This is to certify that the
dissertation entitled

GALECTIN-1, GALECTIN-3, AND TFII-I
IN PRE-MRNA SPLICING

presented by

RICHARD MITCHELL GRAY

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Biochemistry and Molecular Biology

Major Professor's Signature

9/14/06

Date

*MSU is an Affirmative Action/Equal Opportunity Institution*

# GALECTIN-1, GALECTIN-3, AND TFII-I IN PRE-MRNA SPLICING

By

Richard Mitchell Gray

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry and Molecular Biology

2006

## ABSTRACT

## GALECTIN-1, GALECTIN-3, AND TFII-I IN PRE-MRNA SPLICING

By

Richard Mitchell Gray

Galectin-1 (Gal1) and galectin-3 (Gal3) are two members of a family of galactose-specific carbohydrate-binding proteins. The polypeptide of Gal1 consists of a single domain, designated as the carbohydrate recognition domain (CRD). The polypeptide of Gal3 contains two domains, an $NH_2$-terminal domain rich in proline and glycine residues and a COOH-terminal CRD. Both proteins are found in the nucleus of cells. Previous studies had shown that Gal1 and Gal3 are involved in nuclear splicing of pre-mRNA. The experiments in this thesis were performed to investigate the role of the domains in the splicing activity of the polypeptides.

Gal3 was expressed and purified as a fusion protein with glutathione S-transferase (GST). When nuclear extracts of HeLa cells were subjected to adsorption on GST-Gal3 beads, the general transcription factor II-I (TFII-I) was identified as one of the polypeptides specifically bound. Lactose, a saccharide ligand of the galectins, inhibited GST-Gal3 pull-down of TFII-I from nuclear extract. Antibodies directed against TFII-I inhibited the splicing reaction in a dose-dependent fashion and co-precipitated spliceosomal RNA and proteins. These results suggest that TFII-I associates with galectin-1- or galectin-3-containing spliceosomal complexes.

Three site-directed mutants of Gal1 were expressed and purified as fusion proteins with GST. These mutants, designated as GST-Gal1(N46D), GST-Gal1(C60S), and GST-Gal1(E71Q), were compared with the wild-type Gal1 construct, GST-Gal1(WT), in three

assays: (a) binding to asialofetuin-Sepharose as a measure of carbohydrate-binding activity; (b) pull-down of TFII-I from nuclear extract; and (c) reconstitution of splicing in nuclear extract depleted of galectins. GST-Gal1(N46D) exhibited a marked decrease in carbohydrate-binding activity compared to GST-Gal1(WT). Both GST-Gal1(WT) and GST-Gal1(N46D) were equally efficient, however, in pull-down of TFII-I and in reconstitution of splicing activity. Together, these results suggest that the Gal1 saccharide-binding activity *per se* is not required for the splicing activity.

Monoclonal antibody NCL-GAL3 reacts with an epitope in the $NH_2$-terminal 14 amino acids of the Gal3 polypeptide. Addition of this antibody to splicing competent nuclear extract inhibited the splicing reaction. In contrast, the epitope of a second monoclonal antibody, anti-Mac-2, maps to residues 48-100, containing PGAYPGXXX repeats. This antibody had no effect on splicing. A synthetic peptide containing three perfect repeats of the sequence PGAYPGQAP (27-mer) inhibited the splicing reaction. In contrast, addition of peptides corresponding to a single iteration (9-mer) or two repeats (18-mer) of this motif failed to yield the same effect. . One interpretation of these results is that the portion of the Gal3 polypeptide bearing the PGAYPGXXX repeats is sequestered through interaction with the splicing machinery and is inaccessible to the anti-Mac2 antibody.

Together, the results indicate that both the proline- and glycine-rich domain, as well as the CRD, interact with the spliceosomal machinery. Although the CRD contains the saccharide-binding site of the Gal1 and Gal3 polypeptides, the carbohydrate-binding activity is not required for splicing activity.

To my family.
Their support has made this possible.

## AKNOWLEDGEMENTS

I wish to thank my mentor Dr. John Wang. I will forever hold his patience, dedication, optimism, and enthusiasm as standards to emulate in my own personal and professional life. Dr. Wang taught me how to identify the important questions embedded in complex systems and results, and more importantly, how to use solid evidence to direct thinking about any question. His constant observation of, and speculation about, everyday events outside the lab have shown me the value of applying careful obervation and keen questioning to phenomena beside experimental results--and have stimulated lunchtime discussions I will not soon forget.

I would also like to thank the members of my guidance committee: Drs. Ron Patterson, Zach Burton, Bill Henry, and Min-Hao Kuo. They provided many good suggestions which contributed immensely to this thesis and to my development as a scientist.

Thank you to the current and former members of the Wang and Patterson laboratories. Patty Voss provided invaluable technical advice as well as daily reminders that there is a world outside the Biochemistry Building. Ron Patterson's proclivity for debate with anyone on any subject has shown me the excitement to be had in spirited discussion. Kevin Haudek and Weizhong Wang were generous in their indulgence of my frequent trips to the Patterson Lab seeking company, advice, and reagents. Eric Arnoys was an excellent mentor for a young scientist and has become an excellent friend--I will always have fond memories of my introduction to galectin-3 during the summer of 1999. I owe a particular thanks to Dr. Peter Davidson, whose companionship has added immeasurably to this experience. I could not have asked for a better colleague.

Finally, I wish to thank my parents, Ruth and Gary Gray, and my sister, Dr. Catherine Robertson. Their unflagging support, incredible patience, and vigorous encouragement during this endeavor have meant more to me than I can say. Thank you.

## TABLE OF CONTENTS

## LIST OF TABLES

**Chapter 1.**

## LIST OF FIGURES

**Chapter 1.**



**Chapter 2.**



**Chapter 3.**

**Chapter 4.**

## LIST OF ABBREVIATIONS

AdML, adenovirus major late

anti-M2, rat monoclonal anti-galectin-3

ASF, asialofetuin

ASF/SF2, alternate splicing factor/splicing factor2

BAP-135, Bruton's tyrosine kinase associated protein of 135 kD

BHK, baby hamster kidney

Btk, Bruton's tyrosine kinase

CAT, chloramphenicol acetyl transferase

CD, carboxyl-terminal domain

Chrp, cysteine- and histidine-rich protein

CID, collision induced dissociation

CoAA, coactivator activator

CRD, carbohydrate recognition domain

CTD, carboxyl-terminal domain of RNA polymerase II

DTT, dithiothreitol

ESE, exonic splicing enhancer

EGF, epidermal growth factor

ERK, extracellular signal-regulated kinase

ERSE, endoplasmic reticulum stress response element

ERSF, endoplasmic reticulum stress response factor

Gal-1, galectin-1

Gal-3, galectin-3

GFP, green fluorescence protein

G-kinase Iβ, cGMP-dependent protein kinase Iβ

Grp, glucose-regulated protein genes

GST, glutathione S-transferase

HDAC, histone deacetylase

HRP, horseradish peroxidase

hnRNP, hetergeneous nuclear ribonucleoprotein

IL, interleukin

JAK2, Janus kinase 2

Lac, lactose

MAPK, mitogen-activated protein kinase

ND, amino-terminal domain

NE, nuclear extract

NLS, nuclear localization signal

PBS, phosphate-buffered saline

PDGF, platelet-derived growth factor

PGC-1, peroxisome proliferator-activated receptor-γ coactivator-1

PH, pleckstrin homology

PIAS, protein inhibitor of activated STATs

PMSF, phenylmethylsulfonyl fluoride

pre-mRNA, pre-messenger RNA

PSF, PTB-associated splicing factor

PTB, polypyrimidine-tract binding protein

RNP, ribonucleoprotein complex

RRM, RNA recognition motif

SCAF, SR-like CTD associated factors

SF1, splicing factor 1

SIE, c-sis/PDGF-inducible element

SKIP, Ski-interacting protein

SMN, survival of motor neurons

snRNA, small nuclear RNA

snRNP, small nuclear ribonucleoprotein

SPIN, serum response factor-phox 1 interacting protein

SR, serine-arginine-rich

SRE, serum response element

SRF, serum response factor

STAT, signal transducer and activator of transcription

Sufu, suppressor of fused

$TAFII_{68}$, 68 kD subunit of the TATA binding protein associated factor II

TDG, thiodigalactoside

TFII-I, transcription factor II-I

TRBP, thyroid hormone receptor binding protein

T-TBS, tris-buffered saline containing 0.05% Tween 20

TTF-1, thyroid transcription factor-1

U2AF, U2 associated factor

USF, upstream stimulatory factor

Vβ, T cell receptor variable region-derived promoters

WT1, Wilms' tumor gene

# CHAPTER 1

## Literature Review

## I. Galectins

### A. The galectin family

The galectins are a family of carbohydrate-binding proteins that share two key properties: (a) binding affinity for β-galactosides; and (b) conserved sequence elements in the carbohydrate-binding site (1). To date, fifteen mammalian galectins have been identified (Figure 1, panel A) and numbered sequentially according to the accepted numbers for their genes in the Genome Database. Galectins have also been identified in many non-mammalian species, including birds, amphibians, fish, worms, sponges and fungi (2). Screening the databases of genomic DNA sequences and expressed sequence tags has revealed additional candidates for membership in the mammalian galectin family, as well as putative galectins in plants and viruses (3).

Each member of the galectin family contains at least one domain of about 130 amino acids; this domain binds to saccharides and is therefore designated the Carbohydrate Recognition Domain (CRD). Based on the number and organization of domains in the polypeptides, the galectins have been classified into subfamilies (Figure 1, panel A) (4): (a) the Prototype group (galectins-1, -2, -5, -7, -10, -11, -13, -14, and -15) contains one domain, the CRD; (b) the Chimera group (galectin-3) contains an unusual proline- and glycine-rich domain (also about 130 amino acids) fused onto the CRD; and (c) the Tandem Repeat group (galectins-4, -6, -8, -9, and -12) contains two CRDs.

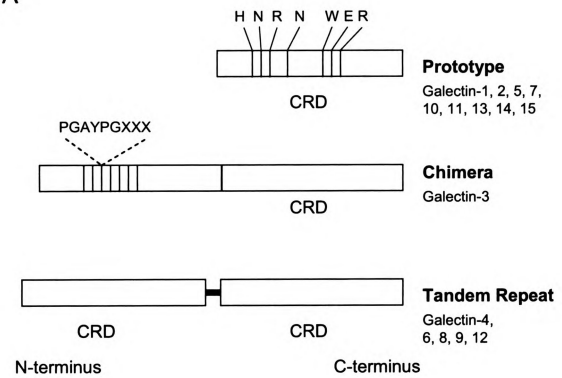
The three-dimensional structures of the CRDs derived from galectins-1, -2, -3, -7 and -10 have been elucidated by X-ray crystallography. They all show a highly conserved tight fold, with two anti-parallel β-pleated sheets forming a sandwich-like structure (5). Amino acid side chains on one of these sheets form the core carbohydrate-

**Figure 1. Schematic diagram illustrating the polypeptide architecture of the galectin family and the backbone folding of a typical carbohydrate recognition domain.**

A) Diagram illustrating the domain organization of each galectin subfamily. CRD, carbohydrate recognition domain. Conserved amino acid residues in the CRD are highlighted. The N-terminal domain of galectin-3, the sole representative of the Chimera group, contains a repeating motif rich in proline and glycine residues. The single letter amino acid code is used; X denotes any amino acid. B) Illustration of the overall folding of the polypeptide backbone of a typical CRD, featuring the two β-pleated sheets, derived from the X-ray crystallographic structure of human galectin-7 (6). Residues (numbering based on the human galectin-7 sequence) involved in binding of saccharide ligands are highlighted.

Figure 1

A


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

B


R-49
N-47
R-74
E-72
H-45
N-62
K-64

binding site. Figure 1B illustrates this typical CRD, derived from the crystal structure of human galectin-7 (6). The highly conserved residues, responsible for saccharide-binding, include H49, N51, R53, N62, W69, E72, and R74 (Figure 1, panel B). The interaction between a galectin CRD and the monosaccharide ligand galactose is actually rather weak ($K_d$ values in the mM range). For most galectins tested, the disaccharide lactose binds with about 100-fold higher affinity than galactose alone (7). Some larger oligosaccharides (e.g., polylactosamine glycans) exhibit even higher affinity than lactose, suggesting that the carbohydrate-binding site extends beyond the core binding site for galactose. The amino acid residues forming this extended binding site are much less conserved among the galectins than those of the core binding site and different galectins show different affinity and specificity for longer oligosaccharides.

Each individual galectin is expressed in some tissue-specific or developmentally regulated fashion. The composite picture derived from these studies highlights three important facts (3): (a) any given organism usually expresses multiple members of the galectin family; (b) different cells within an organism usually contain a different complement of galectins; and (c) almost all cells have at least one galectin. Of the 15 galectins identified so far, eleven have been documented or explicitly stated to be found in the nucleus, as well as in the cytoplasm, of cells (4). On the other hand, nine of the same 15 galectins have been shown to be found outside of cells. Thus, in addition to binding galactose-containing glycoconjugates, some members of the galectin family share another property in terms of their cell biology. They exhibit dual localization, being found in both the intracellular (cytoplasm and nucleus) as well as the extracellular (cell surface and medium) compartments (8). It is generally assumed that all known

galectins are synthesized on cytoplasmic ribosomes, so observations of such proteins in the cytoplasm might not be surprising. However, the finding of any individual galectin both in the nucleus and at the cell surface seems unusual. Following synthesis, there appears to be selective intracellular targeting of specific galectins to subcompartments of the cytosol, to the nucleus, and even to membranes and membrane-bounded vesicles. The mechanism of externalization also appears to be unusual because none of the galectins contains an obvious signal sequence for directing the polypeptide into the classical endomembrane pathway for secretion (9).

**B. Galectin-1**

**1. Structure and chemical properties of galectin-1**

Galectin-1 ($M_r$ ~14 kD) is promiscuously expressed in adult mammalian tissues, including muscle, liver, lung, heart, skin, cells of the immune system, and olfactory neurons. Its spatial and temporal patterns of expression during mouse embryogenesis have also been studied in detail (10). The galectin-1 transcript is first detected on day 4 of development, immediately before implantation. Its expression is restricted to the outer cells of the hatched blastocyst. By day 9, galectin-1 can be found within the myotomic portion of the somites. Later in embryogenesis, most organs express galectin-1 with high levels in mesodermal cells, especially muscle and liver.

The galectin-1 polypeptide consists of a single CRD, which can form non-covalently associated dimers (Figure 1, panel A). It is synthesized on free ribosomes (11). Amino acid sequence analysis showed that the polypeptide, purified from tissue extracts, is acetylated at the N-terminus (12). Both of these observations are more typical of cytosolic proteins than secreted proteins. Indeed, there are numerous reports on the

cytoplasmic localization of galectin-1; these include, for example: (a) myoblasts (13) (14); (b) Chinese hamster ovary cells (15), (c) follicle cells of fetal thyroids (16); and (d) intrahepatic cholangiocarcinoma cells (17).

However, the studies on myoblasts highlight the fact that the cytoplasmic localization can change as a function of differentiation (13, 14). The intracellular staining of the protein decreased as myoblasts fused into myotubes; this is accompanied by its externalization via a novel secretory mechanism involving small evaginations or blebs of the plasma membrane (13). Similarly, the human leukemia cell line K562 expresses galectin-1 in the cytosol. Treatment of K562 cells with erythropoietin induced an erythroid phenotype and led to the externalization of cytosolic galectin-1 (18).

**2. Binding partners and biological activities of galectin-1**

In the cytoplasm of H-Ras(12V)-transformed Rat-1 (EJ) cells, galectin-1 mediates the membrane anchorage of the H-Ras oncogene product, thereby allowing cell transformation (19). Direct interaction between galectin-1 and H-Ras(12V) was demonstrated and this interaction of galectin-1 was not sensitive to lactose inhibition but appeared to be specific for H-Ras(12V) when compared to other Ras isoforms. Moreover, galectin-1 was shown to be essential for the membrane localization of Ras. An inhibitor of Ras membrane anchorage reduced the amount of H-Ras(12V) and galectin-1 found in the plasma membrane. Co-expression of an antisense RNA for galectin-1 and a Green Fluorescence Protein (GFP)-H-Ras(12V) construct resulted in less membrane-associated Ras. Consistent with the known significance of membrane anchorage of Ras to malignant transformation, galectin-1 overexpression resulted in cell transformation, while its antisense RNA inhibited such transformation (19). In BHK

(baby hamster kidney) cells expressing an antisense RNA for galectin-1, there was a reduction of H-Ras(12V) clustering in plasma membrane nonraft microdomains, as well as a decrease in galectin-1 expression. It is proposed that galectin-1 may help stabilize interactions of H-Ras with these nonraft microdomains (20). Thus, it appears that an important interacting partner of galectin-1 in the cytosol is H-Ras(12V) (Table I).

There are also reports that describe galectin-1 both in the cytoplasm and nucleus; these include, for example: (a) Langerhans cells and fibroblasts (21); (b) osteoblasts (22); (c) smooth muscle cells of the respiratory and digestive tracts (23); (d) Sertoli cells of the testis (24, 25); and (e) HeLa cervical carcinoma cells (26). In the latter study, it was also shown that, in the nucleus, a significant portion of the galectin-1 co-localized with the Sm epitopes on core polypeptides of small nuclear ribonucleoproteins (snRNPs) and the serine- arginine-rich (SR) protein SC35 in speckled structures associated with the nuclear matrix (26, 27). Both snRNP core proteins and SC35 have been documented as *bona fide* components of the nuclear machinery for the splicing of pre-messenger RNA (pre-mRNA). Using nuclear extracts derived from HeLa cells, depletion and reconstitution experiments showed that, indeed, galectin-1 (and galectin-3, see below) is a required factor in the splicing of pre-mRNA as assayed in a cell-free system (27). Consistent with this activity, a yeast two-hybrid screen identified the carboxyl-terminal 50 amino acids of the protein Gemin4 (designated hereafter as Gemin4(C50)) as an interacting partner of galectin-1 (28). A direct interaction between galectin-1 and Gemin4(C50) was demonstrated in pull-down assays using the fusion protein containing glutathione S-transferase (GST) and Gemin4(C50) (Table II).

## Table I

Galectin ligands and associated activities in the cytoplasm

| Galectin | Ligand | Associated Activity | Reference |
|---|---|---|---|
| 1 | Ras | Membrane anchorage | (19,20) |
| 3 | Bcl-2 | Apoptosis inhibition | (45, 46) |
| | Synexin | Apoptosis inhibition | (82) |
| | Chrp | ? | (86, 87) |
| | Cytokeratin | ? | (88) |

## Table II

Galectin ligands and associated activities in the nucleus

| Galectin | Ligand | Associated Activity | Reference |
|---|---|---|---|
| 1 | Gemin4 | pre-mRNA splicing | (28) |
| 3 | Gemin4 | pre-mRNA splicing | (28) |
| | TTF-1 | activation | (72) |
| | CBP 70 | ? | (85) |

Gemin4 is found in both the cytoplasm and nucleus, as a member of either the SMN (Survival of Motor Neuron protein) complex or the microRNA particle (29, 30). SMN is encoded by the gene identified for the disease spinal muscular atrophy (31); it was the first identified component of a set of macromolecular complexes containing ~15 polypeptides, including Gemin2 through Gemin7, as well as the core polypeptides of snRNPs (32). The SMN complex plays a role in biogenesis of snRNPs in the cytoplasm, before their entry into the nucleus (33). Nuclear SMN-containing complexes are thought to recycle/resupply snRNPs to the early (H/E) complexes in the spliceosome assembly pathway (34). Nuclear extracts depleted of galectins-1 and -3 assemble pre-mRNAs into H/E complexes, but show no conversion of these complexes into higher-order active splicing structures (35). Thus, the findings that galectin-1 interacts with Gemin4 offer mechanistic insights regarding its role in the splicing pathway by implicating the H/E complex as the locus of action of galectin-1 in spliceosome assembly.

Although the above discussion has highlighted the activities of galectin-1 in the cytosol and nucleus, it should be noted that numerous studies have implicated a host of activities mediated by the protein in the extracellular compartment, including, for example, cell adhesion, establishment and maintenance of synaptic connectivity in olfactory neurons, and induction of apoptosis in T-lymphocytes (for reviews, see (2)). The relationship between any of the intracellular and extracellular activities of galectin-1 (e.g., between nuclear RNA processing and apoptosis induced by exogenously added galectin-1) remains to be explored. In assigning any specific function to galectin-1, it is important to acknowledge that mutant mice with a disrupted galectin-1 gene have been generated (36). These mice were viable and fertile, thus implying that there is no

absolute requirement for the protein in the fundamental activities assigned. Rather, the protein seems to optimize certain processes such as axon guidance in olfactory bulbs (37).

## C. Galectin-3

### 1. Structure and chemical properties of galectin-3

Galectin-3 ($M_r$ ~30 kD) consists of a single polypeptide, whose amino acid sequence suggests that it is a chimera of two distinct domains: a very unusual $NH_2$-terminal domain (ND) fused onto the COOH-terminal CRD (CD) (Figure 1, panel A). The ND is characterized by multiple internal sequence homologies, each of which consists of a 9-residue repeat with a consensus sequence of Pro-Gly-Ala-Tyr-Pro-Gly, followed by three additional amino acids. Thus, the ND is rich in proline and glycine residues. The carbohydrate-binding CD contains a core sequence of about 130 amino acids whose crystal structure revealed a folding pattern typical of a CRD (Figure 1, panel B) (38).

Physico-chemical studies have been carried out on the mouse (39) and hamster (40, 41) homologs of galectin-3, as well as purified ND and CD preparations. These studies have indicated that the two domains of galectin-3 are structurally, as well as functionally, distinct. Differential scanning calorimetry of murine 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 (39). Although this suggested that the two domains are folded independently, more recent nuclear magnetic resonance experiments suggest that portions of the ND may interact with the CD (41). These studies, as well as work on the human homolog (42, 43), clearly show that the CD

bears the carbohydrate-binding activity of the protein. The binding of galectin-3 or its CD to saccharide ligands is accompanied by a conformational change (39), with rearrangement of the backbone loops near the binding site (44).

Within the CD of galectin-3 is the sequence NWGR. This sequence motif is found in the Bcl-2 family of apoptosis repressors, responsible for the homo- and hetero-dimerization of the polypeptides. Indeed, galectin-3 interacts with Bcl-2 and exhibits anti-apoptotic activity (45, 46) (see below). The CD can also self-associate through this NWGR motif and this mode of homodimerization is inhibitable by the saccharide ligands of the lectin (47). A site-directed mutant, with a tryptophan to leucine replacement in the NWGR motif (the W181L mutant), can no longer undergo self-association through the CD. Thus, it appears that the CD is responsible for the oligomerization of galectin-3 in the absence of saccharide ligands.

Several studies have documented that galectin-3 can also self-associate through the ND. Both hemagglutination and positive cooperativity in the binding of galectin-3 to solid phase immunoglobulin E were demonstrated for the full-length polypeptide but not the CD, implying that the ND was critical for self-association in the presence of saccharide ligands (42). Electron microscopic imaging of ND fragments revealed the presence of fibrils formed by intermolecular interactions (41). Cross-linking studies also showed self-association of galectin-3 (42) and the ND (40). Finally, it has been reported that the W181L mutant of galectin-3, which can no longer undergo protein-protein interactions through the NWGR motif in the CD, can still bind to wild-type galectin-3 through interactions of the ND (47).

**2. Expression and localization of galectin-3**

During mouse embryogenesis, galectin-3 has been detected in the trophectoderm of the 4.5-day old blastocyst (48). Although this initial localization overlaps with galectin-1, the spatial patterns of expression for the two proteins diverge thereafter. Between 8.5 and 11.5 gestational days, galectin-3 is observed only in notochord cells. During later stages of development (12 days and beyond), its expression is restricted to the cartilage of vertebrae, ribs, and facial bones, the suprabasal layer of epidermis, the endodermal lining of the esophagus, larynx, and bladder, and in macrophages. In adults, galectin-3 is ubiquitously expressed.

Like galectin-1, the N-terminus of the galectin-3 polypeptide, isolated from cell extracts, is blocked by acetylation (49). The localization of galectin-3 in both the nucleus and the cytoplasm of various cell types has been documented at both light microscopy and ultrastructural levels (50-53). Moreover, the nuclear versus cytoplasmic distribution of the protein can be altered by a number of conditions as seen in the following examples. First, adaptation of murine peritoneal macrophages to *in vitro* culture reduces the nuclear content of galectin-3 (54). It has also been shown that nuclear galectin-3 is elevated in macrophages derived from tumor-bearing hosts, relative to those from normal hosts (55). Second, galectin-3 expression and its intracellular distribution varies along the crypt-to-surface axis of human colonic epithelia. The protein is concentrated in nuclei of differentiated colonic epithelial cells. The progression from normal mucosa to adenoma to carcinoma is characterized by a striking absence of galectin-3 in the nuclei of adenoma and carcinoma cells (56, 57). This general trend of shifting the nuclear localization of galectin-3 in favor of the cytoplasmic compartment during neoplastic progression has

also been reported for tongue, prostate, and follicular and papillary thyroid cancer (58-60).

Third, in fibroblasts, the nuclear versus cytoplasmic distribution of the protein was dependent on the proliferation state of the cells under analysis. In quiescent cultures (serum-starved or density inhibited), galectin-3 was predominantly cytoplasmic; proliferating cultures of the same cells showed intense nuclear staining (61). Parallel nuclear run-off transcription assays and Northern blotting for accumulated mRNA levels showed that galectin-3 is an immediate-early gene, whose activation upon serum stimulation of quiescent fibroblasts does not depend on *de novo* protein synthesis (62). Finally, human diploid fibroblasts have a finite replicative lifespan when subjected to *in vitro* culture. While galectin-3 could be found in both the nucleus and cytoplasm of young, proliferating cells, the protein was predominantly cytoplasmic in senescent human fibroblasts that have lost replicative competence (63, 64).

Galectin-3 shuttles between the nucleus and cytoplasm (65). Gong *et al.* (66) reported that deletion of the first 11 amino acids of galectin-3 resulted in a mutant exhibiting cytoplasmic (and no nuclear) localization. Moreover, when the first 11 amino acids were fused to Green Fluorescent Protein, a mainly nuclear distribution of the reporter was observed. In contrast to these results, Gaudin *et al.* (67) transfected Cos-7 cells with cDNAs encoding mutants of galectin-3 containing N-terminal or internal deletions and showed that nuclear localization does not require the $NH_2$-terminal 100 residues or so. Consistent with conclusions of Gaudin *et al.* (67), more recent studies have identified an ILXT sequence (residues 253-256 of the murine polypeptide) as critical for nuclear localization (68). Nuclear galectin-3 is rapidly and selectively

exported via a leptomycin-inhibitable pathway, suggesting the involvement of the CRM1 export receptor (69). Consistent with this notion, a leucine-rich nuclear export signal has been identified (residues 241-249 of the murine galectin-3 sequence); this nuclear export signal is conserved in the amino acid sequences of galectin-3 from various species (70).

**3. Binding partners and biological activities of galectin-3**

Galectin-3 in the nucleus is associated with the ribonucleoprotein components of the nuclear matrix (51, 71). Under both light microscopy and electron microscopy, treatment of unfixed, permeabilized fibroblasts with ribonuclease A removed the nuclear staining for galectin-3 whereas parallel treatment with deoxyribonuclease failed to yield the same effect (51, 52). When nucleoplasm was subjected to sedimentation in a cesium sulfate gradient, galectin-3 was found in fractions with densities (1.3-1.35 g/ml) matching those reported for the heterogeneous nuclear RNPs (hnRNPs) and snRNPs. On this basis, depletion and reconstitution experiments were carried out to document that galectin-3 was a required factor in the splicing of pre-mRNA in a cell-free assay (35). Although it has been reported that galectin-3 binds to single-stranded DNA in a lactose-independent manner (51, 71), other evidence suggests that the interaction between galectin-3 and pre-mRNA is indirect, most likely mediated via protein complexes containing Gemin4 and SMN. Like galectin-1 (see above), galectin-3 binds to Gemin4(C50) in GST pull-down assays (28). Three lines of evidence converge at the H/E complex to implicate it as the locus of action of galectin-3 in the splicing pathway: (a) nuclear extracts depleted of galectins are arrested in spliceosome assembly at the H/E complex and are inhibited in terms of splicing activity; (b) the addition of the Pro- and Gly-rich ND of galectin-3 to a splicing competent extract results in a dominant negative effect, arresting spliceosome

assembly at the H/E complex and inhibiting splicing activity; and (c) galectin-3, along with galectin-1, is found in SMN-containing complexes which supply functional snRNPs to the H/E complex in the pathway of spliceosome formation. Thus, Gemin4 is an important interacting partner of galectin-3 in the nucleus (Table II).

In the nuclei of papillary thyroid cancer cells, galectin-3 also interacts with thyroid transcription factor 1 (TTF-1) (72). This conclusion is derived from GST pull-down assays which demonstrated a direct interaction between galectin-3 and the homeodomain of TTF-1 (Table II). In addition, gel retardation assays showed that this interaction stimulated the DNA-binding activity of TTF-1. Thus, galectin-3 can up-regulate the transcriptional activity of TTF-1, contributing to the proliferation of the thyroid cells. In a similar vein, galectin-3 can also exert a growth promoting effect through induction of cyclin $D_1$ (73). In the nuclei of human breast epithelial cells, it appears that galectin-3 can enhance or stabilize nuclear protein-DNA complex formation, as revealed by gel retardation assays, at the cAMP-responsive element of the cyclin $D_1$ promoter.

The positive effect of galectin-3 on cell growth has been demonstrated in other systems. First, transfectants of human T lymphoma Jurkat cells expressing galectin-3 were found to increase in cell numbers more rapidly than control transfectants under suboptimal conditions (e.g., low serum) (45). Second, human breast cancer MDA-MB435 cells transfected with anti-sense galectin-3 cDNA suppressed the expression of endogenous galectin-3; these cells exhibited decreased cell proliferation compared to control transfectants (74). Finally, when T-lymphocytes were induced to proliferate by anti-CD3 antibody or IL-2, IL-4, and IL-7, the expression of galectin-3 was elevated.

When specific antisense oligonucleotides were used to suppress the expression of galectin-3, the proliferative response of the lymphocytes to the mitogenic stimuli was decreased (75). These observations could be interpreted either in terms of a growth promoting effect of galectin-3 or in terms of an effect of the protein in protecting cells from death.

Indeed, galectin-3 exhibits anti-apoptotic activity. Jurkat cell transfectants ectopically expressing galectin-3 survived longer than control transfectants when subjected to a variety of apoptosis-inducing agents, including anti-Fas receptor antibody and staurosporine (45). It has also been shown that galectin-3 inhibits apoptosis caused by the loss of cell anchorage (anoikis) (46). The apoptosis inhibiting activity of galectin-3 has been documented in a number of other cell culture systems (46, 76, 77), as well as in cells derived from galectin-3 deficient mice (78). Peritoneal macrophages are susceptible to apoptosis when challenged with inteferon $\gamma$ or lipopolysaccharide. Compared to macrophages derived from wild-type mice, those cells derived from galectin-3 deficient mice died more rapidly. In an independent line of galectin-3 null mutant mice, Colnot *et al.* (79) also observed that there was a reduced number of granulocytes in the peritoneal cavity, compared to wild-type controls, upon injection of thioglycolate broth. It was noted, however, that the granulocytes in the mutant mice did not exhibit an accelerated rate of apoptosis and their uptake by macrophages appeared to be unaffected by the galectin-3 mutation.

A number of studies have provided hints regarding the mechanisms and signaling pathways by which galectin-3 antagonizes apoptosis. First, galectin-3 and the apoptosis repressor Bcl-2 share similarities in features of their polypeptides; both proteins are rich

in proline, glycine, and alanine at the N-terminal region and both contain an NWGR quartet in the C-terminal portion (45). Substitution of glycine to alanine in this NWGR motif abrogates the anti-apoptosis activity (46). Galectin-3 binds to Bcl-2 *in vitro*, mimicking the ability of the Bcl-2 family members to form heterodimers (45). Thus, Bcl-2 is an important ligand of galectin-3 in the cytoplasm (Table I).

Second, Kim *et al.* (80) demonstrated that galectin-3 inhibits anoikis by down-regulating cyclins E and A and up-regulating their inhibitory proteins, $p21^{WAF/Cip1}$ and $p27^{KIP1}$. It has also been reported that galectin-3, $p21^{WAF/Cip1}$, and proliferating cell nuclear antigen (an auxiliary subunit of DNA polymerase δ) were each up-regulated during a distinct period of repair in hepatocytes injured by administration of $CCl_4$ (81). Moreover, galectin-3 was found to be phosphorylated at a tyrosine, suggesting the possibility that it may function as a signaling protein downstream of a tyrosine kinase, as a part of the hepatocytes' mechanism to undergo repair and escape from cell death.

Third, in apoptosis of BT549 human breast carcinoma cells induced by administration of cisplatin, galectin-3 is translocated to mitochondrial membranes, where it prevents mitochrondrial damage, cytochrome c release, and the consequent apoptosome activity (82). This translocation is dependent on synexin, a $Ca^{2+}$- and phospholipid-binding protein. Direct interaction between galectin-3 and synexin was documented by a yeast two-hybrid assay, as well as by GST pull-down assays. Thus, synexin is another significant cytoplasmic ligand of galectin-3 related to the latter's apoptosis inhibiting activity (Table I).

Finally, a proteomic analysis of phagosomes derived from macrophage ingestion of latex beads has identified, along with many hydrolases normally associated with

phago-lysosomes, several proteins related to apoptosis including galectin-3, the 14-3-3 protein, and Alix/AIP-1 (83). This same set of apoptosis-related proteins was also identified when exosomes were analyzed by proteomics (84).

In addition to the ligands related to apoptosis, three other proteins that interact with galectin-3, either in the cytoplasm (Table I) or in the nucleus (Table II), deserve mention. CBP70 was isolated from HL60 cell nuclei on the basis of its binding to glucose/N-acetylglucosamine affinity beads (85). It was found that galectin-3 was co-purified with CBP70 (Table II). The interaction between CBP70 and galectin-3 was disrupted by lactose binding to the latter protein, suggesting that the binding of CBP70 and saccharide ligands to galectin-3 was mutually exclusive. In contrast, it appears that saccharides and the cysteine- and histidine-rich protein (Chrp) can bind to galectin-3 simultaneously (86, 87). Chrp was initially identified in a yeast two-hybrid screen of a murine 3T3 cell cDNA library using galectin-3 as the bait. Direct interaction between galectin-3 and Chrp was confirmed by immunoprecipitation and *in vitro* binding assays. Immunofluorescence analysis revealed that, in 3T3 cells, Chrp was distributed throughout the cytoplasm but was especially concentrated in a concentric ring at the nuclear envelope (86). Thus, while the cytoplasm contained both galectin-3 and Chrp, the latter protein appeared to be strikingly excluded from the nucleus where, in fact, galectin-3 is found prominently. Chrp binds to the CD of galectin-3. Nevertheless, galectin-3, in complex with Chrp, can still bind to carbohydrate-bearing ligands, including laminin (87). Therefore, the CD of galectin-3 can simultaneously accommodate two ligands, carbohydrate and Chrp (Table I).

Most of the ligands listed in Tables I and II interact with galectin-3 via protein-protein, rather than lectin-glycoconjugate, interactions. The two exceptions are cytokeratin (Table I) and CBP70 (Table II). Goletz *et al.* (88) showed that some cytokeratins are glycosylated and that the glycans include structures containing terminal α1-3 linked N-acetylgalactosamine residues, which potentially can serve as high-affinity ligands for galectins. Moreover, *in vitro* binding of galectin-3 to cytokeratins was documented. The binding was inhibitable by glycoconjugates bearing terminal N-acetylgalactosamine and was sensitive to periodate oxidation or α-N-acetylgalactosaminidase treatment of the cytokeratins. On this basis, cytokeratins appear to be a natural carbohydrate ligand for cytoplasmic galectin-3 (Table I).

**4. Phosphorylation of galectin-3**

In mouse 3T3 fibroblasts, galectin-3 exists in two isoelectric variants: (a) a non-phosphorylated species corresponding to the native polypeptide (pI ~8.7); and (b) a phos-phorylated derivative (pI ~8.2). The nonphosphorylated form is found exclusively in the nucleus while phosphorylated galectin-3 can be found both in the nucleus and cytoplasm (89). Mass spectrometric analysis of the canine homolog of galectin-3 identified serine-6 as the major site of phosphorylation *in vivo*, with a minor site at serine-12 (90) Human galectin-3 was phosphorylated *in vitro* by casein kinase I; phosphorylation significantly reduced its binding to carbohydrate ligands while dephosphorylation fully restored the saccharide-binding activity (91). In addition, cDNAs containing site-directed mutants of galectin-3 at serine-6 were generated and used to transfect human BT549 breast carcinoma cells. Both serine-6 mutants (S6A and S6E) failed to protect cells from cisplatin-induced apoptosis (92). The mutants also failed to protect cells from anoikis

with $G_1$ arrest when the BT549 cells were cultured in suspension. These results suggest that phosphorylation of galectin-3 regulates its anti-apoptotic activity (92). The relationship of serine-6 phosphorylation and the tyrosine phosphorylation mentioned above (in association with hepatocyte survival from $CCl_4$ intoxication (81)) remains to be elucidated.

As was noted with galectin-1, it is important to acknowledge that mutant mice have been generated in which the galectin-3 gene or both the galectin-1 and galectin-3 genes were disrupted (78, 79, 93). The strains of mutant mice were viable and fertile, thus implying that there is no absolute requirement for these proteins in the activities assigned to them.

## II. Pre-mRNA Splicing

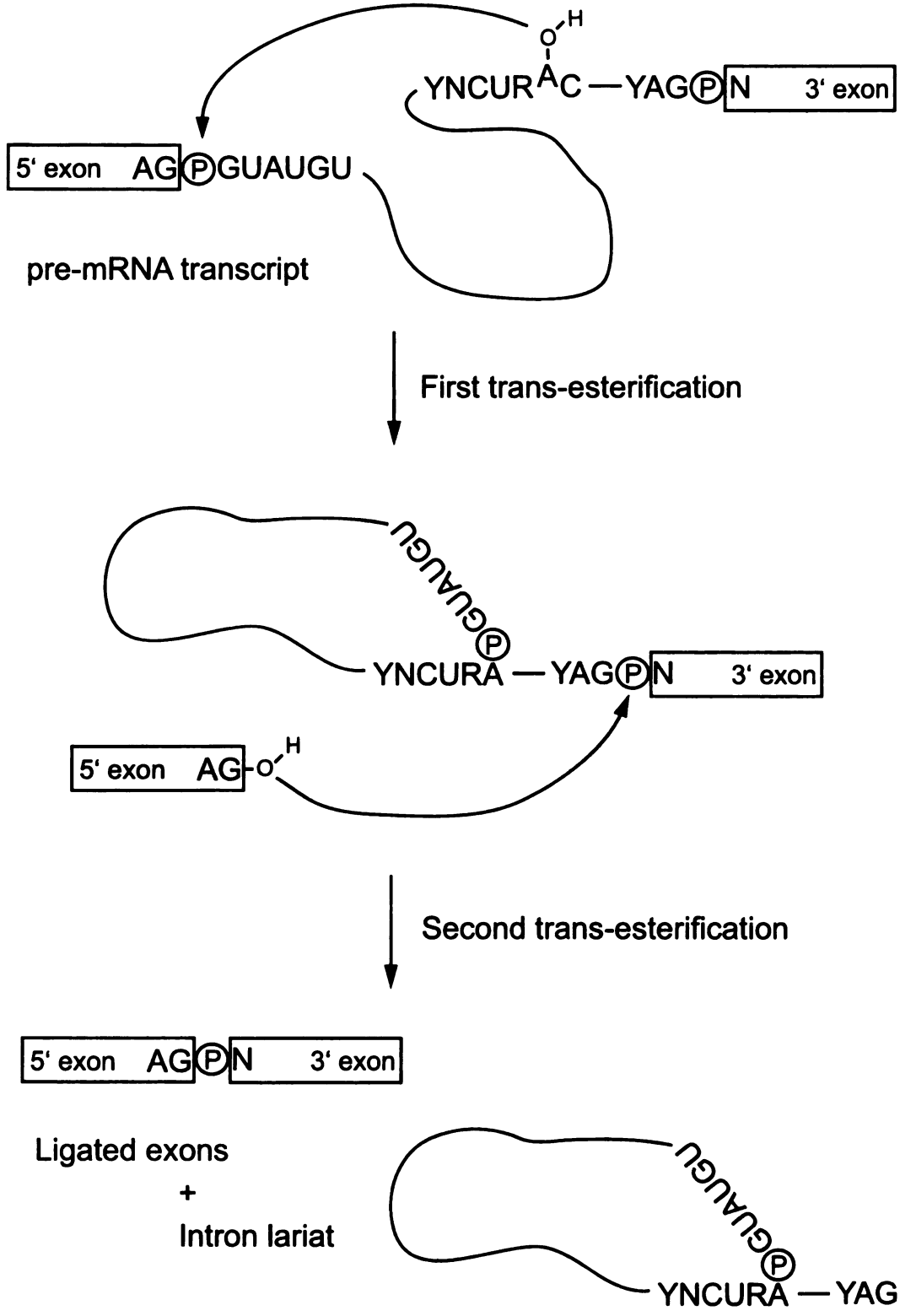
### A. Chemical steps in pre-mRNA splicing

Pre-mRNA splicing of coding sequences (exons) is an essential post-transcriptional modification for most eukaryotic transcripts (94). The process occurs in the nucleus and involves removal of noncoding sequences, termed intervening sequences or introns, present in the transcript. Accurate splice site selection is essential for production of functional proteins and is effected by four consensus sequences. In mammals, the 5' splice site is characterized by the nucleotide sequence AG/GUAUGU (where / is the exon/intron junction). The 3' splice site is marked by the sequence YAG/N (where Y=pyrimidine and N=any nucleotide) and is typically preceded by a polypyrimidine tract. The branch point is located 18-40 nucleotides upsteam of the 3' splice site and is characterized by the sequence YNCURAC (where R=purine, and the site of branch formation is shown in bold) (95). However, sequences which do not match the consensus may still be spliced, and many splicing consensus sequences are not used as splice sites. This causes alternative splicing, in which exons may be left out the the mRNA (skipped). Certain introns may be included in the mRNA, and 5' and 3' splice sites can be shifted to change exon length (96, 97).

The mechanism of splicing includes two ATP-independent trans-esterification reactions (98-100). The first involves a nucleophilic attack on the phosphodiester bond at the 3' end of the 5' exon by a 2' hydroxyl group of the adenosine residue at the intron branch point (Figure 2). This reaction generates the splicing intermediates consisting of the free 5' exon and the 3' exon attached to the intron, which exhibits a closed loop formation termed the lariat. The second step consists of a nucleophilic attack on the

**Figure 2. Schematic diagram illustrating the chemical steps in pre-mRNA splicing.** The splicing process consists of two trans-esterification reactions. The first results in formation of free 5' exon and and 3' exon-intron lariat. The second results in formation of ligated exons and the intron lariat. Consensus sequences are indicated using the single-letter nucleotide code where Y=pyrimidine and N=any nucleotide. P=phosphodiester bond.

Figure 2


YNCURAC—YAGPN 3' exon
5' exon AGPGUAUGU
pre-mRNA transcript
First trans-esterification
GUAUGP
YNCURA—YAGPN 3' exon
5' exon AG
Second trans-esterification
5' exon AGPN 3' exon
Ligated exons
+
Intron lariat
GUAUGP
YNCURA—YAG

phosphodiester bond at the 5' end of the 3' intron by the 3' hydroxyl group at the end of the free 5' exon. This reaction generates the splicing products consisting of the ligated exons (mRNA) and the free intron lariat (101) (Figure 2).

**B. Proteins involved in pre-mRNA splicing**

In metazoans, this chemical mechanism is mediated by a host of splicing factor proteins which are involved in orientation, stabilization, and disruption of RNA within a large complex designated the spliceosome (102, 103). However, two lines of evidence suggest that RNA plays a substantial role in the splicing process. First, a group of introns (termed group II) present in organelles of lower eukaryotes and bacteria can undergo splicing in the absence of protein components and ATP *in vitro*. The chemical mechanism of this process is identical to nuclear pre-mRNA splicing (104, 105). Second, mutagenesis and UV cross-linking studies have shown that the base-pairing interactions between small nuclear RNA (snRNA) and pre-mRNA are necessary for specific stages of spliceosome assembly during nuclear splicing (106, 107).

The spliceosome includes various groups of well-characterized proteins which function during splicing (reviewed in (107-110). hnRNPs associate *in vivo* with nascent RNA polymerase II transcripts (111, 112). This binding is independent of temperature, ATP, and functional splice sites, suggesting that the interactions are not splicing-specific. Direct roles for hnRNP proteins during the splicing process have not been determined, but certain members of the family appear to function in alternative splice-site selection (113-118). The SR proteins are characterized by an RNA recognition motif (RRM) within the amino terminal domain, a glycine-rich "hinge" region, and a carboxyl terminal domain enriched in arginine and serine (119). Members of this family include SC35,

alternate splicing factor/splicing factor 2 (ASF/SF2), U2 associated factor (U2AF)35 and 65, polypyrimidine-tract binding protein (PTB), and PTB-associated splicing factor (PSF) (120-129). This family is among the first proteins to interact with pre-mRNA and recruit additional splicing factors to the assembling spliceosome (130, 131). The snRNPs are essential components of the spliceosome and have been designated U1, U2, U4, U5, and U6 (132-136). Each snRNP is composed of a snRNA (137), a set of snRNP core proteins (designated Sm proteins), and snRNP-specific proteins. Each snRNA contains a uridine-rich sequence which serves as the binding site for the Sm core proteins (138, 139) and additional sequence elements which interact with pre-mRNA. There are eight Sm proteins: B' (29 kD), B (28 kD), D1 (16 kD), D2 (16.5 kD), D3 (18 kD), E (12 kD), F (11 kD), and G (9 kD). U1, U2, U4, and U5 snRNAs are trancribed by RNA polymerase II and acquire a monomethyl cap. They are subsequently exported to the cytoplasm (140) where they bind the Sm core proteins. This allows trimethylation of the guanosine cap (141), and together the Sm core proteins and methylated cap form a nuclear localization signal which allows transport of the snRNPs into the nucleus (142-144) for association with the spliceosome. U6 snRNA is transcribed by RNA polymerase III (145) and remains in the nucleus (146) where it associates with eight Sm-like (LSm) proteins (147) before participation in pre-mRNA splicing.

Various splicing inhibitors and enhancers can act to control spliceosome formation and splice site selection to generate desired alternatively-spliced transcripts. For example, a yeast homolog of the human U5-220 kD protein, called Prp8, is capable of stabilizing the U4/U6 association to prevent formation of the catalytically active spliceosome. In contrast, prp28, a yeast homolog of human U5-100 kD protein,

destabilizes the U4/U6 complex in conjunction with Brr2 (U5-200 kD), thus acting as a stimulator of splicing activity (148, 149). There also exist sequence motifs designated exonic splicing enhancers (ESEs), which have been shown to affect splice site selection in conjunction with trans-acting proteins (150).

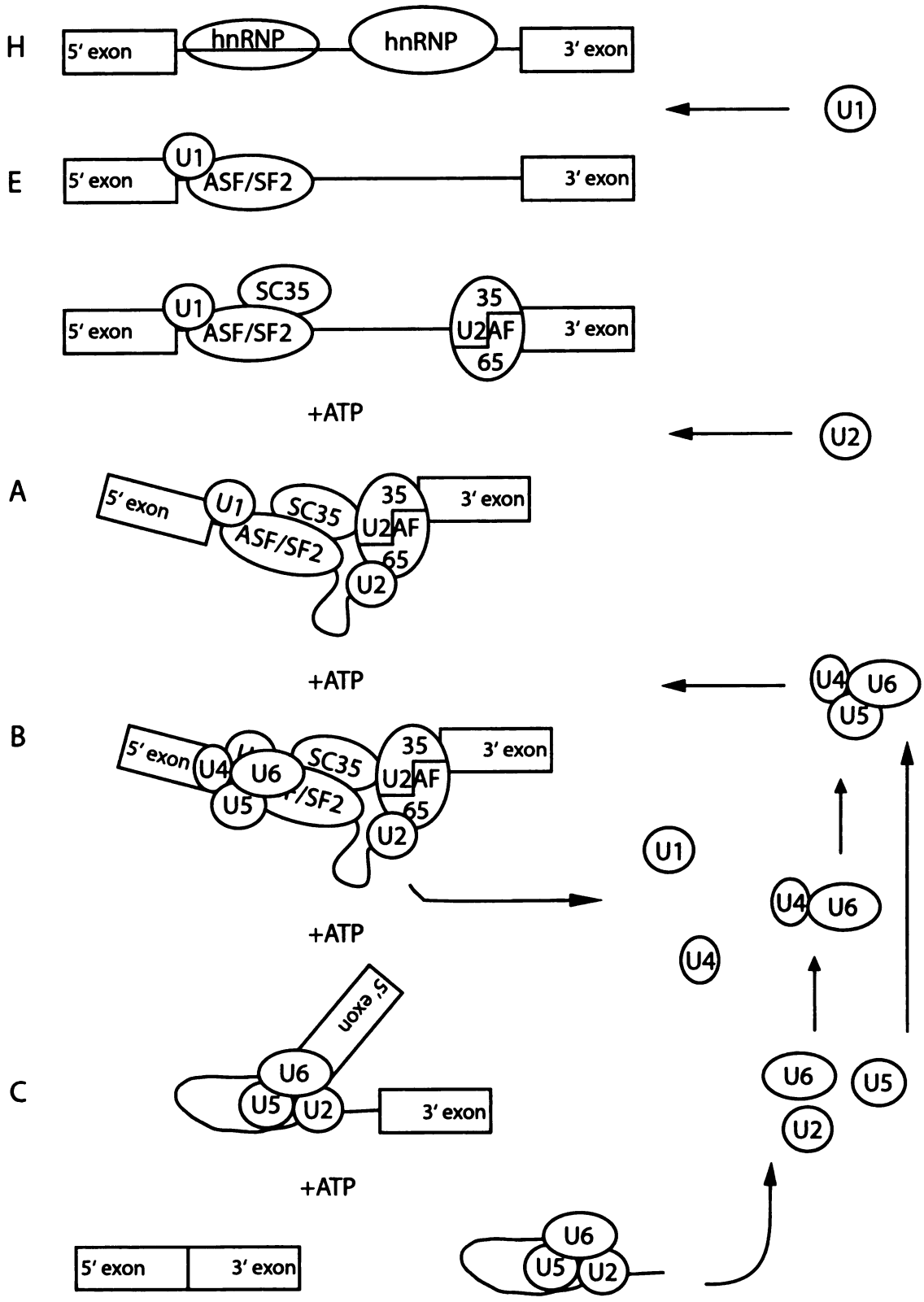
The combined use of tandem mass spectrometry with automated database searches has permitted a rapid identification of a large number of peptides assembled on pre-mRNA (151-154). These studies have lead to the discovery of many proteins which were not expected to be found in the spliceosome. Of particular interest is the discovery of proteins involved in other gene expression steps including RNA processing, transcription and translation. Neither galectin-1 nor galectin-3 has been identified in these studies. One possible reason is that galectin-1 and -3 dissociate from the spliceosome during the purification procedures used. They may also be present at sub-stoicheometric amounts, making their identification within a complex sample more difficult. The problem may also be the small size of the proteins (14 and 28 kD), which limits the number of tryptic fragments for mass spectrometric analysis. Indeed, Rappsilber *et al.* (153) speculate that the small size of known splicing factors Lsm5 (9.8 kD), U5 snRNP 15 kD peptide (16.7 kD), and U-snRNP-associated cyclophilin (19.2 kD), prevented identification of each protein in their study. Mass spectrometers can be programmed with an inclusion list, which designates specific peptides to be subjected to MS/MS analysis. Use of such a list containing the theoretical tryptic peptide sequences of galectin-1 and -3 would aid in their identification, if present, during future mass spectrometric analyses.

**C. Spliceosome assembly**

In order for these proteins to effect splicing of pre-mRNA, they must be assembled stepwise into the active spliceosome (155). The first complex to form on pre-mRNA is designated the H complex and is comprised of hnRNP proteins bound to the RNA in an ATP-, temperature-, and splice site-independent manner (Figure 3). Formation of the early (E) complex is the next step in spliceosome formation and is marked by ATP-independent (156) binding of U1 snRNP to the pre-mRNA. This interaction is mediated through base pairing bewteen the 5' splice site consensus sequence and a highly conserved sequence at the 5' end of the U1 snRNA (157, 158). The interaction is stabilized by ASF/SF2 which may act to recruit U1 to the 5' splice site (159-162). The next addition consists of the heterodimer U2AF65-U2AF35. The 65 kD subunit binds the polypyrimidine tract (116) and the 35 kD subunit binds the 3' splice site (163). U2AF35 brings both the 5' and 3' splice sites into proximity by interacting with ASF/SF2 via the bridging protein SC35 (164). Formation of the A complex is characterized by recruitment of U2 snRNP to the branch point by U2AF. Upon binding, U2 snRNA hybridizes with the branch site (165). This requires ATP and causes the bulging of the branch point adenosine, helping to set up its nucleophilic attack of the 5' splice site (166). Formation of B complex is marked by association of the U4/U5/U6 tri-snRNP particle to the A complex. The tri-snRNP is formed independently of the spliceosome and consists of U4-U6 snRNA base pairing interactions and protein-protein interactions between U4/U6 and U5. Association of the tri-snRNP is accompanied by significant rearangement of the spliceosome which requires ATP and DEXD/H box helicase proteins and leads to the active, or catalytic (C) complex (167). First, U4 is released from the spliceosome; next, U1 snRNA association with the 5' splice site is

**Figure 3. Schematic diagram illustrating the complexes formed during stepwise assembly of the spliceosome.**
H complex is the first to form on nascent pre-mRNA and is comprised of hnRNP proteins. Formation of E complex is marked by binding of U1 snRNP to the pre-mRNA. Binding of U2 snRNP indicates formation of A complex. B complex forms upon association of the U4/U5/U6 tri-snRNP. C complex is the catalytically-active spliceosome and includes release of U1 and U4 snRNP. Following the splicing reaction the intron lariat is released and the snRNPs are recycled for use in additional rounds of splicing.

Figure 3


H
5' exon
hnRNP
hnRNP
3' exon
U1
E
5' exon
U1
ASF/SF2
3' exon
SC35
35
U2AF
65
+ATP
U2
A
B
C
U4
U6
U5
U2
5' exon
3' exon

replaced by hybridization of the 5' end of the U6 snRNA to the 5' splice site. Finally, U1 and U4 are released from the spliceosome prior to catalysis of the trans-esterification reactions (148, 149).

## III. Transcription Factor II-I

### A. Initial Identification as a general transcription factor

TFII-I was originally discovered during attempts to identify factors binding to pyrimidine-rich initiator elements within basal promoters. Fractionation of HeLa cell nuclear extracts and tracking initiator-binding and initiator-dependent transcription activities resulted in the purification of a $M_r$ ~120 kD band which was designated TFII-I (168). Although the role for TFII-I in transcription has been carefully studied, a growing body of literature indicates that pre-mRNA splicing occurs co-transcriptionally. Therefore it is of interest to further investigate its role in pre-mRNA splicing. TFII-I was cloned by multiple independent groups at virtually the same time, resulting in the additional designations SPIN (serum response factor-phox1 interacting protein) (169) and BAP-135 (Bruton's tyrosine kinase-associated protein of 135 kD) (170). Four distinct lines of evidence suggested that SPIN was identical to TFII-I. First, an amino acid sequence comparison showed that the two proteins were identical. Second, SPIN was capable of binding the identical AdML initiator sequence as TFII-I. Third, antiserum raised against a peptide of TFII-I reacted with HPLC fractions containing SPIN activity in western blot analysis. Fourth, this same antiserum abolished SPIN-DNA interaction (169). BAP-135 was determined to be identical to TFII-I by amino acid sequence alone (171).

### B. Structure and chemical properties of TFII-I

The open reading frame for the initially-cloned TFII-I transcript predicted a protein of 957 amino acids and molecular mass of 107.9 kD. The discrepancy between this calculated molecular mass and the observed electrophoretic mobility may be due to

an abundance of acidic residues in the amino terminal portion of the protein. TFII-I contains six iterations of a 95-residue motif (170) later designated an I-repeat domain (172-178). Each motif codes for a putative helix-loop-helix domain with a basic region comprising amino acids 301-306, preceding the second of these repeats (178). In addition to these regions, TFII-I contains two motifs similar to the Src autophosphorylation site, one within amino acids 244-248, and the other within 273-277 (170)). A consensus mitogen-activated protein kinase (MAPK) interaction domain (D box) is located within amino acids 282-293, and two consensus MAPK substrate sites at 627 and 633 (179, 180). Amino acids 23-44 encode a putative leucine zipper (176, 181) and 277-304 encode a functional nuclear localization signal (182). To date four alternatively spliced isoforms of TFII-I have been identified (176): α (977 amino acids), β (978 amino acids), and γ (998 amino acids) (182). Each isoform contains all six helix loop helix motifs, the putative leucine zipper, the nuclear localization signal, the D box, and the MAPK substrate motifs. TFII-I was detected in spleen, thymus, prostate, testes, uterus, intestine, peripheral blood leukocytes, brain, muscle, liver, kidney, lung, and pancreas tissues (170, 176) by northern hybridization.

**C. Binding partners and activities of TFII-I**

Accurate transcription initiation depends on core promoter elements including a TATA-box, the initiator element, and the downstream promoter element (183-187). A given promoter can contain these elements in combination or individually (188) to direct preinitiation complex formation through interaction with a variety of transcription factors (185, 186, 189)

Electrophoretic gel mobility shifts demonstrated that TFII-I bound specifically to the initiator sequences of the AdML, HIV-1, TdT, and T cell receptor variable region-derived (Vβ) promoters and to the upstream binding site (E box) for the helix-loop-helix activator protein USF (also known as upstream stimulatory factor) *in vitro* (168, 190) and *in vivo* (168, 178, 191).

In addition to its DNA binding activity, TFII-I was also found to have transcriptional activity. Addition of TFII-I or TFII-A to a reaction containing RNA polymerase II and TFIIB, TFIID, and TFIIE/F, activated transcription activity, indicating that TFII-I was indeed a transcription factor (168). *In vivo* studies were carried out using transient transfection of cells with TFII-I and the AdML core promoter fused to the chloramphenicol acetyl transferase (CAT) reporter. Cells co-transfected with plasmids containing TFII-I and the AdML-CAT reporter gene exhibited more CAT activity than cells co-transfected with empty vector and AdML-CAT, further supporting the role of TFII-I as a transcription factor (178).

TFII-I was found to bind USF to exert a synergistic effect on both DNA binding and transcription activity. Gel shift experiments demonstrated an interaction between TFII-I and USF which resulted in enhanced DNA binding by USF *in vitro* (168). In addition, cells co-transfected with AdML-CAT and both TFII-I and USF exhibited significantly more CAT activity than cells co-transfected with AdML-CAT and either TFII-I or USF alone (178). Interaction between TFII-I and Myc was also observed. E-box or initiator probes were used in gel mobility shift assays. Myc alone did not bind to the either probe, TFII-I formed a complex on each probe, and TFII-I and Myc together formed a distinct complex on each probe. Interestingly, when Myc and TFII-I were

added to transcription assays containing RNA polymerase II and TFIIB, TFIID, and TFIIE/F, transcription activity was inhibited. However, when the same system was supplemented with Myc and TFIIA, transcription activity was not affected. Therefore, it appears that Myc is able to inhibit transcription by interaction with TFII-I at initiator elements (192).

An independent line of investigation also showed the transcriptional activity of TFII-I. Phox1 was identified as an interacting partner which enhanced the rate of binding and dissociation of serum response factor (SRF) to the serum response element (SRE) (193). Fractionation of HeLa nuclear extract yielded a protein complex, containing Phox1, SRF, and SPIN (TFII-I), that bound to the SRE in the c-fos promoter. GST-Phox1 and GST-SRF pulled down SPIN protein and DNA binding activity from nuclear extract while GST-phospholipase C-γ did not, indicating binding activity between SPIN/TFII-I and Phox1 and SRF took place independent of DNA. DNase protection assays showed that SPIN/TFII-I protected two distinct regions of the *c-fos* promoter. These regions corresponded to the SRE and the upstream c-sis/platelet-derived growth factor (PDGF)-inducible element (SIE). A plasmid containing the *c-fos* basal promoter fused to the CAT reporter gene was used in *in vivo* assays to test the ability of SPIN/TFII-I to cooperate with Phox1 during serum responsive transcription. Cells co-transfected with *c-fos*-CAT as well as GST-SPIN/TFII-I and Phox1 exhibited significantly higher CAT activity than did cells transfected with c-fos-CAT and either GST-SPIN/TFII-I or Phox1 alone, corroborating the findings of Roy *et al.* in a distinct system (169).

Additional studies confirmed the requirement for the SIE and SRE elements, isolated a complex including TFII-I, SRF, signal transducer and activator of transcription-1 (STAT1), and STAT3, and demonstrated a synergistic interaction between TFII-I and the STAT transcription factors during transcription activation (194). TFII-I has also been shown to interact with the transcription activator ATF6 and the endoplasmic reticulum stress response element (ERSE) (195). This element is present in the promoters of glucose-regulated protein (Grp) genes and is activated upon $Ca^{2+}$ depletion or glycosylation blocking (196, 197). Gel mobility shift assays with ERSE sequence probe were used to identify TFII-I as an interactor. Anti-TFII-I antibodies inhibited this interaction. Cells co-transfected with ATF6 and TFII-I were used in immunoprecipitation assays and demonstrated that ATF6 and TFII-I were both present in a complex, but whether they bind directly to one another is not clear. In addition, it has been shown that ER stress induces a marked increase in TFII-I phosphorylation and binding to the Grp promoter *in vivo* (198).

Interestingly, interaction with the protein inhibitor of activated STATs (PIAS) was also revealed using the yeast two-hybrid assay. Co-expression of PIASxβ with TFII-I resulted in augmentation of transcription activity (199). In an extension of this study, TFII-I was found to interact with histone deacetylase 3 (HDAC3). Co-expression caused inhibition of TFII-I transcription which could be overcome with increased expression of TFII-I (200).

The DNA binding activity of TFII-I was determined to reside within an amino-terminal region of the protein. Thrombin digestion separated the protein into two distinct domains. The 70 kD fragment (p70) corresponding to the amino terminal 677 amino

acids (containing 4 of the 6 I-repeats) exhibited DNA binding activity in gel shift assays while the 43 kD fragment (p43) did not. p70 was unable to activate transcription in transfection assays; however, fusion of the GAL4 activation domain to p70 rescued its ability to activate transcription. Transfection of cells with p43 fused to the GAL4 DNA binding domain failed to activate transcription, indicating that p43 is necessary but not sufficient for TFII-I's activation function (201). Deletion analysis suggests that the basic region, consisting of amino acids 301-306 within p70, is required for DNA binding (202).

The alternatively spliced forms of TFII-I exhibit similar DNA binding and localization characteristics individually. However, co-immunoprecipitation and GST-pull-down experiments indicate that the various isoforms exhibit homomeric and heteromeric interactions with themselves and that these complexes exhibit differential transcriptional activation. Co-transfection of cells with Vβ promoter-luciferase reporter plasmid and a single isoform of TFII-I exhibited less luciferase activity than cells co-transfected with the promoter-reporter and multiple isoforms of TFII-I. However, the effect was reversed in the case of signal-responsive transcription. In response to epidermal growth factor, cells co-transfected with c-fos promoter-luciferase reporter plasmid and a single isoform of TFII-I exhibited more luciferase activity than cells co-transfected with the promoter-reporter and multiple isoforms of TFII-I (182).

Two putative nuclear localization signals (NLS1 and NLS2) were found by computer analysis of the TFII-I sequence. The NLS residing within residues 278-284 (NLS1) was determined to be the sole functional NLS on the basis of localization studies using TFII-I-GFP fusion constructs to transfect cells. TFII-I-GFP and TFII-IΔNLS2-GFP exhibited nuclear localization while TFII-IΔNLS1 exhibited cytoplasmic localization.

GFP alone was present in the nucleus and the cytoplasm. Both homomeric and heteromeric complexes of TFII-I isoforms demonstrate preferential nuclear localization. When cells were co-transfected with TFII-IΔNLS1-GFP and each of the TFII-I isoforms (lacking GFP), fluorescence was visualized primarily in the nucleus, with some signal remaining in the cytoplasm. TFII-IΔNLS1-GFP alone exhibited cytoplasmic localization exclusively (182).

It has been proposed that the leucine zipper acts as a primary site of homomeric interaction while the I-repeats function as secondary sites of interaction (202). A deletion mutant lacking the amino terminal 90 amino acids (and therefore, the leucine zipper motif), designated TFII-IΔN90, fails to engage in homomeric interaction, but allows heteromeric interaction and homomeric interaction between TFII-IΔN90 proteins (202). Further support for this conclusion arose from pull-downs using the amino terminal 90 amino acids fused to GST. GST-N90 bound to multiple TFII-I isoforms, but failed to bind to TFII-IΔN90. In this same study, point mutations within the putative leucine zipper sequence failed to abolish the TFII-I-Bruton's tyrosine kinase (Btk) interaction, suggesting that regions outside this specific motif are required this binding (203).

**D. Phosphorylation of TFII-I**

Btk is a tyrosine kinase expressed primarily in B cell lineages. It belongs to a family of Src-like kinases called the Tec family, characterized by a pleckstrin homology (PH) domain (204, 205). The interaction between Btk and TFII-I was discovered during an effort to identify downstream targets for Btk. Lysates of RAMOS cells were subjected to immunoprecipitation by antibodies raised against Btk. Sequence analysis of two bands co-precipitated with Btk showed identity to TFII-I. GST pull-downs with truncated

forms of Btk demonstrated a requirement for the PH domain for TFII-I binding. In addition, it was found that association of Btk and TFII-I occurs *in vivo* and that TFII-I tyrosine phosphorylation is a result of Btk activation by B cell receptor engagement (170). Not only did activation of RAMOS cells stimulate phosphorylation, but also release of TFII-I from Btk. Anti-Btk immunoprecipitation from activated or resting RAMOS cell extracts followed by western blotting with anti-TFII-I showed a marked decrease in the amount of TFII-I co-immunoprecipitated from the activated cells while equivalent amounts of Btk were precipitated from both groups (171). Western blotting analysis shows a marked increase in nuclear localization of total and phosphorylated TFII-I concomitant with release from Btk (171). Site-directed mutagenesis and phosphopeptide mapping were used to locate sites of TFII-I phosphorylation by Btk *in vitro*. These sites matched the major phosphorylation sites *in vivo*, and consisted of Y248, Y357, and Y462 (206). Phosphorylation of TFII-I is not required for DNA binding activity (207), but is necessary for activation of transcription (171, 207). Purified, dephosphorylated TFII-I was found to bind Vβ initiator sequence nearly as efficiently as purified phosphorylated TFII-I. However, addition of dephosphorylated TFII-I to an *in vitro* transcription reaction depleted of TFII-I, failed to reconstitute transcription activity. Conversely, addition of phosphorylated TFII-I did reconstitute transcription activity (207). Individual point mutations of TFII-I Y248F, Y357F, or Y462F, impaired transcription activation at the *c-fos* promoter *in vivo*, supporting the conclusion of the *in vitro* study (206).

TFII-I was also found to be phosphorylated on tyrosines 248 and 611 by Src. This phosphorylation was particularly pronounced following activation of cells by

epidermal growth factor (EGF). As in the case with Btk-induced phosphorylation, TFII-I localization shifted from primarily cytoplasmic to primarily nuclear upon activation of NIH 3T3 cells with PDGF. This shift in localization was abolished in the presence of a Src-specific inhibitor, but persisted in the presence of an inactive inhibitor analog. Interestingly, this nuclear localization correlated with TFII-I-dependent transcription activation. Cells containing a stably integrated HA-tagged c-fos gene driven by the SRE promoter were immunostained for localization of TFII-I and presence of HA. Following stimulation by PDGF, TFII-I localized to the nucleus concomitant with HA appearance. Introduction of antibodies against TFII-I or Src abolished both nuclear localization and HA production while pre-immune serum exerted no effect (208). Similarly, phosphorylation of Y248 by Src during ER stress results in nuclear localization and activation of transcription from the Grp78 promoter (198).

In addition to Btk and Src, and ERK, JAK2 (Janus kinase 2) has been implicated in phosphorylation of TFII-I tyrosine 248 and 277. This interaction was initially discovered by treatment of fibroblasts with an inhibitor of JAK2 which abolished *c-fos* promoter activation by TFII-I. Remarkably, under these conditions, a Src-specific inhibitor showed no effect. *In vitro* phosphorylation experiments were carried out on wild type TFII-I and on TFII-I containing mutations in two consensus tyrosine phosphorylation sites. The result of these experiments indicated that both tyrosine 248 and 277 could be phosphorylated by JAK2, but that the tyrosine 248 was a more significant substrate site. *In vivo* experiments showed that mutation of tyrosine 248 diminished phosphorylation, but did not abolish it, suggesting the presence of additional tyrosine phosphorylation (171, 207). Interestingly, JAK2 can regulate a TFII-I-ERK

(extracellular signal-regulated kinase) interaction (described below). Western blots showed a significant reduction in the amount of ERK bound to TFII-I in cells co-transfected with GST-TFII-I, ERK, and a dominant negative JAK2 mutant, as compared to cells transfected with wild type JAK2. This led to additional investigation that indicated tyrosine 248 is required for interaction with ERK. A Y248F mutation in TFII-I completely abolished complex formation between TFII-I and ERK. As expected, this mutation also substantially impaired phosphorylation of TFII-I by ERK and was unable to activate the *c-fos* promoter upon serum activation in co-transfection assays (209).

TFII-I was shown to be a substrate for ERK. TFII-I contains a consensus MAPK interaction domain within residues 282-293 with significant homology to the Elk-1 kinase interaction domain bound by ERK. This prompted GST pull-downs from cells co-transfected with GST-TFII-I and HA-tagged ERK. GST-TFII-I pulled down ERK while GST-TFII-I-L289A did not. The L289A mutation also abolished transcription activation of the *c-fos* promoter in transfection assays. Mutation of both serines within the MAPK substrate sequence (amino acids 627-634) greatly reduced phosphorylation by ERK, suggesting that S627 and S633 are the major MAPK phosphorylation sites on TFII-I (179).

A second serine kinase, cGMP-dependent protein kinase Iβ (G-kinase Iβ), has also been shown to interact directly with TFII-I. GST-pull-downs using truncated mutants of each protein revealed that the association occurs through the amino terminal 110 amino acids of G-kinase Iβ and the fourth I-repeat domain of TFII-I. G-kinase Iβ phosphorylates S371 and S743, each found within the I-repeat region of TFII-I. Two distinct functional consequences of this interaction were discovered using co-transfection

experiments. First, G-kinase Iβ enhanced TFII-I-dependent transactivation of an SRE-controlled reporter in the presence of wild type TFII-I, but not in the presence of TFII-I (S371A/S743A). Second, G-kinase Iβ transcriptional activation of a fos promoter was synergistically enhanced by TFII-I (210).

**E. Transcription factors involved in pre-mRNA splicing**

Pre-mRNA splicing appears to occur both co-transcriptionally and post-transcriptionally (reviewed in 211-215). The former phenomenon was first suggested on the basis of electron micrographs of actively-transcribed *Drosophila* genes which showed nascent pre-mRNA shortening due to intron removal (216, 217). Two lines of circumstantial evidence have supported this suggestion. First, use of different promoters by RNA polymerase II results in different splicing patterns for the same transcript (218, 219). Second, reduction in speed of transcription (220), or insertion of RNA polymerase II pause sites (221, 222), causes differential splicing of transcripts.

In addition to these transcription-related effects on splicing products, a number of proteins have been identified which implicate a connection between transcription and pre-mRNA splicing. Here it is useful to delineate these proteins into two distinct groups. The first consists of transcription factors discovered in association with the spliceosome or bound to splicing factors within transcription complexes, but whose effect on the splicing process is as yet unknown. This group includes SR-like RNA polymerase II carboxyl-terminal domain (CTD) associated factors (SCAFs) (223), WT1 (224), CA150 (225, 226), and p54nrb (227). The SCAF proteins were discovered during the initial characterization of the CTD's role in splicing using a yeast two hybrid assay with the distal CTD repeats as bait. Four of the interacting proteins had sequences similar to SR

proteins, resulting in the designation SCAFs. WT1 is a transcription factor which has been shown to activate or repress GC-rich promoters, depending on context (228). Genetic defects within the WT1 gene are associated with Wilms' tumor (from which the name WT1 is derived), Denys Drash, and Frasier syndromes (229, 230). WT1 exhibits a specked localization within the nucleus and co-localizes with coiled bodies and splicing factors (231). This observation stimulated its use in a yeast two-hybrid assay which revealed an interaction with U2AF65. In addition, it was found assembled on splicing-competent pre-mRNA with Sm core protein B" (224). CA150 has been shown to act as an RNA polymerase II repressor by inhibition of transcript elongation (232, 233). Interaction between CA150 and splicing factor 1 (SF1) was discovered by Far-Western blotting. Interestingly, SF1 tethered to a Gal4 DNA binding domain caused repression of transcription from a promoter containing Gal4 binding sites (234). This raises the possibility that recruitment of SF1 by CA150 can result in repression of transcription (225). Further studies have shown that CA150 exhibits a speckled localization in the nucleus and co-localizes with splicing factors Sm core proteins, U2AF65, and ASF/SF (226). p54nrb was shown to interact with the 5' splice site of pre-mRNA using protein-RNA crosslinking assays. These assays in fact revealed the presence of two complexes at the 5' splice site: one containing RNA polymerase II, p54nrb, U1 snRNP, and PSF, and the other containing each of these proteins with the addition of U2 snRNP (227). This set of interactions between proteins involved in splicing and proteins involved in transcription compliments the initial evidence for co-transcriptional pre-mRNA splicing, but is still largely circumstantial. More concrete evidence has come to light with use of experiments designed to indicate how transcriptional proteins affect splicing.

This second group of proteins consists of transcription factors and splicing factors which interact and have been shown to exert an effect on pre-mRNA splicing during this interaction. Members of this group include the CTD itself (235, 236), human papillomavirus E2 (237), peroxisome proliferator-activated receptor (PPAR)-γ coactivator-1 (PGC-1) (238), Ski-interacting protein (SKIP) (239, 240), and coactivator activator (CoAA) (241). The CTD binds to SR proteins (223), and its overexpression in mammalian cells causes transcription complexes to migrate to nuclear speckles containing SR proteins (242, 243). The CTD can stimulate splicing in vitro (235, 236) and was found to interact with snRNPs (244), PSF (245), and the survival of motor neurons protein, which has been implicated in snRNP biogenesis and recycling (246). Recent studies suggest that the CTD exerts its effect by stimulating a bypass of the H complex step of spliceosome assembly. In an *in vitro* system capable of transcription and pre-mRNA splicing, mature spliceosomal complexes formed concomitantly with the appearance of RNA polymerase II transcripts following a 6 minute reaction. T7 transcripts formed primarily non-catalytic H complexes even after a 20 minute reaction. The increased kinetics of spliceosome formation was accompanied by faster formation of mRNA product. Both these effects were inhibitable by α amanitin (247). The E2 protein functions as a regulator of papillomavirus gene expression. The protein consists of an activation domain and a DNA binding domain with a linker region between the two which is serine/arginine/glycine-rich. This E2 linker region bound four SR proteins, as indicated by far-western assays. In addition, transcripts transactivated by full-length E2 or by E2 without the hinge region exhibited a 2-fold difference in the ratio of spliced to unspliced transcript. The same gene driven by an E2-independent promoter exhibited

similar amounts of spliced transcript in the presence of E2 or hinge-deleted E2. This finding suggested the necessity of E2 within the transcription complex for efficient splicing activity (237). PGC-1 is a transcriptional coactivator of many nuclear receptors involved in adaptive thermogenesis (248, 249). The protein was shown to co-localize with both U1-70 kD and SC35 in nuclear speckles. In similar experiments as those outlined above, PGC-1 was found to stimulate splicing activity when present within the transcription complex (238). SKIP was shown to interact with the vitamin D receptor (250) and the v-Ski oncoprotein (251) as a nuclear receptor coactivator. Interaction between SKIP and multiple components of the U5 snRNP were discovered by GST pull-down. SKIP also exhibited an effect on splicing. Expression of a dominant negative SKIP construct caused accumulation of unspliced mRNA of a co-transfected reporter gene. This condition was rescued by coexpression of wild type SKIP (239). Further studies have shown that SKIP interacts with HIV-1 Tat and facilitates recognition of a Tat-dependent alternative splice site, and is in fact required for Tat transactivation in vivo. in vitro, SKIP stimulates HIV-1 transcription elongation (252). CoAA is an hnRNP-like protein which was shown to interact with the thyroid hormone receptor binding protein (TRBP) (253-255) coregulator to enhance transcription activity (256, 257) and affect alternative splicing (241, 258). It was later shown that the CoAA effects on alternative splicing were not a direct result of its transcription activity. It enhanced transcription of genes driven by TRBP-independent promoters, but only affected splicing of genes driven by TRBP-dependent promoters (241). These results outlining direct consequences for splicing activity dependent upon the presence of certain transcription

factors, offer convincing functional evidence for cooperation between proteins involved in these two biological processes.

Proteomic analyses of the spliceosome have confirmed association of CA150, SKIP, p54nrb, and PIAS (151-154). In addition to these factors, Tat SF1, subunits of RNA polymerase II, TFII-I, basic helix-loop-helix protein SHARP, XPA binding protein 2 (XAB2), TATA binding protein associated factor II 68 kD subunit ($TAFII_{68}$), and putative transcription factors have also been found associated with purified spliceosomes (152-154, 259).

Although evidence for the temporal and spatial overlap of RNA transcription and processing is abundant and convincing, there is still much to learn about the mechanisms of these coordinated processes. The steps of assembly for each complex, what degree of control they exert on one another, and how the protein compositions change during the processes which they catalyze, are all questions which must be addressed to gain a comprehensive understanding of these complex molecular functions.

## IV. References


1. **Barondes, S. H., V. Castronovo, D. N. W. Cooper, R. D. Cummings, K. Drickamer, T. Feizi, M. A. Gitt, J. Hirabayashi, C. Hughes, and K. Kasai.** 1994. Galectins: a family of animal beta-galactoside-binding lectins. Cell **76:**597-8.

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

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

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

5. **Rini, J. M., and Y. D. Lobsanov.** 1999. New animal lectin structures. Curr. Opin. Struct. Biol. **9:**578-84.

6. **Leonidas, D., E. Vatzaki, H. Vorum, J. Celis, P. Madsen, and K. Acharya.** 1998. Structural basis for the recognition of carbohydrates by human galectin-7. Biochemistry **37:**13930-13940.

7. **Leffler, H., and S. H. Barondes.** 1986. Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. J. Biol. Chem. **261:**10119-26.

8. **Liu, F.-T., R. J. Patterson, and J. L. Wang.** 2002. Intracellular functions of galectins. Biochim. Biophys. Acta. **1572:**263-73.

9. **Hughes, R. C.** 1999. Secretion of the galectin family of mammalian carbohydrate-binding proteins. Biochim. Biophys. Acta. **1473:**172-85.

10. **Colnot, C., M. A. Ripoche, F. Scaerou, D. Foulis, and F. Poirier.** 1996. Galectins in mouse embryogenesis. Biochem. Soc. Trans. **24:**141-6.

11. **Wilson, T. J., M. N. Firth, J. T. Powell, and F. L. Harrison.** 1989. The sequence of the mouse 14 kDa beta-galactoside-binding lectin and evidence for its synthesis on free cytoplasmic ribosomes. Biochem. J. **261:**847-52.

12. **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:**692-9.

13. **Cooper, D. N. W., and S. H. Barondes.** 1990. Evidence for export of a muscle lectin from cytosol to extracellular matrix and for a novel secretory mechanism. J. Cell Biol. **110:**1681-91.

14. **Harrison, F. L., and T. J. Wilson.** 1992. The 14 kDa beta-galactoside binding lectin in myoblast and myotube cultures: localization by confocal microscopy. J. Cell Sci. **101 ( Pt 3):**635-46.

15. **Cho, M., and R. D. Cummings.** 1995. Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. II. Localization and biosynthesis. J. Biol. Chem. **270:**5207-12.

16. **Savin, S. B., D. S. Cvejic, and M. M. Jankovic.** 2003. Expression of galectin-1 and galectin-3 in human fetal thyroid gland. J. Histochem. Cytochem. **51:**479-83.

17. **Shimonishi, T., K. Miyazaki, N. Kono, H. Sabit, K. Tuneyama, K. Harada, J. Hirabayashi, K. Kasai, and Y. Nakanuma.** 2001. Expression of endogenous galectin-1 and galectin-3 in intrahepatic cholangiocarcinoma. Hum. Pathol. **32:**302-10.

18. **Lutomski, D., M. Fouillit, P. Bourin, D. Mellottee, N. Denize, M. Pontet, D. Bladier, M. Caron, and R. Joubert-Caron.** 1997. Externalization and binding of galectin-1 on cell surface of K562 cells upon erythroid differentiation. Glycobiology **7:**1193-9.

19. **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:**7486-93.

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

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

22. **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 ( Pt 20):**3035-43.

23. **Kaltner, H., K. Seyrek, A. Heck, F. Sinowatz, and H. J. Gabius.** 2002. Galectin-1 and galectin-3 in fetal development of bovine respiratory and digestive tracts. Comparison of cell type-specific expression profiles and subcellular localization. Cell Tissue Res. **307:**35-46.

24. **Wollina, U., G. Schreiber, M. Gornig, S. Feldrappe, M. Burchert, and H. J. Gabius.** 1999. Sertoli cell expression of galectin-1 and -3 and accessible binding sites in normal human testis and Sertoli cell only-syndrome. Histol. Histopathol. **14:**779-84.

25. **Dettin, L., N. Rubinstein, A. Aoki, G. A. Rabinovich, and C. A. Maldonado.** 2003. Regulated expression and ultrastructural localization of galectin-1, a proapoptotic beta-galactoside-binding lectin, during spermatogenesis in rat testis. Biol. Reprod. **68:**51-9.

26. **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:**419-28.

27. **Vyakarnam, A., S. F. Dagher, J. L. Wang, and R. J. Patterson.** 1997. Evidence for a role for galectin-1 in pre-mRNA splicing. Mol. Cell. Biol. **17:**4730-7.

28. **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. Nucleic Acids Res. **29:**3595-602.

29. **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.

30. **Mourelatos, Z., J. Dostie, S. Paushkin, A. Sharma, B. Charroux, L. Abel, J. Rappsilber, M. Mann, and G. Dreyfuss.** 2002. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev. **16:**720-8.

31. **Melki, J.** 1999. Molecular basis of spinal muscular atrophy: recent advances. J. Child Neurol. **14:**43.

32. **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.

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

34. **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.

35. **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. U.S.A. **92:**1213-7.

36. **Poirier, F., and E. J. Robertson.** 1993. Normal development of mice carrying a null mutation in the gene encoding the L14 S-type lectin. Development **119:**1229-1236.

37. **Puche, A. C., F. Poirier, M. Hair, P. F. Bartlett, and B. Key.** 1996. Role of galectin-1 in the developing mouse olfactory system. Dev. Biol. **179:**274-87.

38. **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.

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

40. **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.

41. **Birdsall, B., J. Feeney, I. D. Burdett, S. Bawumia, E. A. Barboni, and R. C. Hughes.** 2001. NMR solution studies of hamster galectin-3 and electron microscopic visualization of surface-adsorbed complexes: evidence for interactions between the N- and C-terminal domains. Biochemistry **40:**4859-66.

42. **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.

43. **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:**4455-60.

44. **Umemoto, K., H. Leffler, A. Venot, H. Valafar, and J. H. Prestegard.** 2003. Conformational differences in liganded and unliganded states of Galectin-3. Biochemistry **42:**3688-95.

45. **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. U.S.A. **93:**6737-6742.

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

47. **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:**4086-92.

48. **Fowlis, D., C. Colnot, M. A. Ripoche, and F. Poirier.** 1995. Galectin-3 is expressed in the notochord, developing bones, and skin of the postimplantation mouse embryo. Dev. Dyn. **203:**241-51.

49. **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:**26704-26711.

50. **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.

51. **Laing, J. G., and J. L. Wang.** 1988. Identification of carbohydrate binding protein 35 in heterogeneous nuclear ribonucleoprotein complex. Biochemistry **27:**5329-34.

52. **Hubert, M., S. Y. Wang, J. L. Wang, A. P. Seve, and J. Hubert.** 1995. Intranuclear distribution of galectin-3 in mouse 3T3 fibroblasts: comparative analyses by immunofluorescence and immunoelectron microscopy. Exp. Cell Res. **220:**397-406.

53. **Craig, S. S., P. Krishnaswamy, A. M. Irani, C. L. Kepley, F.-T. Liu, and L. B. Schwartz.** 1995. Immunoelectron microscopic localization of galectin-3, an IgE binding protein, in human mast cells and basophils. Anat. Rec. **242:**211-9.

54. **Dumic, J., G. Lauc, M. Hadzija, and M. Flogel.** 2000. Transfer to in vitro conditions influences expression and intracellular distribution of galectin-3 in murine peritoneal macrophages. Z. Naturforsch [C] **55:**261-6.

55. **Askew, D., C. J. Burger, and K. D. Elgert.** 1993. Tumor growth and adherence change the expression of macrophage Mac-2. Cancer Lett. **69:**67-74.

56. **Lotz, M. M., C. W. Andrews, Jr., C. A. Korzelius, E. C. Lee, G. D. Steele, Jr., A. Clarke, and A. M. Mercurio.** 1993. Decreased expression of Mac-2 (carbohydrate binding protein 35) and loss of its nuclear localization are associated with the neoplastic progression of colon carcinoma. Proc. Natl. Acad. Sci. U.S.A. **90:**3466-70.

57. **Sanjuan, X., P. L. Fernandez, A. Castells, V. Castronovo, F. A. van den Brule, F.-T. Liu, A. Cardesa, and E. Campo.** 1997. Differential expression of galectin 3 and galectin 1 in colorectal cancer progression. Gastroenterology **113:**1906-15.

58. **Honjo, Y., H. Inohara, S. Akahani, T. Yoshii, Y. Takenaka, J. Yoshida, K. Hattori, Y. Tomiyama, A. Raz, and T. Kubo.** 2000. Expression of cytoplasmic

galectin-3 as a prognostic marker in tongue carcinoma. Clin. Cancer Res. **6:**4635-40.

59. **van den Brule, F. A., D. Waltregny, F.-T. Liu, and V. Castronovo.** 2000. Alteration of the cytoplasmic/nuclear expression pattern of galectin-3 correlates with prostate carcinoma progression. Int. J. Cancer **89:**361-7.

60. **Orlandi, F., E. Saggiorato, G. Pivano, B. Puligheddu, A. Termine, S. Cappia, P. De Giuli, and A. Angeli.** 1998. Galectin-3 is a presurgical marker of human thyroid carcinoma. Cancer Res. **58:**3015-20.

61. **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. U.S.A. **84:**6452-6.

62. **Agrwal, N., J. L. Wang, and P. G. Voss.** 1989. Carbohydrate-binding protein 35. Levels of transcription and mRNA accumulation in quiescent and proliferating cells. J. Biol. Chem. **264:**17236-17242.

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

64. **Openo, K. P., M. M. Kadrofske, R. J. Patterson, and J. L. Wang.** 2000. Galectin-3 expression and subcellular localization in senescent human fibroblasts. Exp. Cell Res. **255:**278-90.

65. **Davidson, P. J., M. J. Davis, R. J. Patterson, M. A. Ripoche, F. Poirier, and J. L. Wang.** 2002. Shuttling of galectin-3 between the nucleus and cytoplasm. Glycobiology **12:**329-37.

66. **Gong, H. C., Y. Honjo, P. Nangia-Makker, V. Hogan, N. Mazurak, R. S. Bresalier, and A. Raz.** 1999. The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells. Cancer Res. **59:**6239-6245.

67. **Gaudin, J. C., B. Mehul, and R. C. Hughes.** 2000. Nuclear localisation of wild type and mutant galectin-3 in transfected cells. Biol. Cell **92:**49-58.

68. **Davidson, P. J., S.-Y. Li, A. G. Lohse, R. Vandergaast, E. Verde, A. Pearson, R. J. Patterson, J. L. Wang, and E. J. Arnoys.** 2006. Transport of galectin-3 between the nucleus and cytoplasm. I. Conditions and signals for nuclear import. Glycobiology **16:**602-611.

69. **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:**250-61.

70. **Li, S.-Y., P. J. Davidson, N. Y. Lin, R. J. Patterson, J. L. Wang, and E. J. Arnoys.** 2006. Transport of galectin-3 between the nucleus and cytoplasm. II. Identification of the signal for nuclear export. Glycobiology **16:**612-622.

71. **Wang, L., H. Inohara, K. J. Pienta, and A. Raz.** 1995. Galectin-3 is a nuclear matrix protein which binds RNA. Biochem. Biophys. Res. Cummun. **217:**292-303.

72. **Paron, I., A. Scaloni, A. Pines, A. Bachi, F.-T. Liu, C. Puppin, M. Pandolfi, L. Ledda, C. Di Loreto, G. Damante, and G. Tell.** 2003. Nuclear localization of galectin-3 in transformed thyroid cells: a role in transcriptional regulation. Biochem Biophys Res Commun **302:**545-53.

73. **Lin, H.-M., R. G. Pestell, A. Raz, and H. R. Kim.** 2002. Galectin-3 enhances cyclin D(1) promoter activity through SP1 and a cAMP-responsive element in human breast epithelial cells. Oncogene **21:**8001-10.

74. **Honjo, Y., P. Nangia-Makker, H. Inohara, and A. Raz.** 2001. Down-regulation of galectin-3 suppresses tumorigenicity of human breast carcinoma cells. Clin. Cancer Res. **7:**661-668.

75. **Joo, H.-G., P. S. Goedegebuure, N. Sadanaga, M. Nagoshi, W. von Bernstorff, and T. J. Eberlein.** 2001. Expression and function of galectin-3, a {beta}-galactoside-binding protein in activated T lymphocytes. J. Leukoc. Biol. **69:**555-564.

76. **Matarresea, P., N. Tinari, M. L. Semeraroa, C. Natolib, S. Iacobelli, and W. Malorni.** 2000. Galectin-3 overexpression protects from cell damage and death by influencing mitochondrial homeostasis. FEBS Lett. **473:**311-5.

77. **Lin, H.-M., B.-K. Moon, F. Yu, and H.-R. C. Kim.** 2000. Galectin-3 mediates genistein-induced G2/M arrest and inhibits apoptosis. Carcinogenesis **21:**1941-1945.

78. **Hsu, D. K., R.-Y. Yang, Z. Pan, L. Yu, D. R. Salomon, W.-P. Fung-Leung, and F.-T. Liu.** 2000. Targeted Disruption of the Galectin-3 Gene Results in Attenuated Peritoneal Inflammatory Responses. Am. J. Pathol. **156:**1073-1083.

79. **Colnot, C., M. A. Ripoche, G. Milon, X. Montagutelli, P. R. Crocker, and F. Poirier.** 1998. Maintenance of granulocyte numbers during acute peritonitis is defective in galectin-3-null mutant mice. Immunology **94:**290-6.

80. **Kim, H.-R. C., H.-M. Lin, H. Biliran, and A. Raz.** 1999. Cell Cycle Arrest and Inhibition of Anoikis by Galectin-3 in Human Breast Epithelial Cells. Cancer Res. **59:**4148-4154.

81. **Yamazaki, K., A. Kawai, M. Kawaguchi, Y. Hibino, F. Li, M. Sasahara, K. Tsukada, and K. Hiraga.** 2001. Simultaneous induction of galectin-3

phosphorylated on tyrosine residue, p21(WAF1/Cip1/Sdi1), and the proliferating cell nuclear antigen at a distinctive period of repair of hepatocytes injured by CCl4. Biochem. Biophys. Res. Cummun. **280:**1077-84.

82. **Yu, F., R. L. Finley, Jr., A. Raz, and H.-R. C. Kim.** 2002. Galectin-3 translocates to the perinuclear membranes and inhibits cytochrome c release from the mitochondria. A role for synexin in galectin-3 translocation. J. Biol. Chem. **277:**15819-15827.

83. **Garin, J., R. Diez, S. Kieffer, J.-F. Dermine, S. Duclos, E. Gagnon, R. Sadoul, C. Rondeau, and M. Desjardins.** 2001. The phagosome proteome: insight into phagosome functions. J. Cell Biol. **152:**165-180.

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

85. **Seve, A. P., M. Felin, M. A. Doyennette-Moyne, T. Sahraoui, M. Aubery, and J. Hubert.** 1993. Evidence for a lactose-mediated association between two nuclear carbohydrate-binding proteins. Glycobiology **3:**23-30.

86. **Menon, R. P., M. Strom, and R. C. Hughes.** 2000. Interaction of a novel cysteine and histidine-rich cytoplasmic protein with galectin-3 in a carbohydrate-independent manner. FEBS Lett. **470:**227-31.

87. **Bawumia, S., E. A. Barboni, R. P. Menon, and R. Colin Hughes.** 2003. Specificity of interactions of galectin-3 with Chrp, a cysteine- and histidine-rich cytoplasmic protein. Biochimie **85:**189-94.

88. **Goletz, S., F. G. Hanisch, and U. Karsten.** 1997. Novel alphaGalNAc containing glycans on cytokeratins are recognized invitro by galectins with type II carbohydrate recognition domains. J. Cell Sci. **110 ( Pt 14):**1585-96.

89. **Cowles, E. A., N. Agrwal, R. L. Anderson, and J. L. Wang.** 1990. Carbohydrate-binding protein 35. Isoelectric points of the polypeptide and a phosphorylated derivative. J. Biol. Chem. **265:**17706-12.

90. **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.

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

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

93. **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:**306-13.

94. **Sharp, P.** 1994. Split genes and RNA splicing. Cell **77:**805-15.

95. **Stephens, R., and T. Schneider.** 1992. Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. J. Mol. Biol. **228:**1124-36.

96. **Lopez, A.** 1998. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. Annu. Rev. Genet. **32:**279-305.

97. **Smith, C., and J. Valcarcel.** 2000. Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem. Sci. **25:**381-8.

98. **Padgett, R., M. Konarska, P. Grabowski, S. Hardy, and P. Sharp.** 1984. Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science **225:**898-903.

99. **Ruskin, B., A. Krainer, T. Maniatis, and M. Green.** 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell **38:**317-31.

100. **Maschhoff, K. L., and R. A. Padget.** 1993. The stereochemical course of the first step of pre-mRNA splicing. Nucleic Acids Res. **21:**5456-5462.

101. **Moore, M., and P. Sharp.** 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature **365:**364-8.

102. **Frendewey, D., and W. Keller.** 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell **42:**355-67.

103. **Grabowski, P., S. Seiler, and P. Sharp.** 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell **42:**345-53.

104. **Padgett, R., M. Podar, S. Boulanger, and P. Perlman.** 1994. The stereochemical course of group II intron self-splicing. Science **266:**1685-1688.

105. **Michel, F., and J. Ferat.** 1995. Structure and activities of group II introns. Annu. Rev. Biochem. **64:**435-61.

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

107. **Staley, J., and C. Guthrie.** 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell **92:**315-26.

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

109. **Will, C., and R. Luhrmann.** 2001. Spliceosomal UsnRNP biogenesis, structure and function. Curr. Opin. Cell Biol. **13:**290-301.

110. **Newman, A.** 1997. The role of U5 snRNP in pre-mRNA splicing. EMBO J. **16:**5797-800.

111. **Dreyfuss, G., M. Matunis, S. Pinol-Roma, and C. Burd.** 1993. hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. **62:**289-321.

112. **Krecic, A., and M. Swanson.** 1999. hnRNP complexes: composition, structure, and function. Curr. Opin. Cell Biol. **11:**363-71.

113. **Mayeda, A., D. M. Helfman, and A. R. Krainer.** 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. Mol. Cell. Biol. **13:**2993-3001.

114. **Mayeda, A., S. Munroe, J. Caceres, and A. Krainer.** 1994. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. EMBO J. **13:**5483-95.

115. **Caceres, J., S. Stamm, D. Helfman, and A. Krainer.** 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science **265:**1706-1709.

116. **Singh, R., J. Valcarcel, and M. Green.** 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science **268:**1173-1176.

117. **Yang, X., M. Bani, S. Lu, S. Rowan, Y. Ben-David, and B. Chabot.** 1994. The A1 and A1B proteins of heterogeneous nuclear ribonucleoparticles modulate 5' splice site selection in vivo. Proc. Natl. Acad. Sci. U.S.A. **91:**6924-6928.

118. **Min, H., R. Chan, and D. Black.** 1995. The generally expressed hnRNP F is involved in a neural-specific pre- mRNA splicing event. Genes Dev. **9:**2659-2671.

119. **Birney, E., S. Kumar, and A. R. Krainer.** 1993. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res. **21:**5803-5816.

120. **Vellard, M., A. Sureau, J. Soret, C. Martinerie, and B. Perbal.** 1992. A potential splicing factor is encoded by the opposite strand of the trans-spliced c-myb exon. Proc. Natl. Acad. Sci. U.S.A. **89:**2511-2515.

121. **Fu, X., and T. Maniatis.** 1992. Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. Science **256:**535-538.

122. **Lazar, G., T. Schaal, T. Maniatis, and H. Goodman.** 1995. Identification of a Plant Serine-Arginine-Rich Protein Similar to the Mammalian Splicing Factor SF2/ASF. Proc. Natl. Acad. Sci. U.S.A. **92:**7672-7676.

123. **Ge, H., P. Zuo, and J. Manley.** 1991. Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. Cell **66:**373-82.

124. **Tacke, R., A. Boned, and C. Goridis.** 1992. ASF alternative transcripts are highly conserved between mouse and man. Nucleic Acids Res. **20:**5482-.

125. **Zamore, P., J. Patton, and M. Green.** 1992. Cloning and domain structure of the mammalian splicing factor U2AF. Nature **355:**609-14.

126. **Gil, A., P. Sharp, S. Jamison, and M. Garcia-Blanco.** 1991. Characterization of cDNAs encoding the polypyrimidine tract-binding protein. Genes Dev. **5:**1224-1236.

127. **Patton, J., S. Mayer, P. Tempst, and B. Nadal-Ginard.** 1991. Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. Genes Dev. **5:**1237-1251.

128. **Ghetti, A., S. Pinol-Roma, W. M. Michael, C. Morandi, and G. Dreyfuss.** 1992. hnRNP 1, the polyprimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. Nucleic Acids Res. **20:**3671-3678.

129. **Patton, J., E. Porro, J. Galceran, P. Tempst, and B. Nadal-Ginard.** 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. Genes Dev. **7:**393-406.

130. **Tarn, W., and J. Steitz.** 1995. Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. Proc. Natl. Acad. Sci. U.S.A. **92:**2504-2508.

131. **Roscigno, R. F., and M. A. Garcia-Blanco.** 1995. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA **1:**692-706.

132. **Luhrmann, R., B. Kastner, and M. Bach.** 1990. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. Biochim. Biophys. Acta. **1087:**265-92.

133. **Zieve, G., and R. Sauterer.** 1990. Cell biology of the snRNP particles. Crit. Rev. Biochem. Mol. Biol. **25:**1-46.

134. **Behrens, S., and R. Luhrmann.** 1991. Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. Genes Dev. **5:**1439-1452.

135. **Behrens, S. E., K. Tyc, B. Kastner, J. Reichelt, and R. Luhrmann.** 1993. Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. Mol. Cell. Biol. **13:**307-319.

136. **Bach, M., G. Winkelmann, and R. Luhrmann.** 1989. 20S Small Nuclear Ribonucleoprotein U5 Shows a Surprisingly Complex Protein Composition. Proc. Natl. Acad. Sci. U.S.A. **86:**6038-6042.

137. **Guthrie, C., and B. Patterson.** 1988. Spliceosomal snRNAs. Annu. Rev. Genet. **22:**387-419.

138. **Branlant, C., A. Krol, J. Ebel, E. Lazar, B. Haendler, and M. Jacob.** 1982. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. EMBO J. **1:**1259-65.

139. **Liautard, J., J. Sri-Widada, C. Brunel, and P. Jeanteur.** 1982. Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. Proteins interact closely with a similar structural domain of U1, U2, U4 and U5 small nuclear RNAs. J. Mol. Biol. **162:**623-43.

140. **Zeller, R., T. Nyffenegger, and E. De Robertis.** 1983. Nucleocytoplasmic distribution of snRNPs and stockpiled snRNA-binding proteins during oogenesis and early development in Xenopus laevis. Cell **32:**425-34.

141. **Mattaj, I.** 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell **46:**905-11.

142. **Mattaj, I., and E. De Robertis.** 1985. Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. Cell **40:**111-8.

143. **Fischer, U., and R. Luhrmann.** 1990. An essential signaling role for the m3G cap in the transport of U1 snRNP to the nucleus. Science **249:**786-790.

144. **Hamm, J., E. Darzynkiewicz, S. Tahara, and I. Mattaj.** 1990. The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. Cell **62:**569-77.

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

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

147. **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:**5789-802.

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

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

150. **Watakabe, A., K. Tanaka, and Y. Shimura.** 1993. The role of exon sequences in splice site selection. Genes Dev. **7:**407-418.

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

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

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

154. **Zhou, Z., L. Licklider, S. Gygi, and R. Reed.** 2002. Comprehensive proteomic analysis of the human spliceosome. Nature **419:**182-5.

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

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

157. **Mount, S., I. Pettersson, M. Hinterberger, A. Karmas, and J. Steitz.** 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell **33:**509-18.

158. **Black, D., B. Chabot, and J. Steitz.** 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell **42:**737-50.

159. **Kohtz, J., S. Jamison, C. Will, P. Zuo, R. Luhrmann, M. Garcia-Blanco, and J. Manley.** 1994. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature **368:**119-24.

160. **Jamison, S. F., Z. Pasman, J. Wang, C. Will, R. Luhrmann, J. L. Manley, and M. A. Garcia-Blanco.** 1995. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. Nucleic Acids Res. **23:**3260-3267.

161. **Staknis, D., and R. Reed.** 1994. SR proteins promote the first specific recognition of Pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. Mol. Cell. Biol. **14:**7670-7682.

162. **Zuo, P., and J. Manley.** 1994. The Human Splicing Factor ASF/SF2 can Specifically Recognize Pre-mRNA 5' Splice Sites. Proc. Natl. Acad. Sci. U.S.A. **91:**3363-3367.

163. **Wu, S., C. Romfo, T. Nilsen, and M. Green.** 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature **402:**832-5.

164. **Wu, J., and T. Maniatis.** 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell **75:**1061-70.

165. **Konarska, M., and P. Sharp.** 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell **49:**763-74.

166. **Query, C., M. Moore, and P. Sharp.** 1994. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. Genes Dev. **8:**587-597.

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

168. **Roy, A., M. Meisterernst, P. Pognonec, and R. Roeder.** 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature **354:**245-8.

169. **Grueneberg, D. A., R. W. Henry, A. Brauer, C. D. Novina, V. Cheriyath, A. L. Roy, and M. Gilman.** 1997. A multifunctional DNA-binding protein that promotes the formation of serum response factor/homeodomain complexes: identity to TFII-I. Genes Dev. **11:**2482-2493.

170. **Yang, W., and S. Desiderio.** 1997. BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. Proc. Natl. Acad. Sci. U.S.A. **94:**604-609.

171. **Novina, C. D., S. Kumar, U. Bajpai, V. Cheriyath, K. Zhang, S. Pillai, H. H. Wortis, and A. L. Roy.** 1999. Regulation of nuclear localization and

transcriptional activity of TFII-I by bruton's tyrosine kinase. Mol. Cell. Biol. **19:**5014-5024.

172. **Bayarsaihan, D., and F. H. Ruddle.** 2000. Isolation and characterization of BEN, a member of the TFII-I family of DNA-binding proteins containing distinct helix-loop-helix domains. Proc. Natl. Acad. Sci. U.S.A. **97:**7342-7347.

173. **Franke, Y., R. Peoples, and U. Francke.** 1999. Identification of GTF2IRD1, a putative transcription factor within the Williams-Beuren syndrome deletion at 7q11.23. Cytogenet. Cell Genet. **86:**296-304.

174. **O'Mahoney, J. V., K. L. Guven, J. Lin, J. E. Joya, C. S. Robinson, R. P. Wade, and E. C. Hardeman.** 1998. Identification of a novel slow-muscle-fiber enhancer binding protein, MusTRD1. Mol. Cell. Biol. **18:**6641-6652.

175. **Osborne, L., T. Campbell, A. Daradich, S. Scherer, and L. Tsui.** 1999. Identification of a putative transcription factor gene (WBSCR11) that is commonly deleted in Williams-Beuren syndrome. Genomics **57:**279-84.

176. **Perez Jurado, L., Y. Wang, R. Peoples, A. Coloma, J. Cruces, and U. Francke.** 1998. A duplicated gene in the breakpoint regions of the 7q11.23 Williams- Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK. Hum. Mol. Genet. **7:**325-334.

177. **Tassabehji, M., M. Carette, C. Wilmot, D. Donnai, A. Read, and K. Metcalfe.** 1999. A transcription factor involved in skeletal muscle gene expression is deleted in patients with Williams syndrome. Eur. J. Hum. Genet. **7:**737-47.

178. **Roy, A., H. Du, P. Gregor, C. Novina, E. Martinez, and R. Roeder.** 1997. Cloning of an inr- and E-box-binding protein, TFII-I, that interacts physically and functionally with USF1. EMBO J. **16:**7091-104.

179. **Kim, D.-W., and B. H. Cochran.** 2000. Extracellular signal-regulated kinase binds to TFII-I and regulates Its activation of the c-fos promoter. Mol. Cell. Biol. **20:**1140-1148.

180. **Yang, S.-H., P. R. Yates, A. J. Whitmarsh, R. J. Davis, and A. D. Sharrocks.** 1998. The Elk-1 ETS-domain transcription factor contains a mitogen-activated protein kinase targeting motif. Mol. Cell. Biol. **18:**710-720.

181. **Wang, Y., L. Perez-Jurado, and U. Francke.** 1998. A mouse single-copy gene, Gtf2i, the homolog of human GTF2I, that is duplicated in the Williams-Beuren syndrome deletion region. Genomics **48:**163-70.

182. **Cheriyath, V., and A. L. Roy.** 2000. Alternatively Spliced Isoforms of TFII-I. Complex formation, nuclear translocation, and differential gene regulation. J. Biol. Chem. **275:**26300-26308.

183. **Burke, T., and J. Kadonaga.** 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev. **10:**711-724.

184. **Burke, T. W., and J. T. Kadonaga.** 1997. The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. Genes Dev. **11:**3020-3031.

185. **Lee, T., and R. Young.** 2000. Transcription of eukaryotic protein-coding genes. Annu. Rev. Genet. **34:**77-137.

186. **Roeder, R.** 1996. The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem. Sci. **21:**327-35.

187. **Smale, S., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57:**103-13.

188. **Novina, C., and A. Roy.** 1996. Core promoters and transcriptional control. Trends Genet. **12:**351-5.

189. **Smale, S.** 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim. Biophys. Acta. **1351:**73-88.

190. **Manzano-Winkler, B., C. D. Novina, and A. L. Roy.** 1996. TFII Is required for transcription of the naturally TATA-less but initiator-containing Vbeta promoter. J. Biol. Chem. **271:**12076-12081.

191. **Roy, A., S. Malik, M. Meisterernst, and R. Roeder.** 1993. An alternative pathway for transcription initiation involving TFII-I. Nature **365:**355-9.

192. **Roy, A., C. Carruthers, T. Gutjahr, and R. Roeder.** 1993. Direct role for Myc in transcription initiation mediated by interactions with TFII-I. Nature **365:**359-61.

193. **Grueneberg, D. A., S. Natesan, C. Alexandre, and M. Z. Gilman.** 1992. Human and drosophila homeodomain proteins that enhance the DNA-binding activity of serum response factor. Science **257:**1089-1095.

194. **Kim, D.-W., V. Cheriyath, A. L. Roy, and B. H. Cochran.** 1998. TFII-I Enhances Activation of the c-fos Promoter through Interactions with Upstream Elements. Mol. Cell. Biol. **18:**3310-3320.

195. **Parker, R., T. Phan, P. Baumeister, B. Roy, V. Cheriyath, A. L. Roy, and A. S. Lee.** 2001. Identification of TFII-I as the endoplasmic reticulum stress response element binding factor ERSF: its autoregulation by stress and interaction with ATF6. Mol. Cell. Biol. **21:**3220-3233.

196. **Roy, B., and A. Lee.** 1999. The mammalian endoplasmic reticulum stress response element consists of an evolutionarily conserved tripartite structure and interacts with a novel stress-inducible complex. Nucleic Acids Res. **27:**1437-1443.

197. **Yoshida, H., K. Haze, H. Yanagi, T. Yura, and K. Mori.** 1998. Identification of the cis-Acting Endoplasmic Reticulum Stress Response Element Responsible for Transcriptional Induction of Mammalian Glucose-regulated Proteins. Involvement of basic leucine zipper transcription factors. J. Biol. Chem. **273:**33741-33749.

198. **Hong, M., M.-y. Lin, J.-m. Huang, P. Baumeister, S. Hakre, A. L. Roy, and A. S. Lee.** 2005. Transcriptional Regulation of the Grp78 Promoter by Endoplasmic Reticulum Stress: role of TFII-I and its tyrosine phosphorylation. J. Biol. Chem. **280:**16821-16828.

199. **Tussie-Luna, M. I., B. Michel, S. Hakre, and A. L. Roy.** 2002. The SUMO Ubiquitin-Protein Isopeptide Ligase Family Member Miz1/PIASxbeta /Siz2 Is a Transcriptional Cofactor for TFII-I. J. Biol. Chem. **277:**43185-43193.

200. **Tussie-Luna, M. I., D. Bayarsaihan, E. Seto, F. H. Ruddle, and A. L. Roy.** 2002. Physical and functional interactions of histone deacetylase 3 with TFII-I family proteins and PIASxbeta. Proc. Natl. Acad. Sci. U.S.A. **99:**12807-12812.

201. **Cheriyath, V., C. D. Novina, and A. L. Roy.** 1998. TFII-I Regulates Vbeta Promoter Activity through an Initiator Element. Mol. Cell. Biol. **18:**4444-4454.

202. **Cheriyath, V., and A. L. Roy.** 2001. Structure-Function Analysis of TFII-I: roles of the N-terminal end, basic region, and I-repeats. J. Biol. Chem. **276:**8377-8383.

203. **Sacristan, C., M. I. Tussie-Luna, S. M. Logan, and A. L. Roy.** 2004. Mechanism of Bruton's Tyrosine Kinase-mediated Recruitment and Regulation of TFII-I. J. Biol. Chem. **279:**7147-7158.

204. **Desiderio, S.** 1997. Role of Btk in B cell development and signaling. Curr. Opin. Immunol. **9:**534-40.

205. **Lemmon, M., K. Ferguson, and J. Schlessinger.** 1996. PH domains: diverse sequences with a common fold recruit signaling molecules to the cell surface. Cell **85:**621-4.

206. **Egloff, A. M., and S. Desiderio.** 2001. Identification of phosphorylation sites for bruton's tyrosine kinase within the transcriptional regulator BAP/TFII-I. J. Biol. Chem. **276:**27806-27815.

207. **Novina, C. D., V. Cheriyath, and A. L. Roy.** 1998. Regulation of TFII-I Activity by Phosphorylation. J. Biol. Chem. **273:**33443-33448.

208. **Cheriyath, V., Z. P. Desgranges, and A. L. Roy.** 2002. c-Src-dependent transcriptional activation of TFII-I. J. Biol. Chem. **277:**22798-22805.

209. **Kim, D.-W., and B. H. Cochran.** 2001. JAK2 activates TFII-I and regulates its interaction with extracellular signal-regulated kinase. Mol. Cell. Biol. **21:**3387-3397.

210. **Casteel, D. E., S. Zhuang, T. Gudi, J. Tang, M. Vuica, S. Desiderio, and R. B. Pilz.** 2002. cGMP-dependent protein kinase Ibeta physically and functionally interacts with the transcriptional regulator TFII-I. J. Biol. Chem. **277:**32003-32014.

211. **Bentley, D.** 2002. The mRNA assembly line: transcription and processing machines in the same factory. Curr. Opin. Cell Biol. **14:**336-42.

212. **Kornblihtt, A. R., M. De La Mata, J. P. Fededa, M. J. Munoz, and G. Nogues.** 2004. Multiple links between transcription and splicing. RNA **10:**1489-1498.

213. **Maniatis, T., and R. Reed.** 2002. An extensive network of coupling among gene expression machines. Nature **416:**499-506.

214. **Neugebauer, K. M.** 2002. On the importance of being co-transcriptional. J. Cell Sci. **115:**3865-3871.

215. **Proudfoot, N.** 2003. Dawdling polymerases allow introns time to splice. Nat. Struct. Biol. **10:**876-878.

216. **Beyer, A., and Y. Osheim.** 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev. **2:**754-765.

217. **Osheim, Y., O. Miller, and A. Beyer.** 1985. RNP particles at splice junction sequences on Drosophila chorion transcripts. Cell **43:**143-51.

218. **Cramer, P., C. G. Pesce, F. E. Baralle, and A. R. Kornblihtt.** 1997. Functional association between promoter structure and transcript alternative splicing. Proc. Natl. Acad. Sci. U.S.A. **94:**11456-11460.

219. **Cramer, P., J. Caceres, D. Cazalla, S. Kadener, A. Muro, F. Baralle, and A. Kornblihtt.** 1999. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol. Cell **4:**251-8.

220. **Chen, Y., D. Chafin, D. H. Price, and A. L. Greenleaf.** 1996. Drosophila RNA Polymerase II Mutants That Affect Transcription Elongation. J. Biol. Chem. **271:**5993-5999.

221. **Roberts, G., C. Gooding, H. Mak, N. Proudfoot, and C. Smith.** 1998. Co-transcriptional commitment to alternative splice site selection. Nucleic Acids Res. **26:**5568-5572.

222. **Robson-Dixon, N. D., and M. A. Garcia-Blanco.** 2004. MAZ Elements Alter Transcription Elongation and Silencing of the Fibroblast Growth Factor Receptor 2 Exon IIIb. J. Biol. Chem. **279:**29075-29084.

223. **Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden.** 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc. Natl. Acad. Sci. U.S.A. **93:**6975-6980.

224. **Davies, R. C., C. Calvio, E. Bratt, S. H. Larsson, A. I. Lamond, and N. D. Hastie.** 1998. WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. Genes Dev. **12:**3217-3225.

225. **Goldstrohm, A. C., T. R. Albrecht, C. Sune, M. T. Bedford, and M. A. Garcia-Blanco.** 2001. The Transcription Elongation Factor CA150 Interacts with RNA Polymerase II and the Pre-mRNA Splicing Factor SF1. Mol. Cell. Biol. **21:**7617-7628.

226. **Sanchez-Alvarez, M., A. C. Goldstrohm, M. A. Garcia-Blanco, and C. Sune.** 2006. Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through Its amino and carboxyl regions. Mol. Cell. Biol. **26:**4998-5014.

227. **Kameoka, S., P. Duque, and M. Konarska.** 2004. p54(nrb) associates with the 5' splice site within large transcription/splicing complexes. EMBO J. **23:**1782-91.

228. **Wang, Z., Q. Qiu, and T. Deuel.** 1993. The Wilms' tumor gene product WT1 activates or suppresses transcription through separate functional domains. J. Biol. Chem. **268:**9172-9175.

229. **Barbaux, S., P. Niaudet, M. Gubler, J. Grunfeld, F. Jaubert, F. Kuttenn, C. Fekete, N. Souleyreau-Thervi..., E. Thibaud, M. Fellous, and K. McElreavey.** 1997. Donor splice-site mutations in WT1 are responsible for Frasier syndrome. Nat. Genet. **17:**467-70.

230. **Little, M., and C. Wells.** 1997. A clinical overview of WT1 gene mutations. Hum.Mutat. **9:**209-25.

231. **Larsson, S., J. Charlieu, K. Miyagawa, D. Engelkamp, M. Rassoulzadegan, A. Ross, F. Cuzin, V. van Heyningen, and N. Hastie.** 1995. Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. Cell **81:**391-401.

232. **Sune, C., and M. A. Garcia-Blanco.** 1999. Transcriptional Cofactor CA150 Regulates RNA Polymerase II Elongation in a TATA-Box-Dependent Manner. Mol. Cell. Biol. **19:**4719-4728.

233. **Sune, C., T. Hayashi, Y. Liu, W. Lane, R. Young, and M. Garcia-Blanco.** 1997. CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. Mol. Cell. Biol. **17:**6029-6039.

234. **Zhang, D., A. J. Paley, and G. Childs.** 1998. The Transcriptional Repressor ZFM1 Interacts with and Modulates the Ability of EWS to Activate Transcription. J. Biol. Chem. **273:**18086-18091.

235. **Hirose, Y., R. Tacke, and J. L. Manley.** 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. Genes Dev. **13:**1234-1239.

236. **Zeng, C., and S. M. Berget.** 2000. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. Mol. Cell. Biol. **20:**8290-8301.

237. **Lai, M.-C., B. H. Teh, and W.-Y. Tarn.** 1999. A Human Papillomavirus E2 Transcriptional Activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. J. Biol. Chem. **274:**11832-11841.

238. **Monsalve, M., Z. Wu, G. Adelmant, P. Puigserver, M. Fan, and B. Spiegelman.** 2000. Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1. Mol. Cell **6:**307-16.

239. **Zhang, C., D. R. Dowd, A. Staal, C. Gu, J. B. Lian, A. J. Van Wijnen, G. S. Stein, and P. N. MacDonald.** 2003. NCoA62/SKIP is a nuclear matrix-associated coactivator that may couple vitamin D receptor-mediated transcription and RNA splicing. J. Biol. Chem. **278:**35325-35336.

240. **Nagai, K., T. Yamaguchi, T. Takami, A. Kawasumi, M. Aizawa, N. Masuda, M. Shimizu, S. Tominaga, T. Ito, T. Tsukamoto, and T. Osumi.** 2004. SKIP modifies gene expression by affecting both transcription and splicing. Biochem. Biophys. Res. Cummun. **316:**512-517.

241. **Auboeuf, D., D. H. Dowhan, X. Li, K. Larkin, L. Ko, S. M. Berget, and B. W. O'Malley.** 2004. CoAA, a Nuclear Receptor Coactivator Protein at the Interface of Transcriptional Coactivation and RNA Splicing. Mol. Cell. Biol. **24:**442-453.

242. **Du, L., and S. L. Warren.** 1997. A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. J. Cell Biol. **136:**5-18.

243. **Misteli, T., and D. Spector.** 1999. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol. Cell **3:**697-705.

244. **Robert, F., M. Blanchette, O. Maes, B. Chabot, and B. Coulombe.** 2002. A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. J. Biol. Chem. **277:**9302-9306.

245. **Emili, A., M. Shales, S. McCracken, W. Xie, P. W. Tucker, R. Kobayashi, B. J. Blencowe, and C. J. Ingles.** 2002. Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD. RNA **8:**1102-1111.

246. **Pellizzoni, L., B. Charroux, J. Rappsilber, M. Mann, and G. Dreyfuss.** 2001. A Functional Interaction between the Survival Motor Neuron Complex and RNA Polymerase II. J. Cell Biol. **152:**75-86.

247. **Das, R., K. Dufu, B. Romney, M. Feldt, M. Elenko, and R. Reed.** 2006. Functional coupling of RNAP II transcription to spliceosome assembly. Genes Dev. **20:**1100-1109.

248. **Puigserver, P., Z. Wu, C. Park, R. Graves, M. Wright, and B. Spiegelman.** 1998. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. Cell **92:**829-39.

249. **Wu, Z., P. Puigserver, U. Andersson, C. Zhang, G. Adelmant, V. Mootha, A. Troy, S. Cinti, B. Lowell, R. Scarpulla, and B. Spiegelman.** 1999. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. Cell **98:**115-24.

250. **Baudino, T. A., D. M. Kraichely, S. C. Jefcoat Jr., S. K. Winchester, N. C. Partridge, and P. N. MacDonald.** 1998. Isolation and Characterization of a Novel Coactivator Protein, NCoA-62, Involved in Vitamin D-mediated Transcription. J. Biol. Chem. **273:**16434-16441.

251. **Dahl, R., B. Wani, and M. Hayman.** 1998. The Ski oncoprotein interacts with Skip, the human homolog of Drosophila Bx42. Oncogene **16:**1579-86.

252. **Bres, V., N. Gomes, L. Pickle, and K. A. Jones.** 2005. A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. Genes Dev. **19:**1211-1226.

253. **Caira, F., P. Antonson, M. Pelto-Huikko, E. Treuter, and J.-A. Gustafsson.** 2000. Cloning and Characterization of RAP250, a Novel Nuclear Receptor Coactivator. J. Biol. Chem. **275:**5308-5317.

254. **Ko, L., G. R. Cardona, and W. W. Chin.** 2000. Thyroid hormone receptor-binding protein, an LXXLL motif-containing protein, functions as a general coactivator. Proc. Natl. Acad. Sci. U.S.A. **97:**6212-6217.

255. **Lee, S.-K., S. L. Anzick, J.-E. Choi, L. Bubendorf, X.-Y. Guan, Y.-K. Jung, O. P. Kallioniemi, J. Kononen, J. M. Trent, D. Azorsa, B.-H. Jhun, J. H.**

**Cheong, Y. C. Lee, P. S. Meltzer, and J. W. Lee.** 1999. A Nuclear Factor, ASC-2, as a Cancer-amplified Transcriptional Coactivator Essential for Ligand-dependent Transactivation by Nuclear Receptors in Vivo. J. Biol. Chem. **274:**34283-34293.

256. **Iwasaki, T., W. W. Chin, and L. Ko.** 2001. Identification and characterization of RRM-containing coactivator activator (CoAA) as TRBP-interacting protein, and Its splice variant as a coactivator modulator (CoAM). J. Biol. Chem. **276:**33375-33383.

257. **Jung, D.-J., S.-Y. Na, D. S. Na, and J. W. Lee.** 2002. Molecular cloning and characterization of CAPER, a novel coactivator of activating protein-1 and estrogen receptors. J. Biol. Chem. **277:**1229-1234.

258. **Auboeuf, D., A. Honig, S. M. Berget, and B. W. O'Malley.** 2002. Coordinate regulation of transcription and splicing by steroid receptor coregulators. Science **298:**416-419.

259. **Deckert, J., K. Hartmuth, D. Boehringer, N. Behzadnia, C. L. Will, B. Kastner, H. Stark, H. Urlaub, and R. Luhrmann.** 2006. Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. Mol. Cell. Biol. **26:**5528-5543.

# CHAPTER 2

## Identification of Transcription Factor II-I in Association with Galectin-containing Spliceosomal Complexes

## ABSTRACT

Galectin-3 (Gal3), a protein involved in nuclear splicing of pre-mRNA, was expressed and purified as a fusion protein with glutathione S-transferase (GST). When nuclear extracts of HeLa cells were subjected to adsorption on GST-Gal3 beads, the general transcription factor II-I (TFII-I) was identified by mass spectrometry as one of the polypeptides specifically bound. Lactose, a saccharide ligand of the galectins, inhibited GST-Gal3 pull-down of TFII-I from nuclear extract. Similar results were also obtained using galectin-1, another member of the same protein family that is also involved in pre-mRNA splicing. Previous proteomic analysis of the spliceosome had identified TFII-I as one of the components of the macromolecular complex. Consistent with this notion, antibodies directed against TFII-I inhibited the splicing reaction in a dose-dependent fashion. Moreover, when a splicing reaction containing $^{32}$P-labeled pre-mRNA substrate was subjected to immunoprecipitation with anti-TFII-I, spliceosomal complexes were coprecipitated with the cognate antigen, a conclusion based on finding radiolabeled RNA species that are produced on the spliceosome during the splicing reaction. These results suggest that TFII-I associates with galectin-1 or galectin-3 containing spliceosomal complexes.

## INTRODUCTION

Galectins are a family of widely distributed proteins that: (a) bind to β-galactoside containing glycoconjugates; and (b) contain characteristic amino acid sequences in the carbohydrate recognition domain (CRD) of the polypeptides (1, 2). In previous studies, we had reported the localization of galectin-1 (Gal1) and galectin-3 (Gal3) in the cell nucleus (3, 4). Several key findings suggest that Gal1 and Gal3 are two of the many proteins involved in the splicing of pre-mRNA, assayed in a cell-free system (5, 6): (a) nuclear extracts (NE) derived from HeLa cells, capable of carrying out splicing of pre-mRNA, contained both Gal1 and Gal3; (b) depletion of both galectins from NE, either by lactose (Lac) affinity chromatography or by antibody adsorption, resulted in the concomitant loss of splicing activity; and (c) either recombinant Gal1 or recombinant Gal3 was able to reconstitute splicing activity in a galectin-depleted extract.

NEs depleted of Gal1 and Gal3 failed to form active spliceosomal complexes and gel mobility shift assays of $^{32}$P-labeled pre-mRNA revealed only bands migrating in the region corresponding to early (H- and E-) complexes (7). 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 Gal3 with similar dose-response curves (5). On the basis of these results, we hypothesized that the galectins may be required at an early stage in the splicing pathway, such as organization of the early complexes for addition of other splicing factors (7-9). More recently, however, we have found that when a splicing reaction containing $^{32}$P-labeled pre-mRNA is subjected to immunoprecipitation with either anti-Gal1 or anti-Gal3, radiolabled RNA species corresponding to the starting substrate, splicing intermediates, and mature RNA products

of active spliceosomes are all co-precipitated with the specific galectin (10). These results, in turn, suggest the Gal1 and Gal3 are associated with spliceosomal complexes throughout the splicing pathway.

To define the stage and the precise role played by Gal1 and Gal3 in the splicing process, it was important to identify the interacting partner(s) of the spliceosome and the complex(es) with which the galectins are associated. In the present communication, we report that a fusion protein containing glutathione S-transferase (GST) and Gal1 (GST-Gal1) or GST-Gal3 can pull-down from NE the general transcription factor II-I (TFII-I), a protein previously identified in the spliceosome by proteomic analysis (11). In the second article of this series, we have taken advantage of this interaction to document that site-directed mutants of Gal1, devoid of carbohydrate-binding activity, retained the association to the TFII-I complex as well as the cell-free splicing activity, thereby dissociating the saccharide-binding of the protein from its spliceosomal function. Finally, in the third article, we report that the $NH_2$-terminal domain (ND) of Gal3, containing unusual Pro- and Gly-rich repeating motifs, also plays a role in the interaction of the polypeptide with the splicing machinery.

## EXPERIMENTAL PROCEDURES

Antibody reagents --- For antibodies directed against TFII-I, we used an affinity purified preparation purchased from Bethyl Labs. This antibody, designated as #558, was derived from serum of rabbits immunized with a peptide sequence contained in exons 32 and 33 of TFII-I. We also used an affinity purified rabbit anti-TFII-I raised against the recombinant human protein (Protein Tech Group, Inc.). Anti-GST antibodies were affinity purified from serum of rabbits immunized with GST-Gal1. Anti-Gal1 was obtained from the same rabbit serum. The details of the immunogen preparation and affinity purification procedure for both anti-GST and anti-Gal1 are described in the accompanying manuscript (12). Two antibodies directed against Gal3 were used: (a) the rat monoclonal antibody designated as anti-Mac-2 (13, 14); and (b) a polyclonal rabbit antiserum directed against Gal3 (#49). Human autoimmune serum reactive against the Sm epitopes (anti-Sm) found on the core polypeptides of small nuclear ribonucleoprotein complexes (snRNPs) was purchased from The Binding Site. For antibodies directed against the Survival of Motor Neuron Protein (SMN), we used a mouse monoclonal antibody (directed against residues 14-174 of the SMN polypeptide) purchased from BD Transduction Labs.

NE and splicing reactions --- Human HeLa S3 cells were grown in suspension culture by the National Cell Culture Center (Minneapolis, MN). NE was prepared in buffer D (20 mM Hepes-KOH, pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride, 0.5 mM dithiothreitol (DTT)) as described by Dignam *et al.*, (15). NEs were frozen as aliquots in a dry ice-ethanol bath and stored at -80°C.

Protein concentrations were determined by the Bradford assay (16). In this study, the protein concentration of NE was ~6 mg/ml.

To test the effect of specific antibodies on the splicing reaction, NaCl was added to NE in Dignam buffer D to 0.5 M and set on ice for 20 minutes. The samples were then dialyzed against 60% buffer D in the presence of various amounts of antibodies, anti-TFII-I (#558) and anti-Mac-2. Dialysis was carried out for 70 minutes at 4 °C in a microdialyzer with a 6-8 kD cutoff dialysis membrane (6). Splicing reaction mixtures, in a total volume of 12 µl, contained dialyzed NE sample (10 µl), [$^{32}$P]MINX pre-mRNA (17), 2.5 mM $MgCl_2$, 1.5 mM ATP, 20 mM creatine phosphate, 0.5 mM DTT, and 20 U RNasin (Promega). Splicing reactions were incubated at 30 °C for 45 minutes. The assay was stopped by addition of proteinase K and SDS to final concentrations of 4 mg/ml and 0.1%, respectively. The samples were then diluted with buffer containing 125 mM Tris, pH 8, 1 mM EDTA, and 0.3 M sodium acetate. RNA was extracted with phenol-chloroform (50:50 (vol/vol)), followed by chloroform, and precipitated with ethanol at -80 °C. The isolated RNA was then subjected to electrophoresis in 13% polyacrylamide-8.3 M urea gels, followed by autoradiography. Quantitation of product formation was carried out by exposing the gel to a Storage Phosphor Screen (Amersham Biosciences), scanning on a Storm 860 scanner (Molecular Dynamics), and using the program Image Quant (Molecular Dynamics) to determine the percentage of radioactivity in specific bands in each lane.

Splicing reaction mixtures were also carried out in a total volume of 50 µl, containing 30 µl of NE and [$^{32}$P]MINX. These splicing reactions were incubated at 31°C for 40 minutes and then subjected to immunoprecipitation. Antibodies were bound to protein

A-Sepharose beads (Pharmacia) in 60% buffer D containing 0.05% Triton X-100. After removal of unbound material and washing, the splicing reaction (containing at least $10^6$ cpm of $^{32}$P-labeled RNA in 50 µl) was added to the antibody beads in 200 µl of buffer D containing 0.05% Triton X-100. Incubation was carried out at 4 °C for 1.5 hours. The unbound material was removed and the beads were washed three times, each with 0.5 ml buffer D containing 0.05% Triton X-100. SDS solution and proteinase K were added to a final concentration of 0.1% and 4 mg/ml, respectively, and the samples were incubated at 37 °C for 20 minutes. The RNA components of the precipitated fraction were extracted and analyzed as described above.

GST fusion proteins and pull-down assay --- The cDNA for human Gal1 (18) was subcloned into the BamHI restriction site of the pGEX-2T vector (Pharmacia) to produce GST-Gal1. The plasmid pWJ31 containing the cDNA for murine Gal3 (19) was subcloned into the EcoRI restriction site of the vector pGEX-5X-1 to produce GST-Gal3. Agrwal *et al.* (20) had described site-directed mutagenesis of the cDNA for murine Gal3 in which Gly 138 and Gly139 were replaced by two stop codons. Thus, the translation open reading frame stops at Pro137, giving an ND covering residues 1-137. This mutant cDNA of Gal3 was subcloned into the expression vector pGEX-5X-1 to express GST-ND.

GST fusion proteins were expressed in *E. coli* BL-21 codon plus (DE3) cells (Stratagene) by induction with 100 µM isopropyl-β-D-galactopyranoside for 2-3 hours at 30°C. Cells were pelleted and stored at -70°C. Thawed bacterial pellets were suspended in phosphate-buffered saline (PBS) containing protease inhibitors (4 µg/ml aprotinin, 5 µg/ml leupeptin, 0.2 µg/ml pepstatin A, and 1 mM Pefabloc (Roche)) and sonicated using

a microtip probe. Triton X-100 was added to a final concentration of 0.1%. After rocking for 1 hour at 4°C, cell debris was removed by centrifugation at 12,000 x g for 10 minutes at 4°C. The supernatant was purified on the basis of GST binding to glutathione-agarose beads (Pierce).

For GST pull-down experiments, ~10 μg of each GST fusion protein was incubated with 20 μl of packed glutathione beads in the presence of 60% buffer D at 4 °C for ~14 hours. Unbound material was removed and the beads were washed three times with 400 μl of 60% buffer D. The beads were then incubated with 36 μl of NE (~144 μg total protein) along with 24 μl of 60% buffer D, with 14.7 mM creatine phosphate, 2.4 mM $MgCl_2$, and 0.4 mM ATP (final concentrations). In experiments to test the effect of saccharides on the pull-down assay, they were included in this addition at a concentration of 100 mM. The incubation was carried at 4 °C for 12 hours. After removal of unbound material, the beads were washed four times with 200 μl of 60% buffer D. The material bound to the beads was then eluted by incubation with glutathione elution buffer (16 mM glutathione, 60 mM HEPES-KOH, pH 7.9, 11.4% glycerol, 57 mM KCl, and 0.114 mM EDTA) at 31 °C for 30 minutes, followed by incubation at room temperature for one hour. The eluted material was then subjected to SDS-PAGE analysis.

SDS gel electrophoresis, silver staining, and immunoblotting --- Samples were subjected to SDS-PAGE as described by Laemmli (21). Proteins were visualized by silver staining as described by Merril *et al.* (22). For immunoblotting, samples were electrophoretically transferred onto Hybond Nitrocellulose paper (Amersham Biosciences) in the presence of buffer containing 25 mM Tris, 193 mM glycine, and 10% methanol, pH 8.3. Following transfer, membranes were incubated overnight in 10% nonfat dry milk in Tris-buffered

saline containing Tween 20 (10 mM Tris, 0.5 M NaCl, 0.05% Tween 20; T-TBS). Antibodies for immunoblotting were diluted in the same buffer and incubated with membranes for one hour at room temperature. This was followed by four washes (15 minutes each) in T-TBS. The membranes were blocked with unconjugated goat anti-rabbit immunoglobulin (Sigma) at a dilution of 1:2000 in T-TBS. The membranes were incubated with horseradish peroxidase-conjugated secondary antibodies for one hour and washed four times. Horseradish peroxidase-conjugated goat anti-rabbit and goat anti-mouse immunoglobulin (Bio-Rad) and goat anti-rat immunoglbulin (Roche) were each used at a dilution of 1:10,000. The proteins were visualized using the Western Lighting Chemiluminescence System (Perkin-Elmer Life Sciences).

Mass spectrometric analysis of selected gel slices derived from GST pull-downs --- After SDS-PAGE, the gel was stained with Brilliant Blue G-colloidal concentration (Sigma). Gel slices corresponding to the ~40 kD and ~135 kD region were digested with trypsin (Promega) following a modified protocol of Shevchencko *et al.* (23). The trypsin digests were fractionated by reverse-phase high pressure liquid chromatography, followed by electrospray ionization mass spectrometry (LC/MS/MS). The mass spectrometry and the subsequent MS/MS ion search (Mascot) were carried out by the Proteomics Core of the Research Technology Support Facility at Michigan State University.

## RESULTS

GST-Gal3 pull-down of TFII-I from NE

Human HeLa cell NE was subjected to adsorption onto glutathione beads containing GST-Gal3 and GST. After washing, the material bound to the beads was eluted with soluble glutathione and analyzed by SDS-PAGE. Comparison of the silver-stained gels revealed two bands present in the material bound to the GST-Gal3 beads but not to GST; these were designated as p40 and p135, corresponding to the approximate molecular weights of their positions of migration (Figure 1). Gel slices containing p40 and p135 were digested with trypsin and the resulting peptides were fractionated by reverse-phase high pressure liquid chromatography, followed by mass spectrometry. The MS/MS ion search on the analysis of the p135 gel slice revealed twelve matches, representing five distinct tryptic peptides (each with a carboxyl terminal lysine or arginine), with the amino acid sequence of human TFII-I. In contrast, the corresponding gel slice from parallel SDS-PAGE of the material derived from the GST beads did not yield any of these TFII-I peptides. The MS/MS ion search on the analysis of the p40 gel slice yielded four matches representing three peptides of actin.

The identification of TFII-I was confirmed by immunoblotting analysis. Using antibodies directed against TFII-I (the 558 antibody), a positive reaction was observed in the material bound to GST-Gal3 but not in the material bound to GST (Figure 2, lanes 2 and 3). In contrast, antibodies directed against the Survival of Motor Neuron (SMN) protein failed to yield a positive immunoblot in the GST-Gal3 pull-down material (Figure 2). Using anti-GST antibodies, we ascertained that approximately equal amounts of GST proteins were used in the pull-down assay. These results suggest that the GST-Gal3

**Figure 1. Comparison of the polypeptides bound to GST-Gal3 and GST on glutathione beads.**

NE was subjected to GST pull-down in 60% buffer D containing 14.7 mM creatine phosphate, 2.4 mM $MgCl_2$, and 0.4 mM ATP. The material bound to the beads was eluted with 16 mM glutathione and subjected to SDS-PAGE. Polypeptides were revealed by silver staining. The positions of migration of p135, GST-Gal3, p40, and GST are highlighted on the left; the positions of migration of molecular weight markers are indicated on the right.

Figure 1

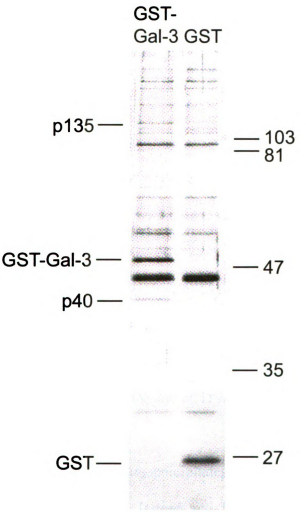
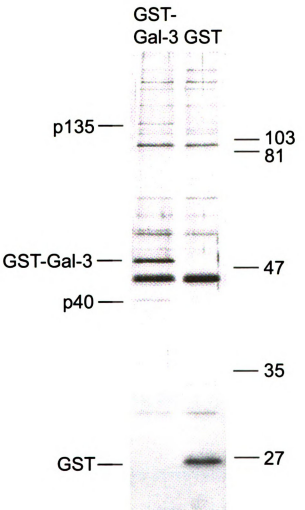


**Figure 1. Comparison of the polypeptides bound to GST-Gal3 and GST on glutathione beads.**

NE was subjected to GST pull-down in 60% buffer D containing 14.7 mM creatine phosphate, 2.4 mM $MgCl_2$, and 0.4 mM ATP. The material bound to the beads was eluted with 16 mM glutathione and subjected to SDS-PAGE. Polypeptides were revealed by silver staining. The positions of migration of p135, GST-Gal3, p40, and GST are highlighted on the left; the positions of migration of molecular weight markers are indicated on the right.

Figure 1



pull-down of TFII-I represented a specific association, either directly or indirectly, between the two proteins.

Effect of the saccharide ligand Lac and involvement of the CRD

Such a conclusion is supported by the observation that Lac, a specific saccharide ligand of Gal3, inhibited the GST-Gal3 pull-down of TFII-I (Figure 2, lane 6). In contrast, cellobiose, which does not bind to Gal3 (24), failed to yield the same effect (data not shown). These saccharide-specific effects suggest that the galectin-TFII-I interaction might be mediated by the CRD of the galectin polypeptides.

The amino acid sequence of the Gal3 polypeptide can be delineated into two domains (a) an ND containing multiple repeats of a 9-residue motif, PGAYPGXXX; and (b) a COOH-terminal CRD that shows sequence similarity with the corresponding CRD of other members of the galectin family (see, for example, reference 2). While full-length GST-Gal3 interacted with TFII-I, GST-Gal3ND failed to yield TFII-I in the pull-down (Figure 2, lanes 3 and 4). On the other hand, GST-Gal1, which contains a single CRD, interacted with TFII-I and this interaction was also sensitive to Lac inhibition (Figure 2, lanes 5 and 7). Therefore, it appears that the site of TFII-I interaction resides within the CRD of either Gal1 or Gal3.

Finally, we tested for Gal3 in the immunoprecipitates of anti-TFII-I as a reciprocal of the GST-galectin pull-down experiments. Indeed, anti-TFII-I immunoprecipitated not only its own cognate antigen but Gal3 as well (Figure 3). The 558 anti-TFII-I antibody is a polyclonal rabbit antiserum affinity purified over the peptide immunogen. For negative control, we used an anti-GST antiserum that went through the same affinity purification

**Figure 2. Immunoblotting analysis of various GST pull-down experiments.** NE (lane 1) was subjected to pull-down by various GST constructs: GST (lane 2), GST-Gal3 (lane 3), GST-Gal3ND (lane 4), GST-Gal1 (lane 5), GST-Gal3 in the presence of 100 mM Lac (lane 6), and GST-Gal1 in the presence of 100 mM Lac (lane 7). Top panel: immuno-blotting by anti-TFII-I (#558); middle panel: immunoblotting by anti-SMN; and bottom panel: immunoblotting by anti-GST to monitor the amount of GST fusion proteins bound to the glutathione beads. The positions of migration of molecular weight markers are indicated on the right.

Figure 2

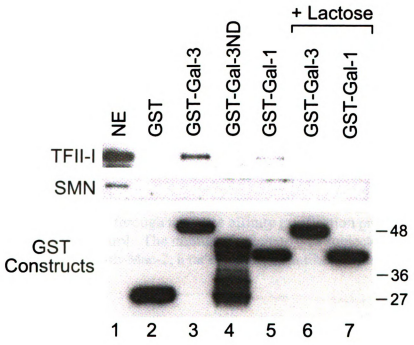
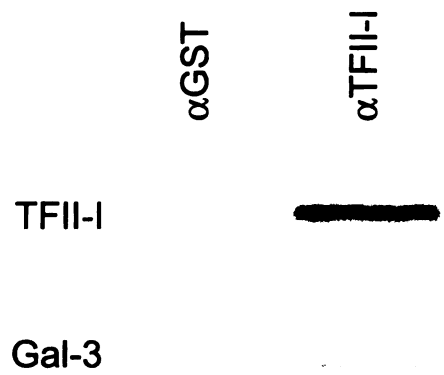

+ Lactose
NE
GST
GST-Gal-3
GST-Gal-3ND
GST-Gal-1
GST-Gal-3
GST-Gal-1
TFII-I
SMN
GST
Constructs
48
36
27
1 2 3 4 5 6 7

**Figure 3. Coimmunoprecitation of Gal3 by anti-TFII-I.**
NE was subjected to immunoprecipitation by anti-TFII-I (Protein Tech Group), a rabbit antiserum affinity purified over the immunogen. Anti-GST, an antiserum that went through the same affinity purification procedure, was used as a negative control. The immunoprecipitate was subjected to blotting with anti-TFII-I and anti-Mac-2, a rat monoclonal antibody directed against Gal3.

Figure 3



procedure. Only a trace amount of Gal3 was observed in the control precipitate (Figure 3).

Inhibition of the splicing reaction by anti-TFII-I

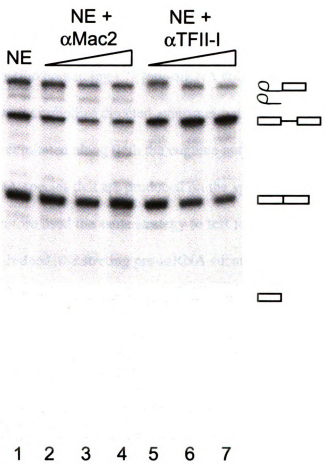
Although it was initially identified as a general transcription factor (25), TFII-I has actually been studied under a wide variety of contexts and therefore, the same polypeptide has acquired a number of different names: (a) Serum Responsive Factor-Phox 1 Interacting Protein (SPIN) (26); and (b) Bruton's tyrosine kinase-associated protein of $M_r$ ~135,000 (BAP135) (27). Inasmuch as we had previously documented that Gal1 and Gal3 are factors involved in pre-mRNA splicing (5, 6, 10), it was of particular interest that a proteomic analysis of the spliceosome identified TFII-I as one of its proteins (11). On this basis, we tested for the association of TFII-I with the splicing machinery and, in particular, the RNA components of the splicing reaction.

When antibodies directed against TFII-I were added to a complete NE during the splicing reaction, a dose-dependent inhibition of product formation was observed. For example, the mRNA product (ligated exons) accounted for ~29% of the radioactivity in the splicing reaction containing NE alone (Figure 4, lane 1). The corresponding values for splicing reactions carried out in the presence of anti-TFII-I were (Figure 4, lanes 5-7): ~31% (12 nM antibody); ~26% (24 nM); and ~22% (30 nM). There was also a dose-dependent decrease in the intermediates of the splicing reaction (e.g. lariat-exon 2). Concomitantly, there was accumulation of the pre-mRNA substrate (Figure 4, compare lanes 1 versus 5-7).

In the accompanying manuscript (28), we have documented that while a monoclonal antibody directed at the $NH_2$-terminal 14 residues of Gal3 inhibited the splicing reaction,

**Figure 4: Effect of anti-TFII-I on the splicing of pre-mRNA.**
Lane 1 - splicing activity of NE (no additions). Lanes 2-4 - splicing activity of NE in the presence of anti-Mac-2, a rat monoclonal specific for Gal3, at 12, 24, and 30 nM. Lanes 5-7 - splicing activity of NE in the presence of anti-TFII-I (#558) at 12, 24, and 30 nM. The cell-free splicing assay was carried out using $^{32}$P-labeled MINX pre-mRNA substrate. Products of the splicing reaction were analyzed by electrophoresis through a 13% polyacrylamide-8.3 M urea gel system, followed by autoradiography. The positions of migration of the pre-mRNA substrate, the splicing intermediates (exon 1 and lariat-exon 2), and the products (ligated exons and lariat intron) are indicated at the right.

Figure 4



the anti-Mac-2 antibody, whose epitope maps to residues 48-100, did not perturb the splicing reaction. On this basis, we used the anti-Mac-2 antibody as negative controls in the present experiments. This anti-Mac-2 antibody did not show inhibition; the percent product formation was 30-32% over the concentration range 12-30 nM (Figure 4, lanes 2-4).

Immunoprecipitation by anti-TFII-I of $^{32}$P-labeled RNA in the splicing reaction

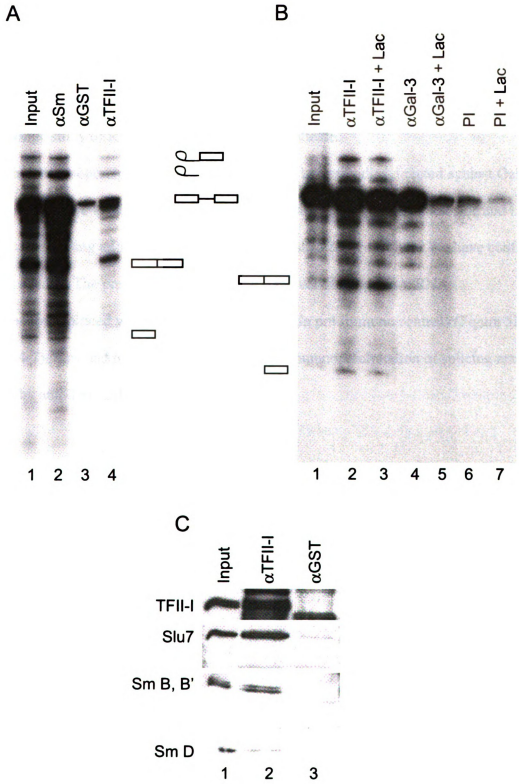
When a splicing reaction containing [$^{32}$P]pre-mRNA substrate was subjected to immunoprecipitation with antibodies directed against either Gal1 or Gal3, spliceosomal complex(es) were coprecipitated along with the cognate antigen, a conclusion based on finding $^{32}$P-labeled RNA species that are produced on the spliceosome during the splicing reaction (10). We therefore used the same strategy to test for an association of TFII-I with the spliceosome. Indeed, the starting pre-mRNA substrate, the products of the splicing reaction (ligated exon and lariat intron), as well as the intermediates (e.g. lariat-exon 2) were all observed in the anti-TFII-I precipitate (Figure 5A, lane 4). In contrast, much less radioactivity, corresponding only to the pre-mRNA substrate species, was observed in the precipitate of the control anti-GST antibody (Figure 5A, lane 3). The profile of RNA precipitation observed with anti-TFII-I was also found with human autoimmune serum reactive against the Sm epitopes of the core polypeptides of snRNPs, which served as a positive control (Figure 5A, lane 2).

Along with the analysis for $^{32}$P-labeled RNA species, parallel samples of the immunoprecipitate were subjected to SDS-PAGE and immunoblotting (Figure 5C). In addition to its own cognate antigen, the anti-TFII-I precipitate also yielded positive reactions with: (a) the Sm epitopes of snRNPs, and (b) Slu7, a factor required for the

**Figure 5. Analysis of spliceosomal RNA species and proteins immunoprecipitated by various antisera.**

Parallel splicing reactions incubated for 40 minutes with $^{32}$P-labeled MINX were subjected to antibody adsorption and the bound RNA was analyzed by electrophoresis through a 13% polyacrylamide- 8.3 M urea gel system, followed by autoradiography. Panel A: lane 1 - the splicing reaction mixture (4%) that was subjected to immunoprecipitation; lane 2 - immunoprecipitate of anti-Sm; lane 3 - immunoprecipitate of affinity-purified anti-GST; lane 4 – immunoprecipiate of anti-TFII-I (#558). Panel B: lane 1 - the splicing reaction mixture (0.8%) that was subjected to immunoprecipitation; lane 2 – immunoprecipitation by anti-TFII-I (#558) carried out in the absence of Lac; lane 3 - immunoprecipitation by anti-TFII-I (#558) carried out in the presence of 100 mM Lac; lane 4 - immunoprecipitation by anti-Gal3 (#49) carried out in the absence of Lac; lane 5 - immunoprecipitation by anti-Gal3 (#49) carried out in the presence of 100 mM Lac; lane 6 - immunoprecipitation by preimmune normal rabbit serum (#49) in the absence of Lac; and lane 7 – immunoprecipitation by preimmune serum (#49) in the presence of 100 mM Lac. Panel C: Following a splicing reaction, the sample (input, lane 1) was subjected to immunoprecipitation by anti-TFII-I (lane 2) or anti-GST (lane 3). Top panel: immunoblotting by anti-TFII-I (#558); middle panel: immunoblotting by anti-Slu7; bottom panel: immunoblotting by anti-Sm.

Figure 5


A
Input
αSm
αGST
αTFII-I
ρ
ρ
1 2 3 4
B
Input
αTFII-I
αTFII-I + Lac
αGal-3
αGal-3 + Lac
PI
PI + Lac
1 2 3 4 5 6 7
C
Input
αTFII-I
αGST
TFII-I
Slu7
Sm B, B'
Sm D
1 2 3

second trans-esterification reaction during splicing (29), representing a late stage splicing complex (Figure 5C, lane 2). In contrast, no reaction was observed when an antibody directed against Ran (which is associated with nuclear transport) was used to immunoblot the anti-TFII-I precipitate. On the basis of analysis of both RNA and proteins species, therefore, it appears that TFII-I is a *bona fide* component of spliceosomes, as the proteomic study of Rappsilber *et al.* (11) had implicated.

The immunoprecipitation of $^{32}$P-labeled RNA by antibodies directed against Gal1 or Gal3 is sensitive to inhibition by saccharide ligands such as thiodigalactoside and Lac (Weizhong Wang and Ronald Patterson, unpublished observations). We have confirmed these results. The presence of Lac reduced the level of radioactive RNA immunoprecipitated with anti-Gal3 to that found in pre-immune controls (Figure 5B, lanes 4-7). Lac had no effect, however, on the immunoprecipitation of splicing reaction RNA by anti-TFII-I (Figure 5B, lanes 2 and 3).

## DISCUSSION

The key findings of the present study include: (a) When NEs of HeLa cells were subjected to pull-down experiments with GST-Gal1 or GST-Gal3, the general transcription factor TFII-I was identified as one of the polypeptides bound. (b) Lac, a saccharide ligand of the galectin family of proteins, inhibited the pull-down of TFII-I from NE. (c) Antibodies directed against TFII-I inhibited the splicing reaction when added to a splicing competent NE. (d) Antibodies directed against TFII-I also immunoprecipitated spliceosomal RNA species from a splicing reaction mixture containing $^{32}$P-labeled pre-mRNA. These results suggest that TFII-I associates with Gal1- or Gal3-containing spliceosomal complexes.

TFII-I had been identified as a component of the spliceosome by proteomic analysis (11). Our present demonstration, that anti-TFII-I can exert a perturbation effect on the splicing assay and that it can coprecipitate RNA species produced during the splicing reaction, complement the chemical studies in implicating TFII-I as a part of the functional spliceosome. There are now several lines of evidence linking transcription and pre-mRNA splicing (see reference 30 for a review). First, there are transcription-related effects on the pattern of splicing products. For example, the use of different promoters by RNA polymerase II results in different splicing patterns for the same transcript (31). In *Drosophila,* mutants that modulate the speed of transcription result in differential splicing of transcripts (32). In addition to these promoter and kinetic effects, there are also direct physical links between components of the transcription and splicing machinery. For example, the carboxyl-terminal domain of RNA polymerase II (CTD) can stimulate splicing *in vitro* (33, 34) and interacts with both snRNPs (35) and the SMN

protein that is responsible for snRNP biogenesis (36). One functional coupling between RNA polymerase II transcription and splicing was revealed by the recent study of Das *et al.* (37), who showed that CTD exerts its effect by stimulating a bypass of the H-complex step of spliceosome assembly.

In addition to the general transcription factor TFII-I, it has also been reported that Gal3 can interact with the thyroid specific transcription factor TTF-1 in the nuclei of papillary thyroid cancer cells (38). GST pull-down assays demonstrated a direct interaction between Gal3 and the homeodomain of TTF-1. In addition, gel retardation assays showed that this interaction stimulated the DNA-binding activity of TTF-1. Thus, Gal3 can up-regulate the transcriptional activity of TTF-1, contributing to the proliferation of the thyroid cells.

Gal3 also interacts with the protein Sufu (Suppressor of fused), a negative regulator of the Hedgehog signal transduction pathway that binds directly with the Gli family of transcription factors (39). The Sufu polypeptide contains a functional leucine-rich nuclear export signal and the fusion protein derived from Sufu and green fluorescent protein is found predominantly in the cytoplasm of transfected HeLa cells. As expected, mutants of Sufu in which the nuclear export signal has been inactivated (Sufu(L383A; L385A)) localized mostly to the nucleus. When co-transfected with Gal3, however, the same Sufu(L383A; L385A) mutant was found in the cytoplasm, colocalized with Gal3. Thus, the possibility was raised that Gal3 plays a role in the nuclear versus cytoplasmic distribution of the transcriptional regulator.

One other ligand of Gal3, the cysteine- and histidine-rich protein (Chrp), deserves mention. Chrp was initially identified in a yeast two-hybrid screen of a murine 3T3 cell

cDNA library using Gal3 as the bait (40). Direct interaction between Gal3 and Chrp was confirmed by *in vitro* binding assays. Immunofluorescence analysis revealed that, in 3T3 cells, Chrp was distributed throughout the cytoplasm but was especially concentrated in a concentric ring at the nuclear envelope. Chrp binds to the CRD of Gal3. Nevertheless, Gal3, in complex with Chrp, can still bind to glycoconjugate ligands, including the glycoprotein laminin (41). Therefore, the data suggest that the CRD of Gal3 can simultaneously accommodate two ligands: saccharide and Chrp.

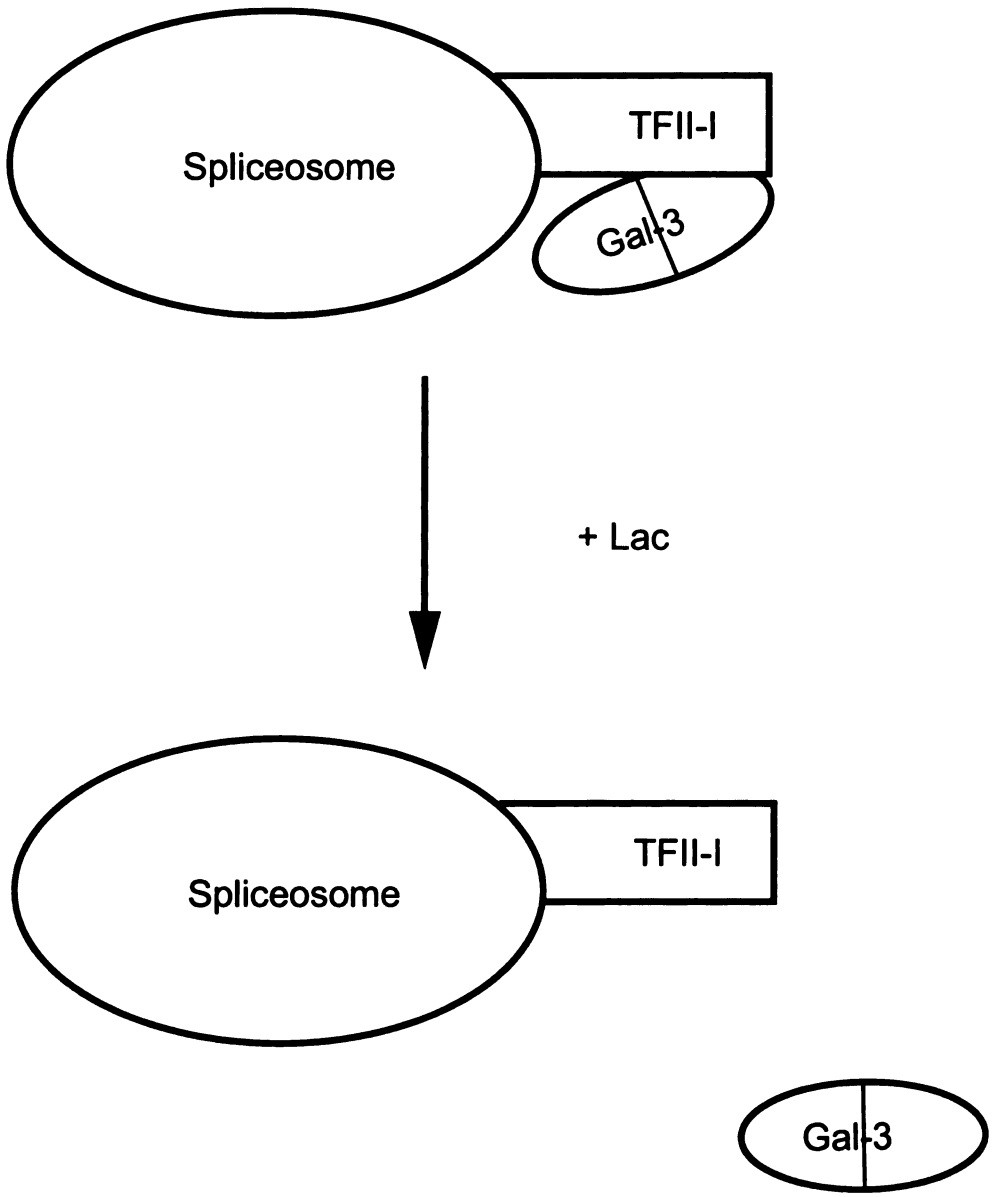
This may also apply to the binding of saccharides and TFII-I to the CRD of galectins as well. In the accompanying article (12), we document that mutants of Gal1 devoid of saccharide-binding activity can still carry out the pull-down of TFII-I from NE as well as reconstitute splicing in NE depleted of the galectins. Thus, it appears that we can dissociate the carbohydrate-binding activity of the CRD in the Gal1 polypeptide from its association with spliceosomal components. Extended to the CRD of Gal3, this would be consistent with the notion that there are separate binding sites for saccharides and for TFII-I within the CRD.

On the other hand, the inhibition by Lac of the GST-Gal3 (and GST-Gal1) pull-down of TFII-I suggests that the binding of TFII-I and saccharide ligands to Gal3 is mutually exclusive and that the two ligands may compete for the same binding surface within the galectin CRD. One possibility is that binding of carbohydrates to the CRD induces a conformational change in the galectin polypeptide that precludes association with TFII-I. Evidence for a conformational change upon saccharide-binding has been reported for Gal3, using differential scanning calorimetry (20) and NMR spectroscopy (42).

In any case, the disruption by Lac of the galectin-TFII-I association provided an opportunity to probe the order with which TFII-I and Gal3 is recruited to the spliceosome. Lac inhibits both the immunoprecipitation of spliceosomal RNA by anti-Gal3 as well as the GST-Gal3 pull-down of TFII-I. In contrast, the saccharide had no effect on the immunoprecipitation of spliceosomal RNA by anti-TFII-I. It appears, therefore, that TFII-I is bound to the spliceosome independent of Gal3 (Figure 6). On the other hand, it appears that Gal3 recruitment to the spliceosome is dependent on its association with TFII-I, which is disrupted by the binding of Lac. These results provide the basis for a more detailed characterization of the galectin-TFII-I association. In particular, the key question remains whether Gal3 (and Gal1) binds directly to TFII-I and if the interaction is indirect, what other components might mediate this association.

**Figure 6. Schematic illustration of the association of TFII-I and Gal3 with the spliceosomal complex.**
The association of TFII-I with the spliceosome appears to be independent of its interaction with Gal3, which is disrupted by Lac binding. On the other hand, the association of Gal3 with the spliceosome may be mediated through its interaction with TFII-I as both are disrupted by Lac binding. Although the present schematic depicts it as direct binding, whether the association between TFII-I and Gal3 is direct or indirect is not known.

Figure 6



## REFERENCES


1. **Barondes, S. H., V. Castronovo, D. N. W. Cooper, R. D. Cummings, K. Drickamer, T. Feizi, M. A. Gitt, J. Hirabayashi, C. Hughes, and K. Kasai.** 1994. Galectins: a family of animal beta-galactoside-binding lectins. Cell **76:**597-8.

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

3. **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:**419-28.

4. **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.

5. **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. U.S.A. **92:**1213-7.

6. **Vyakarnam, A., S. F. Dagher, J. L. Wang, and R. J. Patterson.** 1997. Evidence for a role for galectin-1 in pre-mRNA splicing. Mol. Cell. Biol. **17:**4730-7.

7. **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:**1008-1020.

8. **Hastings, M., and A. Krainer.** 2001. Pre-mRNA splicing in the new millennium. Curr. Opin. Cell Biol. **13:**302-9.

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

10. **Wang, W., J. W. Park, J. L. Wang, and R. J. Patterson.** 2006. Nucleic Acids Res.

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

12. **Voss, P. G., S. W. Dickey, W. Wang, R. M. Gray, K. Kasai, J. Hirabayashi, R. J. Patterson, and J. L. Wang.** 2006. Interactions of Galectin-1 and Galectin-3 in pre-mRNA Splicing II. Dissociation of the Splicing and the Carbohydrate-binding Activities of Galectin-1. accompanying article, paper II.

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

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

15. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475-1489.

16. **Bradford, 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:**248-54.

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

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

19. **Agrwal, N., J. L. Wang, and P. G. Voss.** 1989. Carbohydrate-binding protein 35. Levels of transcription and mRNA accumulation in quiescent and proliferating cells. J. Biol. Chem. **264:**17236-17242.

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

21. **Laemmli, U.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-5.

22. **Merril, C., M. Dunau, and D. Goldman.** 1981. A rapid sensitive silver stain for polypeptides in polyacrylamide gels. Anal Biochem **110:**201-7.

23. **Shevchenko, A., M. Wilm, O. Vorm, and M. Mann.** 1996. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem **68:**850-8.

24. **Leffler, H., and S. H. Barondes.** 1986. Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. J. Biol. Chem. **261:**10119-26.

25. **Roy, A., M. Meisterernst, P. Pognonec, and R. Roeder.** 1991. Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature **354:**245-8.

26. **Grueneberg, D. A., R. W. Henry, A. Brauer, C. D. Novina, V. Cheriyath, A. L. Roy, and M. Gilman.** 1997. A multifunctional DNA-binding protein that promotes the formation of serum response factor/homeodomain complexes: identity to TFII-I. Genes Dev. **11:**2482-2493.

27. **Yang, W., and S. Desiderio.** 1997. BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement. Proc. Natl. Acad. Sci. U.S.A. **94:**604-609.

28. **Gray, R. M., M. J. Davis, K. M. Ruby, P. G. Voss, R. J. Patterson, and J. L. Wang.** 2006. Accompanying article, paper III.

29. **Butcher, S., and D. Brow.** 2005. Towards understanding the catalytic core structure of the spliceosome. Biochem. Soc. Trans. **33:**447-9.

30. **Neugebauer, K. M.** 2002. On the importance of being co-transcriptional. J. Cell Sci. **115:**3865-3871.

31. **Cramer, P., J. Caceres, D. Cazalla, S. Kadener, A. Muro, F. Baralle, and A. Kornblihtt.** 1999. Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. Mol. Cell **4:**251-8.

32. **Chen, Y., D. Chafin, D. H. Price, and A. L. Greenleaf.** 1996. Drosophila RNA Polymerase II Mutants That Affect Transcription Elongation. J. Biol. Chem. **271:**5993-5999.

33. **Hirose, Y., R. Tacke, and J. L. Manley.** 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. Genes Dev. **13:**1234-1239.

34. **Zeng, C., and S. M. Berget.** 2000. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. Mol. Cell. Biol. **20:**8290-8301.

35. **Robert, F., M. Blanchette, O. Maes, B. Chabot, and B. Coulombe.** 2002. A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. J. Biol. Chem. **277:**9302-9306.

36. **Pellizzoni, L., B. Charroux, J. Rappsilber, M. Mann, and G. Dreyfuss.** 2001. A Functional Interaction between the Survival Motor Neuron Complex and RNA Polymerase II. J. Cell Biol. **152:**75-86.

37. **Das, R., K. Dufu, B. Romney, M. Feldt, M. Elenko, and R. Reed.** 2006. Functional coupling of RNAP II transcription to spliceosome assembly. Genes Dev. **20:**1100-1109.

38. **Paron, I., A. Scaloni, A. Pines, A. Bachi, F. Liu, C. Puppin, M. Pandolfi, L. Ledda, C. Di Loreto, G. Damante, and G. Tell.** 2003. Nuclear localization of

galectin-3 in transformed thyroid cells: a role in transcriptional regulation. Biochem. Biophys. Res. Cummun. **302:**545-53.

39. **Paces-Fessy, M., D. Boucher, E. Petit, S. Paute-Briand, and M. Blanchet-Tournier.** 2004. The negative regulator of Gli, Suppressor of fused (Sufu), interacts with SAP18, Galectin3 and other nuclear proteins. Biochem. J. **378:**353-62.

40. **Menon, R. P., M. Strom, and R. C. Hughes.** 2000. Interaction of a novel cysteine and histidine-rich cytoplasmic protein with galectin-3 in a carbohydrate-independent manner. FEBS Lett. **470:**227-31.

41. **Bawumia, S., E. A. Barboni, R. P. Menon, and R. Colin Hughes.** 2003. Specificity of interactions of galectin-3 with Chrp, a cysteine- and histidine-rich cytoplasmic protein. Biochimie **85:**189-94.

42. **Umemoto, K., H. Leffler, A. Venot, H. Valafar, and J. H. Prestegard.** 2003. Conformational differences in liganded and unliganded states of Galectin-3. Biochemistry **42:**3688-95.

# CHAPTER 3

## Dissociation of the Splicing and the Carbohydrate-binding Activities of Galectin-1

## ABSTRACT

Three site-directed mutants of galectin-1 (Gal1) were expressed and purified as fusion proteins with glutathione S-transferase (GST). These mutants, designated as GST-Gal1(N46D), GST-Gal1(C60S), and GST-Gal1(E71Q), were compared with the corresponding fusion protein containing wild-type Gal1, GST-Gal1(WT), in three assays: (a) binding to asialofetuin-Sepharose as a measure of the carbohydrate-binding activity; (b) pull-down of transcription factor II-I (TFII-I) from nuclear extract of HeLa cells; and (c) reconstitution of splicing in HeLa nuclear extract depleted of galectins as a test of the *in vitro* splicing activity. The binding of GST-Gal1(N46D) to asialofetuin-Sepharose was less than 10% of that observed for GST-Gal1(WT), indicating that the mutant was deficient in carbohydrate-binding activity. Both GST-Gal1(WT) and GST-Gal1(N46D) were equally efficient, however, in pull-down of TFII-I and in reconstitution of splicing activity in a galectin-depleted nuclear extract. GST-Gal1(C60S) retained all three activities and behaved similarly to the wild-type protein. Finally, GST-Gal1(E71Q) exhibited drastically reduced activities in all three assays, suggesting that the mutation may have caused misfolding of the polypeptide. Together, all of the results suggest that the carbohydrate-binding and the splicing activities of Gal1 can be dissociated and therefore, saccharide-binding, *per se*, is not required for the splicing activity.

## INTRODUCTION

Using the criteria of depletion and reconstitution, we had documented in previous studies that galectin-1 (Gal1) and galectin-3 (Gal3) are nuclear proteins involved in the splicing of pre-mRNA, assayed in a cell-free system (1-3). In the course of these studies, we observed that saccharides which bind to Gal1 and Gal3 with high affinity, such as lactose (Lac) and thiodigalactoside (TDG), inhibited the splicing reaction when added to a complete nuclear extract (NE) (1). In contrast, saccharides that do not bind to Gal1 and Gal3, such as cellobiose, had no effect. Does inhibition of splicing by Lac and TDG indicate that the nuclear ligand for the galectins interacts with the carbohydrate-binding site and that it is competitively displaced upon Lac/TDG addition? Or, does Lac/TDG addition alter the conformation of the galectins, causing the release or altered interaction with a nuclear ligand bound via protein-protein, rather than protein-carbohydrate interactions?

The availability of site-directed mutants of Gal1 devoid of carbohydrate-binding activity (4, 5) provided the key reagents to test whether saccharide-binding *per se* is necessary for the splicing activity. In any site-directed mutagenesis experiment, however, a critical issue is whether the mutation has caused a disruption of the overall folding of the polypeptide. Our recent discovery (6) of the association of Gal1 and Gal3 with the general transcription factor II-I (TFII-I), in turn, provided the opportunity for an independent monitor of the folding of the wild-type and mutant polypeptides.

In the present communication, we report the identification of a site-directed mutant of Gal1, deficient in carbohydrate-binding activity, that is still capable of functioning in the

splicing assay. The results suggest that the association of Gal1 with TFII-I and its splicing function are independent of the carbohydrate-binding activity of the polypeptide.

## EXPERIMENTAL PROCEDURES

Glutathione S-transferase (GST)-Fusion Proteins --- The cDNAs for wild-type (WT) and three mutants (N46D, C60S, and E71Q) of human Gal1 have been described (4, 5). Each cDNA was subcloned into the BamHI restriction site of the pGEX-2T vector (Pharmacia) to produce a fusion protein between GST and Gal1, hereafter designated, respectively: GST-Gal1(WT), GST-Gal1(N46D), GST-Gal1(C60S), and GST-Gal1(E71Q). Each of the constructs was subjected to DNA sequencing to verify: (a) the juncture of the fusion protein between GST and Gal1; and (b) the wild-type and mutant amino acid at the mutagenized residue.

The constructs were transformed into *E. coli* (strain BL21-codon Plus (DE3); Stratagene) and the GST fusion proteins were purified from 1 liter cultures on the basis of GST binding to glutathione-agarose beads (Pierce). The purity of the protein preparations was assessed by SDS-PAGE (7). The polypeptides were revealed by silver staining (8) or by immunoblotting (6), following procedures that have been described previously. Affinity purified polyclonal rabbit anti-Gal1 and anti-GST antibodies were used for the immunoblotting.

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

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

Assay of Carbohydrate-binding Activity The preparation of the affinity beads, asialofetuin (ASF)-Sepharose 4B, has been described (9). Approximately 180 nmoles of ASF were coupled per ml of Sepharose beads. GST proteins (350 ng each of GST, GST-Gal1 (wild-type and three mutants)) were incubated with 35 μl of ASF-Sepharose for 2 hours at 4 °C. The incubations were carried out in 60% buffer D (buffer D is 20 mM Hepes, pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM phenylmethylsufonyl fluoride, and 0.5 mM dithiothreitol) containing 0.1% NP-40 (Pierce) in the presence and absence of Lac (100 mM). The material not bound to the beads was removed by centrifugation (1000 x g); after resuspension, the beads were washed four times in 60%

buffer D containing 0.1% NP-40. The GST proteins bound to the beads were subjected to SDS-PAGE and immunoblotting with anti-GST and anti-Gal1 antibodies.

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

GST pull-down of TFII-I from NE The assay was carried out as described in the previous article (7). Approximately 10 μg of each GST protein was incubated overnight with 20 μl of packed glutathione-agarose beads (Sigma) in 60% buffer D at 4 °C. After removal of unbound material and washing, 36 μl (~144 μg total protein) of NE was added to the beads with 24 μl of 60% buffer D containing 14.7 mM creatine phosphate, 2.4 mM $MgCl_2$, and 0.4 mM ATP. In experiments to test the effect of Lac, the saccharide was added to a final concentration of 100 mM. The mixture was incubated at 4 °C for 12 hours. After removal of unbound material and washing, the bound fraction was eluted with 20 μl of glutathione elution buffer (16 mM glutathione, 60 mM HEPES-KOH, pH 7.9, 11.4% glycerol, 57 mM KCl, and 114 μM EDTA) at 31 °C for 30 minutes, followed by 1 hour at room temperature. The eluted material was then subjected to SDS-PAGE and immunoblotting analysis with anti-TFII-I (#558, Bethyl Laboratories) (6).

Assay of Splicing Activity HeLa S3 cells were grown in suspension culture by the National Cell Culture Center (Minneapolis, MN). NE was prepared in buffer C (20 mM Hepes, pH. 7.9, 25% glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM phenylmethylsufonyl fluoride, 0.5 mM dithiothreitol), as described by Dignam *et al.* (10). Protein concentrations were determined by the Bradford assay (11); in this study, the protein concentration of the NE was ~4 mg/ml.

NEs were depleted of Gal1 and Gal3 by adsorption on beads covalently coupled with rabbit anti-Gal1 (#55) and rat anti-Mac-2 (2). Anti-Mac-2 (anti-M2) is a rat monoclonal antibody specific against Gal3 (12, 13). The beads (150 µl) were washed with 20 mM Hepes, pH 7.9, 0.5 M NaCl; 30 µl NE were added and incubated on ice for 20 minutes in disposable spin columns (Millipore). The unbound fraction was removed and the beads were washed with 12 µl of 60% buffer D adjusted to 0.42 M NaCl and this wash was combined with the unbound fraction. Aliquots of nondepleted NE and unbound fractions of the depletion were dialyzed in a microdialyzer against 60% buffer D for 40 minutes at 4 °C using a dialysis membrane with a 10 kD cutoff. In reconstitution experiments, GST or GST-Gal1 (wild-type and mutant proteins) were added to the unbound fractions prior to dialysis. The dialyzed fractions were then assayed for splicing activity.

The plasmid used to transcribe the MINX pre-mRNA substrate was obtained from Dr. Susan Berget (Baylor College of Medicine, Houston, TX) (14). The MINX pre-mRNA was labeled with [$^{32}$P]GTP and the monomethyl cap was added during SP6 polymerase (Gibco BRL) transcription. Splicing reaction mixtures contained a total volume of 10 µl: dialyzed NE (4 µl) or unbound fraction (8 µl), [$^{32}$P]MINX pre-mRNA, 2.5 mM $MgCl_2$, 1 mM ATP, 20 mM creatine phosphate, 0.5 mM dithiothreitol, and 20 U of RNasin

(Promega). Splicing reactions were incubated at 30 °C for 45 minutes. Proteinase K-SDS solution was added to a final concentration of 4 mg/ml and 0.1%, respectively. The sample was incubated at 37 °C for 15 minutes. Each sample was then diluted to 100 µl with 125 mM Tris, pH 8, 1 mM EDTA, 0.3 M sodium acetate. RNA was extracted with 200 µl of phenol-chloroform (50:50 (vol/vol)), followed by 200 µl chloroform. RNAs were precipitated with 400 µl of ethanol at -80 °C. The extracted RNAs were subjected to electrophoresis through 13% polyacrylamide (bisacrylamide-acrylamide, 1.9:50 (wt/wt)), 8.3 M urea gels, followed by autoradiography.

Quantitation of the amount of radioactivity was carried out on a STORM phosphorimager (Molecular Dynamics). The percent product formation was calculated by dividing the radioactivity present in the final product (ligated exon 1-exon 2) by the total radioactivity present in the pre-mRNA substrate, the splicing intermediates (lariat-exon 2 and free exon 1), and the products (lariat intron and ligated exon 1-exon 2).

## RESULTS

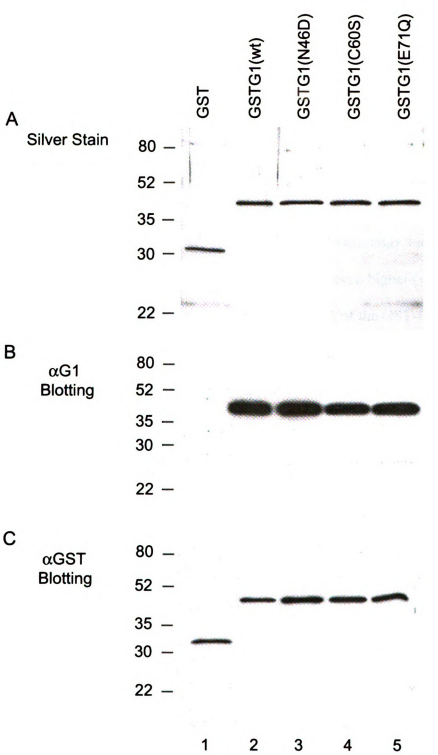
GST-Gal1(N46D) and GST-Gal1(E71Q) are deficient in carbohydrate-binding activity

Hirabayashi and Kasai (4, 5) had demonstrated that substitutions at highly conserved residues of the carbohydrate recognition domain of human Gal1, such as Asn 46 and Glu 71, resulted in loss of saccharide-binding activity. The cDNAs corresponding to WT and the N46D and E71Q mutants were each subcloned into the pGEX-2T vector to produce a fusion protein between GST and Gal1 so that we can take advantage of the easy purification of the proteins on the basis of GST binding to glutathione beads. As a control, we also carried out the same analysis on another mutant, C60S, whose substitution did not abolish the carbohydrate-binding activity (4, 5). Nucleotide sequence analysis of the constructs verified the sequence of the fusion proteins; in particular, residue 46, 60, and 71 in the Gal1 sequence was confirmed to be Asp, Ser, and Gln, respectively, for GST-Gal1(N46D), GST-Gal1(C60S), and GST-Gal1(E71Q). In parallel, we also expressed GST from the same vector to be carried as a control in the functional tests below.

On SDS-PAGE and silver staining (Figure1A), each of the purified GST proteins yielded a single band with a mobility corresponding to the expected molecular weights: GST, ~27 kD; and GST-Gal1 (wild-type and mutants), ~42 kD. Immunoblotting with affinity purified anti-Gal1 antibodies yielded a single band at the same molecular weight for GST-Gal1 (wild-type and mutant proteins) (Figure 1B, lanes 2-5). No reaction was observed between anti-Gal1 and GST (Figure 1B, lane 1). Finally, immunoblotting of the respective GST protein preparations with affinity purified anti-GST antibodies yielded the same single band patterns as were observed by

**Figure 1. Characterization of the preparations of fusion proteins containing wild-type or mutant Gal1 by SDS-PAGE.**
Lane 1: GST; lane 2, GST-Gal1(WT); lane 3, GST-Gal1(N46D); lane 4, GST-Gal1(C60S); and lane 5, GST-Gal1 (E71Q). The proteins (~30 ng in each lane) were electrophoresed through 12.5% acrylamide gels. Panel A: silver staining; Panel B: immunoblotting with affinity purified anti-Gal1 antibodies (#55); and Panel C: immnoblotting with affinity purified anti-GST antibodies. The binding of the primary antibodies in panels B and C were revealed with horseadish peroxidase-conjugated goat anti-rabbit immunoglobulin and the enhanced chemiluminescence system. The positions of migration of molecular weight standards (80 kD, 52 kD, 35 kD, 30 kD, and 22 kD) are indicated on the left.

Figure 1


GST
GSTG1(wt)
GSTG1(N46D)
GSTG1(C60S)
GSTG1(E71Q)
A
Silver Stain
80
52
35
30
22
B
αG1
Blotting
80
52
35
30
22
C
αGST
Blotting
80
52
35
30
22
1
2
3
4
5

silver staining (Figure 1C). All of these results establish the purity of the protein reagents to be compared in the functional assays below.

Purified GST-Gal1(WT) was compared to the mutant counterparts in terms of their binding to ASF-Sepharose beads. After washing, the material bound to the beads was subjected to SDS-PAGE and immunoblotting with anti-GST antibodies. Using known amounts of GST-Gal1 to establish standard curves, we quantitated the Lac-inhibitable binding, as well as the binding not inhibitable by Lac. The latter accounted for about 3% of the total binding observed for both the wild-type and mutant proteins. In terms of Lac-inhibitable binding, 10-15% of the GST-Gal1(WT) added to the assay was bound specifically; the level of binding for GST-Gal1(C60S) was even higher (~23% of the added protein bound specifically). In contrast, less than 1% of the GST-Gal1(N46D) and GST-Gal1(E71Q) added to the assay was bound. The binding of these two mutants was drastically reduced relative to the wild-type protein and was essentially the same as that observed for GST alone (Figure 2). The same overall conclusion was obtained using either anti-GST antibodies or anti-Gal1 antibodies for the quantitation (the latter obviously could not detect GST itself, as was shown in Fig 1B, lane 1).

Comparison of the GST pull-down of TFII-I by wild-type and mutant Gal1

In the preceding article of this series, we had documented that both GST-Gal3 and GST-Gal1 can pull-down TFII-I from NE (6). A comparison was made, therefore, of the GST pull-down of TFII-I by wild-type and mutant Gal1 proteins. GST-Gal1(WT), GST-Gal1(N46D) and GST-Gal1(C60S) yielded good signals for TFII-I in the pull-down assay (Figure 3A, lanes 2-4). The signal for TFII-I in the GST-Gal1(E71Q) pull-down was substantially weaker; however, it was nevertheless still above the background level

**Figure 2. Comparison of the carbohydrate-binding activity of GST-Gal1(WT), GST-Gal1(N46D), GST-Gal1(C60S), and GST-Gal1(E71Q).**
The GST proteins (350 ng) were incubated with ASF-Sepharose for 2 hours at 4 °C in 60% buffer D containing 0.1% NP-40. Parallel incubations were carried out in the presence and absence of 100 mM Lac. Proteins bound to the ASF-beads were quantitated by immunoblotting with anti-GST. The values shown represent Lac inhibitable specific binding.

Figure 2

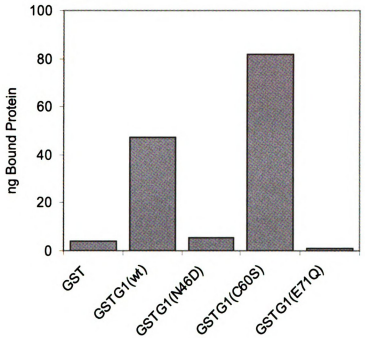
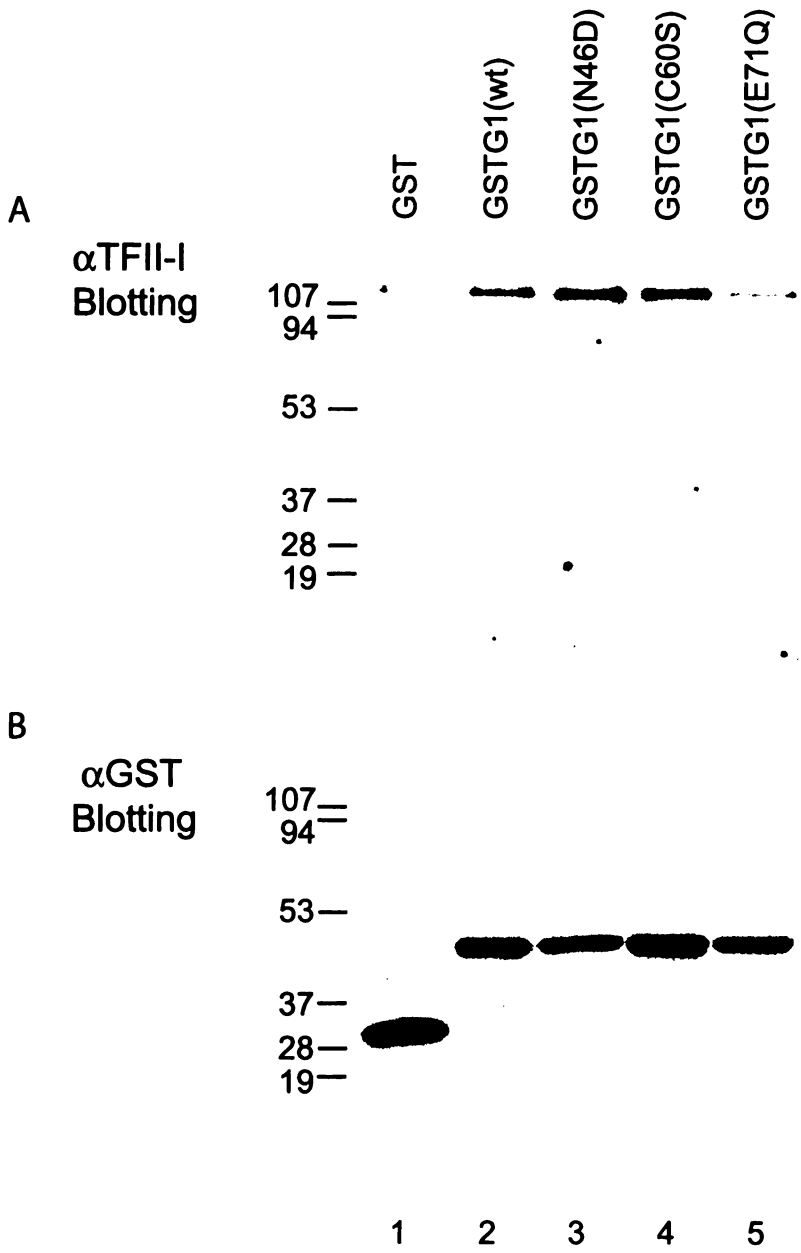

100
80
60
40
20
0
ng Bound Protein
GST
GSTG1(wt)
GSTG1(N46D)
GSTG1(C60S)
GSTG1(E71Q)

**Figure 3: Comparison of the GST pull-down of TFII-I in NE by wild-type and mutant Gal1 polypeptides.**
The GST proteins were incubated overnight with glutathione beads at 4 °C in 60% buffer D. After washing to remove the unbound material, the beads were then incubated with NE (36 µl) for 12 hours at 4 °C. Material bound to the various beads was eluted with glutathione (16 mM) and analyzed by SDS-PAGE and immunoblotting. Panel A: immunoblotting with anti-TFII-I (#558). Panel B: immunoblotting with anti-GST to ascertain that approximately equal amounts of GST proteins were bound to the beads in the pull-down assay.

Figure 3


A
αTFII-I
Blotting
GST
GSTG1(wt)
GSTG1(N46D)
GSTG1(C60S)
GSTG1(E71Q)
107
94
53
37
28
19
B
αGST
Blotting
107
94
53
37
28
19
1
2
3
4
5

observed for GST alone (Figure 3A, lanes 5 versus 1). Control immunoblots with anti-GST showed that approximately equal amounts of GST proteins were bound to the glutathione beads in the pull-down assay (Figure 3B).

Thus, it appears that GST-Gal1(C60S) retained both carbohydrate and TFII-I binding activities observed with GST-Gal1(WT). For GST-Gal1(N46D), on the other hand, the mutant polypeptide must have retained sufficient structure to preserve the association with TFII-I while the saccharide-binding activity was compromised. This argues against the notion that the single amino acid substitution resulted in gross misfolding of the polypeptide. Finally, it appears that the mutation in GST-Gal1(E71Q) reduced both the carbohydrate-binding activity and its association with TFII-I, possibly reflecting unfolding of the polypeptide.

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

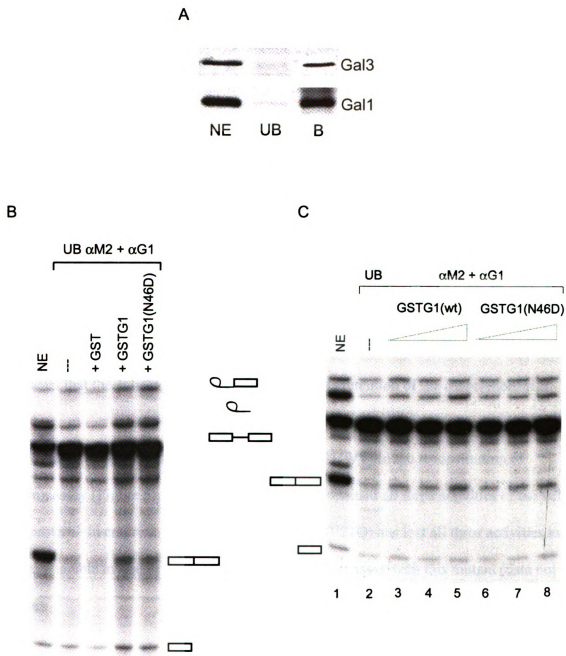
NE was prepared in buffer C, which contained 0.42 M NaCl to dissociate splicing or ribonucleoprotein complexes. This NE was incubated with beads covalently coupled with anti-Gal1 and anti-Gal3 (the anti-M2 monoclonal). Western blotting analysis documented that both proteins were present in the NE and in the bound fraction of the antibody beads. Only trace amounts of either Gal1 or Gal3 remained in the unbound fraction (Figure 4A).

NE depleted of the galectins showed reduced splicing activity (Figure 4B, lanes 1-2). Product mRNA formation was decreased from ~20% to ~5%. We had previously documented that GST, by itself, had little or no effect on the cell-free splicing activity (15) so the effects of purified GST-Gal1(WT) and GST-Gal1(N46D) could be directly

**Figure 4. Comparison of the splicing activities of NE, NE after depletion of Gal1 and Gal3, and depleted NE reconstituted with GST, GST-Gal1(WT), and GST-Gal1(N46D).**

NE was depleted of Gal1 and Gal3 by adsorption on beads covalently coupled with anti-Gal1 and anti-M2 (rat monoclonal antibody against Gal3). Panel A: Immunoblotting for Gal1 and Gal3 in NE, the unbound (UB) fraction of the double antibody adsorption, and the bound (B) fraction. The amount of material electrophoresed in the NE and UB lanes represents ~41% of the amount electrophoresed in the B lane. Panel B: Splicing of $^{32}$P-labled MINX pre-mRNA. Lane 1, the complete (non-depleted) NE; lane 2, the unbound (UB) fraction of the double antibody adsorption; lanes 3-5, depleted extract reconstituted with GST, GST-Gal1(WT), and GST-Gal1(N46D), respectively. The concentration of the GST proteins was 6.5 μM. Panel C: Dose-response of the reconstitution of splicing activity by GST-Gal1(WT) and GST-Gal1(N46D). Lane 1, the complete NE; lane 2, the unbound (UB) fraction of the double antibody adsorption; lanes 3-5, reconstitution with 1 μM, 6.5 μM, and 13 μM GST-Gal1(WT); and lane 6-8, reconstitution with 1 μM, 6.5 μM, and 13 μM GST-Gal1(N46D). In panels B and C, products of the splicing reaction were analyzed by electrophoresis through a 13% polyacrylamide-urea gel and autoradiography. The positions of migration of the pre-mRNA substrate, the splicing intermediates (exon 1 and lariat-exon 2), and RNA products (ligated exon 1-exon 2 and intron lariat) are indicated at the center.

Figure 4


A
Gal3
Gal1
NE
UB
B
B
UB αM2 + αG1
NE
---
+ GST
+ GSTG1
+ GSTG1(N46D)
C
UB
αM2 + αG1
GSTG1(wt)
GSTG1(N46D)
NE
---
1
2
3
4
5
6
7
8

compared. The same preparations of the wild-type and mutant proteins that exhibited drastic differences in saccharide-binding activity (Figure 2) both reconstituted splicing in the galectin-depleted NE (Figure 4B, lanes 4-5). The concentration of GST proteins used in the reconstitution was 6.5 μM. At this concentration, the level of product formation (~10%) in the reconstitution assay was comparable to that achieved previously with recombinant Gal1 (2). In contrast, GST by itself could not reconstitute the splicing activity in the depleted NE (Figure 4B, lane 3).

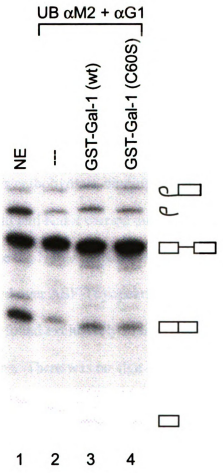
GST-Gal1(WT) and GST-Gal1(N46D) showed similar dose-response curves in reconstituting the splicing activity of a galectin-depleted NE (Figure 4C). Over a concentration range of 1 - 13 μM, product formation rose from ~5% to ~10%. Both products of the splicing reaction, the ligated exons and the liberated intron lariat, yielded the same conclusion. Since GST-Gal1(N46D) has lost its carbohydrate-binding activity, it appears that the splicing activity correlates with its retention of the association with TFII-I.

GST-Gal1(C60S) behaved similarly to GST-Gal1(WT) in the reconstitution of splicing assay (Figure 5). Thus, GST-Gal1(C60S) retained all three activities of the wild-type protein. In contrast, it appears that GST-Gal1(E71Q) has lost all three activities as we were unable to obtain a reproducible reconstitution assay with this mutant (data not shown).

**Figure 5. Comparison of the splicing activities of NE, NE after depletion of Gal1 and Gal3, and depleted NE reconstituted with GST, GST-Gal1(WT), and GST-Gal1(C60S).**

NE was depleted of Gal1 and Gal3 by adsorption on beads covalently coupled with anti-Gal1 and anti-M2 (rat monoclonal antibody against Gal3). Lane 1, the complete (non-depleted) NE; lane 2, the unbound (UB) fraction of the double antibody adsorption; lanes 3-4, depleted extract reconstituted with GST-Gal1(WT), and GST-Gal1(C60S), respectively. Splicing of $^{32}$P-labled MINX pre-mRNA was analyzed by electrophoresis through a 13% polyacrylamide-urea gel and autoradiography.

Figure 5



## DISCUSSION

The key findings of the present study include: (a) the N46D mutant of Gal1, expressed as a fusion protein with GST, is deficient in carbohydrate-binding activity, compared to the wild-type protein; (b) the same GST-Gal1(N46D) protein retained, however, both its association with TFII-I and the reconstitution of splicing activity in NE depleted of galectins, just like GST-Gal1(WT); and (c) the C60S mutation did not appear to affect any activity of the polypeptide as GST-Gal1(C60S) behaved similarly to GST-Gal1(WT). On the basis of the results obtained with GST-Gal1(N46D), compared to a parallel analysis of an innocuous mutant (GST-Gal1(C60S)), it appears that the splicing functions of Gal1 can be dissociated from its carbohydrate-binding activity.

In previous studies, Hirabayashi and Kasai carried out site-directed mutagenesis on the human Gal1 cDNA to generate the N46D mutant (4, 5). When bacterial lysates expressing this mutant were passed over ASF-Toyopearl columns, Gal1(N46D), as detected by immunoblotting, was found exclusively in the flow-through fraction, together with the bulk of the *E. coli* proteins. There was no sign of retardation on the affinity columns that might reflect some retention of weak carbohydrate-binding capacity. Thus, it was concluded that substitution of Asn by Asp at residue 46 resulted in complete loss of saccharide-binding activity (4, 5). Like the parent mutant protein Gal1(N46D), our present fusion construct GST-Gal1(N46D) also resulted in a polypeptide deficient in binding to ASF. The level of binding of GST-Gal1(N46D) was found to be similar to that observed with GST. That residue 46 of human Gal1 is critical in saccharide-binding is consistent with the results of X-ray crystallographic analysis of the carbohydrate

recognition domain of the galectins, in which the Asn residue at this position serves as an acceptor of a hydrogen bond from the hydroxy group at C-4 of galactose (16-19).

In any site-directed mutagenesis experiment, a critical issue is whether the mutation has caused a disruption of the overall folding of the polypeptide. Thus, structural data will ultimately be required to provide a definitive basis for proper interpretation of the results obtained with the mutant polypeptide(s). It should be noted, however, that our present results showing that GST-Gal1(N46D) can associate with TFII-I and reconstitute splicing despite being deficient in carbohydrate-binding suggest that we have been able to dissociate the former two activities from saccharide interactions in the N46D mutant. This, in turn, suggests that the mutant polypeptide must have retained sufficient structure to preserve the splicing-related activities while the carbohydrate-binding activity was compromised. We draw analogy to two previously documented examples. Prostaglandin endoperoxide synthase, which catalyzes the committed step in the synthesis of prostaglandins and thromboxanes, exhibits two activities: (a) cyclooxygenase activity; and (b) peroxidase activity. A site-directed mutant (Y384F) lacked cyclooxygenase activity but retained peroxidase activity, arguing against the notion that the single amino acid substitution resulted in gross misfolding of the polypeptide (20). Similarly, Bicoid, the anterior determinant of *Drosophila*, controls embryonic gene expression by transcriptional activation and translational repression. Replacement of arginine at residue 54 (R54S) shifts the binding properties of the homeodomain to prefer DNA over RNA recognition. This abolishes mRNA translational repression without affecting transcriptional activation (21). On the basis of these comparisons, we interpret our

present data to indicate that the saccharide-binding activity of Gal1 is not required for the splicing activity.

This raises the question of how to reconcile the two apparently disparate findings: while carbohydrate-binding is not required for splicing activity and for association with TFII-I, saccharides such as Lac and TDG nevertheless exert an inhibitory effect when added to the splicing assay (1) or to the GST-Gal1 pull-down of TFII-I (6). It is possible that binding of saccharide ligands to the carbohydrate-binding site results in a conformational change that disrupts the interaction of the galectin polypeptide with a component of the splicing machinery such as TFII-I. Consistent with this notion is the preliminary observation that while GST-Gal1(WT) pull-down of TFII-I is sensitive to Lac inhibition (see Figure 2A of preceding article, reference 6), the pull-down of TFII-I by GST-Gal1(N46D) is not inhibited by Lac, presumably because the mutant protein does not bind the saccharide to induce the conformational change.

To the best of our knowledge, a conformational change in the Gal1 polypeptide upon saccharide binding has not been reported. Moreover, most of the crystallographic structures of the galectin family of polypeptides have been determined in complexes containing glycoconjugate ligands (16-19). The lone exception is that of galectin-7, in which the X-ray structure was determined in both free and carbohydrate-bound forms (22). A comparison of the three-dimensional structures of the two forms showed that the galectin-7 polypeptide does not undergo any significant conformational changes upon saccharide-binding. On the other hand, evidence for a conformational change in the COOH-terminal carbohydrate recognition domain of Gal3 upon Lac binding has been suggested by differential scanning calorimetry, in which the melting temperature during

thermal denaturation of the globular fold is shifted from ~56 °C to ~65 °C by the presence of Lac (23). In addition, more recent NMR studies have shown that binding of saccharide ligands to Gal3 is accompanied by a conformational change, with rearrangements of the backbone loops near the binding site (24). Finally, several studies have documented that Gal3 binds to multivalent ligands with positive cooperativity (25-27) and it was suggested that saccharide-binding might expose hydrophobic surfaces for interactions between the carbohydrate recognition domain of separate Gal3 molecules (28). Thus, a rigorous test of our hypothesis must await physico-chemical studies on the Gal1 polypeptide itself, as well as tests of the effect of Lac or TDG on the interaction between Gal1 (and Gal3) with components of the splicing machinery

## REFERENCES


1. **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. U.S.A. **92:**1213-7.

2. **Vyakarnam, A., S. F. Dagher, J. L. Wang, and R. J. Patterson.** 1997. Evidence for a role for galectin-1 in pre-mRNA splicing. Mol. Cell. Biol. **17:**4730-7.

3. **Patterson, R., W. Wang, and J. Wang.** 2004. Understanding the biochemical activities of galectin-1 and galectin-3 in the nucleus. Glycoconj J **19:**499-506.

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

5. **Hirabayashi, J., and K. Kasai.** 1994. Further evidence by site-directed mutagenesis that conserved hydrophilic residues form a carbohydrate-binding site of human galectin-1. Glycoconj J **11:**437-42.

6. **Gray, R. M., P. G. Voss, R. J. Patterson, and J. L. Wang.** 2006. Interactions of Galectin-1 and Galectin-3 in pre-mRNA Splicing I. Identification of Transcription Factor II-I in Association with Galectin-containing Spliceosomal complexes. accompanying article, paper I.

7. **Laemmli, U.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-5.

8. **Blum, H., H. Berer, and H. J. Gross.** 1987. Electrophoresis **8:**93-99.

9. **Roff, C., and J. Wang.** 1983. Endogenous lectins from cultured cells. Isolation and characterization of carbohydrate-binding proteins from 3T3 fibroblasts. J. Biol. Chem. **258:**10657-10663.

10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475-1489.

11. **Bradford, 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:**248-54.

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

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

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

15. **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. Nucleic Acids Res. **29:**3595-602.

16. **Lobsanov, Y., M. Gitt, H. Leffler, S. Barondes, and J. Rini.** 1993. X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-A resolution. J. Biol. Chem. **268:**27034-27038.

17. **Liao, D., G. Kapadia, H. Ahmed, G. Vasta, and O. Herzberg.** 1994. Structure of S-Lectin, a Developmentally Regulated Vertebrate {beta}- Galactoside-Binding Protein. Proc. Natl. Acad. Sci. U.S.A. **91:**1428-1432.

18. **Bourne, Y., B. Bolgiano, D. Liao, G. Strecker, P. Cantau, O. Herzberg, T. Feizi, and C. Cambillau.** 1994. Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides. Nat. Struct. Biol. **1:**863-70.

19. **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.

20. **Shimokawa, T., R. Kulmacz, D. DeWitt, and W. Smith.** 1990. Tyrosine 385 of prostaglandin endoperoxide synthase is required for cyclooxygenase catalysis. J. Biol. Chem. **265:**20073-20076.

21. **Niessing, D., W. Driever, F. Sprenger, H. Taubert, H. Jackle, and R. Rivera-Pomar.** 2000. Homeodomain position 54 specifies transcriptional versus translational control by Bicoid. Mol. Cell **5:**395-401.

22. **Leonidas, D., E. Vatzaki, H. Vorum, J. Celis, P. Madsen, and K. Acharya.** 1998. Structural basis for the recognition of carbohydrates by human galectin-7. Biochemistry **37:**13930-13940.

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

24. **Umemoto, K., H. Leffler, A. Venot, H. Valafar, and J. H. Prestegard.** 2003. Conformational differences in liganded and unliganded states of Galectin-3. Biochemistry **42:**3688-95.

25. **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.

26. **Massa, S., D. Cooper, H. Leffler, and S. Barondes.** 1993. L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity. Biochemistry **32:**260-7.

27. **Probstmeier, R., D. Montag, and M. Schachner.** 1995. Galectin-3, a beta-galactoside-binding animal lectin, binds to neural recognition molecules. J Neurochem **64:**2465-72.

28. **Barboni, E. A. M., S. Bawumia, K. Henrick, 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.

# CHAPTER 4

## Epitope for the Mac-2 Monoclonal Antibody in the Proline-, Glycine-rich Domain of Galectin-3

## ABSTRACT

Previous depletion and reconstitution experiments had established that galectin-3 (Gal3) is a factor required for cell-free splicing of pre-mRNA. The epitope of one monoclonal antibody, NCL-GAL3, maps to the $NH_2$-terminal 14 amino acids of the Gal3 polypeptide. The addition of this antibody to a splicing competent nuclear extract inhibited the splicing reaction. Native gel electrophoresis showed that NCL-GAL3 exerted its effect early in the spliceosome assembly process, blocking the progression of H-/E-complexes into active spliceosomal complexes. In contrast, the epitope of a second monoclonal antibody, anti-Mac-2, maps to residues 48-100, which contains multiple repeats of a 9-residue motif, PGAYPGXXX. This antibody had no effect on splicing. One interpretation of these results is that the portion of the Gal3 polypeptide bearing the PGAYPGXXX repeats is sequestered through interaction with the splicing machinery and is inaccessible to the anti-Mac2 antibody. Consistent with this notion, a synthetic peptide containing three perfect repeats of the sequence PGAYPGQAP (27-mer) inhibited the splicing reaction, mimicking a dominant-negative mutant. In contrast, addition of peptides corresponding to a single iteration (9-mer) or two repeats (18-mer) of this motif failed to yield the same effect. Finally, GST-hGal3(1-100), a fusion protein containing glutathione S-transferase and a portion of the Gal3 polypeptide bearing the PGAYPGXXX repeats, also exhibited the dominant negative effect on splicing.

## INTRODUCTION

Galectins are a family of widely distributed proteins that bind to β-galactosides and contain characteristic amino acid sequences in the carbohydrate recognition domain(s) (CRD) of the polypeptide (1). At present, 15 mammalian galectins have been reported and classified into three subgroups, according to the content and organization of the domains in their respective polypeptides (for reviews, see references (2) and (3)): (a) prototype subgroup containing a single domain, the CRD; (b) tandem repeat type containing two CRDs joined by a linker region; and (c) chimera type containing two domains, a CRD fused onto a Pro-, Gly-rich domain.

Galectin-3 (Gal3) is, at present, the sole representative of the chimera subgroup. Like most other members of the family, Gal3 exhibits dual localization, being found in both the extracellular compartment (cell surface and medium) as well as the intracellular compartment (cytoplasm and nucleus) (4). In previous studies, we reported the localization of Gal3 to the cell nucleus in the form of a ribonucleoprotein (RNP) complex (5, 6). We also identified it as one of the many proteins required for the splicing of pre-mRNA, assayed in a cell-free system. This conclusion was based on several key observations (7-9): (a) Nuclear extracts (NEs) derived from HeLa cells, capable of carrying out splicing, contain Gal3 and another family member, galectin-1 (Gal1); (b) NEs depleted of Gal1 and Gal3 are deficient in splicing activity and spliceosome formation; (c) The splicing activity and spliceosomal assembly of the galectin-depleted extracts are reconstituted by the addition of either purified recombinant Gal1 or Gal3; and (d) Saccharides that bind the galectins with high affinity inhibit the cell-free splicing

reaction. These results strongly suggest that Gal1 and Gal3 are redundant factors in the splicing of pre-mRNA.

The polypeptide of Gal3 can be delineated into two distinct domains: (a) an $NH_2$-terminal domain (ND) containing multiple repeats of a 9-residue motif, PGAYPGXXX; and (b) a COOH-terminal CRD that shows sequence similarity with the corresponding CRDs of other members of the galectin family (10-13) (see also http://www.ncbi.nlm.nih.gov). Previous studies had documented that the CRD was sufficient to reconstitute splicing activity in a galectin-depleted NE (8). However, the minimum concentration required for activity was four to eight times higher than that required of the full-length Gal3 polypeptide, which contains the PGAYPGXXX motifs. These results suggest that Gal3 uses, at least in part, the ND to interact with components of the splicing machinery. In the present communication, we provide three lines of evidence that implicate the repeating 9-residue motif, PGAYPGXXX, in mediating this interaction.

## EXPERIMENTAL PROCEDURES

Antibodies and peptides used in functional assays --- A rat monoclonal antibody was developed against the Mac-2 antigen (14), which has been shown to be Gal3 (15). The hybridoma line producing this monoclonal antibody (M3/38.1.2.8.HL.2) was obtained from the American Type Culture Collection (TIB 166). The hybridoma cells were cultured in serum-free medium (RPMI 1640 containing Nutridoma SP (Boehringer Mannheim)). After centrifugation to pellet the cells, supernatants from the cultures were pooled, subjected to ammonium sulfate precipitation (45% of saturation), dialyzed against phosphate-buffered saline (PBS; 140 mM NaCl, 2.68 mM KCl, 10 mM $Na_2HPO_4$, 1.47 mM $KH_2PO_4$, pH 7.4) exhaustively, and stored in aliquots at a concentration of 250 μg/ml. This antibody preparation is hereafter designated as rat monoclonal anti-Mac-2. We also obtained an independent preparation of the anti-Mac-2 antibody from a commercial source (Acris Antibodies, GmbH, Hiddenhausen, Germany).

A murine hybridoma, designated as NCL-GAL3, was derived using human Gal3 as the immunogen. The NCL-GAL3 antibody used in this study was purchased from Vector Laboratories (VP-G802; hybridoma clone 9C4). Human autoimmune serum (anti-Sm) reactive with the Sm antigens of small nuclear ribonucleoprotein complexes (snRNPs) was purchased from The Binding Site.

The following peptides were synthesized in the Macromolecular Structure Facility (Michigan State University): (a) 9-mer (PGAYPGQAP) corresponding to residues 42-50 of human Gal3; (b) 18-mer (PGAYPGQAPPGAYPGQAP) corresponding to two iterations of the 9-residue motif; (c) 27-mer (PGAYPGQAPPGAYPGQAPPGAYPG-

QAP), three iterations; (d) 14-mer (MADNFSLHDALSGS) corresponding to residues 1-14 of human Gal3; and (e) mt14-mer (MADNFALHDALSGS), with a S6A substitution.

Assays for pre-mRNA splicing and spliceosome assembly --- HeLa S3 cells were grown in suspension culture by the National Cell Culture Center (Minneapolis, MN). Nuclear extract (NE) was prepared in buffer D (20 mM Hepes-KOH, pH 7.9, 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 0.5 mM dithiothreitol (DTT)), as described by Dignam *et al.* (16). NEs were frozen as aliquots in a liquid nitrogen bath and stored at -80°C. Protein concentrations were determined by the Bradford assay (17). In this study the protein concentration of NE was ~6 mg/ml.

NaCl was added to NE in Dignam buffer D to 0.5 M and set on ice for 20 minutes. Samples of this NE were dialyzed against 60% Dignam buffer D in the presence or absence of the appropriate amounts of antibodies: anti-Mac-2, NCL-GAL3, or anti-Sm. Similarly, recombinant Gal3 (8), glutathione S-transferase (GST) or GST-hGal3(1-100) was added to the NE at this step to test the effect of the recombinant or GST-fusion proteins on the splicing reaction. Dialysis was carried out for 70 minutes at 4°C in a microdialyzer with a 6-8 kD cutoff dialysis membrane (8). Splicing reaction mixtures, in a total volume of 12 µl, contained dialyzed NE sample (10 µl), [$^{32}$P]MINX pre-mRNA (18), 2.5 mM $MgCl_2$, 1.5 mM ATP, 20 mM creatine phosphate, 0.5 mM DTT, and 20 U RNasin (Promega). Splicing reactions were incubated at 30°C for 45-60 minutes. The assay was stopped by 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 minutes, diluted to 100 µl with 125 mM Tris pH 8, 1 mM EDTA, and 0.3 M sodium acetate. RNA was extracted with 200 µl of phenol/chloroform (50:50 v/v) followed by 200 µl of

chloroform. The RNAs were precipitated with 600 µl of ethanol at -80°C. The extracted RNAs were subjected to electrophoresis through 13% polyacrylamide (acrylamide: bisacrylamide 50:1.9 (w/w))-8.3 M urea gels. The radioactive RNA bands were visualized by autoradiography. Quantitation of product formation was carried out by exposing the gel to a Storage Phosphor Screen (Amersham Biosciences), scanning on a Storm 860 scanner (Molecular Dynamics), and using the program Image Quant (Molecular Dynamics) to determine the percentage of radioactivity in specific bands in each lane.

The assembly of spliceosomes was monitored by gel mobility shift assay for complex formation (18, 19). Non-denaturing 4% polyacrylamide gels (acrylamide:bisacrylamide 80:1 (w/w)), 50 mM Tris pH 8.8, 50 mM glycine, 10 mM EDTA pH 8.0) were pre-run at 150V for 30 minutes at 4°C. Heparin (1 µl at 10 mg/ml) was added to the splicing reaction, incubated for 15 minutes at 30°C, and set on ice for 5 minutes. Then, 1.3 µl of 10X loading dye (97% glycerol, 1% bromophenol blue, 1% xylene cyanol) was added. Half of each sample was loaded and electrophoresed at 150V for 90 minutes at 4°C. The gel was overlaid on gel blot paper (Schleicher and Schuell), dried, analyzed by autoradiography, and quantitated using the phosphor imaging screen, scanner and quantitation program as described above.

The effect of peptides on the splicing reaction and on spliceosome assembly was tested in a similar fashion, with the following modifications. Purified peptides were added directly to NE in a final 10 µl volume of 60% Dignam buffer D. Splicing was carried out as above in a total volume of 12 µl and incubated at 30°C for 0-45 minutes. Splicing reactions were processed as above for RNA analysis. For complex formation

experiments, duplicate samples of the splicing reactions were removed at 0-15 minute time points and snap frozen. Upon thawing, 1 μl of heparin (10 mg/ml) was added to each sample. The samples were incubated at 30°C for 15 minutes, and on ice for 5 minutes prior to running on the non-denaturing gels.

Construction of fusion proteins containing GST and Gal3 --- A 5' BamHI restriction site was introduced into the 750 bp human Gal3 cDNA (20) using the 5' primer (ATATATAGGATCCAAATGGCAGACAATTTTTCGCTC). The 3' primer (TAATAAGCGGCCGCACTAGTGATT) includes the 3' NotI restriction endonuclease site. PCR products were purified, digested with BamHI and NotI, ligated into the vector pGEX 5X-2, and transformed into *E. coli* DH5α cells via electroporation. The harvested plasmid derived from an ampicillin-selected colony was then sequenced using a primer (GGGCTGGCAAGCCACGTTTGGTG) complementary to a site just upstream of the multiple cloning region. This confirmed that the human Gal3 insert is present in the correct orientation and reading frame. This plasmid, pGEX-hgal3, expresses the full-length fusion protein, GST-hGal3(1-250) (see Figure 2).

Fusion proteins containing GST followed by portions of the human Gal3 sequence (Figure 2) were generated by site-directed mutagenesis. Complementary oligonucleotide primer sets with mutations that convert a pair of sense codons into a pair of stop codons at the desired location were designed for each desired mutant: (a) GST-hGal3(1-100) --- 5'-CCAAGTGCCCCCGGAGCCTAATAGGCCACTGGCCCCTA-TGG-3' and 3'-CCATAGGGGCCAGTGGCCTATTAGGCTCCGGGGGCACTTGG-5'; (b) GST-hGal3(1-25) --- 5'-GGATGGCCTGGCGCATGATAGAACCGGTCTGC-TGGGGCAGGGGG-3' and 3'- CCCCCTGCCCCAGCAGACCGGTTCTATCATG-

CGCCAGGCCATCC-5'. (Note that the primers code for a mutation that introduces an AgeI restriction endonuclease site downstream of the stop codons.) (c) GST-hGal3(1-14) --- 5'- GCGTTATCTGGGTCTTGATAAGCTTACCCTCAAGGATGGCCTGGC-3' and 3'- GCCAGGCCATCCTTGAGGGTAAGCTTATCAAGACCCAGATAACGC-3'. (Note that these primers also code for a new HindIII restriction endonuclease site downstream of the stop codons.) Following thermocycling with pGEX-hgal3, the template DNA was cleaved by incubation with the Dam methylation dependent endonuclease DpnI. The reaction mix was then used to transform DH5α cells via electroporation. Colonies expressing the proper mutation were screened based on new restriction endonuclease sites introduced during the site directed mutagenisis (AgeI and HindIII respectively). Colonies expressing the proper mutation were further screened by western blotting of transformed bacterial lysates with mouse monoclonal antibody NCL-GAL3, rat monoclonal antibody anti-Mac-2, and rabbit polyclonal antibody against GST. The plasmid DNA for each was sequenced to confirm the mutations.

Constructs expressing the fusion proteins GST-hGal3(1-47) and GST-hGal3(46-250) were generated using the XmaI restriction endonuclease sites coded by the DNA base pairs of amino acid residues 48 and 45, respectively. GST-hGal3(1-47) was created by excising a portion of the Gal3 DNA insert in pGEX-hgal3 using the 5' BamHI and internal Gal3 XmaI restriction sites and ligating into pGEX 5X-2 that had been digested with BamHI and XmaI.. Similarly, GST-hGal3(46-250) was created by excising a portion of the Gal3 DNA insert in pGEX-hgal3 using the internal Gal3 XmaI and 3' NotI restriction sites and ligating into pGEX 5X-2 that had been digested with XmaI and NotI. These ligation reactions were used to transform DH5α bacteria by electroportation.

Colonies were screened based on the size of the insert excised by double digestion of sites remaining on the pGEX multiple cloning region, BamHI and XhoI for the 1-47 construct and EcoRI and NotI for the 46-250 construct. The plasmid DNA from each was sequenced using the same primer as pGEX-hgal3 to confirm the mutations.

Protein expression and purification --- GST fusion proteins were expressed in 500 ml cultures of *E.coli* BL-21 codon plus (DE3) cells (Stratagene) by induction with 100 μM isopropyl-β-D-galactopyranoside (IPTG) for 2-3 hours at 30°C. Cells were pelleted and stored at -70°C. Thawed bacterial pellets were resuspended (one-twentieth of the culture volume) in phosphate buffered saline (PBS) containing protease inhibitiors (4 μg/ml aprotinin, 5 μg/ml leupeptin, 0.2 μg/ml pepstatin A, and 1 mM Pefabloc (Roche)) and sonicated using a microtip probe. Triton X-100 was added to a final concentration of 1% for lysates that would be purified using glutathione beads or to 0.1% for lysates that would be purified using lactose beads. After rocking for 1 hour at 4°C, cell debris was removed by centrifugation at 12,000 x g for 10 minutes at 4°C. The supernatant was aliquotted, snap-frozen, and stored at -70°C.

GST-hGal3(1-250) and GST-hGal3(46-250) were purified by lactose affinity chromatography. All procedures were carried out at 4°C. Frozen stock lysate (5 ml) was thawed and diluted with 90 ml of binding-wash buffer (PBS containing 1 mM DTT and 0.5 mM PMSF) and rotated overnight with 5 ml of lactose-agarose beads (Sigma), The slurry was poured into a poly-prep chromatography column (Bio-Rad) and allowed to flow through. The column was washed with 10 column volumes of binding-wash buffer. The bound protein was eluted with 15 ml of elution buffer (PBS, 0.4 M lactose, 1 mM DTT, 0.5 mM PMSF). In this procedure, the elution buffer was first allowed to flow

into the column (approximately 5 ml), the column was stopped for 1 hour and then 15 x 1 ml fractions were collected. Samples from each fraction were electrophoresed on 12.5% SDS–PAGE, silver stained and screened by western blotting with NCL-GAL3, anti-Mac-2, and anti-GST. Fractions were selected for highest quantity and purity, pooled, and concentrated using a Centricon-10 filter unit (Amicon) allowing buffer exchange and removal of lactose. The amount of protein in each concentrated preparation was quantitated using the Bradford assay (17). Silver staining of the SDS-PAGE, compared to known amounts of standards, provided confirmation of this quantitation.

Because GST and the other fusion proteins lacked the CRD of Gal3 (see Figure 2), they were purified by glutathione affinity chromatography. All procedures were carried out at 4° C. Frozen stock lysate (5 ml) was thawed, diluted with 45 ml of binding buffer (PBS, 0.1% TritonX-100, 1 mM DTT, 0.5 mM PMSF) along with 1 ml of glutathione beads (Pierce) and rotated for 2 hours. The beads were then pelleted by centrifugation at 2,000 x g for 3 minutes. Supernatant was removed and the beads were washed 3 times with 50 ml binding buffer. After removal of the last wash, the beads were washed with 50 ml of wash buffer (PBS, 1 mM DTT, 0.5 mM PMSF). The beads were then resuspended in 10 ml of wash buffer and loaded into a chromatography column. The column was washed with another 10 ml of wash buffer. Purified proteins were eluted with 10 ml of elution buffer (PBS, 10 mM glutathione, 0.5 mM PMSF), by allowing 1 ml of elution buffer to flow into the column and stopping the column for 15 minutes. Fractions (0.5 ml) were collected and analyzed as above. Selected fractions were pooled, dialyzed, and quantitated.

SDS gel electrophoresis, silver staining, and western blotting --- Purified proteins or NE were electrophoresed on 12.5% acrylamide gels in the presence of sodium dodecyl sulfate (SDS-PAGE), as described by Laemmli (21). Proteins were visualized by silver staining as described by Merril *et. al.* (22). Some gels were electrophoretically transferred onto Hybond Nitrocellulose paper (Amersham Biosciences) in the presence of buffer containing 25 mM Tris, 193 mM glycine, and 10% methanol, pH 8.3. Following transfer, membranes were incubated overnight in 10% nonfat dry milk in Tween tris buffered saline (T-TBS; 10 mM Tris, 0.5 M NaCl, 0.05% Tween 20, pH 7.5). Antibodies were diluted in 1% nonfat dry milk T-TBS and incubated with membranes for 1 hr at room temperature. This incubation was followed by four 15 minute washes in T-TBS. Membranes blotted with rabbit primary antibodies were pre-blocked with unconjugated goat anti-rabbit antibody (Sigma) at 1:2000 dilution. The membranes were then incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies for 1 hour and then washed four times for 15 minutes. The proteins were visualized by use of the Western Lightning Chemiluminescence System (Perkin Elmer Life Sciences).

NCL-GAL3 was used at a dilution of 1:3000. Anti-GST antibodies were affinity purified from serum of rabbits immunized with GST constructs and used at a dilution of 1:6000. Anti-Mac-2 antibody was used at 1:1000 dilution. Goat anti-mouse HRP (Bio-Rad), goat anti-rat HRP (Roche), and goat anti-rabbit HRP (Bio-Rad), were each used at a dilution of 1:10,000. Peptide inhibition experiments were carried out by incubating the appropriate amounts of antibody and peptide in 0.5 ml PBS at 4°C overnight, followed by appropriate dilution with 1% milk T-TBS and incubation with membrane as described above.

## RESULTS

Different effects of two monoclonal antibodies directed against Gal3 on pre-mRNA splicing

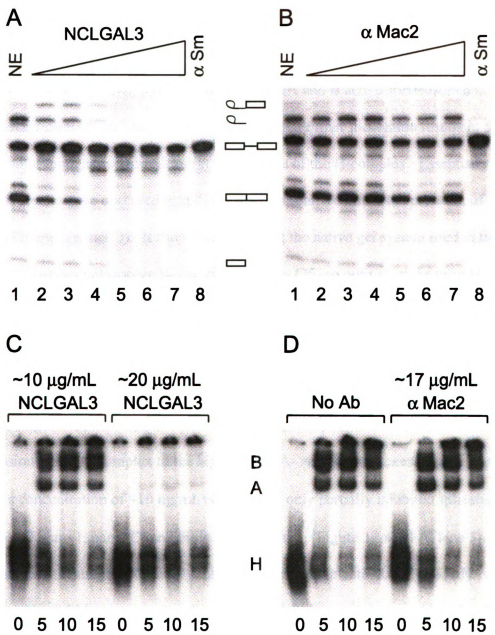
Using human Gal3 as the immunogen, a murine hybridoma, designated as NCL-GAL3, was derived that secretes a monoclonal antibody against the protein. Anti-Mac-2 is a rat monoclonal antibody originally derived through immunization with mouse macrophages expressing Gal3 (14). The effects of these two monoclonal antibodies were tested on splicing competent NE derived from HeLa cells. In this system, human autoimmune serum reactive against the Sm epitopes of snRNPs (anti-Sm) served as the positive control. Anti-Sm inhibited the conversion of the pre-mRNA substrate into the mRNA product (Figure 1, lane 8 in both panels A and B). Tested under the same conditions, the NCL-GAL3 monoclonal antibody also inhibited splicing in a dose-dependent fashion (Figure 1, panel A). Partial inhibition was observed at a concentration as low as ~9 μg/ml (~60 nM) (Figure 1, panel A, lane 2) and complete inhibition was achieved at a concentration of ~20 μg/ml (~140 nM) (Figure 1, panel A, lane 5). In contrast, the anti-Mac-2 monoclonal did not inhibit splicing over the same concentration range of 9 – 23 μg/ml (Figure 1, panel B, lanes 2-7). Both products of the splicing reaction, ligated exons and free intron, were observed at all concentrations tested.

In previous studies (23), we had documented that NCL-GAL3 recognizes a single polypeptide, corresponding to human Gal3, in extracts of HeLa cells (also see below, Figure 3, panel C, lane 1). We had also shown that addition of recombinant Gal3 alone to a splicing competent NE had no effect on the splicing reaction (24). On the other hand, the inhibitory effect of NCL-GAL3 on splicing can be overcome by prior incubation with

**Figure 1. The effect of antibody addition on the splicing of pre-mRNA and on spliceosome assembly.**

Panel A: Effect of NCL-GAL3 on the splicing reaction; Panel B: Effect of anti-Mac2 on the splicing reaction. In both panels, the splicing activity of NE (no additions) is shown in lane 1; the effect of anti-Sm antibodies (1:23 dilution of human autoimmune serum) is shown in lane 8. In both panels, the concentrations of the Gal3-specific antibody tested were: lane2, ~8 μg/ml; lane 3, ~10 μg/ml; lane 4, ~17 μg/ml; lane 5, ~21 μg/ml; lane 6, ~23 μg/ml; and lane 7, ~24 μg/ml. The cell-free splicing assay was carried out using $^{32}$P-labeled MINX pre-mRNA substrate. Products of the splicing reaction were analyzed by electrophoresis through a 13% polyacrylamide-urea gel system, followed by autoradiography. The positions of migration of the pre-mRNA substrate, the splicing intermediates (exon 1 and lariat-exon 2), and the products (mature RNA and lariat intron) are indicated between the two panels. Panel C: Time course of spliceosome assembly in the presence of two different concentrations of NCL-GAL3. Panel D: Time course of spliceosome assembly in the absence and presence of anti-Mac-2. In both panels, splicing reaction mixtures containing $^{32}$P-labeled MINX pre-mRNA were sampled at various times indicated and were anaylzed by electrophoresis through non-denaturing gel system, followed by autoradiography. The regions of migration of early complexes (H- and E-complexes) and active spliceosomes (A- and B-complexes) are indicated between the two panels.

Figure 1


A
NCLGAL3
NE
α Sm
ρ ρ
1 2 3 4 5 6 7 8
B
α Mac2
NE
α Sm
1 2 3 4 5 6 7 8
C
~10 μg/mL NCLGAL3
~20 μg/mL NCLGAL3
0 5 10 15 0 5 10 15
D
No Ab
~17 μg/mL α Mac2
B
A
H
0 5 10 15 0 5 10 15

recombinant Gal3 (data not shown). These results suggest that the effect of NCL-GAL3 on splicing was due to specific recognition of its antigen.

At various early time points, aliquots of the splicing reaction mixture were analyzed by non-denaturing gel electrophoresis to assess progress in spliceosome assembly. In the absence of antibody addition, the radiolabeled pre-mRNA initially found in the region labeled as the H-complex (t=0) is converted to the A and B active spliceosomal complexes within five minutes (Figure 1, panel D). Further incubation results in the formation of more active complexes at the expense of the H-complex. Essentially identical results were observed in splicing reactions carried out in the presence of anti-Mac-2 (Figure 1, panel D). It should be noted that the native gel system used in the present study does not resolve H- and E-complexes (25) so our use of the term H-complex is meant only to indicate a region of the gel rather than a distinction between the two early complexes of the spliceosome assembly pathway.

At a concentration of ~24 μg/ml, the NCL-GAL3 antibody inhibited the splicing reaction (Figure 1, panel A, lane 7). This is paralleled by an almost complete arrest of the progression of the H-complex to the higher order A- and B-complexes (Figure 1, panel C). At a concentration of ~10 μg/ml, NCL-GAL3 only partially inhibited spliceosome assembly (Figure 1, panel C) and the splicing reaction (Figure 1, panel A, lane 3).

These results raise the possibility that the epitope of the anti-Mac-2 antibody is buried in interactions with components of the splicing machinery and those Gal3 molecules assembled into the splicing complexes are inaccessible to this monoclonal antibody. In contrast, the epitope of the NCL-GAL3 antibody is available and the binding of the monoclonal antibody either inhibited the splicing reaction or precipitated those Gal3-

containing spliceosomes out of the reaction mixture. For this reason, it was of interest to map the epitope of the two monoclonal antibodies on the Gal3 polypeptide.

Epitope mapping for the Mac-2 and NCL-GAL3 monoclonal antibodies

The cDNA for human Gal3 was cloned into the pGEX 5X-2 vector and full-length Gal3 was expressed as a fusion protein with GST, designated as GST-hGal3 (1-250). In addition, site-directed mutagenesis was carried out to introduce translation termination codons at various positions so that the expressed fusion protein was truncated at specific residues of the Gal3 polypeptide chain (see Figure 2). Finally, a construct was also engineered to express the GST fusion protein in which the $NH_2$-terminal 45 amino acids of Gal3 are missing (GST-hGal3 (46-250)). The GST-fusion proteins were purified by glutathione-affinity or lactose-affinity chromatography and subjected to SDS-PAGE analysis. Silver staining provided documentation on the purity of each of the fusion protein preparations (Figure 3, panel A). Each of the fusion proteins was also detectable by immunoblotting with polyclonal rabbit anti-GST (Figure 3, panel B). Both of these techniques also ascertained that the molecular weight of the predominant polypeptide was in agreement with the expected size of the fusion protein (Figure 3, panel A).

When the various GST-fusion proteins were subjected to immunoblotting with NCL-GAL3, a positive reaction was observed with each except GST-hGal3 (46-250) (Figure 3, panel C). GST itself also failed to react with NCL-GAL3 (Figure 3, panel C, lane 6). These results suggest that the epitope of NCL-GAL3 lies in the first 14 residues of the galectin-3 polypeptide. This notion is consistent with the observation that a Gal3 S6A mutant (serine to alanine mutation at residue 6) resulted in the loss of reactivity with the NCL-GAL3 monoclonal antibody (data not shown).

**Figure 2. Fusion proteins containing glutathione S-transferase and galectin-3 sequences of varying lengths.**
Human galectin-3 cDNA (hGal3) was engineered into the pGEX 5X-2 vector bearing the *Schistosoma japoncium* glutathione S-tranferase (GST) sequence. The numbers along the left-hand side indicate lane assignments in panels A-D of Figure 3. The names of each construct are listed beside the lane assignments. Numbers in parentheses indicate the galectin-3 amino acids included in each construct. N denotes the amino terminus; C denotes the carboxyl terminus. The rectangles represent the GST protein, and bars indicate the portions of galectin-3 included in the fusion protein (with amino acid residues listed above each bar). The approximate molecular weights of each construct are listed at the right. Each construct was expressed and the fusion protein was purified, subjected to SDS-PAGE, and analyzed as documented in Figure 3.

Figure 2

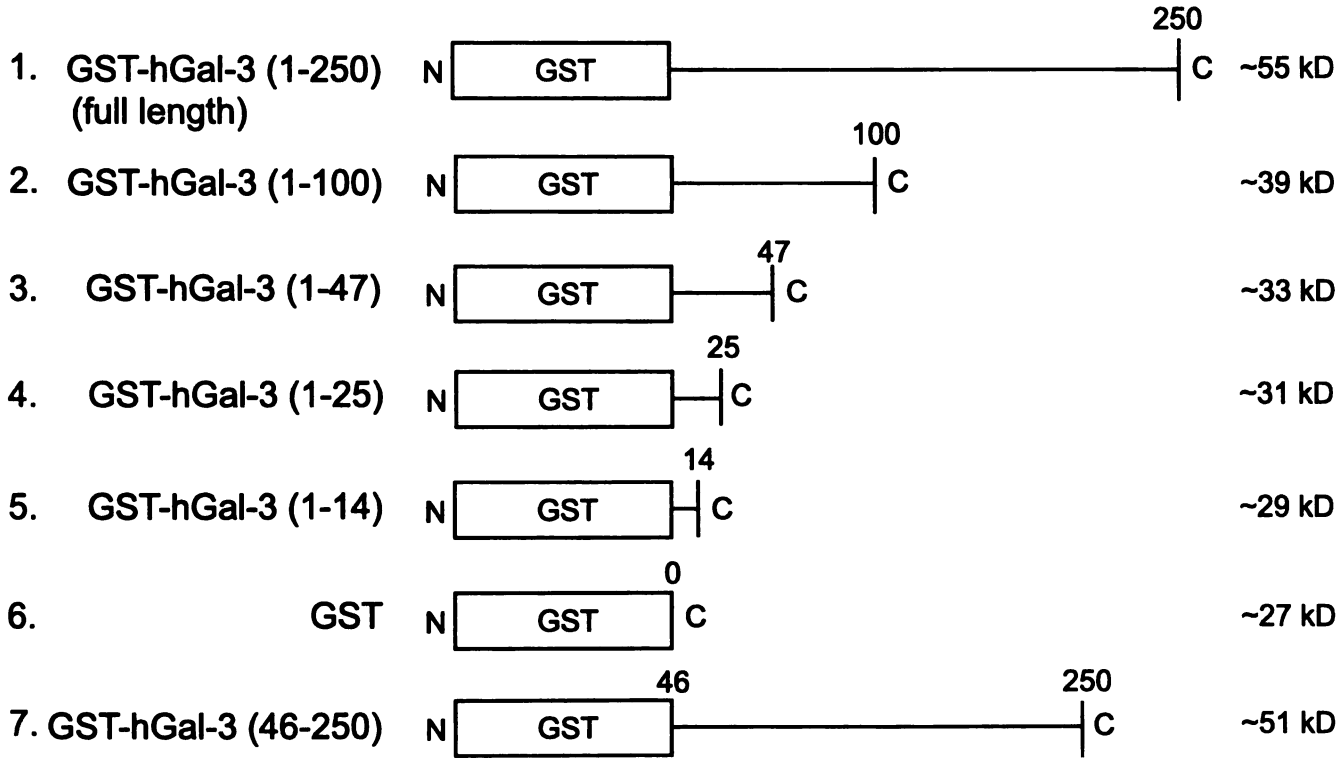
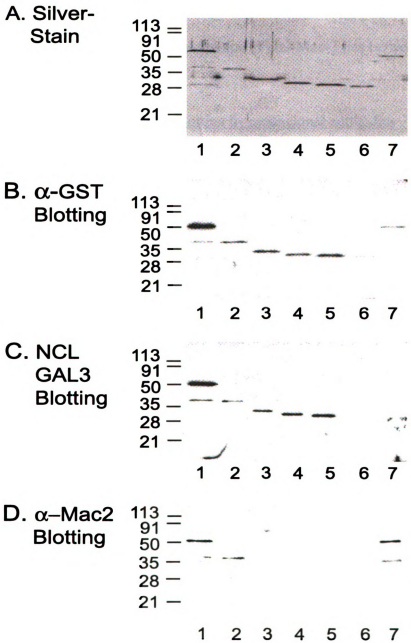

1. GST-hGal-3 (1-250) (full length)
N
GST
250
C
~55 kD
2. GST-hGal-3 (1-100)
N
GST
100
C
~39 kD
3. GST-hGal-3 (1-47)
N
GST
47
C
~33 kD
4. GST-hGal-3 (1-25)
N
GST
25
C
~31 kD
5. GST-hGal-3 (1-14)
N
GST
14
C
~29 kD
6. GST
N
GST
0
C
~27 kD
7. GST-hGal-3 (46-250)
N
GST
46
250
C
~51 kD

**Figure 3. SDS-PAGE, silver staining, and western blotting analysis of the GST-hGal3 fusion proteins.**

Each of the constructs listed in Figure 2 was expressed and the fusion protein was purified, subjected to SDS-PAGE, and analyzed by silver staining (panel A) and by immunoblotting with anti-GST antibodies (panel B), with mouse monoclonal NCL-GAL3 (panel C), and with rat monoclonal anti-Mac-2 (panel D). Approximately 35 ng of each purified fusion protein was electrophoresed. The numbers on the left indicate the positions of migration of molecular weight markers.

Figure 3


A. Silver-Stain
113
91
50
35
28
21
1 2 3 4 5 6 7
B. α-GST Blotting
113
91
50
35
28
21
1 2 3 4 5 6 7
C. NCL GAL3 Blotting
113
91
50
35
28
21
1 2 3 4 5 6 7
D. α–Mac2 Blotting
113
91
50
35
28
21
1 2 3 4 5 6 7

When the various GST-fusion proteins were subjected to immunoblotting with anti-Mac-2, the antibody reacted with GST-hGal3 (1-250), GST-hGal3 (1-100), and GST-hGal3 (46-250) (Figure 3, panel D, lanes 1, 2, and 7). In contrast, the fusion proteins missing residues 48-100 of the Gal3 polypeptide all failed to react (Figure 3, panel D, lanes 3-5). These results suggest that the epitope of anti-Mac-2 lies between residues 48 and 100.

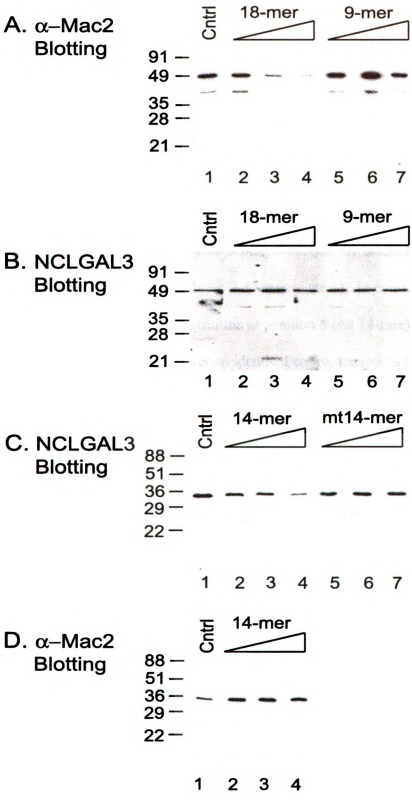
Specific peptide inhibition of immunoblotting by monoclonal antibodies

The region of the Gal3 polypeptide bearing the anti-Mac2 epitope (between residues 48 and 100) contains multiple repeats of a 9-residue motif, with a consensus sequence PGAYPGXXX (10-13). For example, residues 41-67 of the murine Gal3 sequence contain three perfect tandem repeats of PGAYPGQAP. On this basis, three peptides were synthesized containing three iterations (27-mer), two iterations (18-mer), and a single iteration (9-mer) of this sequence and were tested for their ability to block immunoblotting of recombinant GST-hGal3(1-250) ($M_r$ ~55 kD) by the anti-Mac-2 antibody.

Over a concentration range of 2.5-250 nM (about 1.3- to 130-fold excess over antibody), the 18-mer peptide inhibited the reaction between anti-Mac-2 and Gal3 in a concentration dependent fashion (Figure 4, panel A, compare lanes 2-4 vesus lane 1). Similar results were obtained with the 27-mer (data not shown). In contrast, the 9-mer showed no effect at the corresponding molar concentrations (Figure 4, panel A, lanes 5-7). Neither the 9-mer nor the 18-mer affected the immunoblotting by the NCL-GAL3 monoclonal antibody (Figure 4, panel B). These results suggest that the epitope of the anti-Mac-2 antibody: (a) requires two iterations of the PGAYPGQAP sequence; or (b)

**Figure 4. Peptide inhibition of immunoblotting of galectin-3 by two monoclonal antibodies directed against galectin-3.**
Equal amounts of purified recombinant GST-human galectin-3 (panels A and B) or HeLa cell nuclear extract (panels C and D) were subjected to SDS-PAGE. The antibodies used for immunoblotting are listed on the left. The triangle above each panel signifies increasing concentrations of peptide tested as inhibitors of blotting by an antibody. 18-mer, PGAYPGQAPPGAYPGQAP; 9-mer, PGAYPGQAP. 14-mer, MADNFSLHDALSGS; mt 14-mer, MADNFALHDALSGS.

Figure 4


A. α–Mac2 Blotting
Cntrl
18-mer
9-mer
91
49
35
28
21
1 2 3 4 5 6 7
B. NCLGAL3 Blotting
Cntrl
18-mer
9-mer
91
49
35
28
21
1 2 3 4 5 6 7
C. NCLGAL3 Blotting
Cntrl
14-mer
mt14-mer
88
51
36
29
22
1 2 3 4 5 6 7
D. α–Mac2 Blotting
Cntrl
14-mer
88
51
36
29
22
1 2 3 4

overlaps the end of one repeat with the beginning of the second one (e.g. $^{46}$PGQAPPGAY$^{54}$). In light of our observation that GST-hGal3 (46-250), which contains such a sequence, reacted positively with anti-Mac-2 (Figure 3, panel D, lane 7), we favor the latter hypothesis.

Similarly, a peptide corresponding to the sequence of the first 14 residues of human Gal3, MADNFSLHDALSGS, was also synthesized. This peptide, designated 14-mer, was tested for the ability to inhibit blotting of endogenous Gal3 ($M_r$ ~30 kD) present in NE of HeLa cells by the NCL-GAL3 antibody. Over a concentration range of 0.07-7 nM (3-300-fold excess over antibody), the 14-mer peptide inhibited the immunoblotting by NCL-GAL3 (Figure 4, panel C, compare lanes 2-4 versus lane 1). In contrast, the 14-mer peptide containing a serine to alanine substitution at position 6 (mt 14-mer) failed to inhibit immunoblotting by NCL-GAL3 over an identical concentration range (Figure 4, panel C, lanes 5-7). These findings support the conclusion that the NCL-GAL3 epitope lies within the first 14 amino acids of Gal3 and also indicate that serine 6 constitutes an important determinant in the epitope. The 14-mer peptide did not inhibit blotting by the anti-Mac-2 antibody (Figure 4, panel D, lanes 2-4): (a) indicating that it does not inhibit antibody-antigen interactions nonspecifically, and (b) lending additional support to the conclusion that the anti-Mac-2 epitope lies outside of the first 14 amino acids of Gal3.

Effect of addition of PGAYPGQAP peptides on the *in vitro* splicing reaction

We had rationalized the failure of the anti-Mac-2 monoclonal antibody to inhibit splicing (Figure 1) on the basis that its epitope is buried in protein-protein interactions of the spliceosome. Thus, we wanted to test whether peptides bearing its epitope, the PGAYPGQAP repeating motif, can perturb the splicing reaction. Control NE exhibited

good splicing activity, converting ~35% of the pre-mRNA substrate into the mature RNA product (Figure 5, panel A, lane 5 and panel B). Addition of the 27-mer synthetic peptide containing the PGAYPGQAP motif inhibited the splicing reaction. At a concentrations of 300 μM and 600 μM, product formation was reduced to ~15% and <5%, respectively (Figure 5, panel A, lanes 2 and 3 and panel B). There were also lower levels of the intermediates of the splicing reaction (free exon 1 and lariat-exon 2) as well as the lariat intron. Complete inhibition was observed at 1 mM; there were barely detectable levels of products and splicing intermediates and substantially higher levels of the starting substrate (Figure 5, panel A, lane 4 and panel B).

In contrast, parallel addition of the 9-mer (Figure 5, panel A, lanes 10-13) and the 18-mer (Figure 5, panel A, lanes 6-9) synthetic peptides did not yield the same result over the identical concentration range. A slight decrease in the percentage product formed was observed (Figure 5, panel B) but bands corresponding to the products and intermediates were clearly found for all concentrations tested (Figure 5, panel A). These results suggest that multiple repeats (more than two) are necessary to perturb the endogenous Gal3 interaction with components of the splicing reaction.

Consistent with this notion, GST-hGal3 (1-100), which contains seven repeats of the PGAYPG motif (some repeats are imperfect such as PGVYPGPPSG), inhibited the splicing reaction (Figure 6, panel A, lanes 2-4 and panel B). On the other hand, GST alone did not inhibit splicing over the same concentration range (Figure 6, panel A, lanes 5-7 and panel B). In fact, a comparison of the dose-response indicated that the concentration required to achieve full inhibition was observed at a much lower concentration of GST-hGal3 (1-100) (Figure 6, panel B) than the 27-mer synthetic

**Figure 5. Comparison of the effect of addition of synthetic peptides containing the PGAYPGQAP motif on the splicing activity of nuclear extract.** 27-mer, 18-mer, and 9-mer contain three iterations, two iterations, and a single iteration of the 9-residue motif PGAYPGQAP, respectively. Panel A: autoradiogram of the splicing assay. The peptides were tested at concentrations of 100, 300, 600 and 1000 μM. All reactions contained $^{32}$P-labeled MINX pre-mRNA substrate (5000 c.p.m.) and products of the splicing reaction 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 intron and ligated exon 1-exon 2) are highlighted on the right. Panel B: Dose-response curve of the effects of peptides on product formation, derived from the experiment shown in panel A.

Figure 5

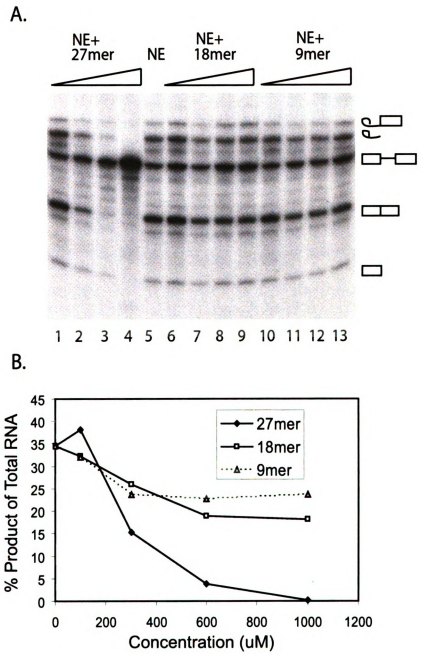
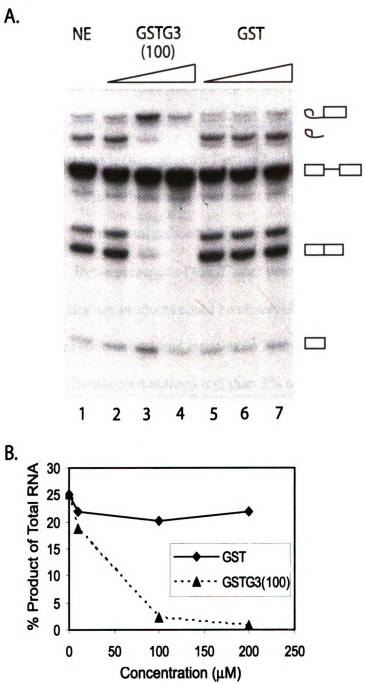

A.
NE+ 27mer
NE
NE+ 18mer
NE+ 9mer
1 2 3 4 5 6 7 8 9 10 11 12 13
B.
% Product of Total RNA
Concentration (uM)
27mer
18mer
9mer
0
5
10
15
20
25
30
35
40
45
0
200
400
600
800
1000
1200

**Figure 6. Comparison of the effect of GST-hGal3 (1-100) and GST on pre-mRNA splicing.**

Panel A: autoradiogram of the splicing assay. The proteins were tested at concentrations of 10, 100, and 200 μM. All reactions contained $^{32}$P-labeled MINX pre-mRNA substrate (5000 c.p.m.) and products of the splicing reaction were analyzed by electrophoresis through a 13% polyacrylamide-urea gel, followed by autoradiography. The positions of migration of pre-mRNA substrate, splicing intermediates, and RNA products are highlighted on the right. Panel B: Dose-response curve of the effects of the proteins on product formation, derived from the experiment shown in panel A.

Figure 6


A.
NE
GSTG3
(100)
GST
1 2 3 4 5 6 7
B.
% Product of Total RNA
30
25
20
15
10
5
0
0 50 100 150 200 250
Concentration (μM)
GST
GSTG3(100)

peptide (Figure 5, panel B). Together, the results provide confirmation of the dominant negative effect and suggest that the PGAYPGXXX motif contributes to the interaction of the Gal3 polypeptide with the splicing machinery.

Effect of the 27-mer on the kinetics of spliceosomal assembly and product formation

When the splicing reaction was carried out with the control NE, the products of the first cleavage reaction, free exon 1 and lariat-exon 2, were observed after 10 minutes (Figure 7, panel A). Product bands (ligated exon 1-exon 2 and intron) increased monotonically as a function of time, becoming prominent after 20 minutes. Each of these bands accounted for 15-20% of the total radioactivity of a splicing reaction at 30 minutes (Figure 7, panels B and C). In the presence of the 27-mer peptide inhibitor (600 μM), however, neither intermediates nor products could be observed at 15 minutes (Figure 7, panel A). The intermediates (free exon 1 and lariat-exon 2) appeared to accumulate at 20-30 minutes, but product formation remained less than 3% over the time course (Figure 7, panels B and C).

In the same manner, we monitored the kinetics with which early complexes (H- and E-complexes) are chased into the active spliceosomal A-, B-, and C-complexes (Figure 8, panel A). Both in the absence and presence of the 27-mer peptide inhibitor, the H-complex disappeared with the same kinetics (Figure 8, panels A and B). However, while ~50% of the radioactive pre-mRNA in the reaction without the inhibitor has progressed into B-complexes within the first 5 minutes (Figure 8, panel D), ~50% of the pre-mRNA in the reaction containing the peptide inhibitor is still in the A-complex at the same time point (Figure 8, panel C). Therefore, the peptide appears to slow the progression into the B-complex, resulting in an accumulation of the A-complex at 5 minutes and a persistently

**Figure 7. The effect of the 27-mer peptide on the kinetics of the splicing reaction.** Panel A: autoradiogram of the splicing assay. The 27-mer was tested at a concentration of 600 μM. All reactions contained $^{32}$P-labeled MINX pre-mRNA substrate (5000 c.p.m.) and productrs of the splicing reaction were analyzed by electrophoresis in polyacrylamide-urea gels, follwed by autoradiography. Panel B: Quantitation of the data shown in panel A, using one product, the ligated exons, as a measure of the reaction. Panel C: Quantitation of the data shown in panel A, using the other product, the excised lariat, as a measure of the reaction.

Figure 7

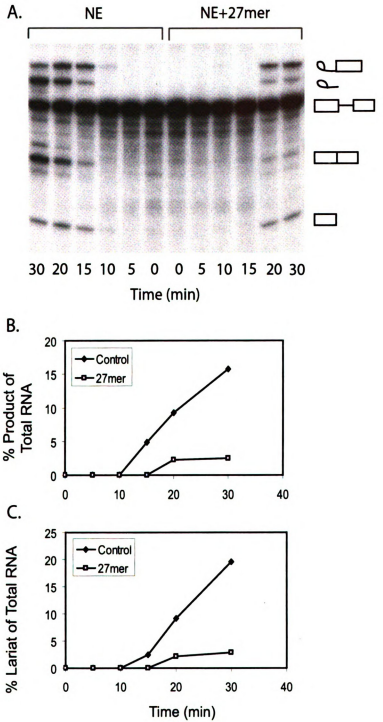
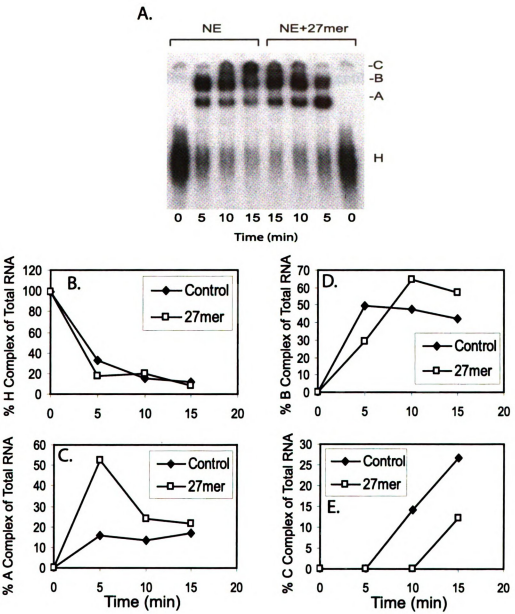

A.
NE
NE+27mer
30 20 15 10 5 0 0 5 10 15 20 30
Time (min)
B.
% Product of Total RNA
Control
27mer
0 5 10 15 20
0 10 20 30 40
C.
% Lariat of Total RNA
Control
27mer
0 5 10 15 20 25
0 10 20 30 40
Time (min)

**Figure 8. The effect of the 27-mer peptide on the kinetics of spliceosome assembly.** Panel A: autoradiogram of the non-denaturing gel. The 27-mer was tested at a concentration of 600 μM. All reactions contained $^{32}$P-labeled MINX pre-mRNA substrate (5000 c.p.m.) and were analyzed in non-denaturing polyacrylamide gels, followed by autoradiography. The positions of migration of the H-, A-, B-, and C-spliceosomal complexes are highlighted on the right. Panel B-E: Quantitation of the data shown in panel A, for H-, A-, B, and C-complexes, respectively.

Figure 8


A.
NE
NE+27mer
-C
-B
-A
H
0 5 10 15 15 10 5 0
Time (min)
B.
Control
27mer
% H Complex of Total RNA
D.
% B Complex of Total RNA
C.
% A Complex of Total RNA
Time (min)
E.
% C Complex of Total RNA
Time (min)

higher amount of the A-complex compared to controls (Figure 8, panel C). Thus, a steady increase in the active C-complex in the reaction without the inhibitor precedes the first signs of an active C-complex in the reaction containing the peptide (Figure 8, panel E). This is consistent with the observation that, in the presence of the peptide, there is not even a hint of the products of the first cleavage reaction until 20 minutes have elapsed (Figure 7, panel A).

## DISCUSSION

The key findings of the present study include: (a) The epitope of the NCL-GAL3 monoclonal antibody resides in the first 14 residues of the Gal3 polypeptide. (b) The epitope of the anti- Mac-2 monoclonal antibody maps to residue 48-67 of the Gal3 polypeptide, corresponding to one or two repeats of the motif PGAYPGQAP; (c) addition of NCL-GAL3 to a splicing competent NE inhibits the *in vitro* splicing reaction whereas parallel addition of anti-Mac-2 fails to yield the same effect; and (d) addition of the 27-mer synthetic peptide, bearing three iterations of PGAYPGQAP, to NE inhibits the splicing reaction.

Inhibition of *in vitro* splicing has been demonstrated with other antibodies and synthetic peptides to implicate the involvement of specific proteins in spliceosome assembly and the splicing reaction. For example, Yuryev *et al.* (26) monitored the effect on splicing by a monoclonal antibody directed against the carboxyl-terminal domain of the large subunit of RNA polymerase II. At an antibody concentration of ~33 μM, there was a concomitant loss of the products, as well as intermediates, of the splicing reaction. In the same study, Yuryev *et al.* also used a peptide consisting of eight consensus repeats in the carboxyl-terminal domain of RNA polymerase II large subunit to inhibit splicing. Partial inhibition was observed at a peptide concentration as low as ~20 μM and complete inhibition at ~40 μM.

Four peptides that inhibit $Ca^{2+}$-dependent calmodulin kinase II were shown to block spliceosome assembly and pre-mRNA splicing *in vitro* (27). One of the peptides (designated GS) was derived from the sequence of glycogen synthase and competitively inhibited the kinase from binding its substrate. This GS peptide inhibited splicing at a

concentration of ~300 μM. More interestingly, Parker and Steitz (27) observed splicing products after prolonged incubation. This delay (hours) in appearance of splicing products is consistent with the observation of stalled spliceosome assembly at the B-complex stage. These results are similar to our own observations with the 27-mer peptide, in which there is a delay in the appearance of intermediates and products, as well as in slowing the rates of A- to B-complex progression.

Several lines of evidence have now been accumulated to indicate that the interaction of Gal3 with components of the splicing machinery is mediated, at least in part, by the ND of the polypeptide. First, although a CRD (Gal1 alone or the COOH-terminal domain of Gal3) is sufficient to restore splicing activity to a galectin-depleted NE, the minimum concentrations required for reconstitution are four to eight times higher than that of the intact (full-length) Gal3 polypeptide (8). It was hypothesized that the Gal3 ND, containing the proline- and glycine-rich repeats, plays a role in protein-protein interactions, providing the basis for enhanced interactions with the splicing machinery. Second, that such protein-protein interactions occur between the ND and the spliceosome was suggested by the inhibition of splicing observed when the activity of NE is assayed in the presence of exogenously added ND. This could be demonstrated using the corresponding ND sequences of either human (GST-hGal3(1-100)) or murine Gal3 (24). Thus, the ND appears to exert its dominant negative effect by competing for the spliceosomal component with which Gal3 is associated. Alternatively, the ND can interact with Gal3, resulting in a conformational change that precludes the Gal3 molecule from association with the spliceosome. ND interaction with itself (28-31) or with the CRD (31) has been implicated by electron microscopic imaging, by nuclear magnetic

resonance, by cross-linking, and by analysis of positive cooperativity in the binding of Gal3 to multivalent ligands. Finally, we have now shown that inhibition of splicing can also be obtained using synthetic peptides bearing the predominant structural motif of the ND, the PGAYPGXXX repeats. In particular, the most potent inhibition was observed with the 27-mer peptide which contains three such repeats while the 9-mer did not show any inhibition.

The differential effects of the monoclonal antibodies NCL-GAL3 and anti-Mac-2 on the splicing reaction are interpreted in this context. The epitope of anti-Mac-2 resides in the PGAYPGXXX repeats of the ND. Therefore, its failure to inhibit the cell-free splicing reaction is consistent with the notion that this region is buried in interactions with the splicing machinery and is inaccessible to the antibody. On the other hand, it appears that the $NH_2$-terminus of the Gal3 polypeptide, even when associated with the spliceosome, remains accessible to the NCL-GAL3 antibody, which inhibits the splicing reaction.

A key consideration in the interpretation of the antibody effects is the assignment of the epitope location. Gong *et al.* (32) reported that deletion of the $NH_2$-terminal 11 amino acids of human Gal3 resulted in loss of immunoblotting by the anti-Mac-2 antibody. It was concluded that the antigenic recognition site of anti-Mac-2 is at the amino terminus. This is clearly inconsistent with the results of our mapping studies, which showed that the anti-Mac-2 antibody immunoblotted GST fusion proteins containing residues 46-100 of the Gal3 polypeptide but failed to immunoblot GST-hGal3 (1-14), GST-hGal3 (1-26), GST-hGal3 (1-47), all of which contained the $NH_2$-terminal 11 residues. Moreover, the 18-mer and 27-mer peptides bearing at least two iterations of

the PGAYPGQAP motif between residues 46-100 inhibited the immunoblotting of the anti-Mac-2 antibody without any effect on the immunoblotting of the NCL-GAL3 antibody. Finally, the 14-mer peptide containing the $NH_2$-terminus of human Gal3 failed to inhibit the immunoblotting observed with anti-Mac-2.

The original source of our anti-Mac-2 antibody was clone M3/38.1.2.8.HL.2 from ATCC TIB 166, identical to that reported in reference (32). Because of this apparent discrepancy with the results and conclusions of Gong *et al.* (32), we obtained an independent source of the anti-Mac-2 antibody from a commercial source (Acris Antibodies GmbH, Hiddenhausen, Germany). Using our GST fusion protein reagents, we found that the epitope of this preparation of the anti-Mac-2 antibody also mapped to residues 48-66, rather than the $NH_2$-terminal 11 residues of the Gal3 polypeptide. On this basis, we do not understand the discrepancy between our results and conclusions and those of Gong *et al.* (32).

Several intracellular binding partners of Gal3 have been identified: (a) Bcl-2 (33); (b) synexin (34); (c) Chrp (35); (d) cytokeratin (36); (e) Gemin4 (24); (f) Sufu (37); (g) TTF-1 (thyroid-specific transcription factor) (38); (h) β-catenin (39); and (i) the general transcription factor II-I (40). Where there is available evidence, only the COOH-terminal CRD of Gal3 has been implicated in these interactions. To the best of our knowledge, there has been no report of an interaction involving the ND with an identified partner. In this regard, it may be interesting to note that differential scanning calorimetry studies suggest that the ND of Gal3 has a very low melting temperature (~39°C), compared to the globular CRD which has a melting temperature of ~56°C (41). This implies that the ND may not be folded tightly until it interacts with a ligand. On this basis, the identification

of a binding partner that interacts with the PGAYPG motifs of the ND will be of great interest, not only in terms of the splicing reaction but also in terms of the possibility of determining the structure of the ND.

## REFERENCES


1. **Barondes, S. H., V. Castronovo, D. N. W. Cooper, R. D. Cummings, K. Drickamer, T. Feizi, M. A. Gitt, J. Hirabayashi, C. Hughes, and K. Kasai.** 1994. Galectins: a family of animal beta-galactoside-binding lectins. Cell **76:**597-8.

2. **Kasai, K.-i., and J. Hirabayashi.** 1996. Galectins: A Family of Animal Lectins That Decipher Glycocodes. J. Biochem. (Tokyo) **119:**1-8.

3. **Houzelstein, D., I. R. Goncalves, A. J. Fadden, S. S. Sidhu, D. N. W. Cooper, K. Drickamer, H. Leffler, and F. Poirier.** 2004. Phylogenetic Analysis of the Vertebrate Galectin Family. Mol. Biol. Evol. **21:**1177-1187.

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

5. **Laing, J. G., and J. L. Wang.** 1988. Identification of carbohydrate binding protein 35 in heterogeneous nuclear ribonucleoprotein complex. Biochemistry **27:**5329-34.

6. **Hubert, M., S. Y. Wang, J. L. Wang, A. P. Seve, and J. Hubert.** 1995. Intranuclear distribution of galectin-3 in mouse 3T3 fibroblasts: comparative analyses by immunofluorescence and immunoelectron microscopy. Exp. Cell Res. **220:**397-406.

7. **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. U.S.A. **92:**1213-7.

8. **Vyakarnam, A., S. F. Dagher, J. L. Wang, and R. J. Patterson.** 1997. Evidence for a role for galectin-1 in pre-mRNA splicing. Mol. Cell. Biol. **17:**4730-7.

9. **Patterson, R. J., 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.

10. **Robertson, M., K. Albrandt, D. Keller, and F. Liu.** 1990. Human IgE-binding protein: a soluble lectin exhibiting a highly conserved interspecies sequence and differential recognition of IgE glycoforms. Biochemistry **29:**8093-100.

11. **Cherayil, B., S. Chaitovitz, C. Wong, and S. Pillai.** 1990. Molecular Cloning of a Human Macrophage Lectin Specific for Galactose. Proc. Natl. Acad. Sci. U.S.A. **87:**7324-7328.

12. **Oda, Y., H. Leffler, Y. Sakakura, K. Kasai, and S. Barondes.** 1991. Human breast carcinoma cDNA encoding a galactoside-binding lectin homologous to mouse Mac-2 antigen. Gene **99:**279-83.

13. **Raz, A., P. Carmi, T. Raz, V. Hogan, A. Mohamed, and S. Wolman.** 1991. Molecular cloning and chromosomal mapping of a human galactoside-binding protein. Cancer Res. **51:**2173-2178.

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

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

16. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475-1489.

17. **Bradford, 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:**248-54.

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

19. **Konarska, M., and P. Sharp.** 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell **46:**845-55.

20. **Kadrofske, M., K. Openo, and J. Wang.** 1998. The human LGALS3 (galectin-3) gene: determination of the gene structure and functional characterization of the promoter. Arch. Biochem. Biophys. **349:**7-20.

21. **Laemmli, U.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-5.

22. **Merril, C., M. Dunau, and D. Goldman.** 1981. A rapid sensitive silver stain for polypeptides in polyacrylamide gels. Anal Biochem **110:**201-7.

23. **Davidson, P. J., M. J. Davis, R. J. Patterson, M. A. Ripoche, F. Poirier, and J. L. Wang.** 2002. Shuttling of galectin-3 between the nucleus and cytoplasm. Glycobiology **12:**329-37.

24. **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. Nucleic Acids Res. **29:**3595-602.

25. **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:**1008-1020.

26. **Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden.** 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc. Natl. Acad. Sci. U.S.A. **93:**6975-6980.

27. **Parker, A. R., and J. A. Steitz.** 1997. Inhibition of mammalian spliceosome assembly and pre-mRNA splicing by peptide inhibitors of protein kinases. RNA **3:**1301-1312.

28. **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.

29. **Massa, S., D. Cooper, H. Leffler, and S. Barondes.** 1993. L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity. Biochemistry **32:**260-7.

30. **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.

31. **Birdsall, B., J. Feeney, I. D. Burdett, S. Bawumia, E. A. Barboni, and R. C. Hughes.** 2001. NMR solution studies of hamster galectin-3 and electron microscopic visualization of surface-adsorbed complexes: evidence for interactions between the N- and C-terminal domains. Biochemistry **40:**4859-66.

32. **Gong, H. C., Y. Honjo, P. Nangia-Makker, V. Hogan, N. Mazurak, R. S. Bresalier, and A. Raz.** 1999. The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells. Cancer Res. **59:**6239-6245.

33. **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. U.S.A. **93:**6737-6742.

34. **Yu, F., R. L. Finley, Jr., A. Raz, and H.-R. C. Kim.** 2002. Galectin-3 translocates to the perinuclear membranes and inhibits cytochrome c release from the mitochondria. A role for synexin in galectin-3 translocation. J. Biol. Chem. **277:**15819-15827.

35. **Menon, R. P., M. Strom, and R. C. Hughes.** 2000. Interaction of a novel cysteine and histidine-rich cytoplasmic protein with galectin-3 in a carbohydrate-independent manner. FEBS Lett. **470:**227-31.

36. **Goletz, S., F. G. Hanisch, and U. Karsten.** 1997. Novel alphaGalNAc containing glycans on cytokeratins are recognized invitro by galectins with type II carbohydrate recognition domains. J. Cell Sci. **110 ( Pt 14):**1585-96.

37. **Paces-Fessy, M., D. Boucher, E. Petit, S. Paute-Briand, and M. Blanchet-Tournier.** 2004. The negative regulator of Gli, Suppressor of fused (Sufu), interacts with SAP18, Galectin3 and other nuclear proteins. Biochem. J. **378:**353-62.

38. **Paron, I., A. Scaloni, A. Pines, A. Bachi, F. Liu, C. Puppin, M. Pandolfi, L. Ledda, C. Di Loreto, G. Damante, and G. Tell.** 2003. Nuclear localization of galectin-3 in transformed thyroid cells: a role in transcriptional regulation. Biochem. Biophys. Res. Cummun. **302:**545-53.

39. **Shimura, T., Y. Takenaka, S. Tsutsumi, V. Hogan, A. Kikuchi, and A. Raz.** 2004. Galectin-3, a Novel Binding Partner of {beta}-Catenin. Cancer Res. **64:**6363-6367.

40. **Gray, R. M., P. G. Voss, R. J. Patterson, and J. L. Wang.** 2006. Interactions of Galectin-1 and Galectin-3 in pre-mRNA Splicing I. Identification of Transcription Factor II-I in Association with Galectin-containing Spliceosomal complexes. accompanying article, paper I.

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