(C50) (Fig. 2B, lane 6) but not to GST (Fig. 2B, lane 5). In these experiments, galectin-1 and galectin-3 were

Query: 45 tcacctgctatgagactttgagcaagaccaacccttctgtcagctccttgctccagaggg 104

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Gemin4:3049 tcacctgctatgagactttgagcaagaccaacccttctgtcagctccttgctccagaggg 3108

Query: 105 cacacgagcagcgcttcttaaagtccattgctgagggcattggccctgaagaacggcgcc 164

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Gemin4: 3109 cacacgagcagcgcttcttaaagtccattgctgagggcattggccctgaagaacggcgcc 3168

Query: 165 aaaccctgttgcagaagatgagcagcttctgacttggcgtggggagctgggccccaacat 224

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Gemin4: 3169 aaaccctgttgcagaagatgagcagcttctgacttggcgtggggagctgggccccaacat 3228

Query: 225 ggcg 228

||||

Gemin4: 3229 ggcg 3232

Figure 1: Comparison of the nucleotide sequence of Gemin4 and the corresponding sequence of the clone identified as a ligand of galectin-1. DNA from the putative interacting clone was amplified by the polymerase chain reaction and then sequenced. When the sequence of this clone was subjected to a BLAST search against the Genbank genome database, there was 100% identity with the carboxyl-terminal 50 amino acids of Gemin4 (accession number: AF173856).

visualized by immunoblotting. The membranes used for immunoblotting were also stained with Coomassie blue to reveal GST (M_r ~27,000) and GST-Gemin4(C50) (M_r ~32,000). This was done as a control to ascertain that approximately the same amount of GST and GST-Gemin4(C50) was adsorbed onto the glutathione beads (Fig. 2C, lanes 7 and 8). Finally, we also tested and found no binding of bovine serum albumin to the GST and GST-Gemin4(C50) beads (data not shown). Thus, it can be inferred that galectin-1, as well as galectin-3, interacts specifically with Gemin4.

Co-immunoprecipitation of galectin-1 and galectin-3 in complexes containing SMN

Gemin4 is one of approximately 15 polypeptides of a macromolecular complex containing SMN, Gemin2, Gemin3, and the core polypeptides of snRNPs bearing the Sm epitope (Sm B/B', Sm D1-D3, and Sm E, F, G) (Charroux et al., 2000; Meister et al., 2000). Thus, our demonstration of a direct interaction between galectin-1 and Gemin4 prompted us to test whether galectin-1 is also a component of such nuclear complexes containing SMN. Polyclonal rabbit anti-galectin-1 was used to immunoprecipitate complexes from NE derived from HeLa cells. Proteins immunoprecipitated with anti-galectin-1 were separated by SDS-PAGE and subjected to immunoblotting with various antibodies 1 (Fig. 3). First, anti-galectin-1

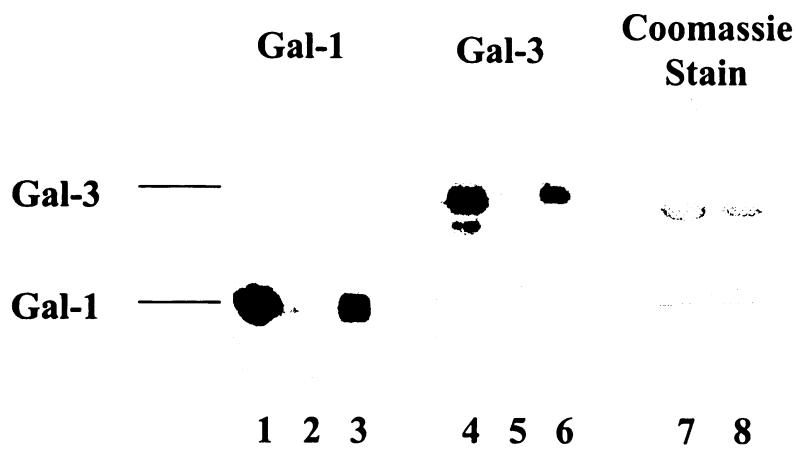
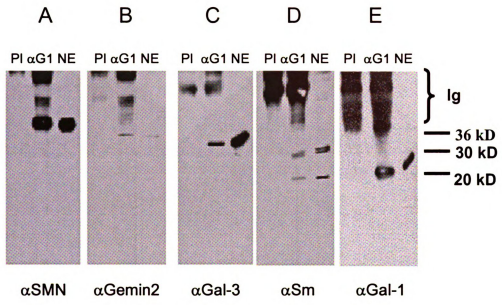

Gal-1
Gal-3
Coomassie Stain
Gal-3
Gal-1
1 2 3 4 5 6 7 8

Figure 2: *In vitro* binding assay between GST-Gemin4(C50) and galectin-1 and galectin-3. Approximately 1 µg of GST (lanes 2, 5, and 7) or GST-Gemin4(C50) (lanes 3, 6, and 8) was bound to glutathione-Sepharose beads that had been pretreated with bacterial lysate from *E. coli* cells used for the production of the recombinant fusion polypeptides. Bovine galectin-1 (0.5 µg, panel A) or recombinant murine galectin-3 (0.5 µg, panel B) was incubated with the GST or GST-Gemin4(C50) beads for 2 hrs. at 4 °C. After washing, the bound material was eluted and subjected to SDS-PAGE. Galectin-1 and galectin-3 were visualized by immunoblotting. In lanes 1 and 4, 10% of the input test protein is shown. After immunoblotting, one of the membranes was stained with Coomassie blue (panel C) to ascertain that equal quantities of GST and GST-Gemin4(C50) were bound to the glutathione beads.

immunoprecipitated galectin-1 from NE (Fig. 3, top panel). In addition, the following polypeptides, for which we had antibody reagents, were also found in the same anti-galectin-1 immunoprecipitate: (i) SMN (~38 kD); (ii) Gemin2. (~36 kD); (iii) galectin-3 (~30 kD); and (iv) Sm B, B', D polypeptides of the snRNPs. Control experiments using preimmune serum (PI) did not precipitate any of the proteins (Fig. 3, labeled PI). Moreover, we also tested and found that irrelevant nuclear
proteins, such as the High Mobility Group (HMG) 14/17, were not detectable in the anti-galectin-1 immunoprecipitate (Fig. 3, bottom panel). These results suggest that the polypeptides identified above represented *bona fide* components of macromolecular complexes co-precipitated by anti-galectin-1.

Effect of the addition of the NH_2-terminal domain of galectin-3 on the *in vitro* splicing reaction

It is interesting to note that galectin-3 is also found in the immunoprecipitate of anti-galectin-1 (Fig. 3). This is consistent with the observation that galectin-3 can interact directly with Gemin4 (Fig. 2B), which may, in turn, incorporate it into the macromolecular complex. The polypeptide of galectin-3 contains two domains: (a) a proline- and glycine-rich ND; and (b) the carbohydrate-binding COOH-terminal domain (CD). We have isolated


A
B
C
D
E
PI αG1 NE
PI αG1 NE
PI αG1 NE
PI αG1 NE
PI αG1 NE
Ig
36 kD
30 kD
20 kD
αSMN
αGemin2
αGal-3
αSm
αGal-1

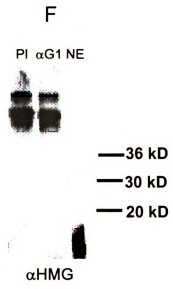

F
PI αG1 NE
36 kD
30 kD
20 kD
αHMG

Figure 3: Analysis of the polypeptides immunoprecipitated from NE by anti-galectin-1. NE (~80 μg protein) was incubated for 2 hrs at 4 °C with protein G-Sepharose beads covalently coupled with polyclonal rabbit anti-rat galectin-1. After washing, the bound material was eluted and subjected to SDS-PAGE. Polypeptides (indicated to the right of each panel) in the immunoprecipitates were identified with the following antibodies: affinity purified rabbit polyclonal antibodies against human galectin-1; mouse monoclonal antibody against SMN; mouse monoclonal antibody against Gemin2; rat monoclonal antibody against galectin-3; human autoimmune serum reactive against the Sm epitopes on the core polypeptides of snRNPs; and rabbit polyclonal antibodies to HMGs 14/17. In each panel, PI represents material precipitated by pre-immune serum; αG1 represents material precipitated by rabbit anti-rat galectin-1; and NE represents 50% of the nuclear extract subjected to immunoprecipitation, except the α galectin-1 immunoblot in which the NE is ~80% of that used for immunoprecipitation.

preparations of ND and CD and have carefully characterized them in terms of purity by SDS-PAGE, immunoblotting, mass spectrometry, and amino acid sequence analysis. The availability of these reagents, representing fragments of the intact galectin-3 polypeptide, suggested the opportunity to test their effects on the cell-free splicing assay. NE derived from HeLa cells exhibited strong splicing activity in our *in vitro* assay, converting more than 50% of the pre-mRNA substrate into the mature mRNA product (Fig. 4, lane 1). When the splicing reaction was carried out in the presence of exogenously added ND, there was a dose-dependent inhibition of splicing (Fig. 4, lanes 5-8). At a concentration as low as 1.2 μM, inhibition of splicing was noticeable (Fig. 4, lane 6). Complete inhibition was observed at 10 μM ND: there were barely detectable levels of products and splicing intermediates and substantially higher levels of the starting substrate (Fig. 4, lane 8).

The splicing reaction was also carried out in the presence of exogenously added full length recombinant galectin-3. Over a concentration range of 3-12 μM, the intact galectin-3 polypeptide had no effect on the *in vitro* splicing reaction (Fig. 4, lanes 2-4). Similarly, no inhibition was observed when the splicing reaction was

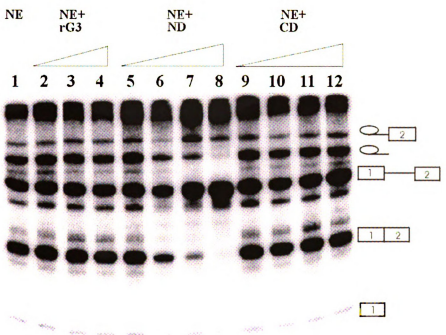

NE
NE+ rG3
NE+ ND
NE+ CD
1 2 3 4 5 6 7 8 9 10 11 12
OP
2
1
2
1 2
1

Figure 4: Comparison of the effect of the addition of recombinant galectin-3 or its NH_2- and COOH-terminal domains on the splicing activity of nuclear extract.
Lanes: 1, nuclear extract (NE); 2-4, recombinant galectin-3 (rG3) at 3, 6, and 12 μM added to NE; 5-8, ND at 0.6, 1.2, 5, and 10 μM added to NE; 9-12, CD at 3, 10, 28, and 57 μM added to NE. All reactions contained ^{32}P-labeled MINX pre-mRNA substrate (5,000 cpm) and products of the splicing reactions were analyzed by electrophoresis through a 13% polyacrylamide-urea gel, followed by autoradiography. The positions of migration of pre-mRNA substrate, splicing intermediates (exon 1 and lariat-exon 2) and RNA products (lariat and ligated exon 1-exon 2) are highlighted on the right.

carried out in the presence of exogenously added CD (Fig. 4, lanes 9-12). The concentration of CD tested ranged from 3-57 μM. Thus, the inhibitory effect of ND on the cell-free splicing assay was specific, inasmuch as parallel addition of either full-length galectin-3 or the CD failed to yield the same effect. Several possibilities were considered to account for the observed inhibition of splicing by ND. For example, the ND preparation could be contaminated by a protease that degraded essential components of the splicing machinery. On the basis of our previous experience showing that depletion of NE with antibodies directed against the Sm antigens of snRNPs resulted in the loss of splicing activity in this assay, we specifically tested for any effects of ND on the Sm antigens. Incubation of NE with ND, CD, or full-length recombinant galectin-3 under the conditions of the splicing assay did not reduce the levels of the B, B', and D polypeptides of the Sm antigens (data not shown).

We also tested for possible ribonuclease activity in our protein preparations. The ^{32}P-labeled pre-mRNA substrate was incubated with NE in the presence or absence of exogenous ND, CD, and recombinant galectin-3. This incubation was carried out in the absence of ATP. (In the presence of ATP, such incubation would simply correspond to a splicing assay and the pre-mRNA will be converted to the mature RNA product.) Surprisingly, the inclusion of any of

our protein preparations (ND, CD, or the full-length recombinant galectin-3) resulted in a higher recovery of ^{32}P-labeled pre-mRNA than incubation with NE alone. The basis for this better recovery is not known, but these results are consistent with the results of the splicing assay, in which substantially higher levels of the starting pre-mRNA substrate were found in reactions inhibited by ND (Fig. 4, lanes 7-8). In any case, the ND preparation did not degrade the pre-mRNA substrate.

Effect of the addition of galectin-3 ND on spliceosome formation

The MINX pre-mRNA substrate forms several complexes, distinguished by the composition of the RNP components, during the course of the splicing reaction (Zillman et al., 1988). Initially, the pre-mRNA is complexed with hnRNP proteins to form the H-complex. In the presence of ATP, specific snRNPs become associated with the H-complex to form the A and B active spliceosomal complexes. We have used the gel mobility shift assay to detect complexes formed by the splicing extract in the presence of exogenously added ND, CD, or full-length recombinant galectin-3.

Under our electrophoretic conditions (10.7 cm gel, 150 V, 90 min), ^{32}P-labeled pre-mRNA, in the absence of any proteins (no NE added), migrated off the bottom of the gel. When NE is added in the absence of ATP, only a radiolabeled

smear corresponding to the H-complex (Fig. 5, lane 1) can be found, even after 5 min of incubation. In the presence of ATP, the H-complex is also observed at 0 min (Fig. 5, lane 2); after 5 min of incubation, bands corresponding to the higher-order active complexes (A and B) become prominent (Fig. 5, lane 3). The summation of complexes A and B accounted for ~70% of the radioactivity in lane 3 (Fig. 5). These higher-order complexes are also observed when full-length recombinant galectin-3 (Fig. 5, lane 4) or CD (Fig. 5, lane 6) is added to the NE, pre-mRNA substrate, and ATP for 5 min. In contrast, the addition of ND arrested the predominant portion of the radioactive pre-mRNA at the H-complex (Fig. 5, lane 5). The H-complex accounted for 83% of the radioactivity in lane 5 (Fig. 5). Although a small amount of radioactivity (~16%) could be found at a position corresponding to the A-complex, no B-complex formation was evident. Finally, there was essentially no difference in the total amount of radioactivity between the samples shown in Figure 5.

It should perhaps be noted that using the electrophoretic conditions described for native gels, the position in the gel indicated as the H-complex region actually consists of two complexes, the H- and E-complexes (Michaud and Reed 1993). Both are formed in the absence of

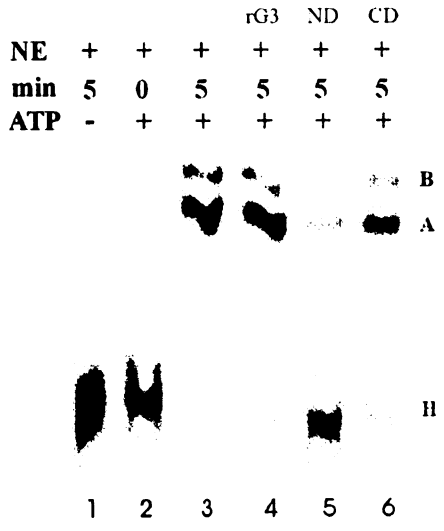

rG3 ND CD
NE + + + + + +
min 5 0 5 5 5 5
ATP - + + + + +
B
A
H
1 2 3 4 5 6

Figure 5: Comparison of the effect of the addition of recombinant galectin-3 or its NH_2- and COOH-terminal domains on the formation of splicing complexes. All samples contained nuclear extract (NE) and ^{32}P-labeled MINX pre-mRNA substrate (5,000 cpm). The sample in lane 1 was incubated for 5 minutes the absence of ATP; the samples in lanes 2-6 were incubated in the presence of ATP. Lanes: 2, no additions to NE - 0 min; 3, no additions to NE - 5 min; 4, 12 μM recombinant galectin-3 (rG3) - 5 min; 5, 10 μM ND - 5 min; 6, 10 μM CD - 5 min. Splicing complexes were analyzed by nondenaturing gel electrophoresis and autoradiography. The positions of the H-, A-, and B-complexes (Zillman et al., 1988) are highlighted on the right.

ATP. H-complexes contain the splicing substrate and hnRNP proteins while E-complexes additionally incorporate U1 snRNP at the 5′ splice site (Kramer 1996). Therefore, our present use of the term H-complex is meant to indicate a region of the gel rather a distinction between H- and E-complexes.

Effect of the addition of GST-Gemin4(C50) on the splicing assay

Since the ND fragment of galectin-3 exerts a dominant negative effect when added to a splicing reaction, we tested whether a fragment of Gemin4, the other component of the galectin-3:Gemin4 interacting partners, might yield a similar effect. Two lines of evidence indicate that the carboxyl-terminal 50 amino acids of Gemin4 possessed sufficient structural information for proper folding and binding to other proteins: (a) galectin-1 interacted with this 50-amino acid fragment to activate the transcriptional reporter in the yeast two-hybrid screen (Fig. 1); and (b) *in vitro* GST pull-down assays showed that Gemin4(C50) can bind both galectin-1 and galectin-3 (Fig. 2). Thus, Gemin4(C50) represented a fragment with retention of at least some of the binding properties of the full length Gemin4 polypeptide.

The NE used in this experiment converted about 30% of the pre-mRNA substrate into the mature mRNA product (Fig. 6,

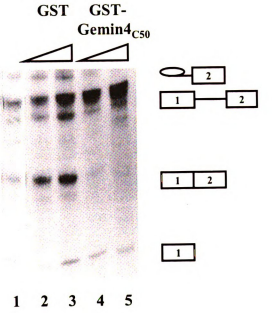

GST
GST-
Gemin4C50
2
1
2
1
2
1
1 2 3 4 5

Figure 6: The effect of Gemin4(C50) addition on the splicing activity of nuclear extract.

Lanes: 1, nuclear extract (NE); 2 and 3, GST at 10 and 30 μM added to NE; 4 and 5, GST-Gemin4(C50) at 10 and 30 μM added to NE. All reactions contained ^{32}P-labeled MINX substrate (5,000 cpm) and products of the splicing reactions were analyzed by electrophoresis through a 13% poly-acrylamide-urea gel, followed by autoradiography. The positions of migration of pre-mRNA substrate, splicing intermediates and RNA product are high-lighted on the right.

lane 1). When the splicing reaction was carried out in the presence of GST-Gemin4(C50), less than 2% conversion into the final product was observed (Fig. 6, lanes 4 and 5). No such inhibition was observed with corresponding concentrations of GST (mature mRNA product accounted for 30-40% of the radioactive RNA species [Fig. 6, lanes 2 and 3]). The purity of the GST and GST-Gemin4(C50) protein preparations was monitored during our GST pull-down studies (e.g., Fig. 2C). Neither preparation appeared to contain nonspecific proteases inasmuch as little or no degradation of galectin-1 or galectin-3 was observed in those binding assays. Finally, neither the GST nor the GST-Gemin4(C50) protein preparations appeared to degrade the pre-mRNA substrate.

DISCUSSION

The key findings of the present study include: (a) Gemin4 is an interacting partner of galectin-1 as identified through a yeast two-hybrid screen and it binds galectin-1 and galectin-3 directly in GST pull-down assays; (b) galectin-1 and galectin-3 are associated with macromolecular complexes that also contain polypeptides previously identified with one or more steps of the pre-mRNA splicing process; (c) fragments of either one of the interacting

partners, Gemin4: galectin-3, exert a dominant negative effect when added exogenously to the cell-free splicing assay; and (d) the inhibition of splicing by the galectin-3 ND appears to be due to a block at the H-complex in the assembly of the active spliceosome. The key points of these findings are summarized in the schematic diagram of Figure 7.

Three different lines of experiments now converge at the H-complex to implicate it as a site of action of galectins-1 and -3 in the splicing of pre-mRNA. First, we have identified galectin-1 and galectin-3 as novel components of macromolecular complexes containing SMN, Gemin2, and the core snRNP polypeptides bearing the Sm epitope (Fig. 7, left side). 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) (Dietz 1998; Matera and Frey 1998). In the former, the SMN complex is associated with snRNP core proteins and plays a critical role in the biogenesis of snRNPs (Fischer et al., 1997; Mattaj 1998). In the nucleus, SMN is required for pre-mRNA splicing by supplying snRNPs to the H-complex in the assembly of spliceosomes (Pellizzoni et al., 1998b; Meister et al., 2000) (Fig. 7, right side). This H-complex juncture is indeed where galectins appear to

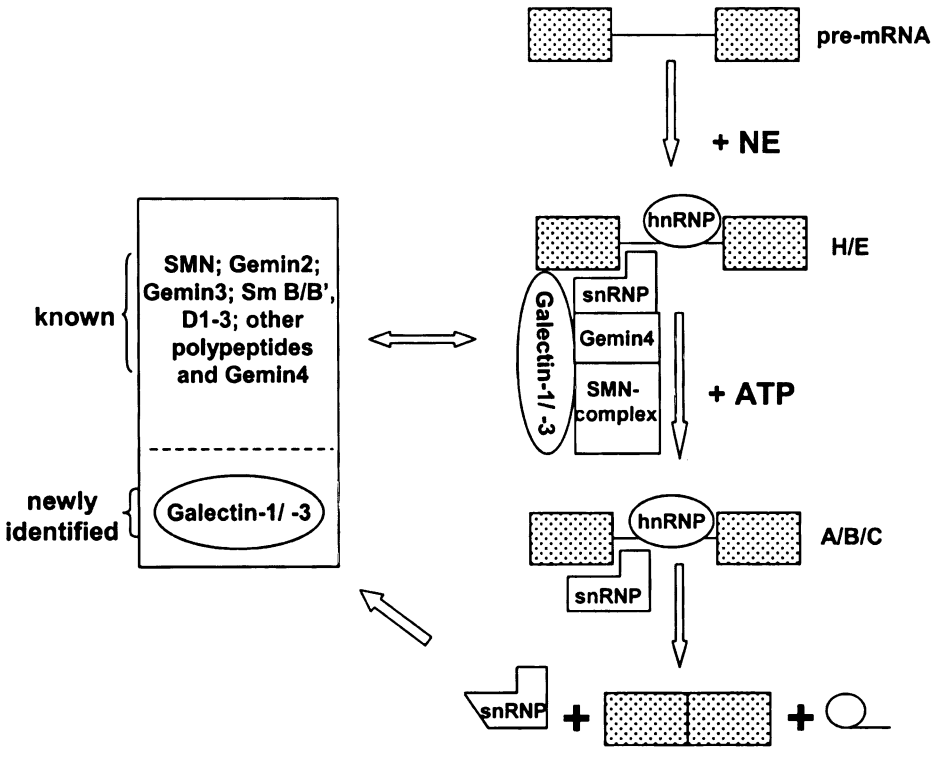

pre-mRNA
+ NE
hnRNP
H/E
snRNP
Gemin4
SMN-
complex
Galectin-1/ -3
+ ATP
known
SMN; Gemin2;
Gemin3; Sm B/B',
D1-3; other
polypeptides
and Gemin4
newly
identified
Galectin-1/ -3
hnRNP
A/B/C
snRNP
snRNP

Figure 7: Schematic diagram summarizing the components of SMN-containing complexes and the intermediates in spliceosome assembly. The known and newly identified polypeptides of SMN complexes are delineated on the left. The steps in active spliceosome formation are shown on the right. The dotted rectangles represent exons and the horizontal line represents the intron on the pre-mRNA. The ATP-independent H- and E-complexes migrate to a position designated as the H-complex in our native gel electrophoretic system. The SMN-containing complexes are illustrated here to supply snRNPs to the H-complex, leading to the formation of ATP-dependent active spliceosomal A-, B-, and C-complexes. This delivery may also involve galectin-1 and galectin-3 inasmuch as galectin depletion results in arrest of spliceosome assembly at the H-complex.

be required, as demonstrated by the effect of galectin depletion and by the effect of galectin-3 ND addition.

Second, in previous studies (Dagher et al., 1995), we had documented that NEs depleted of the galectins by affinity adsorption on lactose-agarose beads failed to form active spliceosomal complexes and gel mobility shift assays of ^{32}P-labeled pre-mRNA revealed only bands migrating in the H-complex region. Moreover, the activities of the galectin-depleted extract, in forming splicing complexes and in performing the *in vitro* splicing reaction, were reconstituted by the addition of recombinant galectin-3 with similar dose-response curves. On the basis of these results, we hypothesized that the galectins may be required at an early stage in the splicing pathway, such as the organization of the H-complex for the addition of other splicing factors (Dagher et al., 1995; Patterson et al., 1997). Interpreted under the context of our current findings, the effect of galectin-depletion is most likely to be exerted at either the assembly of SMN-containing complexes (Fig. 7, left side) or the docking of such complexes at the H-complex (Fig. 7, right side).

Finally, when the splicing assay was carried out in the presence of exogenously added ND of galectin-3, we also observed a dose-dependent inhibition of product formation. Again this inhibition was associated with the arrest of

spliceosome assembly at the H-complex. In both the *in vitro* splicing reaction and in the assembly of active spliceosomes, parallel additions of either the full-length galectin-3 polypeptide or the CD failed to yield the same effect.

Recently, the ND of galectin-3 has been expressed as a polypeptide fused onto GST. Addition of this GST-ND fusion protein to our splicing assay also resulted in inhibition of splicing, just as addition of ND. In contrast, GST itself had no effect when added to the same splicing assay (see also Fig. 6). Thus, two preparations of the galectin-3 ND, expressed in distinct systems and purified by different procedures, both inhibited the splicing reaction.

Physico-chemical studies have been carried out on the mouse (Agrwal et al., 1993) and hamster (Mehul et al., 1994) homologs of galectin-3, as well as the isolated ND and CD preparations derived from the full-length polypeptide. These studies have indicated that the two domains of galectin-3 are structurally, as well as functionally, distinct and independent. For example, differential scanning calorimetry of mouse galectin-3 yielded distinct transition temperatures for ND (~40 °C) and CD (~55 °C), both in the full-length polypeptide and as isolated preparations of individual domains (Agrwal et al., 1993). Similarly, the circular dichroic spectrum of intact hamster

galectin-3 is reproduced by an appropriately weighted summation of the spectra of the ND and CD fragments, suggesting that the isolated domains retain the structure that they had in the intact molecule (Mehul. et al., 1994).

Both of these studies, as well as studies on the human homolog (Hsu et al., 1992; Ochieng et al., 1993), clearly show that the CD bears the carbohydrate-binding activity of the protein. In addition, the CD of galectin-3 has been shown to interact with a cytoplasmic cysteine- and histidine-rich protein (Menon et al., 2000). In previous studies (Vyakarnam et al., 1997), we had found that polypeptides corresponding to the CD (the CD of galectin-3 or galectin-1, which contains a single domain homologous to it) were sufficient to reconstitute splicing activity in a galectin-depleted NE. However, the minimum concentrations required for activity were four to eight times higher than that of the full-length galectin-3 polypeptide, which contains the proline- and glycine-rich ND. This differential concentration effect on splicing reconstitution by intact galectin-3 versus its CD (or galectin-1) implies that both domains of galectin-3 interact with the splicing machinery. The nature of these interactions is unclear. Regardless, the observed splicing inhibition by excess ND likely occurs, then, through perturbation of one set of these interactions and supports this suggestion.

The identification of Gemin4 as an interacting partner of galectin-1 and galectin-3 represents an important advance in our investigation of the role of galectins in the cell nucleus. First, it provides evidence that galectins-1 and -3 interact with splicing components *in vivo*, as had been inferred previously on the basis of depletion-reconstitution and co-localization experiments. The dominant negative effect on splicing exerted by Gemin4(C50) strengthens the notion that direct binding between the galectins and Gemin4 are functionally relevant in the splicing pathway. Second, it supports our previous notion that, mechanistically, the locus of galectin action is at the H-complex during the assembly of spliceosomes. Finally, it offers new insights on the analysis of the composition and identities of polypeptides in the macromolecular complexes required for the splicing machinery and how galectin-1 or galectin-3 may facilitate this process.

Acknowledgement

We thank Kyle Openo for the preparation of recombinant galectin-3 and its domains and Joe Leykam of the Macromolecular Structure, Sequencing, and Synthesis Facility at Michigan State University for his help in characterizing

the protein samples by mass spectrometry and amino acid sequence analysis. We also thank Drs. S. Barondes, T. Feizi, J. Hirabayashi, and K. Kasai for various galectin and antibody reagents detailed in the text. This work was supported by grants GM-38740 from the National Institutes of Health and MCB 97-23615 from the National Science Foundation.

References

Agrwal, N., Q. Sun, S.-Y. Wang, and J.L. Wang. 1993. Properties of the recombinant polypeptide and the individuality of the domains. *J Biol Chem* **268**: 14931-14939.

Arenas, J.E. and J.N. Abelson. 1997. Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc Natl Acad Sci U S A* **94**: 11798-802.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Charroux, B., L. Pellizzoni, R.A. Perkinson, J. Yong, A. Shevchenko, M. Mann, and G. Dreyfuss. 2000. Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J Cell Biol* **148**: 1177-86.

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

Childs, R.A. and T. Feizi. 1979. b-galactoside-binding muscle lectins of man and monkey show antigenic cross-reactions with those of bovine origin. *Biochem J* **183**: 755-758.

Cooper, D.N., S.M. Massa, and S.H. Barondes. 1991. Endogenous muscle lectin inhibits myoblast adhesion to laminin. *J Cell Biol* **115**: 1437-48.

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

Dietz, H. 1998. Polishing the cutting edge of gems. *Nature gen* **20**: 321-322.

Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1993. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammlian nuclei. *Nucl Acids Res* **11**: 1475-1489.

Fields, S. and O. Song. 1989. A novel genetic system to detect protein-protein interactions. *Nature* **340**: 245-247.

Fischer, U., Q. Liu, and G. Dreyfuss. 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**: 1023-9.

Fu, X.-D. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**: 663-680.

Hirabayashi, J. and K. Kasai. 1991. Effect of amino acid substitution by site-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa β-galactoside-binding lectin. *J Biol Chem* **266**: 23648-23653.

Ho, M.K. and T.A. Springer. 1982. Mac-2, a novel 32,000 Mr mouse macrophage subpopulation-specific antigen defined by monoclonal antibodies. *J Immunol* **128**: 1221-1228.

Hsu, D.K., R.I. Zuberi, and 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.

Konarska, M.M. and P.A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNA. *Cell* **46**: 845-855.

Kramer, A. 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu Rev Biochem* **65**: 367-409.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**.

Laing, J.G. and J.L. Wang. 1988. Identification of carbohydrate binding protein 35 in heterogeneous nuclear ribonucleoprotein complex. *Biochemistry* **27**: 5329-5334.

Leffler, H. 1997. Introduction of galectins. *Trends Glycosci. Glycotechnol.* **9**: 9-19.

Matera, A.G. and M.R. Frey. 1998. Nuclear structure 98. Coiled bodies and gems: janus or gemini? *Am J Hum Genet* **63**: 317-321.

Mattaj, I.W. 1998. Ribonucleoprotein assembly: clues from spinal muscular atrophy. *Curr Biol* **8**: 93-95.

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

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

Meister, G., D. Buhler, B. Laggerbauer, M. Zobawa, F. Lottspeich, and U. Fischer. 2000. Characterization of a nuclear 20S complex containing the survival of motor neurons (SMN) protein and a specific subset of spliceosomal Sm proteins. *Hum Mol Gen* **9**: 1977-1986.

Menon, R.P., M. Strom, and C. Hughes. 2000. Interaction of a novel cysteine and histidine-rich cytoplasmic protein with galectin-3 in a carbohydrate-independent manner. *FEBS Lett* **470**: 227-231.

Michaud, S. and R. Reed. 1993. A functional association between the 5' and 3' splice site is established in the earliest prespliceosome complex (E) in mammals. *Genes Dev* **7**: 1008-20.

Moore, M.J., C.C. Query, and P.A. Sharp. 1993. Splicing of precursors to mRNA by the spliceosome. *In The RNA World, R.F. Gesteland and J.F. Atkins, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press)*: 303-357.

Moutsatsos, I.K., J.M. Davis, and J.L. Wang. 1986. Endogenous lectins from cultured cells: subcellular localization of carbohydrate-binding protein 35 in 3T3 fibroblasts. *J Cell Biol* **102**: 477-483.

Ochieng, J., D. Platt, L. Tait, V. Hogan, T. Raz, P. Carmi, and A. Raz. 1993. Structure-function relationship of a recombinant human galactoside-binding protein. *Biochemistry* **32**: 4455-4460.

Patterson, R., S.F. Dagher, A. Vyakarnam, and J.L. Wang. 1997. Nuclear galectins: functionally redundant components in processing of pre-mRNA. *Trends Glycosci. Glycotechnol.* **9**: 77-85.

Pellizzoni, L., N. Kataoka, B. Charroux, and G. Dreyfuss. 1998. A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* **95**: 615-24.

Raker, V.A., K. Hartmuth, B. Kastner, and R. Luhrmann. 1999. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Mol Cell Biol* **19**: 6554-65.

Reed, R. 2000. Mechanisms of fidelity in pre-mRNA splicing. *Curr Opin Cell Biol* **12**: 340-5.

Sharp, P.A. 1994. Split genes and RNA splicing. *Cell* **77**: 805-815.

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

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

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

Chapter 3

Galectin-1 is a component of the early and active spliceosome

Abstract

Galectin-1 (gal-1) and galectin-3 (gal-3) have been shown to be functionally redundant splicing factors using the criteria of depletion and reconstitution of cell free splicing extracts. Additionally, both proteins interact with gemin4, a member of the survival of motor neurons (SMN) complex. Here, we provide direct evidence that gal-1 is a component of both H and E (early) splicing complexes and active spliceosomes. Antibodies specific for gal-1 immunoprecipitate complexes that contain splicing competent pre-mRNA under conditions in which only H or E complexes are formed. These complexes also contain the SMN polypeptide. Anti-gal-1 antiserum does not co-immunoprecipitate H complexes formed on an RNA lacking sequence elements recognized by splicing factors. The association of gal-1 with these complexes is confirmed by reciprocal immunoprecipitation experiments using antibodies against the U1-specific 70K protein and the Sm core proteins. These data are consistent with the idea that gal-1, perhaps with

the SMN complex through its interaction with gemin4, is involved in early spliceosome assembly by delivering snRNPs to the H or E complex.

Using conditions permitting formation of active splicing complexes, anti-gal-1 co-immunoprecipitated splicing substrate, splicing intermediates and ligated exons, although gal-1 showed preferential association with the excised intron lariat. SMN and hnRNP C1/C2 are found in gal-1-containing active splicing complexes. In fact, anti-SMN antiserum is also able to immunoprecipitate splicing substrate. Since the SMN complex functions in the assembly and/or recycling of several RNP particles in the nucleus, gal-1 may function in these processes through its interaction with gemin4. Finally, analysis of snRNPs in the gal-1-containing complexes shows all 5 snRNPs in approximately equal molar amounts, defining gal-1 as a component of preassembled penta-snRNPs. Thus, gal-1 is an integral component of pre-mRNA processing complexes from initial formation to disassembly. In conjunction with the SMN complex, gal-1 may supply snRNPs, possibly as a penta-snRNP particle, to splicing substrate forming H and/or E complexes and assist in the (i) release of intron lariat from the spliceosome and (ii) dissociation of snRNPs from the active complex to be reassembled for subsequent rounds of splicing.

Introduction

Galectins are a family of proteins that bind to ß-galactosides and share sequence similarity, particularly in the carbohydrate recognition domain (Barondes et al., 1994). We previously have shown that gal-1 and gal-3 are required splicing factors using a cell-free splicing assay (Dagher et al., 1995; Vyakarnam et al., 1997). Depletion of both gal-1 and -3 from HeLa nuclear extracts (NE) either by lactose-sepharose affinity or double antibody adsorption chromatography abolished splicing. Splicing activity was restored by the addition of either galectin. These data suggest that the splicing activity attributed to gal-1 and gal-3 is functionally redundant. Consistent with this notion, gal-1 and -3 colocalize with known splicing factors such as Sm core proteins and SC35 in subnuclear speckles (Vyakarnam et al., 1997; Vyakarnam et al., 1998).

We recently identified gemin4 as a putative partner of gal-1 using a yeast two-hybrid genetic screen (Park et al., 2001). The interaction was confirmed by an *in vitro* binding assay in which it was found that gal-3 also interacted with gemin4. Polyclonal antibody against gal-1 co-immunoprecipitated components of the SMN complex (e.g., SMN, gemin2, and Sm polypeptides) implying that gal-1 (and likely gal-3) associates with the SMN complex through a direct interaction with gemin4. The functional role of this

interaction appears in the early assembly of the spliceosome. The addition of the N-domain of gal-3 or the C-terminal 50 amino acids of gemin4 to splicing extracts inhibited splicing and blocked an early step in spliceosome formation (i.e., no conversion of H/E complexes to active spliceosomes) in a dominant negative manner (Park et al., 2001).

SMN-containing complexes are found in the cytoplasm and in discrete nuclear bodies called gems (gemini of coiled bodies) (Liu and Dreyfuss, 1996; Matera and Frey, 1998). The SMN complex is involved in the biogenesis of snRNPs in the cytoplasm (Fischer et al., 1997; Meister et al., 2001; Jablonka et al., 2002; Massenet et al., 2002) and implicated in pre-mRNA splicing in the nucleus (Charroux et al., 2000; Meister et al., 2000; Massenet et al., 2002; Pellizzoni et al., 1998).

Although the function of gemin4 in pre-mRNA splicing is not known, the dominant negative effect on splicing by the carboxy 50 amino acids of gemin4 strongly supports the notion that direct binding between galectins and gemin4 may be functionally relevant in an early step in spliceosome assembly. How gal-1 and/or -3 facilitate assembly of the early spliceosome remains unknown. Whether gal-1 and/or -3 associate with the spliceosome is critical to elucidate the role of galectins in this assembly process. In the present

communication, we provide evidence from co-immunoprecipitation experiments that gal-1 is associated with both early commitment complexes and active spliceosomes. Co-immunoprecipitation experiments further reveal a preferential binding of gal-1 with the excised intron lariat suggesting a potential role of gal-1 in snRNP recycling or spliceosome disassembly. In addition, we provide evidence that the SMN protein is also a component both of early and active splicing complexes, presumably through gemin4-gal-1 interactions. Taken together, these data provide a foundation, as well as an impetus, for future studies on the composition of galectin-containing nuclear complexes and their role in pre-mRNA splicing and snRNP metabolism.

Materials and Methods

Antibodies

Polyclonal rabbit antiserum against recombinant rat gal-1 (Cooper et al., 1991) was a gift from Doug Cooper (University of California, San Francisco). Additionally, we generated a second polyclonal rabbit antiserum against GST-human gal-1 fusion protein (rabbit #55). The specificity of this antiserum is shown in Fig. 1. Western blotting shows that anti-gal-1 specifically recognizes a single polypeptide

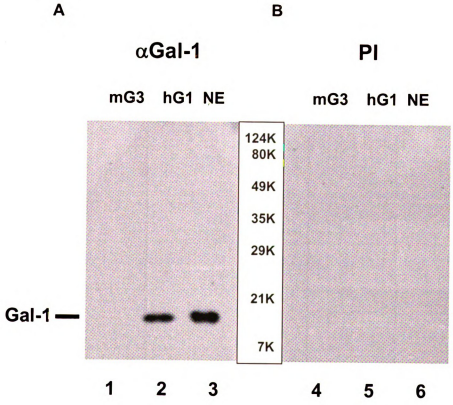

A
B
αGal-1
PI
mG3
hG1
NE
mG3
hG1
NE
124K
80K
49K
35K
29K
21K
7K
Gal-1
1
2
3
4
5
6

Figure 1: Specificity of anti-gal-1 antiserum (rabbit #55). Gal-1, gal-3 and NE were subjected to SDS-PAGE (12.5%) followed by western blotting with rabbit anti-gal-1 antiserum. Lanes 1 and 4, 0.5 µg recombinant mouse gal-3 (mG3); lanes 2 and 5, ~20 ng of recombinant human gal-1 (hG1); and lanes 3 and 6, ~35 µg HeLa NE protein. Lanes 1-3 were blotted with anti-gal-1 antibody and lanes 4-6 were blotted with pre-immune serum.

(~14 kD) in HeLa nuclear extracts (NE) (lane 3), with a mobility corresponding to that of recombinant human gal-1, which is also immunoblotted by the same antiserum (lane 2). In contrast, recombinant gal-3, even at high concentrations, was not detected by the anti-gal-1 antibody (lane 1). Pre-immune serum did not detect any proteins in the HeLa nuclear extract (lane 6) and did not react with gal-1 (lane 5) or gal-3 (lane 4). When used in immunoprecipitation experiments, anti-gal-1 precipitates gal-1, but not an irrelevant nuclear protein (HMG 14/17) from HeLa NE as was previously shown for the "Cooper" anti-gal-1 serum (Park et al., 2001). Autoimmune serum (ENA anti-Sm) was purchased from The Binding Site. A monoclonal antibody against hnRNP C1/C2, 4F4, was provided by Gideon Dreyfuss (University of Pennsylvania). Anti-SMN antibodies against a peptide corresponding to the amino terminal 19 amino acids (N-19) are affinity-purified polyclonal antibodies purchased from Santa Cruz Biotechnology (sc-7804). All antibodies were covalently cross-linked to protein G-sepharose fast flow 4B beads (Sigma) as previously described (Vyakarnam et al., 1997).

Co-immunoprecipitation of ^{32}P-MINX and detection of snRNA

Splicing reaction mixtures (20 µl) consisting of HeLa NE (12 µl in buffer D, ~60 µg protein), uniformly-radiolabeled *in vitro* transcribed MINX pre-mRNA (Dagher et al., 1995) or control RNA (see below), 2.5 mM $MgCl_2$, 0.5 mM dithiothreitol, and 20 U of RNasin (MBR, Milwaukee, WI) were incubated with or without ATP (1.5 mM) and creatine phosphate (20 mM) at 30°C for the times indicated. The template for MINX RNA was kindly provided by Susan Berget (Baylor University). The control RNA is transcribed from the plasmid originally designed as a positive control template for SP6 RNA polymerase (New England Biolabs). Two physical properties, a size similar to MINX RNA and the absence of any sequence elements for pre-mRNA splicing, make this RNA a suitable negative control substrate. The reaction mixture was incubated with 20 µl bed volume antibody (~400 µg IgG) cross-linked protein G-sepharose beads that had been pre-washed 3 times with 500 µl of buffer D (20 mM Hepes, pH 7.9, 20% glycerol, 100 mM KCl and 0.2 mM EDTA) containing 0.05% Triton X-100. The incubation was at 4°C for 2 hrs in the same buffer. After extensive washing with the same 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 4 mg/ml final concentration of proteinase K, 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 of chloroform. RNA was precipitated with 300 µl of ethanol at -80°C. The extracted RNAs were dissolved in 10 µl of sample buffer (9:1 / formamide: bromophenol blue), subjected to electrophoresis through 13% polyacrylamide (bisacrylamide-acrylamide 1.9:50 [wt/wt])-8.3 M urea gels. The gel was stained with ethidium bromide (final concentration, 3 µg/ml) to visualize snRNAs. The snRNAs were quantitated using densitometry (Scion Image). The radioactive RNA bands were revealed by autoradiography and/or quantitated by phosphorimage analysis (Molecular Dynamics). The other aliquot was subjected to electrophoresis in 15% SDS-PAGE (bisacrylamide-acrylamide 0.9:30 [wt/wt]) and analyzed by western blotting.

Glycerol gradient fractionation of splicing reactions

Splicing reaction mixtures were layered onto 5.0 ml 10-30% glycerol gradients in 60% buffer D. The gradients were centrifuged at 100,000 x g (45,000 RPM) using an SW50.1 rotor for 4 hrs at 4°C. The gradients were collected from the top in 18 fractions of 250 µl. 20 µl of each fraction

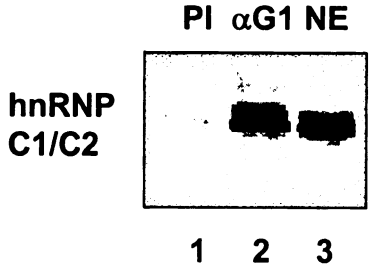
were subjected to native gel electrophoresis (1.5% agarose gel in 0.5X TBE) for analysis of spliceosomal complex formation (Das et al., 2000). The remainder of the sample was frozen at -80°C. Pooled fractions from these frozen samples were analyzed by coimmunoprecipitation experiments as described above.

Results

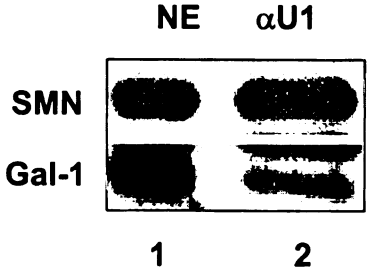
Galectin-1 is a component of H and E complexes

Previously, gal-1 was shown to be associated with Sm core polypeptides along with members of the SMN complex in HeLa NE (Park et al., 2001). This finding prompted us to identify other proteins in complexes immunoprecipitated by anti-gal-1. As seen in Fig. 2A, anti-gal-1, but not pre-immune serum, co-immunoprecipitated hnRNP C1/C2 from NE (lanes 2 and 1, respectively). Since hnRNP C1/C2 proteins have been shown to associate with spliceosomes (Choi et al., 1986), this finding suggests a direct association of gal-1 with spliceosomes. Reciprocal immunoprecipitation experiments were performed to confirm this suggestion. Under two different conditions (i.e., without substrate or with pre-mRNA splicing substrate plus ATP), gal-1 was detected in complexes precipitated by anti-U1 70K (Fig. 2B, lane 2) or anti-Sm (Fig. 2C, lane 3) antiserum. Goat anti-

A



B



C

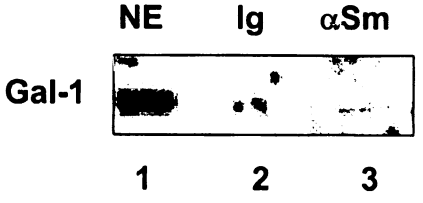


Figure 2: Coimmunoprecipitation experiments to determine gal-1's association with hnRNP C1/C2 and U1 snRNP. Panel A: Anti-gal-1 or pre-immune serum-linked beads were incubated with NE, bound material subjected to SDS-PAGE (12.5%) and blotted with 4F4 antibody (specific for hnRNP C1/C2 proteins). NE represents 50% of the NE used for antibody affinity adsorption. Panel B: NE was subjected to anti-U1 70K affinity adsorption and the bound material subjected to SDS-PAGE (12.5%) and analyzed by western analysis using anti-gal-1 and anti-SMN antiserum. NE represents 50% of the NE used for antibody selection. Panel C: NE was subjected to anti-Sm and goat anti-rabbit IgG affinity adsorption and bound material subjected to SDS-PAGE (12.5%) and blotted with anti-gal-1 antiserum.

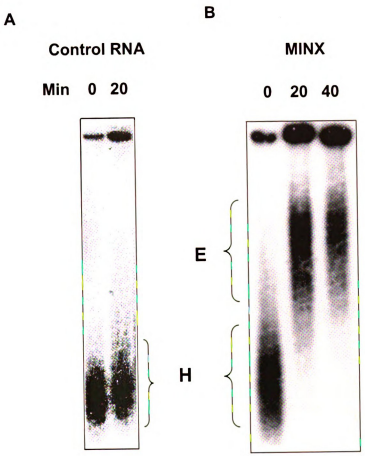

A
Control RNA
Min 0 20
B
MINX
0 20 40
E
H

Figure 3: Native gel analysis of reaction mixtures containing MINX RNA and control RNA. Control RNA lacking sequence elements for pre-mRNA splicing or MINX RNA containing splice sites, branch point, and polypyrimidine track was incubated at 30 °C in the absence of ATP for the indicated times. The reaction mixtures were loaded onto 1.5% agarose gels for native gel electrophoresis. Following electrophoresis, autoradiography was performed to detect RNA-containing complexes.

rabbit IgG cross-linked to protein G-sepharose failed to precipitate gal-1 (Fig. 2C, lane 2).

While these studies suggest that gal-1 is a component of complexes containing other known splicing factors, a direct demonstration of splicing substrate in complexes containing gal-1 would confirm the functional identity of such complexes. Two different *in vitro* transcribed substrates were tested: (i) the MINX pre-mRNA substrate and (ii) an RNA of similar size which lacks sequence elements recognized by splicing cofactors. The RNA substrate lacking donor and acceptor splice sites formed only complexes with mobilities similar to H complexes after 20 minute incubation at 30 °C (Fig. 3A). In contrast, splicing competent substrate formed an H complex upon addition to NE (Fig. 3B, lane 1) that was chased into an E complex (lanes 2 and 3) following incubation at 30°C in the absence of ATP.

Using these parameters, MINX pre-mRNA was incubated in NE without ATP at 4°C and 30°C to assemble H or E complexes, respectively. Fractionation of reactions containing H or E complexes on glycerol gradients followed by native gel analysis confirmed that these reactions consisted predominately of either H or E (Fig. 4A) complexes. Isolated fractions enriched in each complex (fraction 9-11 for H complex, fraction 13-15 for E complex) were

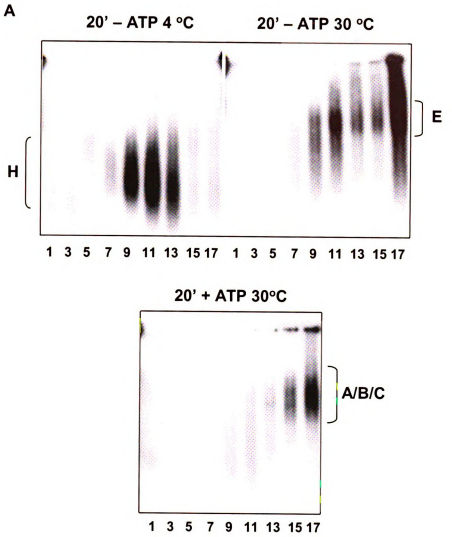

A
20' – ATP 4 °C
20' – ATP 30 °C
H
E
1 3 5 7 9 11 13 15 17 1 3 5 7 9 11 13 15 17
20' + ATP 30°C
A/B/C
1 3 5 7 9 11 13 15 17
B
αGal-1
NE
H
E
A/B/C
SMN
Gal-1

Figure 4: Glycerol gradient analysis of splicing complexes. Panel A: ^{32}P-MINX was incubated with NE minus ATP at 4°C (H complex), 30°C (E complex) for 20 min or incubated with ATP at 30°C (A/B/C complexes) for 20 min and fractionated on 10-30% glycerol gradients. Alternate fractions (numbered at bottom of lanes) were subjected to native gel electrophoresis (1.5% agarose) and splicing complexes detected by autoradiography. Heparin (final concentration of 300 μg/ml) was added to glycerol gradients fractions of the A/B/C complexes before native gel electrophoresis. Panel B: Fractions from the glycerol gradients enriched for H, E and A/B/C complexes were immunoprecipitated with anti-gal-1 antiserum and bound material blotted with anti-SMN or anti-gal-1 antibodies. NE is 5 μl of NE for comparison.

immunoprecipitated with anti-gal-1 antiserum and the bound material blotted for gal-1 and SMN. Each gradient fraction contained gal-1 and SMN (Fig. 4B).

We used these conditions to assemble H and E complexes on MINX RNA and H complexes on control RNA. The reaction mixtures were then subjected to immunoprecipitation with anti-gal-1 and the material bound by the antibody was analyzed in denaturing polyacrylamide gels to determine whether anti-gal-1 antiserum coimmunoprecipitates these substrate RNAs. Anti-gal-1 antibodies did not immunoprecipitate H complexes containing control RNA (Fig 5A, lane 2), but did immunoprecipitate splicing competent pre-mRNA assembled in H complexes at 0 time (Fig. 5B, lane 5) and E complexes at 20 through 60 min of incubation (Fig. 5B, lanes 6-8). To further confirm the identity of the MINX-containing H and E complexes, snRNAs co-immunoprecipitated by anti-gal-1 were characterized (Fig. 5C). The fact that U1 and U2 snRNPs are predominantly found in the bound material of anti-gal-1 antiserum is indicative of the affinity-purification of H and E complexes which also contain gal-1.

Thus, splicing complexes with H and E characteristics (i.e., mobility in glycerol gradients and native gels,

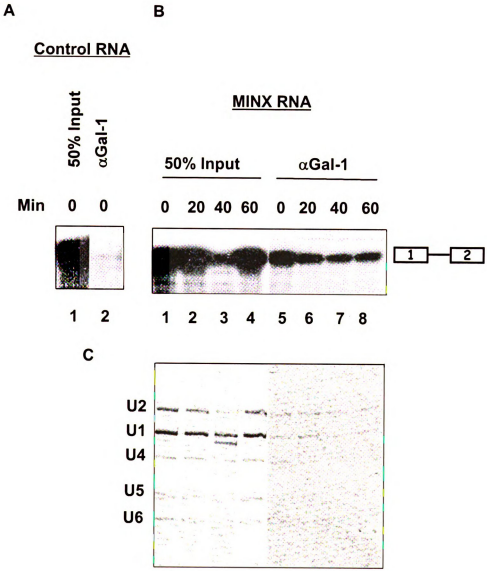

A
B
Control RNA
50% Input
αGal-1
MINX RNA
50% Input
αGal-1
Min 0 0
0 20 40 60 0 20 40 60
1
2
1 2
1 2 3 4 5 6 7 8
C
U2
U1
U4
U5
U6

Figure 5: Analysis of RNAs in H and E splicing complexes precipitated by anti-gal-1 antiserum. Panel A: Anti-gal-1-linked beads were mixed with a splicing reaction containing control RNA in the absence of ATP. The bound material was analyzed by denaturing gel electrophoresis and autoradiography. Lane 1 represents 50% of reaction used for immunoprecipitation and lane 2 is the control RNA bound to anti-gal-1 antibody. Panel B: As above except the splicing reaction contained MINX RNA and aliquots were analyzed following 0, 20, 40 and 60 min incubation. Lanes 1-4 represent 50% of the splicing reaction used for antibody selection and lanes 5-8 represent the antibody-selected material. Panel C: The denaturing gel used for RNA analysis (above) was stained with ethidium bromide to visualize the snRNAs in the input and antibody-bound fractions. Lanes 5-8 are a longer exposure than lanes 1-4. The decreased MINX (panel B, lane 3) and snRNAs (panel C, third lane) were due to degradation during RNA extraction.

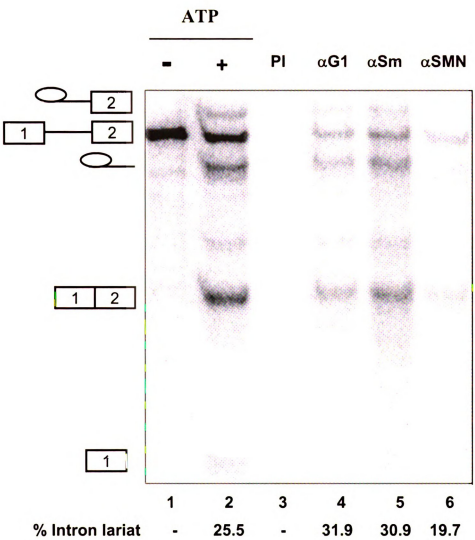
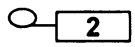
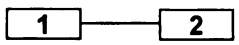
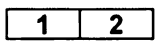

ATP
-
+
PI
αG1
αSm
αSMN
2
1
2
1
2
1
1
2
3
4
5
6
% Intron lariat
-
25.5
-
31.9
30.9
19.7

Figure 6: Analysis of spliceosomal RNA species immunoprecipitated by various antisera. 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, splicing reaction incubated without ATP; lane 2, splicing reaction with ATP representing 33% of the splicing reaction used for affinity selection; lane 3, RNA bound to pre-immune serum; lane 4, RNA bound to anti-gal-1; lane 5, RNA bound to anti-Sm; and lane 6, RNA bound to anti-SMN. The positions of the various splicing intermediates are diagrammed at the left. The percent of intron lariat RNA either in the input or immunoprecipitated RNA is indicated.

presence of U1 and U2 snRNPs and assembly exclusively on splicing competent substrate) contain gal-1 and SMN.

Galectin-1 complexes share characteristics of active spliceosomes

We next determined whether gal-1 (and SMN) remained associated with splicing complexes during the active phases of cleavage and ligation. Initial experiments characterized splicing substrate species precipitated by antisera specific for various splicing factors following 45 min incubation with ATP (Fig. 6). Anti-gal-1 antiserum immunoprecipitated all splicing RNA species, with a preference for intron lariat compared to the input RNA (compare lane 4 to lane 1). While the intron lariat in the input reaction constitutes 25.5% of the total RNA species, the percent of the intron lariat among all RNA species coimmunoprecipitated by anti-gal-1 moderately increased to 31.9% (Fig 6). The preferential coimmunoprecipitation of the excised intron lariat by anti-gal-1 was consistently observed in five independent experiments with two different anti-gal-1 antisera (Table 1). The preferential coimmunoprecipitation of the intron lariat was also observed by anti-Sm antisera (30.9%), which immunoprecipitated all spliceosomal RNAs. In contrast, pre-immune serum did not precipitate any ^{32}P-labeled MINX RNA (lane 3). Interestingly, an antibody

	Input(%)[1]	αG-1(%)	Fold Change[2]	S.D.
	2.7	2.2	-1.2	ND[3]
	19.0	29.6	+1.6	0.13
	42.5	31.7	+0.8	0.14
	35.0	35.2	+1.0	0.14
	1.6	2.0	+1.3	ND

[1] % was calculated by averaging phosphoimage data from 5 independent experiments and represents the percent for each RNA species of the input (spliceosomes) or the gal-1-containing complexes (αG-1)

[2] Fold change was determined by averaging the fold change in 5 independent experiments

[3] N.D. = not determined as the value were too low for accurate quantitation

Table 1: Quantitative analysis of spliceosomal RNAs coimmunoprecipitated by anti-gal-1 antiserum. The percent spliceosomal RNA present in the input splicing reaction and anti-gal-1 immunoprecipitates were compared. The data are based on 5 independent experiments using two different antisera. The fold change represents the average of the individual fold change calculated for each experiment. The standard deviation of the fold change for intron-exon 2 intermediate and exon 1 was not determined as these two RNA species were too low for accurate quantitation in 3 of the experiments.

raised against the amino terminal 19 amino acids of the SMN protein (anti-SMN) precipitated some of the splicing substrate. The coimmunoprecipitation of spliceosomal RNAs by anti-SMN is demonstrated for the first time here and suggests that the SMN protein is associated with spliceosomes, implicating a direct role of the SMN in pre-mRNA splicing. Unlike anti-gal-1 and anti-Sm antibodies, anti-SMN did not preferentially immunoprecipitate the excised intron lariat. The fact that anti-SMN did not show this preferential immunoprecipitation rules out a possibility that the preferential coimmunoprecipitation of the intron lariat is a general phenomenon of all antisera.

To gain a more detailed analysis of gal-1-containing splicing complexes, splicing reactions were subjected to antibody selection during a time course experiment. Splicing activity was detected at 20 min with approximately 40% product formation at 60 min (Fig. 7A). As expected, Sm core polypeptides and hnRNP C1/C2 proteins were found in the complexes throughout the splicing reaction along with gal-1 (Fig. 7C). Although it appears that there is a moderate change in the amount of hnRNP C1/C2 and Sm proteins coimmunoprecipitated by anti-gal-1 over time, this was not further analyzed since the amount of immunoprecipitated gal-

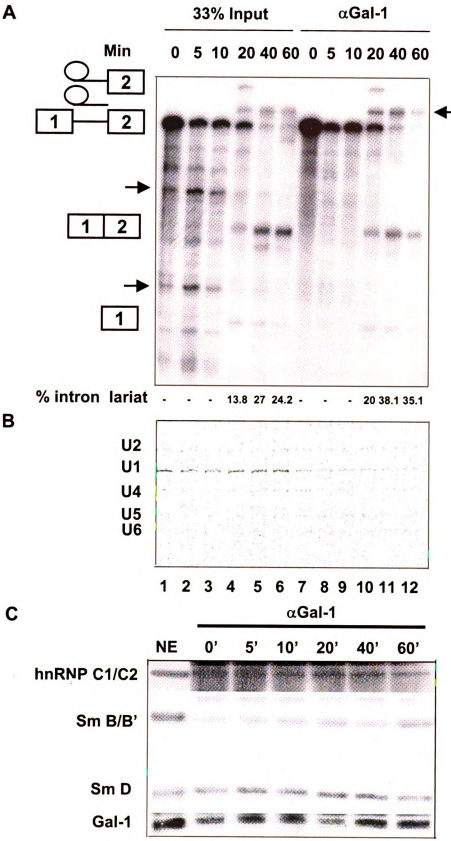

A
33% Input
αGal-1
Min
0 5 10 20 40 60 0 5 10 20 40 60
2
1 2
1 2
1
% intron lariat - - - 13.8 27 24.2 - - - 20 38.1 35.1
B
U2
U1
U4
U5
U6
1 2 3 4 5 6 7 8 9 10 11 12
C
αGal-1
NE 0' 5' 10' 20' 40' 60'
hnRNP C1/C2
Sm B/B'
Sm D
Gal-1

Figure 7: A time course analysis of spliceosomal RNAs and proteins precipitated by anti-gal-1 antiserum. ^{32}P-MINX was incubated with ATP in NE for the times indicated. Aliquots were subjected to anti-gal-1 affinity adsorption and the bound fraction separated into two parts. RNA was extracted from one portion and analyzed by denaturing gel electrophoresis (panels A and B). The other part was subjected to SDS-PAGE and blotted for various proteins (panel C). Panel A: Analysis of MINX at 0, 5, 10, 20, 40 and 60 min (input represents 33% of the reaction subjected to antibody selection) and MINX selected by anti-gal-1 adsorption. Arrows at left indicate degraded RNA species in the input material and arrow at right indicates intron lariat. Panel B: The above gel was stained with ethidium bromide to visualize snRNAs in the splicing reaction (lanes 1-6) and snRNAs in the complexes precipitated by anti-gal-1 (lanes 7-12). Individual snRNAs are indicated at left. Panel C: Proteins from the anti-gal-1 selected complexes were subjected to SDS-PAGE and blotted with 4F4 antibodies (hnRNP C1/C2), anti-Sm antiserum (Sm B/B' and Sm D) and anti-gal-1 (gal-1).

1 varied and there was no obvious correlation between the amounts of Sm or hnRNP C1/C2 and gal-1 to establish a functional significance.

Anti-gal-1 antibody coimmunoprecipitated splicing substrate throughout the reaction as well as splicing intermediates and ligated exons as they appeared during the kinetic analysis (Fig. 7A, lanes 7-12). An internal specificity standard is apparent in this analysis. Degraded RNA (indicated by arrows in the input lanes at 0, 5 and 10 min) is not observed in the immunoprecipitated complexes at these times. Quantitative analysis of these RNA species shows that the percent intron lariat present in the bound material of anti-gal-1 is ~30% higher than pre-mRNA and ~50% higher than ligated exons when compared to the percent intron lariat in the input reactions (indicated with an arrow at the right of the 60 min. lane). This is a strong indication of the preferential coimmunoprecipitation of the excised intron lariat over pre-mRNA or ligated exons as observed in the previous figure. It is possible that gal-1 preferentially interacts with all complexes containing intron RNA. This is not likely since intron intermediate (intron-exon 2) did not appear to be preferentially immunoprecipitated (see Table 1).

	Input[1]	αG-1[2]
U2	0.39	0.95
U1[3]	1.0	1.0
U4	0.3	0.81
U5	0.43	1.08
U6	0.28	0.72

[1] Represents the average density of each snRNA from the 6 input lanes in Fig. 7.
[2] Represents the average density of each snRNA for the gal-1-containing complexes in Fig. 7.
[3] The average density of U1 was set at 1.0

Table 2: Quantitative analysis of snRNAs coimmunoprecipitated by anti-gal-1 antiserum. Data from Fig. 7B were analyzed by densitometric analysis (Scion Corp.) to quantitate the snRNAs in the input fraction and in gal-1 containing complexes (anti-gal-1). The data represent an average of the 6 input and the 6 anti-gal-1 precipitated fractions.

Next, we examined the composition and quantity of snRNPs in the immunoprecipitated complexes (Fig. 7B). The denaturing gel for resolving exogenously added ^{32}P-RNA species was stained with ethidium bromide to visualize snRNAs prior to autoradiography. While mostly U1 and U2 snRNAs were present in the gal-1 immunoprecipitates of H/E complexes, all snRNAs (U1, U2, U4, U5, and U6) were coimmunoprecipitated by anti-gal-1 antibody from active splicing complexes (Table 2). Compared to the snRNAs in the splicing extract in which U1 is ~5-fold greater than U4, U5, or U6 (Fig. 7C, lanes 1-6), snRNAs immunoprecipitated by anti-gal-1 throughout the splicing reaction were approximately equal molar (lanes 7-12). These data suggest that the association of gal-1 with the spliceosome is through an interaction with a component of the spliceosome rather than a direct interaction with snRNPs or pre-mRNA.

Based on the data presented, we conclude that gal-1 is an integral component of early and active spliceosomes. In addition, we present evidence that the SMN protein is involved in pre-mRNA splicing by directly associating with the spliceosome, presumably through gal-1. Gal-1 is stably bound to active spliceosomes throughout the processing of pre-mRNA and is preferentially associated with the intron

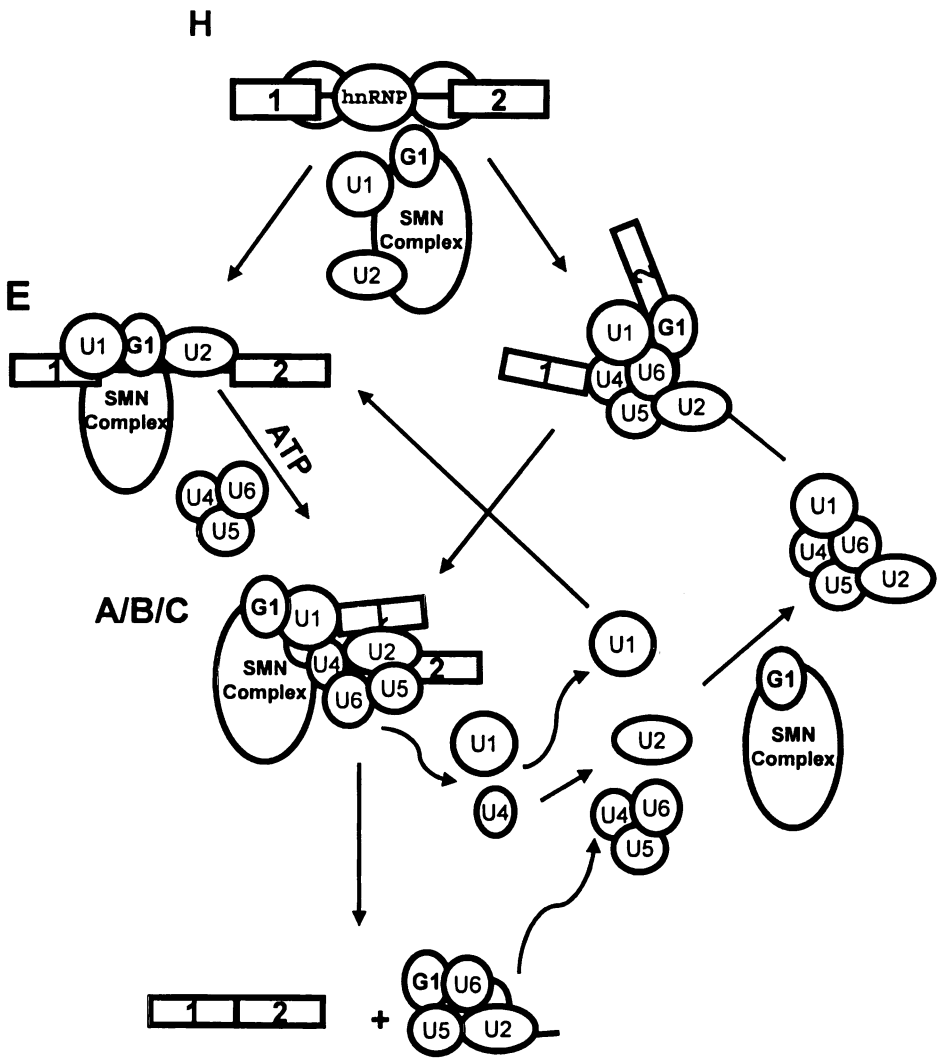

H
1
hnRNP
2
G1
U1
SMN
Complex
U2
E
U1
G1
U2
1
2
SMN
Complex
ATP
U4
U6
U5
U1
G1
U4
U6
U5
U2
U1
U4
U6
U5
U2
A/B/C
G1
U1
U2
U4
U6
U5
SMN
Complex
U1
U2
U4
U4
U6
U5
G1
SMN
Complex
1
2
+
G1
U6
U5
U2

Figure 8: Schematic diagram illustrating the involvement of gal-1 during spliceosome assembly and disassembly. Gal-1 in conjunction with SMN is associated with early splicing complexes, possibly in delivering or supplying snRNPs particles. Gal-1 is stably bound in active spliceosome assembly and preferentially associates with intron lariat to potentially mediate snRNP disassembly and reassembly for subsequent splicing events.

lariat, suggesting a role in snRNP disassembly or reassembly possibly by assisting the SMN complex. Nuclear gal-1 complexes contain snRNAs in ratios that differ from the snRNA ratios observed in nuclear extracts.

Discussion:

We previously have shown by depletion-reconstitution experiments that galectins-1 and -3 are functionally required splicing factors. Data obtained in this study, however, document a direct association of gal-1 in spliceosomal complexes. The significance of this work can be described in two aspects (and is depicted in our working model shown in Fig. 8). First, these data support our hypothesis that gal-1 is a general splicing factor. The fact that anti-gal-1 antibody coimmunoprecipitated pre-mRNA associated with H and E complexes in an RNA-dependent manner illustrates the initial entry of gal-1 during early complex assembly. Gal-1 appears to be stably bound throughout the assembly of the spliceosome and catalysis. While a precise molecular function of gal-1 during spliceosome assembly is not known, it appears to be in the recruitment or supplying of snRNPs to pre-mRNA (nascent primary transcript) based on two observations: (i) depletion of galectins inhibits transition of H/E complexes to A complexes (Dagher et al., 1995), and (ii) gal-1 is associated with gemin4 in SMN

complexes (Park et al., 2001) which are implicated in rejuvenating snRNPs to re-supply them to pre-mRNA in the early commitment complex (Pellizoni et al., 1998, Meister et al., 2000).

The second important contribution is the possibility that gal-1 might also be involved in the disassembly and recycling of snRNPs for subsequent rounds of splicing. Since the intron lariat is the last juncture where snRNPs bind before they are disassembled and recycled, the preferential association of gal-1 with the intron lariat implicates a potential role of gal-1 in these processes. While the assembly of the spliceosome has been studied both biochemically and genetically in detail (for review see Staley and Guthrie, 1998), disassembly of the spliceosome has received less attention. In the yeast *Saccharomyces cerevisiae*, several DEAH helicases including Prp43 have been identified as factors responsible for the release of mature mRNA and intron lariat from the spliceosome (Martin et al., 2002). These factors seem to be preferentially associated with the intron lariat and are essential for splicing activity. By promoting RNA unwinding, DEAH helicases facilitate the release of product and intron lariat from the spliceosome (Schwer and Meszaros, 2000). Although a dramatic modification of snRNP particles during splicing has not been identified, some reassembly or recycling is

required for re-utilization in the splicing pathway (Makarova et al., 2002; Martin et al. 2002). After snRNP particles are dissociated from the intron lariat, U4 snRNP must re-associate with U5/U6 snRNP particles to form a functional tri-snRNP complex. We hypothesize that gal-1 is involved in this process, possibly through its association with the SMN complex.

The SMN complex has been shown to be a mediator of assembly of snRNP complexes in the cytoplasm (Meister et al., 2001; Massenet et al., 2002; Yong et al., 2002). In the nucleus, the SMN complex is concentrated in Cajal bodies and gems. When newly made Sm polypeptides tagged with green fluorescence protein (GFP) were expressed in mammalian cells, they first localized in Cajal bodies and nucleoli, and later in nuclear speckles (Sleeman and Lamond 1999). In contrast, mature snRNPs localized immediately in nuclear speckles. Therefore, it is postulated that snRNPs go through a maturation process in a regulated manner from one nuclear body to another. The ability of the SMN complex to assemble several macromolecular complexes (Terns and Terns 2001; Paushkin et al., 2002) makes the nuclear SMN complex a potential candidate for the nuclear maturation or re-assembly of snRNPs.

Although a precise role of the SMN complex in pre-mRNA splicing has not been demonstrated, there are several lines

of evidence that the SMN protein is involved in pre-mRNA splicing. First, antibodies against the SMN protein inhibit splicing activity (Pellizzoni et al., 1998; Meister et al., 2000). Second, a deletion mutant of the amino-terminal 27 amino acids of SMN inhibits splicing in a dominant negative manner (Pellizzoni et al., 1998). However, the mutant SMN protein is only inhibitory if the splicing extract is pre-incubated with ATP for 30 min prior to the addition of splicing substrate. In contrast, addition of full length SMN protein during preincubation stimulates splicing. These results suggest that the SMN protein or the SMN complex has a critical role during the dissociation of endogenous splicing complexes prior to the next round of splicing, possibly in the reassembly of snRNPs complexes. We have demonstrated that the SMN protein is directly involved in splicing as antibodies specific for SMN precipitate splicing substrate. Taken together, the SMN complex may have dual, although not mutually exclusive, roles in which it mediates the reassembly of snRNP particles after a round of splicing and supplies or delivers snRNPs to the early spliceosomes.

The association of gal-1 with snRNP particles and hnRNP C1/C2 in nuclear extracts without exogenously added splicing substrate suggests the presence of endogenous RNP complexes. Two likely candidates are endogenous spliceosomes or multi-snRNP particles. Conway et al., (1989) observed large snRNP

complexes in HeLa NE that were splicing deficient when incubated with pre-mRNA. These snRNP complexes shifted into slower sedimenting glycerol gradient fractions that were now splicing competent upon substrate addition. These authors suggest that these complexes are not endogenous spliceosomes, as snRNPs and hnRNP C1/C2 are not coimmunoprecipitated by various antisera. In contrast, Lerner et al., (1980) have reported endogenous complexes containing both snRNPs and hnRNA (and presumably hnRNP C1/C2), similar to our findings. Recently Abelson and colleagues (Stevens et al., 2002) reported that *S. cerevisiae* whole cell extracts contain a 45S fully assembled and functional complex composed of all 5 snRNPs (in equal molar amounts) and more than 60 splicing factors that they termed a penta-snRNP. They suggest, then, that splicing need not proceed in a step-wise manner as others predict. Rather, the fully assembled penta-snRNP complex would bind to pre-mRNA and be remodeled throughout the splicing pathway.

Our data show that U1 and U2 were preferentially coimmunoprecipitated from splicing reactions incubated without ATP while the 5 spliceosomal snRNAs (U1, U2, U4, U5, U6) were coimmunoprecipitated in approximately equal molar ratios from splicing reactions incubated with ATP. These findings support the canonical model of stepwise assembly of the spliceosome, although we forward an alternative

explanation suggesting the existence of a penta-snRNP complex that initially binds exogenous splicing substrates. In the presence of ATP (i.e., splicing conditions both *in vivo* and *in vitro*) these complexes are stable and remain intact. Thus, anti-gal-1 precipitated complexes, even at 0 time, contain pre-mRNA and all 5 snRNPs. In contrast, the penta-snRNP/pre-mRNA complex is not stable in the absence of ATP. Thus, gal-1-containing H and E complexes have only U1 and U2: the unstable tri-snRNP having been released. In either case, the association of gal-1 with splicing competent RNA is through factors, possibly the SMN complex, in a dynamic yet intact macromolecular machine.

Recent advances in proteomic techniques using tandem mass spectrometry with an ever growing database have identified an increasing number of proteins associated with the spliceosome and improved our understanding of this complex and dynamic machine (Jurica et al., 2002; Rappsilber et al., 2002; Zhou et al., 2002). Gal-1 and -3 have not been detected in the splicing complexes used in these studies. Other protein components that are involved in U4/U6 and U4/U5/U6 snRNPs functions in splicing also have not been found possibly due to low abundance, weak or transient interactions, or other technical reasons (Rappsilber et al., 2002). Although the identification of galectins in spliceosomes using proteomic techniques would

corroborate our findings, it is inconceivable to question whether galectins are splicing factors based on the plethora of evidence indicating otherwise.

In summary, we have demonstrated that gal-1 is a component of the splicing machinery. In addition, we have provided evidence that suggests gal-1 has dual, although not mutually exclusive, functions in supplying snRNP particles for early spliceosomal assembly and in facilitating reassembly or recycling of snRNPs at the end of the splicing reaction. These functions could be facilitated by the interaction with the SMN complex, presumably through gal-1/gemin4 interactions. We also demonstrated for the first time that the SMN complex is directly associated with the spliceosome supporting our hypothesis that the association of gal-1 with the SMN complex is functionally significant for pre-mRNA splicing. For future studies, the functional role(s) of gal-1 and the significance of its interaction with the SMN complex in snRNP recycling will be mechanistically studied.

References

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, H. Leffler, F.-T. Liu, R. Lotan, A.M. Mercurio, M. Monsigny, S. Pillai, F. Poirier, A. Raz, P.W. Rigby, J.M. Rini, and J.L. Wang. 1994. Galectins: A family of animal β-galactoside-binding lectins. *Cell* **76**: 597-598.

Charroux, B., L. Pellizzoni, R.A. Perkinson, J. Yong, A. Shevchenko, M. Mann, and G. Dreyfuss. 2000. Gemin4: A novel component of the SMN complex that is found in both Gems and Nucleoli. *J Cell Biol* **148**: 1177-1186.

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

Cooper, D.N., S.M. Massa, and S.H. Barondes. 1991. Endogenous muscle lectin inhibits myoblast adhesion to laminin. *J Cell Biol* **115**: 1437-48.

Conway, G.C., A.R. Krainer, D.L. Spector, and R.J. Roberts. 1989. Multiple splicing factors are released from endogenous complexes during in vitro pre-mRNA splicing. *Mol Cell Biol* **9**: 5273-5280

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

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

Fischer, U., Q. Liu, and G. Dreyfuss. 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**: 1023-9.

Jablonka, S., B. Holtmann, G. Meister, M. Bandilla, W. Rossoll, U. Fischer, and M. Sendtner. 2002. Gene targeting of Gemin2 in mice reveals a correlation between defects in the biogenesis of U snRNPs and

motoneuron cell death. *Proc Natl Acad Sci U S A* **99**: 10126-31.

Jurica, M.S., L.J. Licklider, S.P. Gygi, N. Grigorieff, and M.J. Moore. 2002. Purification and characterization of native spliceosome suitable for three-dimensional structural analysis. *RNA* **8**: 426-439.

Lerner, M.R., J.A. Boyle, S.M. Mount, S.L. Wolin, and J.A. Steitz. 1980. Are snRNPs involved in splicing? Nature (London) **283**:220-224

Liu, Q. and G. Dreyfuss. 1996. A novel nuclear structure containing the survival of motor neurons protein. EMBO J **14**:3555-3565

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

Matera, A.G. and M.R. Frey. 1998. Coiled bodies and gems: Janus or gemini? *Am J Hum Gen.* **63**: 317-321

Martin, A., S. Schneider, and B. Schwer. 2002. Prp43 is an essential RNA-depedent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* **277**: 17743-17750

Massenet, S., L. Pellizzoni, S. Paushkin, I.W. Mattaj, and G. Dreyfuss. 2002. The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* **22**: 6533-41.

Meister, G., D. Buhler, B. Laggerbauer, M. Zobawa, F. Lottspeich, and U. Fischer. 2000. Characterization of a nuclear 20S complex containing the survival of motor neurons (SMN) protein and a specific subset of spliceosomal Sm proteins. *Hum Mol Gen* **9**: 1977-1986.

Meister, G., D. Buhler, R. Pillai, F. Lottspeich, and U. Fischer. 2001. A multiprotein complex mediates the ATP-dependent assembly of spliceosomal U snRNPs. *Nat Cell Biol* **3**: 945-9.

Park, J.W., P.G. Voss, S. Grabski, J.L. Wang, and R.J. Patterson. 2001. Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. *Nucl Acids Res* **27**: 3595-3602.

Paushkin, S., A.K. Gubitz, S. Massenet, and G. Dreyfuss. 2002. The SMN complex, an assemblyosome of ribonucleoproteins. *Curr Opin Cell Biol* **14**: 305-12.

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

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

Schwer, B., and T. Meszaros. 2000. RNA helicase dynamics in pre-mRNA splicing. *EMBO J* **19**:6582-6591

Sleeman, J.E. and A.I. Lamond. 1999. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr Biol* **9**: 1065-74.

Staley, J.P. and C. Guthrie. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**: 315-326

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

Terns, M.P. and R.M. Terns. 2001. Macromolecular complexes: SMN - the master assembler. *Cur Biol* **11**: 862-864.

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

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

Yong, J., L. Pellizzoni, and G. Dreyfuss. 2002. Sequence-specific interaction of U1 snRNA with the SMN complex. *EMBO J* **21**: 1188-96.

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

Chapter 4

Concluding statements

Pre-mRNA splicing is an essential post-transcriptional process mediated by a large ribonucleoprotein complex called spliceosome. Accurate assembly and disassembly of spliceosome are important for the fidelity in pre-mRNA splicing. Dynamic interactions between snRNP particles and protein factors facilitate these processes. Galectin-1 and -3 had been initially isolated in nuclear RNP complexes. Our laboratory has made a considerable progress in understanding the significance of such observation. Previous studies showed that gal-1 and -3 were required splicing factors using cell-free splicing assays. Data from depletion-reconstitution experiments proved that gal-1 and -3 were both necessary and sufficient for their function in splicing. In addition, double labeling immunostaining experiments provided *in vivo* evidence that gal-1/-3 colocalize with Sm core proteins and a splicing factor, SC35.

Significance of my present work is two-fold. First, the identification of Gemin4 as an interacting partner for gal-1 using the yeast two hybrid provides a mechanistic insight regarding the role of gal-1 in splicing. Since Gemin4, a component of the SMN complex, which has been implicated in delivery or supplying snRNPs to the early

complex, data demonstrating Gemin4-galectin-1 interaction not only validates previous studies showing gal-1 is a splicing factor, but also suggests that galectin-1 is involved in the assembly of early spliceosome. Depletion of gal-1 and/or -3 resulting in arresting the complex formation at the H complex, addition of the amino terminal domain of gal-3 and carboxy terminal 50 amino acids of Gemin4 to splicing extracts, inhibiting splicing and formation of active spliceosomes, and coimmunoprecipitation of splicing RNA substrate contained in E complex by anti-gal-1 support the hypothesis.

Second, my work has demonstrated the association of gal-1/-3 and SMN in the active spliceosome. Coimmunoprecipitation experiments showing association of gal-1 and SMN with splicing RNA substrates, intermediates, and ligated exons under conditions forming active spliceosomes document direct evidence that gal-1 is a component of the spliceosome. Furthermore, preferential association of galectin-1 with the intron lariat may suggest an additional role in pre-mRNA splicing; snRNP reassembly or recycling. Additional works are needed to test these hypothes. For example, a reliable *in vitro* recycling assay can be useful technique to gain an insight if gal-1 is, indeed, involved in facilitating snRNP recycling through interaction with the SMN complex. Also, proteomics

approaches identifying other interacting proteins of galectin-1 from complexes containing gal-1 and the intron lariat can provide other information to delineate the role of gal-1 in snRNP recycling. Present works contained in this thesis, thus, lay a foundation for future studies in which functions of Gemin4-galectin-1 interactions and the role of Gemin4 in pre-mRNA splicing can be determined.