MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 01046 0537

This is to certify that the

dissertation entitled

The Characterization of Nuclear Lectins as Novel Splicing Factors

presented by

Sue Fay Dagher

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Genetics

Ronald J. Patterson
Major professor

Date 7-12-94

MSU is an Affirmative Action/Equal Opportunity Institution

O-12771

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE

MSU Is An Affirmative Action/Equal Opportunity Institution
c:\circ\datedue.pm3-p.1

THE CHARACTERIZATION OF NUCLEAR LECTINS AS NOVEL SPLICING FACTORS

By

SUE FAY DAGHER

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Genetics Program

1994

ABSTRACT

THE CHARACTERIZATION OF NUCLEAR LECTINS
AS NOVEL SPLICING FACTORS

By

SUE FAY DAGHER

Galactose/lactose specific lectins (galectins) have been identified in the nuclear compartment of a number of mammalian cells. Immunofluorescent studies have shown that galectin-3 is one such lectin. Previous results have shown galectin-3 is a component of ribonucleoprotein complexes, which has lead to testing for its possible role in pre-mRNA splicing.

The addition of various saccharides to nuclear extracts was used to evaluate whether galectin-3 specific saccharides perturbed pre-mRNA splicing. Galactose containing saccharides as well as neoglycoconjugates with a high affinity for galectin-3 perturbed *in vitro* pre-mRNA splicing reactions. Mono- and disaccharides without affinity for galectin-3 had little effect on splicing activity. To specifically implicate galectin-3 to the saccharide inhibition results, nuclear extracts were immunodepleted of galectin-3. Although greater than 95% of galectin-3 is removed, splicing activity is only slightly diminished compared to a control depleted extract. However, when lactose is added to the galectin-depleted extracts, splicing is still inhibited. These results suggest another lactose binding protein(s) may be influenced by the addition of exogenous saccharides. To remove all lactose binding proteins from nuclear extracts, depletions were performed using saccharide-immobilized affinity matrices.

When tested for splicing activity, lactose depleted extracts were unable to process pre-mRNA while control cellobiose depleted extracts retained full splicing activity. Most importantly, splicing activity is restored to lactose depleted extracts by the inclusion of recombinant galectin-3 (rCBP35) or recombinant galectin-1 (rL-14).

We suggest that several lactose binding proteins exist in splicing extracts which are functionally redundant. Removal of one lactose binding protein permits the others to act in pre-mRNA processing, while removal of all of them abolishes splicing activity.

DEDICATION

TO MY PARENTS

Who always stood by me, always had faith in me and understood my love for science.

ACKNOWLEDGMENTS

First of all I would like to say that words alone cannot describe the patience shown and guidance given by Ron Patterson, my mentor, throughout my training. Ron was always ready to listen to my ideas and advise me when necessary. Over the years we worked as a team while performing experiments and designing experimental strategies to tackle new questions. All of this and much more has made science and its study the most exciting and fulfilling field to be involved in.

I want to thank my committee members Sue Conrad, Donna Koslowsky, Natasha Raihkel and Richard Schwartz for their guidance and helpful advise. It has also been a privilege to collaborate with John Wang and the members of his laboratory who provided advise and materials that enhanced my research.

Finally, I want to thank two good friends Trifon Adamidis and Georg Hausner for tolerating my continued tardiness when I was delayed with an experiment.

TABLE OF CONTENTS

ADDENDUM

LIST OF TABLES

LIST OF FIGURES

CHAPTER IV

LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
B	bound
CBP35	carbohydrate binding protein 35
Cello	cellobiose
CELLO-A	beaded agarose derivitized with cellobiose
CM-sepharose	carboxymethyl-sepharose
CP	creatine phosphate
CRD	carbohydrate recognition domain
DME	Dulbecco's modified Eagle's medium
DNase	deoxyribonuclease
FITC	Fluorescein isothiocyanate
Gal	galactose
Glu	glucose
HMG	high mobility group protein
hnRNP	heterogeneous nuclear ribonucleoprotein complex
Lac	lactose
LAC-A	beaded agarose derivitized with lactose
NE	nuclear extract

PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline + Tween 20
PMSF	phenylmethylsulfonyl fluoride
RNase	ribonuclease
SBA	soybean agglutinin
SDS	sodium dodecyl sulfate
snRNP	small nuclear ribonucleoprotein particle
TDG	thiodigalactoside
UB	unbound
WGA	wheat germ agglutinin

CHAPTER I

LITERATURE REVIEW

INTRODUCTION TO LECTINS

Lectin overview

Lectins are non-enzymatic non-immunoglobin proteins that bind carbohydrates. Lectins also agglutinate cells and precipitate polysaccharides and glycoproteins. These latter properties are due to the polyvalent nature of lectins; each lectin molecule has at least two carbohydrate binding sites to allowing crosslinking between cells or between glycoproteins (Sharon and Lis, 1989). Lectins have been found in plants, mammals, microorganisms and viruses. They are structurally diverse and in general are found as oligomeric proteins composed of subunits, usually with one carbohydrate recognition domain (CRD) per subunit. They vary, however, in amino acid composition, molecular weight, metal requirement, and three dimensional structure. Despite this variability, they can be grouped into families of structurally homologous proteins. For more detailed information refer to a comprehensive review by Liener et al. (1986).

The purpose of this review is to give an overview of the properties of mammalian lactose binding lectins (galectins, see below) emphasizing CBP35 (carbohydrate binding protein Mr=35,000). A variety of lactose binding proteins have been classified as soluble lactose binding lectins (S-Lac) (Drickamer, 1988). S-lac lectins are a family of vertebrate proteins which share homologous lactose CRD, do not require divalent cations for carbohydrate binding and are soluble without detergents (Drickamer, 1988; Leffler et al., 1989). Although grouped together partially on the basis of their ability to bind lactose, subtle differences in carbohydrate binding

specificities do exist (Leffler and Barondes, 1986). Recently a more comprehensive nomenclature for S-Lac lectins has been established. S-Lac lectins can be grouped into one of four galectin families (Barondes et al., 1994). For example, the galectin-3 family includes CBP35, Mac-2, IgEBP, CBP-30, RL-29, L-29, L-31, L-34, and LBL (Barondes et al., 1994). The galectin-1 family includes L-14-1, L-14, RL-14.5, galaptin, MGBP, GBP, BHL, CHA, HBP, HPL, HLBP 14, and rIML-1 (Barondes et al., 1994). A number of investigators have attached putative functions to many of the galectins (Harrison, 1991; Hamann et al., 1986, Frigeri and Liu, 1992, Mercurio and Shaw 1991, Cooper et al., 1991) which are presented later in more detail.

A historical prospective of the galectin-3 S-lac proteins

Over the past 14 years members of galectin-3 proteins have been identified in a number of mammalian cells. Although independently characterized and named based on discovery, all possess 80-90% homology to the originally described murine Mac-2 protein (Ho and Springer, 1982) and contain the galectin β-galactose binding motif (Drickamer et al., 1988). The chronological list includes: rat lung lectin (RL-29; Cerra et al., 1985), mouse tumor lectin (mL-34; Raz et al., 1986), human lung lectin (HL-29; Sparrow et al., 1987), rat immunoglobulin E-binding protein (IgE BP; Albrandt et al., 1987), carbohydrate binding protein (CBP35; Moutsatsos et al., 1987), mouse macrophage antigen (Mac-2; Cherayil et al., 1989), rat, mouse and human 29 kDa galactose binding lectin (L-29; Oda et al., 1991), human breast carcinoma galactoside binding lectin (hL-29; Oda et al., 1991) and human tumor-associated 31 kDa galactoside-binding lectin (hL-31; Raz et al., 1991).

Mac-2 was first observed as an inflammation responsive protein on the surface of macrophages which increases during inflammation suggesting an important role in the inflammatory response (Ho and Springer, 1982). In addition Mac-2 is the major nonintegrin laminin-binding protein synthesized by murine inflammatory macrophages, pointing to a potential role for Mac-2 in macrophage-extracellular interactions (Woo et al., 1990). The human homologue of Mac-2 (hMac-2) has been cloned and shown to have a highly conserved primary structure when compared to murine Mac-2 with 85% of the amino acids being similar (Cherayil et al., 1989, 1990). The *in vitro* synthesized hMac-2 has the galactose specific binding property of other L-29 lectins, and binds to purified laminin (Cherayil et al., 1989). A galactose binding protein which localizes on the cell surface, in the cytoplasm and the nucleus of 3T3 cell mouse fibroblasts has been identified and named carbohydrate binding protein 35 (CBP35) (Moutsatsos et al., 1987). Immunofluorescent analysis of 3T3 cells indicated CBP35 is a component of ribonucleoprotein complexes (Jia and Wang, 1988). The same protein was identified and named L-34 and suggested to be a tumor cell surface lectin that enhanced tumor metastasis by promoting the formation of multicellular emboli (Raz et al., 1989). An IgE binding protein found in the cytosol of rat cells (Liu et al., 1985) also has extensive homology to the galectin-3 lectins found thus far. CBP35 (Laing et al., 1989) and Mac-2 (Cherayil et al., 1989) also have the ability to bind murine IgE and this binding can be inhibited by galactose, which raises the possibility that galectin-3 species may be involved in the biosynthetic or functional regulation of IgE (Liu et al., 1985). The classic definition of a lectin is a class of proteins that bind carbohydrates to their CRD. However, we

should not rule out the possibility that glycoproteins and nonglycoproteins could interact with lectins at regions other than the CRD and this interaction may be altered by carbohydrates specific for the lectin. Such a situation may exist between IgE and IgE binding protein (Liu et al., 1985) since it is not known whether their interaction is through a galactose containing side chain or an IgE binding site on IgE binding protein (Cherayil et al., 1989).

The galectin-3 proteins cloned and sequenced so far possess two-domains. The carboxyl-terminal carbohydrate binding domain is particularly conserved in which long, uninterrupted stretches of amino acid identity can be found. The amino-terminal domain contains PGAYPG repeats whose number varies between Mac-2, hMac-2, L-34 and CBP35. The function of these amino-terminal repeats is unclear, however their evolutionarily conserved nature suggests an important role or function for the lectin. The following common characteristics have been summarized for the galectin-3 family of galectins: i) molecular weights between 26,200 and 30,300, ii) composed of two structural domains; the amino-terminal domain contains a number of proline glycine-rich tandem repeats, and the carboxyl-terminal domain contains a β-galactoside CRD, iii) expression by various cell types - both intracellular and as surface molecules, iv) highly conserved amino acid sequence among different species, v) absence of known transmembrane domain and signal sequence.

Putative functions for galectin-3 surface proteins

Although the galectin-3 proteins have been characterized for only 14 years the amount of information gathered has been bountiful as well as diverse.

Immunohistochemical, cell surface labeling, and ligand binding studies have shown galectin-3 proteins to localize in the cytosol, nucleus, and on the cell surface. As cell surface moieties they are assumed to function by interacting with complementary glycoconjugates (Frigeri and Liu, 1992; Gritzmacher et al., 1988; Moutsatsos et al., 1987; Raz et al., 1984; Weis et al., 1991). Woo et al. (1990) have shown Mac-2 may have an important role in the inflammatory process. Mac-2 is the major non-integrin laminin-binding protein expressed by murine peritoneal macrophages exposed to thioglycollate. Consequently, Mac-2 may aid in facilitating adhesion to the basement membrane laminin. A number of studies have provided further information on interactions between lectins and laminin and their functional significance (Bouzon et al., 1990; Cooper et al., 1991; Dean et al., 1990; Mercurio and Shaw, 1991; Sato and Hughes, 1992; Zhou and Cummings, 1990) Briefly, these studies suggest that lectin-laminin interactions may actually impair cell-cell adhesion by blocking laminin interaction with a cell surface receptor such as an integrin near an oligosaccharide. The metastatic state of melanoma and fibrosarcoma cells appears to correlate with the level of expression of L-34 (Raz et al., 1986; 1989) and metastasis of these cells could be inhibited by the addition of monoclonal antibody to L-34 (Meromsky et al., 1986). Cells transfected with cloned cDNA provided further evidence for the role of this lectin in tumor metastasis (Raz et al., 1991). These findings along with those obtained by Woo et al. (1990) demonstrate that Mac-2 may play a role in crossing the basement membrane via galactose containing sugar side chains of the basement laminin. Cherayil et al. (1990) suggest using the monoclonal antibody M3/38 which cross reacts with both murine and human Mac-2 as a tool to investigate the

involvement of human Mac-2 in human diseases.

Nuclear lectins

The existence of nuclear lectins has been reported in mammalian cells (Bourgeois et al., 1987; Facy at el., 1990; Laing et al., 1988; Moutsatsos et al., 1986, 1987; Olins et al., 1988; Seve et al., 1985, 1986), reptilian cells (Hubert et al., 1985) and in the macronucleus of protozoans (Olins et al., 1988). Strong evidence indicates nuclear lectins may be involved in cell proliferation (Hubert et al., 1989; Wang et al., 1991). Quantitative changes in nuclear and cytoplasmic lectins have been reported in conjunction with the proliferative state of the cell (Bourgeois et al., 1987; Moutsatsos et al., 1986, 1987; Seve et al., 1986) and to the cell cycle (Moutsatsos et al., 1986; Facy et al., 1990). Lower amounts of nuclear lectins were found in differentiated cells compared to undifferentiated cells (Facy et al., 1990). Further examination of nuclear lectins has revealed a preference for localization in regions enriched for ribonucleoprotein complexes (RNP) (Bourgeois et al., 1987; Facy et al., 1990; Hubert et al., 1985; Hubert et al 1989; Laing et al., 1988; Jia et al., 1988; Seve et al., 1985; Seve et al., 1986) and at the site of DNA replication in the macronucleus of *Euplotes eurystomus* (Olins et al., 1988). In higher eucaryotes these could be sites of transcriptional events, posttranscriptional modifications and regions RNPs are readied for export to the cytoplasm.

Analysis of binding of D-glucose and N-acetylglucosamine containing neoglycoproteins to nuclei from the the human tumor cell line HL-60 has been shown (Facy et al., 1990). A glucose binding protein (CBP67) isolated from rat liver

nuclei is present in nuclear RNP but absent in polysomal RNP indicating a possible role in nucleocytoplasmic export of mRNA (Schroder et al.,1992). It is possible that nuclear lectins mediate posttranscriptional processing and perhaps export of mRNAs specific for synthesis of proteins required for cell proliferation and DNA replication (Facy et al., 1990). In support of this idea, a set of hnRNPs was found to modulate cell proliferation during differentiation of 3T3 cells and normal keratinocyes in this manner (Minoo et al., 1989).

Nuclear glycoproteins

Theoretically, one would expect nuclear glycoproteins to serve as ligands for nuclear lectins. The existence of nuclear glycoproteins is well established (for reviews refer to Hart et al., 1989; Wang et al., 1991). It has been postulated that specific nuclear functions may be modulated by interactions between nuclear lectins and nuclear glycoproteins (Hubert et al., 1989) A series of nuclear glycoproteins have been identified and may serve as potential ligands [for example, the O-linked glycoproteins poly(A) polymerase (Kurl et al., 1988), several RNA polymerase II transcription factors (Jackson and Tjian, 1989; Lichtsteiner and Schibler, 1989) as well as N-linked glycoproteins; high mobility group proteins HMG14 and HMG17 (Reeves and Chang, 1983)]. In addition, an hnRNP (p43, Mr 43,000) has been characterized as a glycoprotein, bearing N-aceylglucosamine oligosaccharide residues (Soulard et al., 1991).

Glycoproteins have been found on the cytoplasmic and nucleoplasmic surfaces of nuclear pore complexes (Davis and Blobel, 1986; Hanover et al., 1987; Holt et al.,

1987) and associated with the nuclear matrix (Reeves and Chang, 1983). Nuclear protein import is blocked by the lectin wheat germ agglutinin which recognizes residues within the pore complex bearing O-linked N-acetylglucosamine (Finlay et al., 1987; Yoneda et al., 1987). On the other hand wheat germ agglutinin does not inhibit RNA export from isolated nuclei (Schroder et al., 1986). Nuclear export of mRNA and nuclear envelope nucleoside triphosphatase activity is inhibited by the lectin (Gerardia Savagia) specific for D-mannose bearing pore complex glycoproteins (Kljajic et al., 1987).

Although the list of nuclear glycoproteins continues to grow rapidly, no one has shown that any of these glycoproteins act as ligands for nuclear lectins.

Carbohydrate Binding Protein 35 (CBP35): galectin-3

About 10 years ago NIH 3T3 cell extracts were fractionated over an asialofetuin column (Roff and Wang, 1983). Three bound proteins could be eluted from the column with lactose (galactose, β1-4 glucose). Two of the proteins had molecular weights of 14 kDa (galectin-1) and 16 kDa. The third had a molecular weight of 35 kDa and was designated CBP35 (now galactin-3). The research that followed focused on this lectin. As previously mentioned galectin-3 proteins possess N-terminal domains not present in galectin-1 proteins. The carboxyl-domain of galectin-3 contains sequences homologous to other β-galactoside binding proteins, while the amino-terminal domain possesses a distinct repeated pattern of PGAYPG (Jia and Wang, 1988). Galectin-3, originally identified as an extracellular component known as Mac-2 (Ho and Springer, 1982), has been found in the cytoplasm, nucleus and on

the surface of numerous mammalian cells (Moutsatsos et al., 1986). Human Mac-2 has been detected in fibroblast cell lines and various epithelial cell lines, however it has not been detected in the lymphoid cell line BJAB, (Cherayil et al., 1990). Galectin-3 has been detected in all cell lines tested with the exception of the lymphoid cell line D1C1 (S. D., unpublished observations).

The pI of galectin-3 has been determined both experimentally and by calculation from the deduced 264 amino acid sequence (Cowles et al., 1990). When extracts of NIH 3T3 cells were subjected to NEPHGE and immunoblotted, two spots corresponding to pI 8.7 and 8.2 were observed (Cowles et al., 1990). The pI 8.2 form is the result of a post-translational modification of the pI 8.7 by the addition of a single phosphate group. Recently the L-29 galectin-3 has been shown to be phosphorylated predominatly (90%) at Ser^{6} and to a lesser extent (10%) at Ser^{12} (Huflejt et al., 1993). These two isoelectic species were found to be differentially expressed in the nucleus and cytoplasm of NIH 3T3 cells during quiescence, serum stimulation, and development (Moutsatsos et al., 1987; Agrwal et al., 1989). The cytoplasm contains predominantly the phosphorylated form of galectin-3. Both species are found in the nucleus. In quiescent cultures of NIH 3T3 cells galectin-3 is mainly found at low levels in the cytoplasm and only the phosphorylated form can be detected in the nucleus. In proliferating cultures the amount of phosphorylated galectin-3 increases both in the nucleus and cytoplasm with the most dramatic increase seen in the nuclear non-phosphorylated species (Cowles et al., 1990).

Immunofluorescent studies have shown that various enzymatic and biochemical treatments of cells such as high salt extraction and DNAase-treatment did not result

in the removal of nuclear galectin-3. However, RNAase-treatment and salt extraction resulted in quantitative removal of galectin-3 from nuclei (Laing and Wang, 1988). Lactose-affinity chromatograpy of nuclear extracts selected galectin-3 and several other proteins whose molecular weights matched those of hnRNP proteins (Wang et al., 1992; Laing and Wang, 1988). The nuclear localization of galectin-3 and possible association with hnRNP proteins has stimulated inquiry into its putative role in pre-mRNA splicing. The first indication that galectin-3 might be involved in the splicing process appeared when Lac and neoglycoconjugates containing Lac were tested for their ability to perturb pre-mRNA splicing. Lac inhibited *in vitro* pre-mRNA splicing in a concentration dependent manner while saccharides that did not have an affinity for galectin-3 had very little effect on splicing (Wang et al., 1992).

More recently a 70 kDa glucose binding lectin (CBP70) was isolated in a complex with galectin-3 from HL60 membrane-depleted nuclei and this association was controlled by the binding of galectin-3 to Lac (Seve et al., 1993). However, we have been unable to coprecipitate CBP70 and galectin-3 using glucose affinity resins (data not shown).

In summary, i) galectin-3 can be found extracellularly, in the cytoplasm and in the nucleus, ii) the amount and phosphorylation state of galectin-3 varies according to the proliferative state of the cell culture, iii) nuclear galectin-3 was found associated with RNA as component of ribonucleoprotein complexes, iv) specific disaccharides with a high affinity for galectin-3 inhibited *in vitro* splicing activity, whereas, mono- and di-saccharides without affinity for galectin-3 had little effect on splicing activity. These preliminary findings are the basis for the continued

examination of the involvement of galectin-3 in pre-mRNA splicing and the subject of this thesis.

REVIEW OF pre-mRNA SPLICING

Mechanism of pre-mRNA splicing

Pre-mRNA splicing proceeds through a two step mechanism involving two separate trans-esterification reactions (Figure 1). The first step includes cleavage at the 5' splice site generating free exon 1 with a 3'-OH terminus and exon 2-intron lariat. The lariat structure results from the formation of a 2'-5' phosphodiester bond linking the 5' terminus of the intron to the ribose 2' OH group of an adenosine residue in the intron creating a "branch point". The second step of the splicing reaction involves cleavage of the 3' splice site and ligation of exon 1 and exon 2 generating the spliced mRNA and the excised lariat intron (for a recent review, refer to Lamond, 1993).

In vitro splicing of pre-mRNA involves the formation of several splicing complexes including spliceosomes by way of ordered assembly of small nuclear ribonucleoprotein particles (snRNPs) (Table 1) and non-snRNP protein splicing factors on the pre-mRNA (Table 2) (for a review, refer to Luhrmann et al., 1990). The snRNPs involved in splicing are referred to as U1, U2, U4/U6 and U5 snRNPs. The snRNAs U1-U5 are transcribed by RNA polymerase II and snRNA U6, by RNA polymerase III. With the exception of U6, newly transcribed monomethyl G-capped U snRNA is transported to the cytoplasm where the cap is modified to a trimethyl

G structure and the snRNA assembles with proteins to form the snRNP (for a review, refer to Mattaj, 1988). Many of the snRNAs complex with a set of core proteins, B, B', D, D', E, F, and G which serve as antigens for human anti-Sm autoimmune antibodies in many patients with connective tissue disease (Lerner et al., 1979). Several snRNPs also have U specific proteins such as the U1 70 kDa, A and C proteins, and the U2 A' and B" proteins (Table 1). Figure 2 displays several known steps that take place leading to spliceosome formation. Prior to formation of the first splicing complex newly transcribed pre-mRNA is bound by hnRNP proteins to form the hnRNP complex or H complex. Formation of the H complex does not require functional 5' or 3' splice sites thus, excluding its formation as a functional intermediate in spliceosome formation (Konarska and Sharp, 1987). The first of several steps leading to spliceosome formation is the ATP independent interaction of U1 snRNP with both the 5' and 3' regions of the intron forming the "commitment complex". Next a pre-spliceosome complex referred to as the "A complex" is formed by the ATP dependent binding of U2 snRNP to the pre-mRNA branch site. Binding of U2 snRNP requires U1 snRNP (Barabino et al., 1990) and three protein factors: SF1, SF3 (Kramer and Utans, 1991) and U2 auxiliary factor (U2AF) (Ruskin et al., 1988; Zamore and Green, 1989). The spliceosome or "B complex" is formed following the addition of a preassembled U4/U6 U5 tri-snRNP particle and at this point the two transesterification reactions take place within the spliceosome complex.

Other putative mammalian splicing factors in addition to SF2/ASF are SF1 (containing U1 and U2 snRNPs), SF3, SF4A and SF4B (Table 2) (Krainer et al.,

1985). SF1, SF2 and SF4B are required prior to the 5' splice site cleavage step and SF3 and SF4A are required for the second catalytic step of splicing (Mayeda and Krainer, 1992). Another set of splicing factors named SF1*, SF2*, SF3* and SF4* are required for 5' splice site cleavage and lariat formation (Table 2) (Kramer et al., 1987; Utans and Kramer, 1990) (The "*" is used to differentiate SF2 and SF3 found by Mayeda and Krainer, 1992). Galectin-3 has not been found associated with snRNPs either by immunoprecipitation with anti-M2 (a monoclonal directed against Mac-2) or with anti-Sm (a polyclonal antibody directed against specific Sm proteins). However Lac affinity chromatography of nuclear extracts co-selected with galectin-3 several proteins whose molecular weights were characteristic of hnRNP proteins (Wang et al., 1992; Laing and Wang, 1988). I have therefore concentrated the remaining review on a set of splicing factors not associated with snRNPs.

Figure 1

The Two Transesterification steps in pre-mRNA Splicing

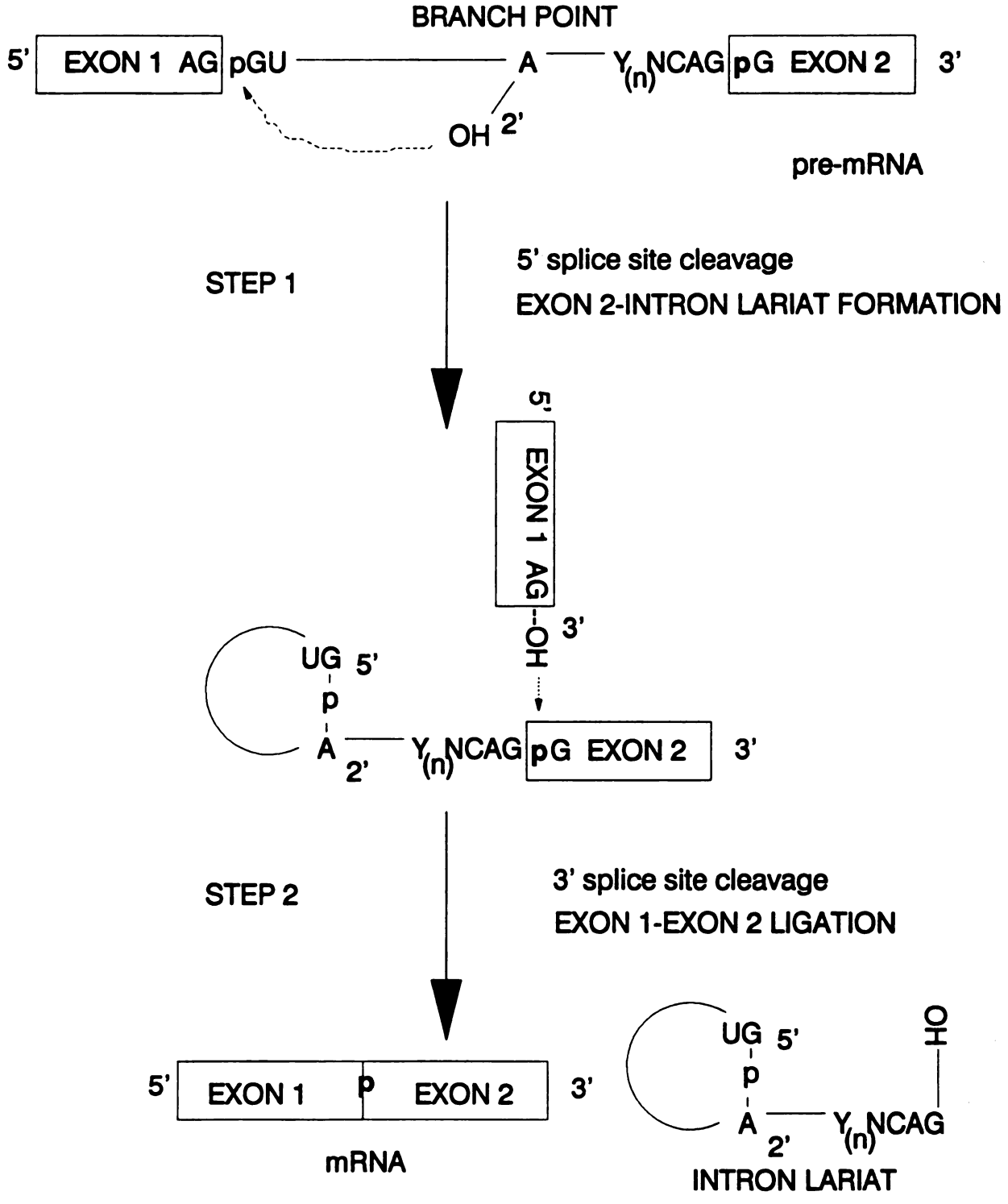


Table 1

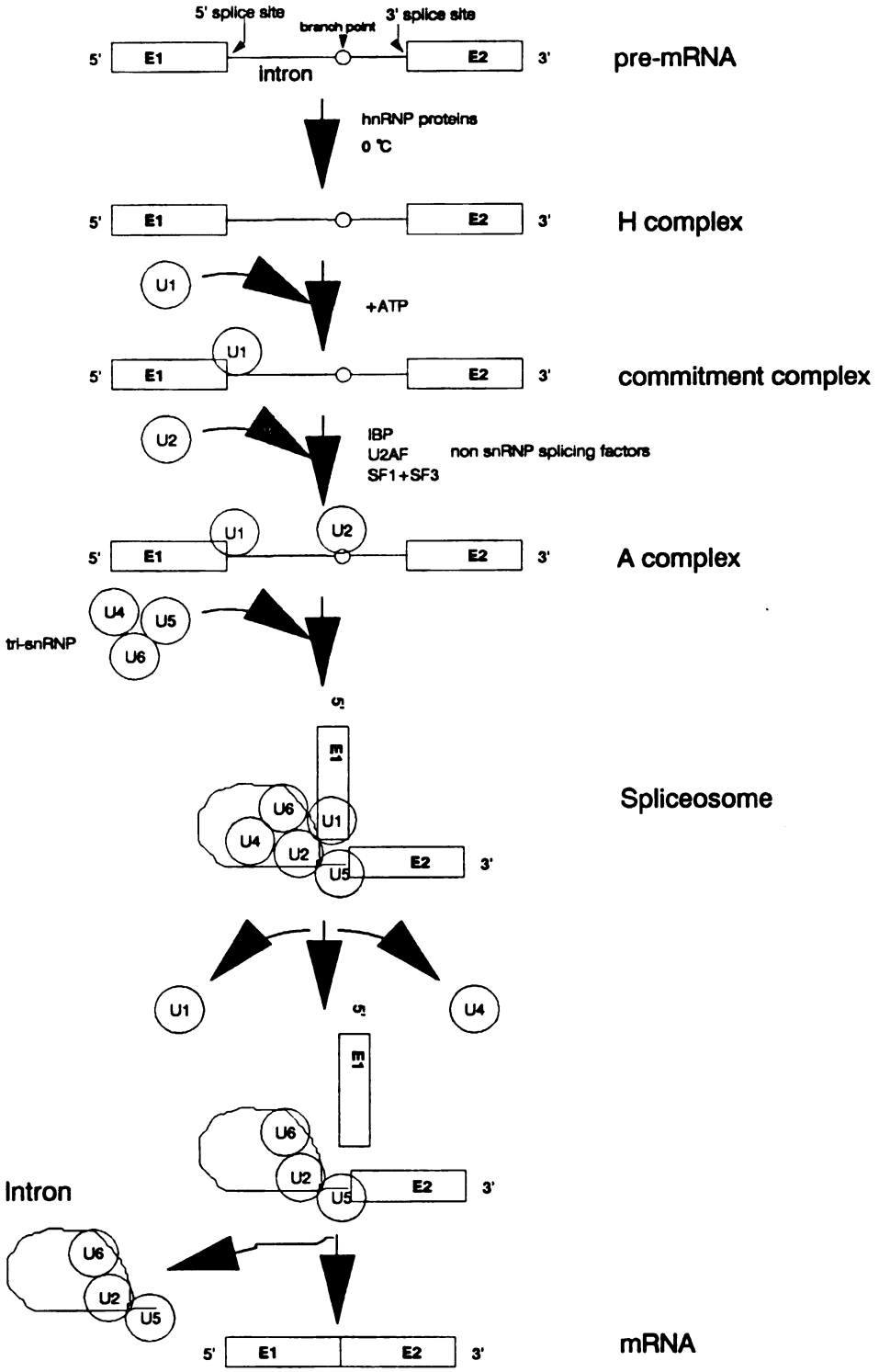
Mammalian snRNP proteins associated with snRNAs involved in pre-mRNA splicing.

The snRNP proteins associated with each snRNA shown to be involved in mammalian pre-mRNA splicing are listed along with protein mass and RNA sizes wherever possible. Core proteins and their respective masses are listed at the end of the table. The data presented is compiled from Reddy, 1986; Bennett et al., 1992 and Lamond, 1993.

RNA	nt length	proteins	molecular mass (kDa)
U1	165	core	
		70K	70
		A	34
		C	22
U2	188	core	
		A′	31
		B′′	28.5
U4	143	core	
U5	106		200
			116
			112
			110
			40
			15
U6	57		
		core	
		B	28
		B′	29
		D	16-18
		E	12
		F	11
		G	9

Figure 2

The pre-mRNA splicing pathway in mammalian cells



Non-snRNP protein splicing factors

Certain non-snRNP proteins have been identified in mammalian cells which are found to be essential for splicing activity *in vitro* and/or are stably associated with splicing complexes (Table 2). A comparison of sequences of these factors indicates they contain two classes of motifs found in other splicing factors: an arginine/serine-rich (RS) motif and an RNA recognition motif (RRM) (Table 2) (Lamm and Lamond, 1993). Certain RS regions of $U2AF^{65}$ (Zamore et al., 1992) and the splicing factor SF2/ASF (Ge et al., 1991; Krainer et al., 1991) were found to be similar to those of U1 snRNP 70k protein (Mancebo et al., 1990), and the Drosophila splicing regulators *transformer,* (Boggs et al., 1987), *transformer-2* (Amrein et al., 1988), *suppressor-of-sable* (Voelker et al., 1991) and *suppressor-of-apricot* (Chou et al., 1987). $U2AF^{35}$ also contains the SR domain, however it lacks the RRM and does not bind directly to RNA (Zhang et al., 1992). The location, length and composition of the SR motif varies among splicing factors. Both SF2/ASF and SC35/PR264 contain the distinctive RRM and RS motif, however with only 31% amino acid sequence identity (Bandziulis et al., 1989; Kenan et al., 1991). SF2/ASF protein is seen to vary in abundance in different rat tissues (Fu et al., 1992) and variable levels of SC35/PR264 mRNAs are expressed in different cell lines and stages of development (Fu and Maniatis, 1992; Vellard et al., 1992). The possible functions of these non-snRNP splicing factors are discussed below.

U2AF (U2 snRNP auxilliary factor) is composed of a 65 kDa protein ($U2AF^{65}$) and an associated 35 kDa protein ($U2AF^{35}$) (Zamore and Green, 1989). So far only the 65 kDa polypeptide is required for *in vitro* splicing of pre-mRNA (Zamore and

Green, 1989). Prior to spliceosome formation $U2AF^{65}$ binds to the pre-mRNA polypyrimidine tract in an ATP independent fashion and is believed to stablize the ATP dependent binding of U2 snRNP to pre-mRNA during splicesome assembly (Ruskin et al., 1988). U2AF has also recently been shown to play a role in alternative splicing (Zamore et al., 1992).

SF2 (splicing factor 2) (Krainer and Maniatis, 1985; Krainer et al., 1990) also known as ASF or ASF-1 (alternative splicing factor) (Ge and Manley, 1990; 1991) is a phosphoprotein with an apparent molecular mass of 27 kDa (Ge et al., 1991; Krainer et al., 1991). SF2/ASF has been shown to be required for the first step in assembly of the first detectable ATP dependent presplicesome complex (A complex) (Krainer et al., 1990; Ge and Manley, 1990). SF2/ASF has been shown to be involved in 5' splice site selection, activating the proximal 5' splice site in pre-mRNAs containing two or more 5' splice sites (Eperon and Krainer, 1993; Ge and Manley, 1990; Harper and Manley, 1992; Krainer et al, 1990). In addition SF2/ASF preferential splice site selection can be counteracted by A1 hnRNP (Mayeda and Krainer, 1992) which is thought to be antagonistic to the binding of U1 snRNP (Buvoli et al., 1992; Eperon et al., 1993). Both SF2/ASF and hnRNP A1 have strand annealing activities which may promote base pairing of snRNAs to the intron at alternative splice sites (Krainer et al., 1990; Kumar et al., 1990; Munroe et al., 1992; Pontius et al., 1990).

SC35 (splicesome component of 35 kDa) (Fu and Maniatis, 1990; 1992a; 1992b) also named (PR264) (Vellard et al., 1992) is a phosphoprotein required for the first ATP dependent step in splicesome assembly and appears to be one of the factors

that mediate U1 snRNP and U2 snRNP interactions at the 3' splice site (Fu and Maniatis, 1992a; 1992b). Similar to SF2/ASF, SC35/PR264 favors proximal over distal 5' splice sites and is counteracted by hnRNP A1 (Fu and Maniatis, 1990; Spector et al., 1991). Unaffected by hnRNP A1 is the ability of both SF2/ASF and SC35/PR264 to influence 3' splice site selection in pre-mRNA containing competing 3' splice sites (Fu et al., 1992). SF2/ASF can compliment extracts immunodepleted of SC35/PR264 (Fu et al., 1992) and both SF2/ASF and SC35/PR264 can complement S100 extracts for splicing (Fu et al., 1992; Krainer and Maniatis, 1985). Recently it has been suggested that both SC35 and SF2/ASF play a role in modulating alternative 3' and 5' splice site selection, favoring proximal splice site selection at both ends of the intron which may be important for preventing exon skipping in complex pre-mRNAs containing multiple introns (Fu et al., 1992). Finally SF2/ASF and SC35/PR264 appear to be redundant *in vitro* splicing factors with highly related but distinct amino acid sequences (Fu et al., 1992).

HnRNP proteins

Several hnRNP proteins also subscribe to a set of non-snRNP protein splicing factors. HnRNP proteins are found on the first detectable splicing complex (Bennett et al., 1992a) and can be detected in purified spliceosomes (Bennett el al., 1992b). A defined function for hnRNP proteins has yet to be established, however there is increasing evidence that hnRNP proteins A1, C1/C2 and I/PTB contribute to efficiency and/or regulation of splicing. The hnRNP A1 protein contains RRMs and a glycine rich domain (Query et al., 1989). The A1 protein has been observed to

preferentially bind the pyrimidine rich tract at the 3' end of introns, (Swanson and Dreyfuss, 1988) which is suggested to aid in discriminating among pre-mRNAs and signal these differences to other components of the splicing complex (Lamm et al., 1993). The A1 protein can also reanneal complementary RNA strands and this activity may alter the interaction of snRNAs with the pre-mRNA (Mayrand and Pederson, 1990). *In vitro* competition between A1 and SF2/ASF or SC-35 has been shown to regulate choice of 5' splice site of pre-mRNAs containing alternative 5' splice sites (Mayeda et al., 1992).

The hnRNP C1 and C2 proteins contain a single RRM domain (Swanson et al., 1987) and bind to poly(U) and to the U rich polypyrimidine region around the 3' splice site of pre-mRNA and in downstream regions important for 3' end cleavage and polyadenylation (Wilusz et al., 1990). A specific monoclonal that recognizes the C proteins (mAb 4F4) has been used to implicate C proteins in the splicing process. For example, addition of mAb 4F4 to an *in vitro* splicing reaction inhibits 5' splice site cleavage but does not block spliceosome formation (Choi et al., 1986).

The hnRNP I, also named PTB, protein contains RRM domains and can be specifically UV crosslinked to the polypyrimidine tract at the 3' end of introns (Garcia et al., 1989; Gil et al., 1991; Patton et al., 1991). Unlike other hnRNP proteins which concentrate in the nucleoplasm, hnRNP/PTB also localizes in the perinuclear space (Ghetti et al., 1992). Antibodies directed against recombinant hnRNP I/PTB coprecipitate the spliceosome (Patton et al., 1991) and hnRNP I/PTB has been detected in purified spliceosomes (Bennett et al., 1992b; Garcia et al., 1989) implicating this protein to the splicing process. Immunodepletion and

complementation studies have shown that hnRNP I/PTB cannot restore splicing activity on its own but also requires a 100 kDa protein which has been shown to coprecipitate with hnRNP I/PTB along with a 33 kDa protein (Patton et al., 1991). HnRNP I/ PTB has been suggested to modulate alternative 3' splice site selection although its exact role in splicing remains to be established (Mullen et al., 1991).

SAPs

A set of proteins termed SAPs (spliceosome associated proteins) associated with purified spliceosmes, have not been previously described as snRNP proteins or non-snRNP splicing factors (Bennett et al., 1992b). The SAPs have been shown to temporally bind to pre-mRNA during spliceosome assembly (Bennett et al., 1992b). A component of U5 snRNP, IBP (intron binding protein), binds to the 3' end of introns and reacts with anti-Sm antibodies (Gerke and Steitz, 1986; Tazi et al., 1986). Immunodepletion of an 88 kDa protein (a putative SAP) from nuclear extracts arrested spliceosome assembly after formation of "A complex" and splicing activity could be restored by addition of affinity purified 88 kDa protein (Ast et al., 1991).

The functional significance of most snRNP and non-snRNP associated proteins in pre-mRNA splicing is not known. Possible roles suggested include splice site selection, positioning of RNA and/or other protein components, stability, and participation in cleavage and/or ligation.

Table 2

Mammalian splicing factors

A comprehensive list of mammalian non-snRNP splicing proteins, splicing factor and SAPs are shown. Included is the size of each factor and numbers of RRM (RNA recognition) and SR (Arginine-serine) motifs found in each factor. This table has been adapted from Table I in Lamm and Lamond, 1993.

Protein	Size (kDa)	RRM	SR
$U2AF^{65}$	53	3	+
$U2AF^{35}$	34	0	+
SF2/ASF	28	1	+
SC-35	26	1	+
SF3	?		not identified
SF4A	?		not identified
SF4B	?		not identified
SF1*	?		not identified
SF2*	50		not identified
SF3*	?		not identified
SF4*	110		not identified
hnRNP A1	34	2	-
hnRNP C1/C2	33/35	1	-
hnRNP 1/PTB	60	4	-
IBP	100		not identified
88-kDa	88		not identified
SAPs	155, 145, 130,115, 114, 102, 92, 90, 88, 82, 72, 68, 62, 61, 60, 57, 55, 49, 42, 33		not identified

CONCLUSION: Nuclear galectins and pre-mRNA splicing: is there a connection?

As previously stated, the major impetus for the research described in this thesis derived from the fact that galectin-3 appeared to be a component of nuclear RNP complexes. While nuclear RNPs are involved in various nuclear events relating to RNA biogenesis, their role in splicing has been unequivocally proved. The fact that mono- and disaccharides specific for galectin-3 perturbed *in vitro* pre-mRNA splicing intensified the need to prove a role for galectin-3 in splicing using the commonly accepted criteria of depletion and reconstitution as the standard for identification of splicing components. This thesis describes such experiments and defines galectin-3 as a new splicing factor.

BIBLIOGRAPHY

Albrandt, K., Orida, N. K., and Liu, F.-T. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 6859-6863.

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

Arumugham, R. G., Hsieh, T. C.-Y., Tanzer, M. L., and Laine, R. A. (1986) *Biochim. Biophys. Acta* 883, 112-126.

Arwal, N., Wang, J. L. and Voss, P. G. (1989) *J. Biol. Chem.* 264, 17236-17242.

Ast. G., Goldblatt, D., Offen, D., Sperling, J. and Sperling, R. (1991) *EMBO J.* 10, 425-432.

Bandziulis, R. J., Swanson, M. S. and Dreyfuss, G. (1989) *Genes Dev.* 3, 431-437.

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

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

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

Bennett, M., Pinol-Roma, S., Staknis, D., Dreyfuss, G. and Reed, R. (1992a) *Mol. Cell Biol.* 12, 3165-3175.

Bennett, M., Michaud, S., Kingston, J. and Reed, R. (1992b) *Genes and Development* 6, 1986-2000.

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

Bourgeois, C. A., Seve, A. P., Monsigny, M. and Hubert, J. (1987) *Exp. Cell Res.* 172, 365-376.

Bouzon, M., Dussert, C., Lissitsky, J.-C., and Martin, P. M. (1990) *Exp. Cell Res.* 190, 47-56.

Buvoli, M., Cobianchi, F. and Riva, S. (1992) *Nuc. Acids Res.* 20, 5017-5025.

Carmo-Fonseca, M., Tollervey, D., Pepperkok, R., Barabino, S. M., Merdes, A.,

Brunner, C., Zamore, P. D., Green, M. R., Hurt, E. and Lamond, A. I. (1991) *EMBO J.* 10, 195-206.

Cerra, R. F., Gitt, M. A. and Barondes, S. H. (1985) *J. Biol. Chem.* 260, 10474-10477.

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

Cherayil, B. J., Chaitovitz, S., Wong, C. and Pillai, S. (1990) *Proc. Natl. Acad. Sci. USA* 87, 7324-7328.

Choi, Y. D., Grabowski, P. J., Sharp P. A. and Dreyfuss, G. (1986) *Science* 231, 1534-1539.

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

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

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

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

Crittenden, S. L., Roff, C. F., and Wang, J. L. (1984) *Mol. Cell. Biol.* 4, 1252-1259.

Davis, L. I. and Blobel, G. (1986) *Cell* 45, 699-709.

Dean, J. W., Chandrasekaran, S., and Tanzer, M. L. (1990) *J. Biol. Chem.* 265, 12553-12562.

Drickamer, K. (1988) *J. Biol. Chem.* 263, 9557-9560.

Eperon, I. C. and Krainer, A. R. (1993) In Hames, B. D. and Higgins, S. J. (eds), *RNA Processing: A practical Approach.*

Facy, P., Seve, A. P., Hubert, M., Monsigny, M. and Hubert, J. (1990) *Exp. Cell Res.* 190, 151-160.

Finlay, D. R., Newmeyer, D. D., Price, T. M. and Forbes, D. J. (1987) *J. Cell Biol.* 104, 189-200.

Frigeri, L. G. and Liu, F.-T. (1992) *J. Immuol.* 148, 861-867.

Fu, X.-D. and Maniatis, T. (1990) *Nature* 343, 437-440.

Fu, X.-D. and Maniatis, T. (1992) *Science* 256, 535-538.

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

Fu, X.-D., Mayeda, A., Maniatis, T. and Krainer, A. R. (1992) *Proc. Natl. Acad. Sci. USA* 89, 11224-11228.

Fujiwara, S., Shinkai, H., Deutzmann, R., Pulsson, M., and Timpl, R. (1988) *Biochem. J.* 252, 453-461.

Garcia, B. M., Jamison, S. F. and Sharp, P. A. (1989) *Genes Dev.* 3, 1874-1886.

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

Ge, H., Zuo, P. and Manley, J. L. (1991) *Cell* 66, 373-382.

Gerke, V. and Steitz, J. A. (1986) *Cell* 47, 973-984.

Ghetti, A., Pinol-Roma, S., Michael, W. M., Morandi, C. and Dreyfuss, G. (1992) *Nuc. Acids Res.* 20, 3671-3678.

Gil, A., Sharp, P. A., Jamison, S. F. and Garcia, M. (1991) *Gens Dev.* 5, 1224-1236.

Green, M. R. *Rev. Cell Biol.* (1991) 7, 559-599.

Gritzmacher, C. A., Robertson, M. W., and Liu, F.-T. (1988) *J. Immunol.* 141, 2801-2806.

Guthrie, C. (1991) *Science* 253, 157-163.

Hamann, K. K., Cowles, E. A., Wang, J. L., and Anderson, R. L. (1991) *Exp. Cell Res.* 196, 82-91.

Hanover, J. A., Cohen, C. K., Willingham, M. C. and Park, M. K. (1987) *J. Biol. Chem.* 262, 9887-9894.

Harper, J. E. and Manley, J. L. (1992) *Gene Expression* 2, 19-29.

Harrison, F. L. (1991) *J. Cell Sci.* 100, 9-14.

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

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

Holt, G. D., Snow, C. M., Senior, A., Haltiwanger, R. S., Gerace, L. and Hart, G. W. (1987) *J. Cell Biol.* 104, 1157-1164.

Hubert, J., Seve, A. P., Bouvier, D., Masson, C., Bouteille, M. and Monsigny, M. (1985) *Biol. Cell* 55, 15-20.

Hubert, J., Seve, A. P., Facy, P. and Monsigny, M. (1989) *Cell Differentiation and Development* 27, 69-81.

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

Jackson, S. P. and Tjian, R. (1988) *Cell* 55, 125-133.

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

Kenan, D. J., Query, C. C. and Keene, J. d. (1991) *Trends Biochem. Sci.* 16, 214-220.

Kljajic, Z., Schroder, H. C., Rottmann, M., Cuperlovic, M., Movsesian, M., Uhlenbruck, G., Gasic, M., Zahn, R. K. and Muller, W. E. G. (1987) *Eur. J. Biochem.* 169, 97-104.

Knibbs, R. N., Perini, F., and Goldstein, I. J. (1989) *Biochemistry* 28, 6379-6392.

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

Krainer, A. R. and Maniatis, T. (1985) *Cell* 42, 725-736.

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

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

Kramer, A., Frick, M. and Keller, W. (1987) *J. Biol. Chem.* 262, 17630-17640.

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

Kumar, A. and Wilson, S. H. (1990) *Biochemistry* 29, 10717-10722.

Kurl, R. N., Holmes, S. C., Verney, E. and Sidransky, H. (1988) *J. Biol. Chem.* 27, 8974-8980.

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

Laing, J., Robertson, M., Gritzmacher, C., Wang, J. and Lui, F.-T. (1989) *J. Biol. Chem.* 264, 1907-1910.

Lamm, G. M. and Lamond, A. I. (1993) *Biochimica et Biophysica Acta* 1173, 247-265.

Lamond, A. I. (1993) *BioEssays* 15, 595-601.

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

Leffler, H., Masiarz, F. R., and Barondes, S. H. (1989) *Biochemistry* 28, 9222-9229.

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

Lichtsteiner, S. and Schibler, U. (1989) *Cell* 57, 1179-1187.

Liener, I. E., Sharon, N. and Goldstein, I. J., eds (1986) *Ann. Rev. Biochem.* 55, 35-67.

Lindstedt, R., Apodaca, G., Barondes, S., Mostov, K., and Leffler, H. (1991) *J. Cell Biol.* 115, 399a (Abstr. 2318).

Liu, F.-T., Albrandt, K., Mendel, E., Kulczycki, A. and Orida, N. K. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4100-4104.

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

Mancebo, R., Lo, P. C. and Mount, S. M. (1990) *Mol. Cell Biol.* 10, 2492-2502.

Maniatis, T. and Reed, R. (1987) *Nature* 325, 673-678.

Massa, S. M., Copper, D. N. W., Leffler, H., and Barondes, S. H. (1992) *J. Cell. Biochem. Suppl.* 16D, 174 (Abstr. P420).

Mattaj, I. W. (1988) In Structure and Function of Major and Minor Small Nuclear RNAs. M. Birnstiel, editor. Springer-Verlag. New York.

Mayeda, A. and Krainer, A. R. (1992) *Cell* 68, 365-375.

Mayeda, A., Zahler, A. M., Krainer, A. R. and Roth, M. B. (1992) *Proc. Natl. Acad. Sci. USA* 89, 1301-1304.

Mayrand, S. H. and Pederson, T. (1990) *Nucleic Acids Res.* 18, 3307-3318.

Mercurio, A. M. and Shaw, L. M. (1991) *Bioassays* 13, 469-473.

Meromsky, L., Lotan, R. and Raz, A. (1986) *Cancer Res.* 46, 5270-5275.

Minoo, P., Sullivan, W., Solomon, L. R., Martin, T. E., Toft, D. O., and Scott, R. E. (1989) *J. Cell Biol.* 109, 1937-1946.

Monsigny, M., Kieda, C., and Roche, A. C. (1983) *Biol. Cell* 47,95-110.

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

Moutsatsos, I. K., Wade, M., Schindler, M. and Wang, J. (1987) *Proc. Natl. Acad. Sci. USA* 84, 6452-6456.

Mullen, M. P., Smith, C. W., Patton, J. C. and Nadal, G. B. (1991) *Genes Dev.* 5, 642-655.

Munroe, S. H., and Dong, X.-F. (1992) *Proc. Natl. Acad. Sci. USA* 89, 895-899.

Oda, Y., Leffler, H., Sakakura, Y., Kasai, K., and Barondes, S. H. (1991) *Gene* 99, 279-283.

Olins, D. A., Olins, A. L., Seve, A. P., Bourgeois, C. A., Hubert, J. and Monsigny, M. (1988) *Biol. Cell* 62, 95-98.

Patton, J. G., Mayer, S. A., Tempst, P. and Nadal-Ginard, B. (1991) *Genes Dev.* 5, 1237-1251.

Pontius, B. W. and Berg, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8403-8407.

Query, C. C., Bentley, R. C. and Keene, J. D. (1989) *Cell* 57, 89-101.

Raz, A., Meromsky, L., Carmi, P., Karkash, R., Lotan, D., and Lotan, R. (1984) *EMBO J.* 3, 2979-2983

Raz, A., Meromsky, L. and Lotan, R. (1986) *Cancer Res.* 46, 3667-3672.

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

Raz, A., Carmi, P., Raz, T., Hogan, V., Mohammed, A., and Wolman, S. R. (1991a) *Cancer Res.* 51, 2173-2178.

Raz, A., Zhu, D., Hogan, V., Shah, N., Raz, T., Karkash, R., Pazerini, G. and Carmi, P. (1991b) *Int. J. Cancer* 46, 871-

Reddy, R. (1986) *Nucl. Acids Res.* 14, r61-r72.

Reeves, R. and Chang, C. (1983) *J. Biol. Chem.* 258, 679-687.

Rio, D. C. (1992) *Curr. Opin. Cell Biol.* 4, 444-452.

Robertson, M. W., Albrandt, K., Keller, D., and Liu, F. T. (1990) *Biochemistry* 29,

8093-8100.

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

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

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

Schroder, H. C., Becker, R., Bachmann, M., Gramzow, M., Seve, A.-P., Monsigny, M. and Muller, W. E. G. (1986) *Biochim. Biophys. Acta 868* 108-118.

Schroder, H. C., Facy, P., Monsigny, M., Pfeifer, K., Bek, A., and Muller, W. E. G. (1992) *Eur. J. Biochem.* 205, 1017-1025.

Seve, A. P., Hubert, J., Bouvier, D., Bouteille, M. and Monsigny, M. (1985) *Exp. Cell Res.* 157, 533-538.

Seve, A. P., Hubert, J., Bouvier, D., Bourgeois, C., Midoux, P., Roche, A. C. and Monsigny, M. (1986) *Proc. Natl Acad. sci. USA* 83, 5997-6001.

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

Sharon, N. and Lis, H. eds (1989) *Lectins* (Chapman and Hall, London).

Sharp, P. A. (1987) *Science* 235, 766-771.

Soulard, M., Barque, J.-P., Valle, V. D., Hernandez-Verdun, H., Masson, C., Danon, F., and Larsen, C.-J. (1991). *Exp. Cell Res.* 193, 59-71.

Sparrow, C. P., Leffler, H. and Barondes, S. H. (1987) *J. Biol. Chem.* 262, 7383-7390.

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

Sparrow, C. P., Leffler, H., and Barondes, S. H. (1987) *J. Biol. Chem.* 262, 7383-7390.

Swanson, M. S., Nakagawa, T. Y., LeVan, K. and Dreyfuss, G. (1987) *Mol. Cell Biol.* 7, 1731-1739.

Swanson, M. S. and Dreyfuss, G. (1988) *EMBO* J. 11, 3519-3529.

Tazi, J., Alibert, C., Temsamini, J., Reveillaud, I., Cathala, G., Brunel, C. and Jeanteur, P. (1986) *Cell* 47, 755-766.

Tracy, B. M., Feizi, T., Abbott, W. M., Carruthers, R. A., Green, B. N., and Lawson,

A. M. (1992) *J. Biol. Chem.* 267, 10342-10347.

Utans, U. and Kramer, A. (1990) *Methods Enzymol.* 181, 3-19.

Vellard, M., Sureau, A., soret, J., Martinerie, C. and Perbal, B. (1992) *Proc. Natl. Acad. Sci. USA* 89, 2511-2515.

Voelker, R. A., Gibson, W., Graves, J. P., Sterling, J. F. and Eisenberg, M. T. (1991) *Mol. Cell Biol.* 11, 894-905.

Wang, J. L., Laing, J. G., and Anderson, R. L. (1991) *Glycobiology* 1, 243-252.

Wang, J. L., Werner, E. A., Laing, J. G. and Patterson, R. J. (1992) *Trans. Biochem. Soc.* 20, 269-274.

Wasano, K., Hirakawa, Y., and Yamamoto, T. (1990) *Cell Tissue Res.* 259, 43-49.

Weis, W. I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W. A. (1991) *Science* 254, 1608-1615.

Wilusz, J. and Shenk, T. (1990) *Mol. Cell Biol.* 10, 6397-6407.

Woo, H. J., Shaw, L. M., Messier, J. M. and Mercurio, A. M. (1990) *J. Biol. Chem.* 265, 7097-7099.

Woo, H. J., Lotz, M. M., Jung, J. U., and Mercurio, A. M. (1991) *J. Biol. Chem.* 266, 18419-18422.

Yoneda, Y., Imamoto-Sonobe, N., Yamaizumi, M. and Uchida, T. (1987) *Exp. Cell Res.* 173, 173, 586-595.

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

Zamore, P. D. and Green, M. R. (1991) *EMBO* 10, 207-214.

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

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

Zhou, Q., and Cummings, R. D. (1990) *Arch. Biochem. Biophys.* 281, 27-35.

CHAPTER II

IDENTIFICATION OF GALECTIN-3 AS A REQUIRED FACTOR IN pre-mRNA SPLICING (CBP35/lectin)

Sue F. Dagher[•], John L. Wang[†] and Ronald J. Patterson[§]

The Departments of Microbiology[§] and Biochemistry[†] and the Genetics Program[•], Michigan State University, East Lansing, MI 48824

FOOTNOTES

Acknowledgement

We thank Richard Schwartz and Betty Werner in the Department of Microbiology for helpful discussions and members of the Wang laboratory for their invaluable assistance. This work was supported by grants MCB 91-22363 from the National Science Foundation (R.J.P.) and GM-38740 from the National Institutes of Health (J.L.W.).

ABSTRACT

Galectin-3 (Mr ~35,000) is a galactose/lactose specific lectin found in association with ribonucleoprotein complexes in many animal cells. Cell-free splicing assays have been carried out to study the requirement of galectin-3 in RNA processing by HeLa cell nuclear extracts using ^{32}P-labeled MINX as pre-mRNA substrate. Addition of saccharides that bind galectin-3 with high affinity inhibited product formation in the splicing assay; parallel addition of carbohydrates that do not bind to the lectin failed to yield the same result. Nuclear extracts depleted of galectin-3 by affinity adsorption on a lactose-agarose column were deficient in splicing. Extracts subjected to parallel adsorption on control cellobiose-agarose remained splicing competent. The activity of the galectin-3-depleted extract could be reconstituted by the addition of purified recombinant galectin-3 whereas the addition of other lectins, either with a similar saccharide binding specificity (soybean agglutinin) or with a different specificity (wheat germ agglutinin), did not restore splicing activity. The formation of splicing complexes was also sensitive to galectin-3 depletion and reconstitution. Together, these results define a requirement for galectin-3 in pre-mRNA splicing and identify it as a new splicing factor.

INTRODUCTION

Galectin-3 (Barondes et al., 1994) is the new name for the galactose(Gal)/lactose(Lac) specific lectin previously known under a number of different designations, including Carbohydrate Binding Protein 35 (CBP35) (Jia and Wang, 1988), Mac-2 (Cherayil et al., 1989), IgE-binding protein (Albrandt et al., 1987), CBP30 (Sato et al., 1993), L-29 (Leffler et al., 1989), and L-34 (Raz et al., 1989). In this communication, we will use the designation galectin-3 when we refer to the gene/protein in general, assuming that studies carried out on the gene/protein under any one of the above names is applicable to all of them. There are instances, however, in which the specific molecule used by one laboratory is slightly (but of significance) different then the corresponding molecule of another laboratory (e.g., the cDNAs reported for murine CBP35, Mac-2, and L-34 are of different lengths). In those cases, we will use the old designation to highlight the specific source of the molecule.

The galectin family of animal lectins is distinguished by the Gal/Lac specificity of its carbohydrate recognition domain (CRD), with highly conserved residues between members of the family (galectin-1, -2, -3, and -4) and between the homologues found in various species for any given single member (for reviews, see Anderson and Wang, 1992 and Hirabayashi and Kasai, 1993). The polypeptide of galectin-3 is delineated into two domains (Anderson and Wang, 1992; Jia and Wang, 1988): an amino-terminal half that is Pro and Gly rich, with limited homology to proteins of the heterogeneous nuclear ribonucleoprotein complexes and a carboxyl-terminal half homologous to the CRD of other members of the galectin family.

Another distinguishing feature of the galectin family is that all members have been found both inside cells as well as at the cell surface. For galectin-3, the majority of the protein is found in the cytoplasm and nucleus of mouse 3T3 fibroblasts, in the form of ribonucleoprotein (RNP) complexes (Laing and Wang, 1988; Wang et al., 1992). For example, treatment of permeabilized cells with ribonuclease A released the lectin from the nucleus, with loss of immunofluorescent staining. Two isoelectric species of galectin-3 have been identified: a phosphorylated (pI 8.2) and a non-phosphorylated (pI 8.7) form (Cowles et al., 1990). The phosphorylated species is found in both the nucleus and cytoplasm whereas the nonphosphorylated form localizes exclusively in the nucleus.

Nuclear extracts (NE), prepared from HeLa cells and capable of carrying out pre-mRNA splicing, contain galectin-3. On the basis of preliminary experiments that showed saccharide ligands with high affinity for galectin-3 can perturb the splicing reaction, we sought conditions to deplete galectin-3 from nuclear extracts. We now report that extracts depleted of galectin-3 do not support spliceosome formation or pre-mRNA splicing. The galectin-3 depleted extracts, however, regain splicing activity upon reconstitution with the recombinant lectin purified from an *E. coli* expression system.

MATERIALS AND METHODS

Preparation of Nuclear Extracts and Their Depletion and Reconstitution. NE were prepared from HeLa S3 suspension cultures in buffer D (20 mM Hepes (pH 7.9), 20% (v/v) glycerol, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM phenyl methylsulfonyl

fluoride (PMSF), and 0.5 mM (DTT)) as described (Dignam et al., 1983). Typically, NE at protein concentrations of approximately 18-22 mg/ml were prepared and stored as frozen aliquots at -80°C. Protein concentrations were determined by Bradford assay (Bradford, 1976, from BioRad).

NE were depleted of carbohydrate binding proteins on the basis of their binding to α-Lac insolubilized on 6% beaded agarose (LAC-A), purchased from Sigma. Agarose derivatized with cellobiose (CELLO-A) was used as a control for depletion. 200-250 μl of packed agarose were washed with 20 volumes of wash buffer (20 mM Hepes (pH 7.9), 0.5 M NaCl) in disposable columns (Pierce). NE were preincubated in buffer 1 (the composition of buffer 1 corresponds to 60% buffer D) containing, in addition, 2.5 mM $MgCl_2$, 1 mM ATP, 5 mM creatine phosphate in a vol of 50 μl for 30 min at 30°C. NaCl was then added to a final concentration of 0.5 M. Preincubated NE were then added to the washed saccharide-agarose and incubated for 30 min at 4°C with rotation. The unbound (UB) fraction was removed from the column. The matrices were next washed with 50 μl of buffer 1 containing 0.5 M NaCl and this wash was added to the UB fraction. The matrices were washed with 10 volumes of wash buffer and the bound (B) material eluted by boiling in 200 μl of Laemmli sample buffer (Laemmli, 1970). Aliquots (20 μl) of nondepleted NE and the UB fractions of the saccharide-adsorption procedure were dialyzed in a microdialyzer (Pierce) for 40 min against buffer 1.

In reconstitution experiments, recombinant CBP35 (rCBP35) (Agrwal et al., 1993) or other lectins (EY Laboratories) were added to the NE or UB fractions prior to dialysis. The dialyzed fractions were then assayed for splicing activity or splicing

complex formation.

Splicing Assay. The plasmid used to transcribe the MINX pre-mRNA substrate (Zillmann et al., 1988) was a kind gift of Dr. Susan Berget (Baylor College of Medicine). The MINX pre-mRNA was labeled with [^{32}P]GTP during Sp6 polymerase (Gibco) transcription and the monomethyl cap was added during transcription (Zillmann et al., 1988).

Splicing reactions (10 μl) contained NE (3 μl) or dialyzed UB or dialyzed non-depleted NE (8 μl) and 2.5 mM $MgCl_2$, 1.5 mM ATP, 20 mM creatine phosphate, 0.5 mM DTT and 20 units RNasin. The final protein concentration of the dialyzed extracts was 7-12 mg/ml. For reactions lacking ATP, both ATP and creatine phosphate were omitted. In experiments testing the effect of exogenously added carbohydrates, the NE were preincubated with saccharides at 30°C for 30 min prior to the addition of ^{32}P-labeled pre-mRNA.

Splicing reactions were carried out at 30°C for 45 min. Each splicing sample was diluted to 100 μl with 125 mM Tris (pH 6.8), 1 mM EDTA, 0.3 M sodium acetate. Proteinase K was added to a final concentration of 2 mg/ml and the sample digested for 1 hr at 37°C. RNA was extracted with 200 μl phenol-chloroform-isoamyl alcohol (50:50:1 (v/v)) followed by 200 μl chloroform. RNAs were precipitated with 300 μl ethanol in a dry ice ethanol bath. The extracted RNAs were subjected to electrophoresis in 13% polyacrylamide (bisacrylamide:acrylamide, 0.8:50 (w/w)) - 8.3 M urea gels, followed by autoradiography.

The intensities of the bands on gels were quantitated by direct ß-counting using an AMBIS system. The percent product formation (ligated exon 1-exon 2) was

calculated by dividing the radioactivity in the product by the total radioactivity in the pre-mRNA substrate, the splicing intermediates (exon 2-lariat, excised lariat, exon 1) and the product.

Gel Mobility Shift Assay for Splicing Complex Formation. The formation of splicing complexes was monitored by native gel electrophoresis (Konarska and Sharp, 1986). Splicing reactions (10 μl) were incubated at 30°C for 15-20 min. Heparin was added (final concentration 0.6 mg/ml) and the mixtures were incubated at 30°C for 15 min. Before electrophoresis, 1 μl of glycerol (containing 0.2% each of bromphenol blue, xylene cyanole and phenol red) was added to each sample. The samples were loaded onto a pre-run polyacrylamide gel (bisacrylamide:acrylamide, 1:80 (w/w)) and electrophoresis was carried out in 0.5 M Tris base - 0.5 M glycine (pH 8.8) at 4°C, 25 V/cm for 105 min. The migration of splicing complexes was determined by autoradiography.

Western Blot Analysis. Polyacrylamide gel electrophoresis in SDS was carried out in 12% acrylamide gels (Laemmli, 1970). The separated components were transferred to PVDF membrane (BioRad) in 25 mM Tris, 193 mM glycine and 20% methanol. Following transfer, the membranes were incubated with 5% nonfat dry milk dissolved in phosphate buffered saline containing 0.05% Tween 20 (PBS-T). Galectin-3 was detected using the rat monoclonal antibody, anti-Mac 2 (Ho and Springer, 1982). Incubation with the primary antibody (freshly diluted 1:5,000 in 5% nonfat dry milk - PBS-T) was carried out at room temperature for 2 hr, followed by five washes in PBS-T, 15 min each. The membranes were incubated with horseradish peroxidase conjugated goat anti-rat antibodies (Pierce) (diluted 1:3000 in PBS-T) for

30 min at room temperature followed by five 15 min washes in PBS-T. The products of the horseradish peroxidase reaction were revealed with the ECL reagents from Amersham, following the manufacturer's protocol.

RESULTS

Effect of Saccharides on pre-mRNA Splicing. The cell-free assay for the splicing of pre-mRNA by NE of HeLa cells was optimized for the MINX substrate. In a typical assay (final protein concentration of reaction ~10 mg/ml), approximately 20-30% product formation was observed (Fig. 1, lane 1). This conversion showed a stringent requirement for ATP (Fig. 1, lane 14). When the NE was preincubated with saccharides (75 mM) prior to the addition of the pre-mRNA substrate, inhibition of product formation was observed for Lac, thiodigalactoside (TDG), and melibiose (Fig. 1, lanes 2-4, respectively). Other mono- and disaccharides tested, as well as the non-carbohydrate inositol, failed to show inhibition of product formation (Fig. 1, lanes 5-13).

The effect of the inhibitory saccharides was concentration dependent; half-maximal inhibition were observed at <50 mM for TDG and Lac (Fig. 2, lanes 3-6 and lanes 7-10, respectively). Both of these saccharides bind to galectin-3 (Leffler and Barondes, 1986) with approximately the same hierarchy of affinities as those seen in the inhibition of splicing. The disaccharide cellobiose (Fig 2, lanes 15-18) and the monosaccharide Gal (Fig. 2, lanes 11-14) did not show an effect on the cell free splicing assay. Although Gal is a saccharide ligand for the galectin family of lectins, its binding affinity for galectin-3 is about two orders of magnitude lower than the

Figure 1

The effect of saccharides on pre-mRNA splicing. Splicing reactions were incubated in the absence of saccharide (lane 1) or in the presence of 75 mM various compounds (lanes 2-13) or in the absence of ATP (lane 14). The positions of migration of pre-mRNA, splicing intermediates and product are indicated on the right.

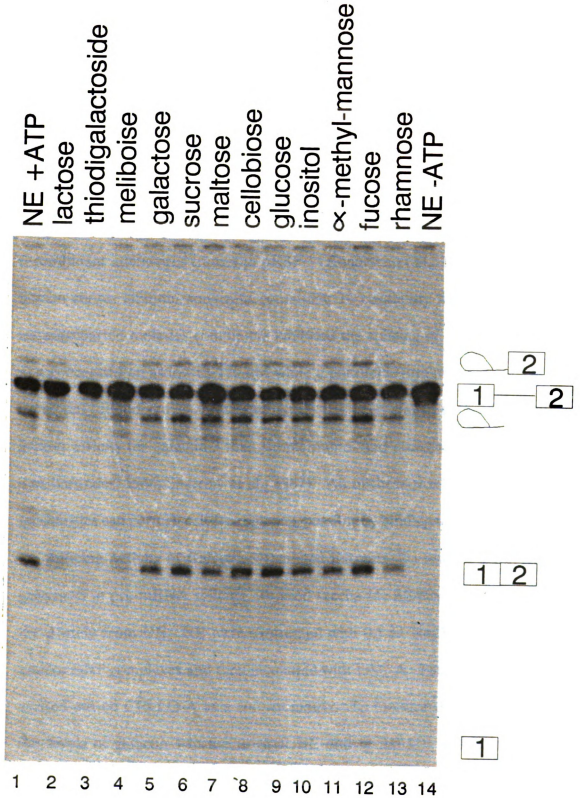

NE +ATP
lactose
thiodigalactoside
meliboise
galactose
sucrose
maltose
cellobiose
glucose
inositol
∝-methyl-mannose
fucose
rhamnose
NE -ATP
2
1
2
1 2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Figure 1

affinity of Lac (Leffler and Barondes, 1986). Thus, the inhibitory versus noninhibitory saccharides observed in the splicing assay correlate with the relative order of their binding affinities for galectin-3.

The inhibitory effects of Lac and TDG were neither extract nor pre-mRNA substrate restricted. NE prepared from HeLa cells over the course of two years were sensitive to inhibition by these saccharides. In addition, similar splicing inhibition was observed using human β-globin pre-mRNA (Krainer et al., 1984) and another recombinant adenovirus transcript pRSP-1 (Konarska et al., 1984). For example, a human serum albumin conjugate containing 14 covalently linked blood group A-tetrasaccharide moieties completely inhibited the splicing of the β-globin substrate at a concentration as low as 170 μM while a parallel serum albumin control yielded no effect. Blood group A-tetrasaccharide has been shown to exhibit significantly greater affinity for galectin-3 than simple mono- and disaccharides ligands (Leffler and Barondes, 1986; Sparrow et al., 1987). All of these results suggest that certain saccharides can perturb cell-free splicing, possibly by binding to galectin-3 in the NE.

Splicing Activity of Galectin-3-Depleted Extracts. To test for a role of nuclear galectin-3 in pre-mRNA splicing, we used saccharide affinity adsorption to deplete the protein from NE. NE were pretreated with 0.5 M NaCl to dissociate splicing and/or RNP complexes and then incubated with LAC-A. Parallel experiments were carried out on CELLO-A as a control matrix. To monitor the extent of depletion, the levels of galectin-3 in the original NE and in the UB and B fractions of the saccharide adsorptions were quantitated using the galectin-3 specific monoclonal antibody, anti-Mac-2 (Ho and Springer, 1982; Cherayil et al., 1989). A comparison

Figure 2

Concentration dependence of the effect of TDG and Lac on pre-mRNA splicing. Splicing reactions were incubated in the presence 50, 75, 125 and 150 mM Lac (lanes 3-6), TDG (lanes 7-10), Gal (lanes 11-14) and Cello (lanes 15-18) as described in Materials and Methods. Control splicing reactions were incubated in the absence (lane 1) or presence (lane 2) of ATP and creatine phosphate. The values shown at the bottom indicate percent product formation.

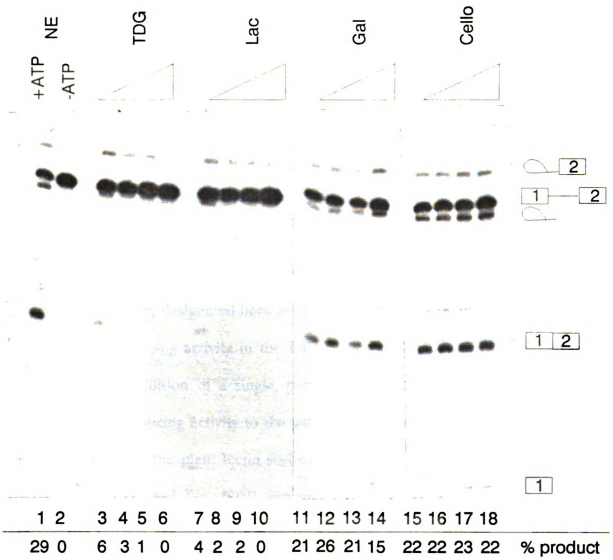

NE
TDG
Lac
Gal
Cello
+ATP
-ATP
2
1
2
1 2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
29 0 6 3 1 0 4 2 2 0 21 26 21 15 22 22 23 22 % product

Figure 2

of the UB and B fractions showed that LAC-A removed >95% of galectin-3 (Fig. 3A, lanes 1 and 2). Control CELLO-A removed <5% of galectin-3 (Fig. 3A, lanes 3 and 4). Protein determinations revealed that the depletion procedure also decreased the total protein content of the UB fractions by about 25%, compared to the non-treated NE. Thus, when the saccharide affinity-depleted extracts were tested in splicing assays, the non-treated extracts were diluted appropriately for comparison.

The UB fraction of the LAC-A adsorption, depleted of galectin-3, exhibited little or no splicing activity (Fig. 3B, lane 3), when compared to the corresponding fraction of the CELLO-A adsorption (Fig. 3B, lane 4) or to the original NE (Fig. 3B, lane 2). CELLO-A depleted extracts showed no change in splicing activity. Thus, the removal of galectin-3 from NE resulted in a complete loss of splicing activity.

Reconstitution of Splicing in Galectin-3-Depleted Extracts. We had previously described the production of a recombinant galectin-3 using an *E. coli* expression system for the cloned cDNA for CBP35 (Agrwal et al., 1993). The availability of this recombinant protein, designated here as rCBP35, provided the unique opportunity to test whether splicing activity in the UB fraction of LAC-A adsorption can be reconstituted by addition of a single, purified protein. Indeed, the addition of rCBP35 restored splicing activity to the galectin-3 depleted extract (Fig. 4, lanes 1 and 2). In contrast, the plant lectin soybean agglutinin, with a saccharide binding specificity (Sharon and Lis, 1989) similar to that of galectin-3, was unable to reconstitute splicing activity (Fig 4, lane 3). Similarly, wheat germ agglutinin, which binds to glycoconjugates containing sialic acids and/or N-acetyl-D-glucosamine (Sharon and Lis, 1989) also failed to restore splicing activity to the galectin-3

Figure 3A

Comparison of the levels of galectin-3 in NE and in the UB and B fractions when NE were subjected to adsorption on LAC-A and CELLO-A. NE were prepared for saccharide affinity chromatography as described in Materials and Methods. B and UB proteins were separated by SDS-PAGE then analyzed by Western blotting with anti-Mac-2 monoclonal antibody. Lanes 1 and 2, UB and B fractions, respectively, from LAC-A; lanes 3 and 4, UB and B fractions, respectively, from CELLO-A and; lane 5, non-depleted NE.

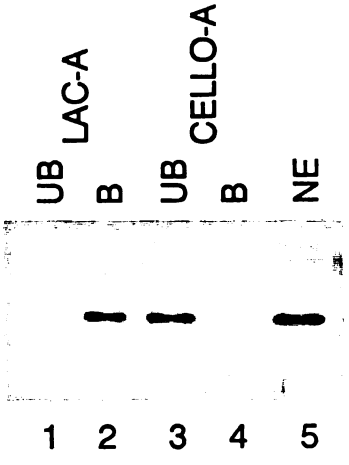


Figure 3A

Figure 3B

Comparison of splicing activity of NE and the UB fractions of LAC- and CELLO-affinity adsorptions. Lanes 1 and 2, nondepleted control splicing reactions in the absence and presence of ATP, respectively; lane 3, LAC-A UB fraction and lane 4, CELLO-A UB fraction.

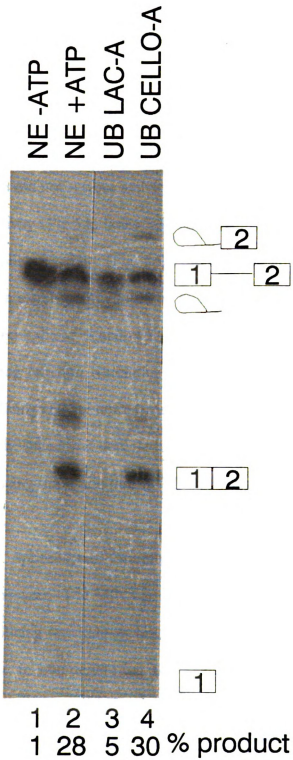


Figure 3B

depleted extract (data not shown). Finally, neither rCBP35 (Fig. 4, lane 4) nor soybean agglutinin (data not shown) had an effect on the splicing assay when added to a non-depleted NE.

The restoration of splicing activity by rCBP35 to the UB fraction of the LAC-A adsorption was dependent on the amount of recombinant protein added. Addition of 0.5 μg rCBP35 did not restore activity (Fig 5A, lane 2) whereas 1 to 4 μg rCBP35 reconstituted activity (Fig 5A, lanes 3-5). Surprisingly, addition of 8 μg of rCBP35 resulted in no splicing activity (Fig. 5A, lane 6). The amount of rCBP35 added in the reconstitution experiments should be put in context of the total protein content in a reconstituted splicing reaction (~60 μg) and of galectin-3 (0.5-1 ng) in an equivalent splicing reaction containing non-depleted NE. The maximal extent of reconstituted splicing activity represented >70% of the activity of non-depleted NE. More importantly, galectin-3 depleted extracts reconstituted with rCBP35 showed inhibition of splicing upon addition of TDG, as was observed with the original NE (Fig 5A, lane 7).

Analysis of Spliceosome Complex Formation in Original, Depleted, and Reconstituted Extracts. The MINX pre-mRNA forms several complexes during the course of the splicing reaction, as revealed by gel mobility shift assays of the ^{32}P-labeled substrate in non-denaturing gels (Konarska and Sharp, 1986). The first complex formed is an ATP-independent complex designated as the H complex. The addition of ATP results in the association of U1 and U2 snRNPs, forming the A and A' complexes, respectively. This, in turn, is followed by the association of U4, U5 and U6 snRNPs to yield the B complex (Zillmann et al., 1988).

Figure 4

The effect of rCBP35 and soybean agglutinin on the splicing activity of the UB fraction of LAC-A. Lanes 1-3, splicing in LAC-A UB fraction and lane 4, splicing in non-depleted extracts. Lane 1, no addition; lanes 2 and 4,addition of rCBP35 (2.0 μg); lane 3 addition of soybean agglutinin (2.0 μg).

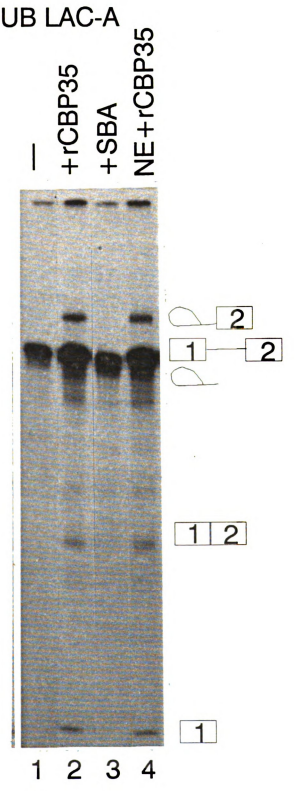


Figure 4

Figure 5A

Concentration dependence of the reconstitution of splicing activity in the UB fraction of LAC-A by rCBP35. NE were depleted of galectins by LAC-A (lanes 1-7) and the indicated amounts of rCBP35 added (lanes 2-6). Lane 1, no addition; lane 7, 2.0 μg rCBP35 and 50 mM TDG added; lane 8, non-depleted NE + ATP.

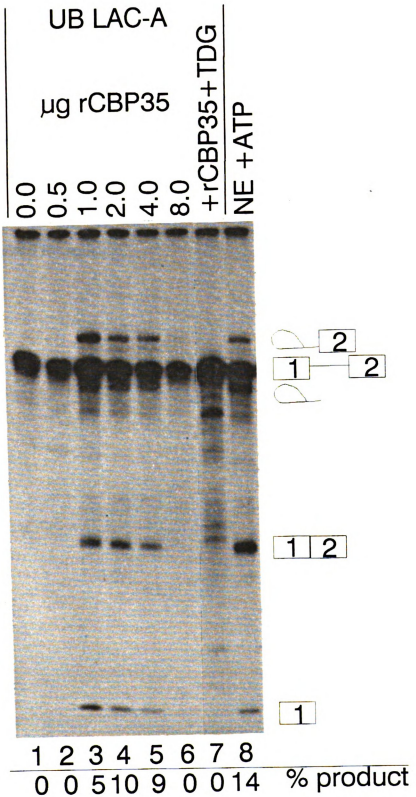

UB LAC-A
µg rCBP35
0.0
0.5
1.0
2.0
4.0
8.0
+rCBP35+TDG
NE +ATP
2
1
2
1 2
1
1 2 3 4 5 6 7 8
0 0 5 10 9 0 0 14
% product

Figure 5A

When NE were depleted of galectin-3 by LAC-A adsorption, the formation of virtually all of the ATP-dependent splicing complexes was inhibited (Fig 5B, lane 1 versus lane 7). NE subjected to adsorption on control CELLO-A retained the ability to form complexes (data not shown). As expected, galectin-3 depleted extracts reconstituted with rCBP35 regained the ability to form splicing complexes (Fig 5B, lanes 3-5). As was observed in assays of pre-mRNA splicing activity, the addition of 8 μg of rCBP35 did not reconstitute splicesome assembly (Fig. 5B, lane 6). All of these results suggest that galectin-3 may play a role in the initial assembly of complexes required for splicing.

Figure 5B

Formation of spliceosomal complexes by the UB fraction by the LAC-A adsorption in the absence and presence of rCBP35. Lanes 1-6, identical to the same lanes in Fig. 5A and lane 7, non-depleted NE + ATP.

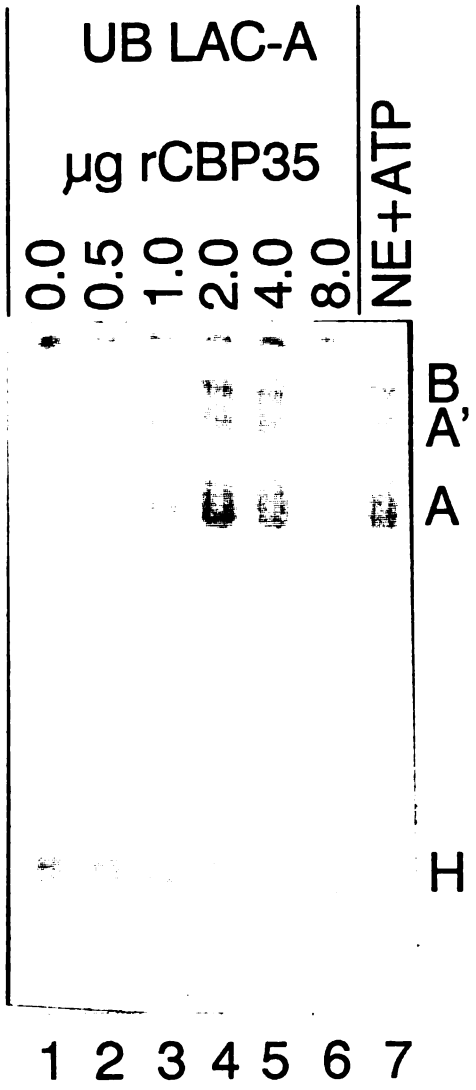


Figure 5B

DISCUSSION

The key findings documented in the present study include: (a) NE, capable of carrying out pre-mRNA splicing, contain galectin-3, a Gal/Lac-specific lectin found in both the cytoplasmic and nuclear fraction of cells; (b) saccharide ligands that bind to galectin-3 inhibit splicing; (c) NE depleted of galectin-3 by affinity adsorption on LAC-A fail to form spliceosomal complexes with pre-mRNA and do not carry out the *in vitro* splicing reaction; and (d) the activities of the galectin-3 depleted extract, in forming splicing complexes and in performing the pre-mRNA processing reaction, can be reconstituted by the addition of recombinant galectin-3, purified from an *E. coli* expression system. On the basis of these results, we believe we have identified a new splicing factor.

Crucial to our conclusion are the experiments involving the restoration of the activities of galectin-3-depleted extracts using rCBP35. Reconstitution of splicing by rCBP35 required significantly more galectin-3 than that found in an equivalent amount of non-depleted extract. This could be due to several factors. One, endogenous nuclear galectin-3 exists as both phosphorylated and unphosphorylated species (Cowles et al., 1990). The rCBP35 expressed in *E. coli* would not be subject to post-translational modifications such as phosphorylation. Perhaps the phosphorylated form is required for splicing activity and such phosphorylation may be rate-limiting or inefficient in HeLa extracts. Second, rCBP35, as produced in the expression vector for secretion out of *E. coli* (Agrwal et al., 1993), contains 3 extra amino acids (Ala-Glu-Phe) added onto the amino terminus of the mature protein. This extension may alter in a subtle manner the three dimensional configuration of

the protein and diminish its activity. Third, the amount and/or nature of the nuclear ligand for galectin-3 may be altered during the extensive dissociation, depletion and reconstitution procedures. Proteolysis or deglycosylation of the ligand are two possible events in this regard. Finally, rCBP35 may require more time to assemble into a splicing competent complex than allowed by the experimental conditions used.

The nuclear component which interacts with galectin-3 has not yet been identified. Three classes of candidates are formally possible. First, galectin-3 may interact via its CRD with a Lac-containing nuclear glycoprotein. Nuclear proteins known to be glycosylated include nucleoporins, transcription factors AP-1, Sp-1, CTF/NF-1, AP-2, and AP-4 (Jackson and Tjian, 1988; Lichtsteiner and Schibler, 1989) and RNA polymerase II (Kurl et al., 1988). It has been suggested, although no rigorously proven, that HMGs 1, 2, 14 and 17 contain N-linked complex oligosaccharides (Reeves and Chang, 1983; Einck and Bustin, 1985) with Gal as one of the constitutents (Reeves et al., 1981). With the possible exception of the HMGs, these nuclear glycoproteins only contain O-linked N-acetyl-D-glucosamine, a carbohydrate that has no detectable affinity for galectin-3. It should be noted that none of these nuclear components have been shown to have a role in splicing, nor have any of the known splicing factors been characterized as glycoproteins. However, an N-acetyl-D-glucosamine bearing glycoprotein (Mr 43,000) has been shown to co-precipitate with hnRNP C proteins and is thought to be a component of the hnRNP complex (Soulard et al., 1991). There is increasing evidence that hnRNP proteins A1, C1/C2 and I/PTP contribute to efficiency and/or regulation of splicing (reviewed in Lamm and Lamond, 1993). It would be interesting to determine whether the 43

kDa glycoprotein can be added to the growing list of hnRNP proteins involved in RNA processing. Second, the CRD domain of galectin-3 may interact with a nonglycosylated nuclear polypeptide containing a structure mimicking Lac. Recently, peptides have been isolated from random hexapeptide libraries which bind to the lectin Concanavalin A with dissociation constants equivalent to mannose, the saccharide ligand for the lectin (Oldenberg et al., 1992; Scott et al., 1992). Alternatively, the amino terminal domain (i.e. the non CRD-containing domain) of galectin-3 could interact with a nonglycosylated nuclear protein. This Pro-, Gly-rich domain has limited homology to some of the hnRNP proteins. Recent studies suggest that galectin-3 is associated with a glucose-binding protein via protein-protein rather than protein-carbohydrate interactions. This association can be disrupted with Lac but not by Gal or glucose (Seve et al., 1993). It should be noted, however, that in our experiments, the control affinity matrix composed of cellobiose (glucose, β1-4glucose)-agarose did not remove galectin-3 from NE and the UB fraction of the CELLO-A adsorption was splicing competent. Third, the pre-mRNA substrate or one of the snRNAs could act as the target for galectin-3 interaction. This interaction could occur through the amino-terminal domain or the saccharide binding region in the carboxyl half of the galectin-3 molecule. In preliminary experiments we have been unsuccessful in showing direct contact with the pre-mRNA by UV-induced crosslinking. Clearly it is critical to identify the nuclear ligand for galectin-3.

Where and how galectin-3 participates in the splicing pathway has not been established. When galectin-3 specific saccharides are added to complete NE, it appears that exon ligation (product formation) is more sensitive to inhibition then

the first cleavage reaction. For example, at 50 mM Lac or TDG (Fig. 2, lanes 3 and 7), product formation is inhibited approximately 80% whereas exon 2-lariat and exon 1 intermediates are only reduced by 5-10% when compared to a non-treated extract. This disparity between inhibition of exon ligation versus the first cleavage event is maintained at the higher concentrations of Lac or TDG tested. Perhaps the interaction of galectin-3 with a splicing factor is sufficiently avid or occurs via the CRD of the galectin so that free ligand cannot act as an effective competitor. As other splicing components enter the active complex, the CRD may become accessible to Lac. The binding of Lac may (i) prevent galectin-3 from binding to its partner and stop further splicing events or (ii) alter the conformation of galectin-3 so that subsequent splicing steps are blocked. On the basis of differential scanning calorimetry studies, a conformational change in galectin-3 has been demonstrated upon Lac binding (Agrwal et al., 1993).

Data from depletion experiments can be interpreted to support this model. In this case, removal of galectin-3 results in a dramatic reduction in spliceosomal complexes without a discernible effect on formation of the H complex. This suggests that galectin-3 is required for assembly of the initial A complex. The simultaneous addition of rCBP35 and free saccharide ligand to a galectin-3-depleted extract would result in complete inhibition of splicing because the CRD of rCBP35 immediately binds saccharide. As stated previously, when rCBP35 binds Lac, required interactions of rCBP35 with a splicing component may be directly or indirectly affected depending on the nature of the splicing component.

A final possibility for the mode of action of galectin-3 is binding to a general

inhibitor of splicing. Upon depletion of galectin-3 or its inactivation (e.g. via carbohydrate binding), the general inhibitor would suppress the processing of the pre-mRNA substrate. This notion is formally distinguished from the other possibilities because it would not require galectin-3 to be present in the splicing complexes. Experiments designed to determine whether galectin-3 is a component of one of the spliceosomal complexes are currently in progress.

Finally, it should be noted that if other members of the galectin family are found in NE, they will also be removed by LAC-A adsorption. In preliminary experiments, we have identified, by Western analysis, galectin-1 (previously designated L-14) in the B fraction of LAC-A (S. D., unpublished observations). However, rCBP35 alone can reconstitute splicing activity to LAC-A adsorbed extracts. Thus, it is possible that galectin-1 or other nuclear galectins are not splicing factors, perhaps because they contain only the CRD but not the Pro-, Gly-rich domain found in galectin-3. Alternatively, all nuclear galectins may be splicing factors, but their mode of action in splicing can be assumed by any one galectin, hinting to a functional redundancy. Experiments are in progress to test these two possibilities.

REFERENCES

Agrwal, N., Sun, Q., Wang, S.-Y. and Wang, J. L. (1992) *J. Biol. Chem.* 268, 14932-14939.

Albrandt, K., Orida, N. K. and Liu, F.-T. (1987) *Proc. Natl. Acad. Sci. USA* 84, 6859-6863.

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

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

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

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

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

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

Einck, L. and Bustin, M. (1985) *Exp. Cell Res.* 156, 295-310.

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

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

Jackson, S. P. and Tjian, R. (1988) *Cell* 55, 125-133.

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

Konarska, M. M., Padgett, R. A., and Sharp, P. A. (1984) *Cell* 38, 731-736.

Konarska, M. M. and Sharp P.A. (1986) *Cell* 46, 845-855.

Krainer, A. R., Maniatis, T., Ruskin, B. and Green, M. R. (1984) *Cell* 36, 993-1005.

Kurl, R. N., Holmes, S. C., Verney, E. and Sidransky, H. (1988) *J. Biol. Chem.* 27, 8947-8980.

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

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

Lamm, M. G. and Lamond, A. I. (1993) *Biochimica et Biophysica Acta* 1173, 247-265.

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

Leffler, H., Masiarz, F. R. and Barondes, S. H. (1989) *Biochemistry* 28, 9222-9229.

Lichtsteiner, S. and Schibler, U. (1989) *Cell* 57, 1179-1187.

Oldenberg, K. R., Loganathan, D., Goldstein, I. J., Schultz, P. G. and Gallop M. A. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5393-5397.

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

Reeves, R., Chang, D. and Chung, S.-C. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6704-6708.

Reeves, R. and Chang, D. (1983) *J. Biol. Chem.* 258, 679-687.

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

Scott, J. K., Loganathan, D., Easley, R. B., Gong, X. and Goldstein, I. J. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5389-5402.

Sharon, N. and Lis, H. eds (1989) *Lectins* (Chapman and Hall, London).

Soulard, M., Barque, J.-P., Valle, V. D., Hernandez-Verdun, D., Masson, C., Danon, F. and Larsen, C.-J. (1991) *Exp. Cell Res.* 193, 59-71.

Sparrow, C. P., Leffler, H. and Barondes, S. H. (1987) *J. Biol. Chem.* 262, 7383-7390.

Wang, J. L., Werner, E. A., Laing, J. G., Patterson, R. J. (1992) *Trans. Biochem. Soc.* 20, 269-274.

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

CHAPTER III

Nuclear Galectins: A Family of Functionally Redundant Splicing Factors

INTRODUCTION

Recently, we identified CBP35(galectin-3), a galactose (Gal)/lactose(Lac) binding protein found in the nucleus of many mammalian cells, as one of the many proteins required for *in vitro* pre-mRNA splicing (Chapter II). This identification was made possible by our observation that addition of saccharide ligands that have high affinity for galectin-3 to splicing extracts resulted in diminished splicing activity. Depletion of galectin-3 from nuclear extracts by saccharide affinity chromatography yielded extracts deficient both in splicing activity and assembly of spliceosomes. These activities could be restored to galectin-3-depleted extracts by recombinant galectin-3 (rCBP35). The reconstituted splicing activity was inhibitable by addition of saccharides which bind avidly to galectin-3.

CBP35 is one member of a family of lectins termed galectins which bind Gal/Lac or Lac containing glycoconjugates (Barondes et al., 1994). The galectins are further subdivided based generally on their number of domains and sequence homology (Barondes et al., 1994). Galectins 1 and 2 are single domain galectins of 14 - 18 kDa; the single domain contains the carbohydrate recognition domain (CRD) (for reviews, see Anderson and Wang, 1992 and Hirabayasi and Kasai, 1993). The galectin-3 types include CBP35 and other similar molecules, have molecular weights between 26-35 kDa and are composed of two distinct domains (Anderson and Wang, 1992; Jia and Wang, 1988). The carboxyl half includes the CRD and the amino-terminal half is a Pro,Gly-rich repeat sequence with limited homology to hnRNP proteins (Anderson and Wang, 1992; Jia and Wang, 1988). The function of this

amino domain is unknown. The galectin-4 subgroup contains one member, a single polypeptide of 36 kDa with two CRDs separated by a linker region (Oda et al., 1993). The CRDs are highly homologous for all the galectins and share similar (yet distinct) carbohydrate binding specificities (Leffler and Barondes, 1986).

The subcellular distribution of the galectins presents interesting dilemmas. These proteins have been identified as cell surface markers (most notably the Mac-2 differentiation antigen of macrophages which is identical to CBP35), soluble proteins released from the cell, cytoplasmic proteins, and unexpectedly, nuclear components (reviewed in Wang et al., 1991). As yet, no signal peptide has been found nor is a consensus nuclear localization signal apparent. The nuclear localization for some of the galectins is either controversial or poorly defined, although the data for nuclear localization for a portion of galectin-3 are firmly established (Laing and Wang, 1988; Wang et al., 1992).

The nuclear localization, coupled with the requirement for galectin-3 for pre-mRNA splicing, enticed us to assess the presence of other galectins in splicing extracts. In this paper we show that one identified galectin (L-14, galectin-1) (Gitt et al., 1992) and an unidentified one are found in HeLa splicing extracts. Depletion of individual galectins by antibody-affinity selection has no effect on splicing and the singly depleted extracts show inhibition of splicing upon addition of specific saccharide ligands. Immunodepletion of both galectins-1 and -3 results in an extract still active in splicing. Moreover, the splicing activity of galectin-1 and galectin-3 depleted extracts is inhibitable by the addition of saccharides with a high affinity for the galectins. Thus, at least one other nuclear galectin exists in HeLa nuclear

extracts. Finally, using reconstitution experiments, we show that the splicing activity involving galectins can be recovered with either galectin-1 or galectin-3, hinting to functional redundancy.

MATERIALS AND METHODS

Refer to Chapter II MATERIALS AND METHODS for preparation of NE and their depletion with LAC-A and reconstitution, splicing assays and Western blot analysis.

Immunodepletion and Reconstitution. Antibodies anti-Mac2 (αM2) (10 μg) and anti-L14 (αL14) (10 μg) were mixed with 50 mg pre-swollen protein G-Sepharose beads (Pharmacia). The mixture was adjusted to 20 mM HEPES, pH 7.6 and 0.5 M NaCl and the antibody covalently cross-linked to protein G with dimethylpimelimidate. The coupled antibody Sepharose mixture was washed 4 times with 1 ml wash buffer (20 mM HEPES (pH 7.9), 0.5 M NaCl). The washed antibody coupled Sepharose was incubated with 50 μl NE adjusted to 0.5 M NaCl, for 30 min at 4°C with gentle rocking. The unbound (UB) depleted fraction was removed using a 25 μl Hamilton syringe. The Sepharose was next washed with 50 μl of buffer 1 (60% Dignam buffer D adjusted to 0.5 M NaCl) and this wash was added to the UB fraction. Fifty μl non-depleted NE were similarly diluted with 50 μl of buffer 1. The Sepharose was washed four times with 1 ml buffer 1 and the bound (B) material eluted by boiling in 200 μl of Laemmli sample buffer (Laemmli, 1970). Aliquots (20 μl) of nondepleted NE and the UB fractions of the immunodepletion were dialyzed in a microdialyzer (Pierce) in the presence and absence of rCBP35 or rL14 for 40

min against 60% Dignam buffer D. The dialyzed fractions were then assayed for splicing activity.

RESULTS

Identification of galectins in splicing extracts. Initially we surveyed which galectins are found in HeLa Dignam splicing extracts. Monospecific antisera for galectins-1 and -3 were used to probe Western blots for galectin identification. To detect galectin-1, affinity purified rabbit polyclonal antiserum (kindly provided by Sam Barondes) was used, and to detect galectin-3, rat monoclonal antibody to the Mac-2 antigen (identical to CBP35, galectin-3) was used. These two galectins are detected in a Dignam splicing extract (Fig. 1). In addition, both of these galectins are identified in the bound fraction of a NE adsorbed to LAC-A (lanes 1 and 2). Also, these data show that each antiserum is monospecific, which is important in subsequent experiments when the antisera are used to singly immunodeplete NE of their respective galectin.

Additional confirmation that the galectins are nuclear components is presented in Fig 2. Immunofluorescent microscopy was used to determine the intracellular localization of the two galectins. For these studies, 3T3 fibroblasts monolayers were incubated with each antisera individually. Appropriate FITC-conjugated secondary antisera were then added for detection. A punctate staining pattern is observed in nuclei for galectin-3, a pattern reminiscent of that observed for many of the known splicing factors and snRNPs (Fu and Maniatis, 1991; Lerner et al., 1981; Deng et al., 1981, Spector et al., 1983; Reuter et al., 1984; Nyman et al., 1986; Spector et al.,

Figure 1

Detection of galectin-1 and galectin-3 in LAC-A bound NE. NE was adsorbed to LAC-A as described in MATERIALS AND METHODS, Chapter II. B proteins were separated by SDS-PAGE then analyzed by Western blotting with αM2 (lane 1) or αL14 (lane 2). The positions of galectin-1 and galectin-3 are indicated on the right.

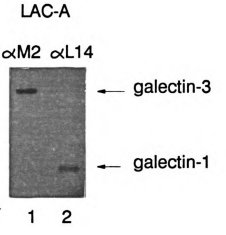


Figure 1

Figure 2

Localization of galectin-1 and galectin-3 by immunofluorescence staining of NIH 3T3 fibroblasts. 3T3 fibroblasts were fixed and permeabilized with Tx-100. The intracellular distribution of galectin-3 (top) and galectin-1 (bottom) was determined by indirect immunofluorescence microscopy. As a control, the middle panel shows a similar analysis of the Sm proteins.

The photographs were kindly provided by Sung-Yuan Wang.

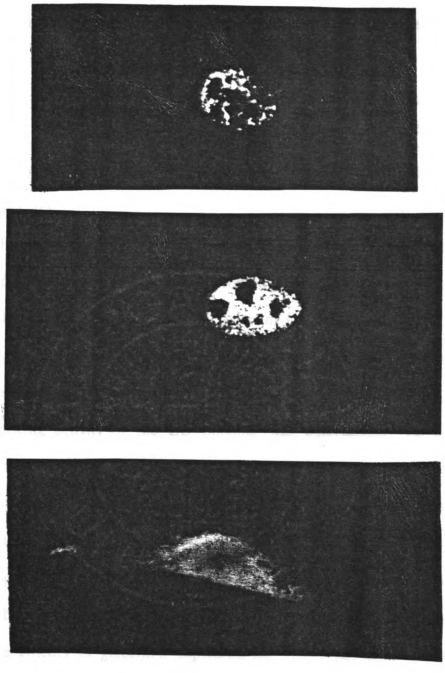


Figure 2

1991). Cells incubated with antiserum to galectin-1 appear to stain the nucleus and diffuse cytoplasmic fluorescence is observed for both galectins.

By these criteria, galectins-1 and -3 appear to be localized to the nucleus of fibroblasts and HeLa cells. While the quantity of each galectin within NE may not reflect the absolute intranuclear distribution, it is clear that both are bona fide nuclear molecules.

Galectin-immunodepletion from splicing extracts. When splicing extracts are depleted of galectins by LAC-A affinity chromatography, the splicing activity of the extract is abolished (Chapter II) and both galectins-1 and -3 are identified bound to the LAC-A matrix (Fig 1). To determine which of these galectins is responsible for the loss in splicing activity, each galectin was individually immunodepleted from NE and the splicing activity of the UB fractions was determined. As a positive immunodepletion control, the snRNPs were removed from NE by polyclonal αSm antiserum. Monoclonal antibody to the transferrin receptor (α-TR) was used as a negative control. Extracts depleted of galectin-3 (UB of αM2, Fig. 3A, lane 4) retain nearly complete splicing activity compared to αTR-depleted extracts (lane 1). As expected, the αSm-depleted extract is deficient in splicing activity (lanes 7). To monitor the efficacy of galectin-3 depletion, the bound (B) and unbound (UB) fractions from the antibody-protein G matrices were used for Western analysis (Fig 3B). Anti-M2 removes greater than 95% of galectin-3 (lane 3). Galectin-3 is not removed by either αTR (lane 1) or αSm (lane 5) antisera. Galectin-1 is not depleted by any of these antisera (data not shown).

Fig. 4A shows the splicing activity of NE depleted of galectin-3 (lane 1) or

Figure 3A

The effect of galectin-3 depletion on pre-mRNA splicing. NE were immunodepleted with αTR, αM2 and αSm immobilized on protein G sepharose as described in MATERIALS AND METHODS. Splicing reactions with αTR, (lanes 1-3) and αM2, (lane 4-6) and αSm (lanes 7-9) immunodepleted extracts were incubated in the absence of saccharide (lanes 1, 4 and 7), presence of 50 mM sucrose (lanes 2, 5, and 8) or in the presence of 50 mM TDG (lanes 3, 6 and 9). The positions of migration of pre-mRNA, splicing intermediates and product are indicated on the right.

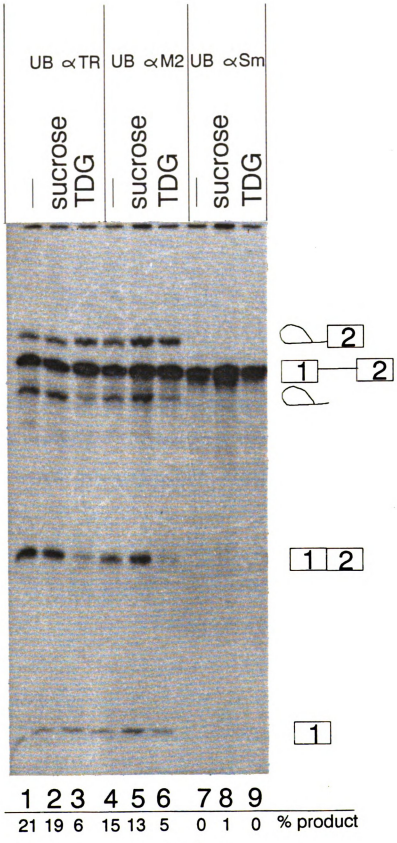

UB αTR
UB αM2
UB αSm
—
sucrose
TDG
—
sucrose
TDG
—
sucrose
TDG
2
1
2
1 2
1
1 2 3 4 5 6 7 8 9
21 19 6 15 13 5 0 1 0 % product

Figure 3A

Figure 3B

Comparison of the levels of galectin-3 in NE and the UB and B fractions of NE after immunodepletion with αTR, αM2 and αSm. UB fractions, lanes 1, 3 and 5, and B fractions, lanes 2, 4 and 6, from NE immunodepleted with αTR (lanes 1 and 2), αM2 (lane 3 and 4) and αSm (lanes 5 and 6). Lane 7, non-depleted NE.

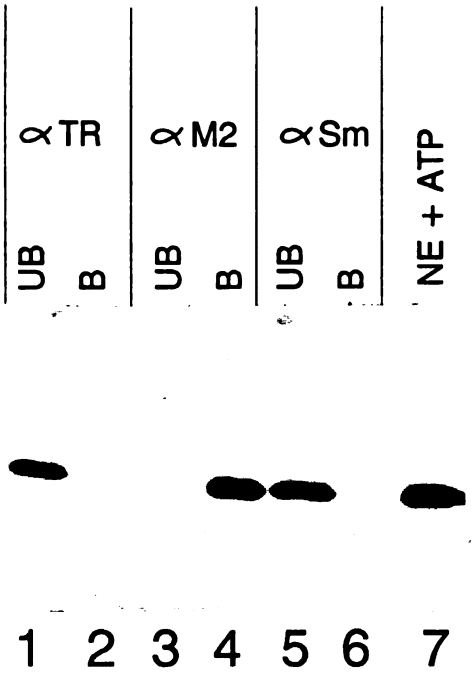


Figure 3B

galectin-1 (lane 5). Neither galectin-3 nor galectin-1 depleted extracts show diminution in splicing activity compared to a non-depleted extract, (lanes 13 and 14), despite the fact that the UB fractions had each galectin quantitatively removed by their respective antiserum (Fig. 4B).

These data are the first indication that the functional role for galectins in pre-mRNA splicing is redundant. Depletion of galectins-1 and -3 by LAC-A adsorption yields a splicing-deficient extract whereas each galectin removed individually resulted in splicing competent extracts. This hypothesis of galectin functional redundancy is strengthened by the fact that extracts singly depleted of galectins are still sensitive to splicing inhibition by the addition of a galectin-specific saccharide (TDG) (Fig. 3A, lane 6 and Fig 4A, lanes 4 and 8).

Double galectin immunodepletion. Nuclear extracts were next immunodepleted of galectins 1 and 3 by adsorption to an affinity matrix containing both galectin antisera and the doubly depleted fraction assayed for splicing activity (Fig. 4A, lane 9). Double galectin depletion had no effect on splicing activity, even though both galectins are quantitatively removed (Fig. 4B, lane 5). These results suggest that yet another lactose-binding protein exists in splicing extracts. To address this possibility, the doubly depleted extract was tested for its sensitivity to galectin-specific saccharide inhibition of splicing. The addition of TDG did, indeed, inhibit splicing activity in the galectin-1 plus galectin-3 depleted extract (Fig. 4A, lane 12).

Galectin reconstitution of splicing activity. If nuclear galectins are functionally interchangeable, a logical prediction is that any galectin individually should be active in restoring activity to an extract depleted of all galectins. To test this prediction,

Figure 4A

Comparison of splicing activity of NE immunodepleted of galectin-1 and/or galectin-3. Galectin-3 depleted (lanes 1-4), galectin-1 depleted (lanes 5-8), galectin-1 + galectin-3 depleted (lanes 9-12) and non-depleted (lanes 13-15) NE were incubated in the presence of recombinant galectin or TDG. Lanes 1, 5 and 9, no addition; lanes 2, 6, 10 and 13 addition of rCBP35 (4 μg); lanes 3, 7, 11 and 14, addition of rL14 (8 μg); lanes 4, 8, 12 and 15, addition of TDG (50mM).

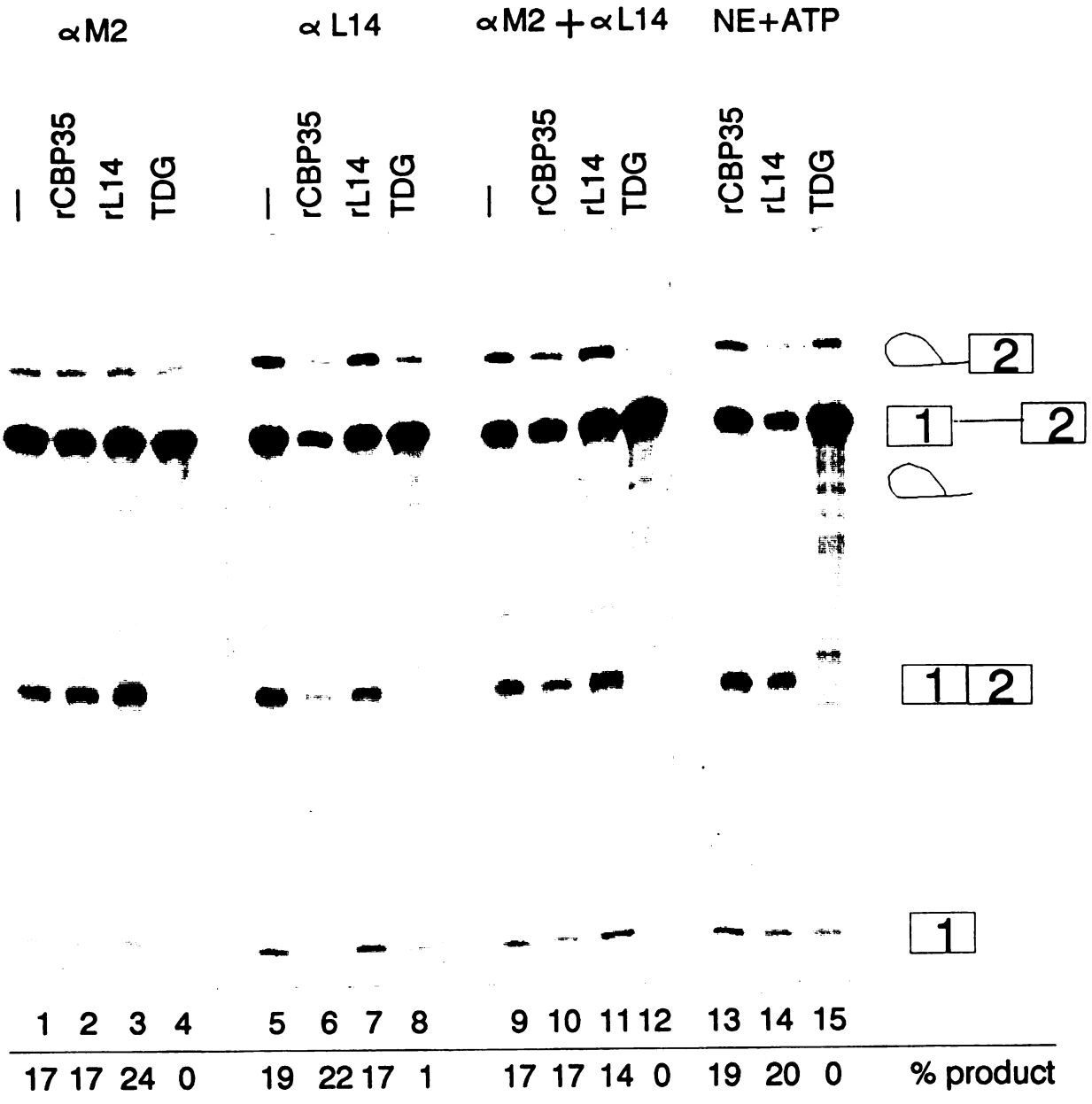

αM2
α L14
αM2 + αL14
NE+ATP
—
rCBP35
rL14
TDG
2
1
2
1 2
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
17 17 24 0 19 22 17 1 17 17 14 0 19 20 0
% product

Figure 4A

Figure 4B

Comparison of the levels of galectins in NE and in the UB and B fractions subjected to immunodepletion by αM2 and αL14 separately and together. Proteins were separated by SDS-PAGE then analyzed by Western blotting with αM2 (1) and αL14 (2). Lanes 1 and 2, UB and B fractions, respectively, from αM2 depleted extract; lanes 3 and 4, UB and B fractions, respectively, from αL14 depleted extract; lanes 5 and 6, UB and B fractions, respectively, from αM2 and αL14 depleted NE; lane 7, non-depleted NE.

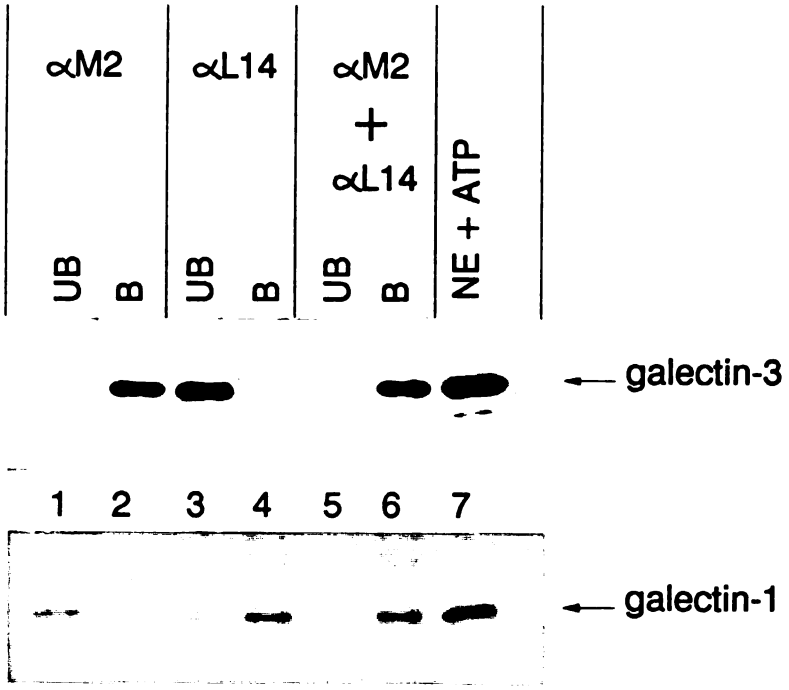

αM2
αL14
αM2
+
αL14
NE + ATP
UB
B
UB
B
UB
B
galectin-3
1 2 3 4 5 6 7
galectin-1

Figure 4B

recombinant forms of galectin-1 (rL-14, kindly provided by Sam Barondes) and galectin-3 (rCBP35, kindly provided by John Wang) were evaluated for their ability to reconstitute splicing activity in NE depleted by LAC-A adsorption (Fig. 5). As expected, the UB fraction from LAC-A adsorption is splicing deficient (lane 1). The addition of either rCBP35 (lane 2) or rL-14 (lane 3) reconstitutes splicing activity, whereas their addition to singly depleted extracts (Fig. 4A, lanes 2, 3, 6 and 7), a doubly depleted extract (Fig. 4A, lanes 10 and 11) or a non-depleted extract (Fig. 4A, lanes 13 and 14) has no observable effect on splicing activity of these splicing competent fractions.

Figure 5

The effect of rCBP35 and rL14 on the splicing activity of the UB fraction of LAC-A. Lanes 1-3, splicing in LAC-A UB fraction and lane 4, splicing in non-depleted extracts. Lane 1, no addition; lane 2, addition of rCBP35 (4 μg); lane 3, addition of rL14 (8 μg).

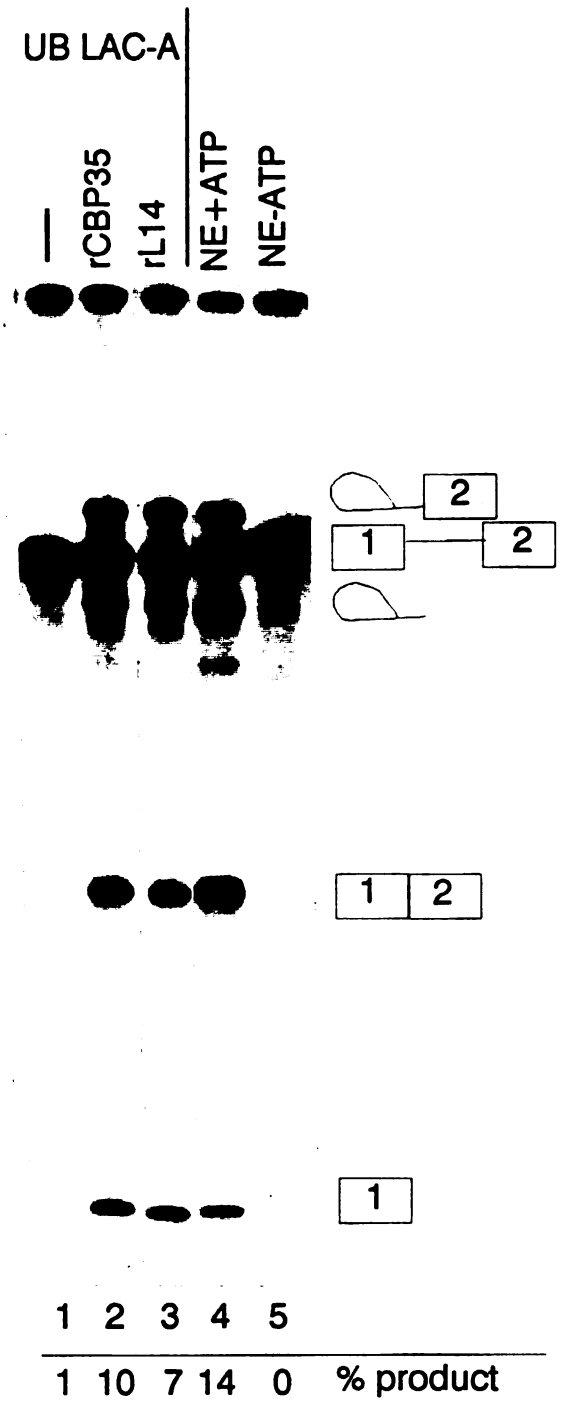

UB LAC-A
—
rCBP35
rL14
NE+ATP
NE-ATP
2
1
2
1
2
1
1 2 3 4 5
1 10 7 14 0 % product

Figure 5

DISCUSSION

Our identification of a new family of splicing factors (nuclear galectins) adds to the growing list of components required for this series of complex reactions (Bennett et al., 1992 and reviewed in Lamm and Lamond, 1993; Lamond, 1993). Many known splicing components are members of protein families whose splicing function is redundant within the group (Lamm and Lamond, 1993). The galectins, because of their novel saccharide binding properties, present possible mechanisms of splicing regulation heretofore unappreciated. Further, this interchangeability (at least as presently defined) suggests a general and critical role in the splicing pathway.

The galectins apparently share no homology with any of the previously identified splicing components. A limited homology between the amino terminal half of CBP35 and the hnRNP A1 protein exists (Wang et al., 1991), but it is unlikely that this domain is active in splicing as the other galectin capable of reconstitution is lacking this domain . The only homology apparent between the galectins is the CRD region and this sequence is not homologous to the RRM, DEAD, DEAH or SR consensus motifs of some of the other known splicing factors (Lamm and Lamond, 1993). We have also been unsuccessful in demonstrating direct galectin-pre-mRNA interactions using UV cross-linking procedures (chapter IV).

We thus speculate that the galectins interact with other proteins, possibly early in the assembly of spliceosomes. We infer this based on the fact that splicing complexes (as assessed by native gel electrophoresis) are greatly diminished in galectin-depleted extracts (chapter II). It is not clear, however, how the galectins are involved in complex assembly.

Several questions are still unanswered. First, what is the identity of the other nuclear galectin(s)? We suspect that galectin-4 is the most likely candidate. Antiserum is available for this galectin and it has been requested from Dr. S. Barondes (Oda et al., 1993). Second, what is the nuclear partner recognized by galectins? No known nuclear glycoproteins seem likely candidates as most (with the possible exception of HMG-1, 2, 14 and 17) have N-acetyl glucosamine as their sole O-linked carbohydrate (Hart et al., 1988, 1989; Davis and Blobel, 1986; Hanover et al, 1987; Holt et al., 1987). It has been suggested, although not rigorously proven, that the HMGs are modified by N-linked complex oligosaccharides with Gal as one of the identified carbohydrates (Reeves and Chang, 1983).

If nuclear glycoproteins are not the partner for nuclear galectins, are any of the known splicing components likely candidates? To answer this question, we must first determine whether galectins are components of spliceosomal complexes. Third, some of the galectins exist as two molecular species; phosphorylated and non-phosphorylated. For galectin-3, both forms are nuclear, where the cytoplasmic species is exclusively phosphorylated (Cowles et al., 1990). What role does phosphorylation play in the splicing capacity of galectin-3? Finally, is it possible that individual galectins function in a pre-mRNA specific manner? The intracellular concentration and distribution of galectin-3 is dependent on cellular proliferation rate and the state of cellular differentiation (Moutsatsos et al., 1987; Agrwal et al., 1989). Do the other galectins exhibit altered intracellular localization dependent on cellular physiology? Experiment are in progress to answer these questions.

REFERENCES

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

Agrwal, N., Wang, J. L. and Voss, P. G. (1989) *J. Biol. Chem.* 264, 17236-17242.

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

Bennett, M., Michaud, S., Kingston, J. and Reed, R. (1992) *Genes and Dev.* 6, 1986-2000.

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

Davis, L. I., and Blobel, G. (1986) *Cell* 45, 699-709.

Deng, J. S., Takasaki, Y. and Tan, E. M.(1981) *J. Cell Biol.* 91, 572-577.

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

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

Hanover, J. A., Cohen, C. K., Willingham, M. C. and Park, M. K. (1987) *J. Biol. Chem.* 262, 9887-9894.

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

Hart, G. W., Holt, G. D. and Haltiwanger, R. S. (1988) *Trends Biochem. Sci.* 13, 380-384.

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

Holt, G. D., Snow, C. M., Senior, A., Haltiwanger, R. S., Gerace, L. and Hart, G. W. (1987) *J. Cell Biol.* 104, 1157-1164.

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

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

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

Lamm, M. G. and Lamond, A. I. (1993) *Biochimica et Biophysica Acta* 1173, 247-265.

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

Leffler, H., Masiarz, F. R. and Barondes, S. H. (1989) *Biochemistry* 28, 9222-9229.

Lerner, E. A., Lerner, M. R., Janway, C. A. and Steitz, J. A. (1981) *Proc. Natl. Acad. Sci. USA.* 78, 2737-2741.

Moutsatsos, I. K., Wade, M., Schindler, M. and Wang, J. L. (1987) *Proc. Natl. Acad. Sci. USA* 84, 6452-6456.

Nyman, U., Hallman, H., Hadlaczky, I., Sharp, G. and Ringertz, N. R. *J. Cell Biol.* 102, 137-144.

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

Reeves, R., Chang, D. and Chung, S.-C. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6704-6708.

Reeves, R. and Chang, D. (1983) *J. Biol. Chem.* 258, 679-687.

Reuter, R., Appel, B., Bringmann, P., Rinke, J. and Luhrmann, R. (1984) *Expl. Cell Res.* 154, 548-560.

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

Spector, D. L., Schrier, W. H. and Busch, H.(1983) *Biol. Cell* 49, 1-10.

Wang, L. W., Laing, J. G. and Anderson, R. L. (1991) *Glycobiology* 1, 243-252.

Wang, J. L., Werner, E. A., Laing, J. G., Patterson, R. J. (1992) *Trans. Biochem. Soc.* 20, 269-274.

CHAPTER IV

PHOSPHORYLATION AND UV CROSS-LINKING

Phosphorylation

INTRODUCTION

3T3 cells contain two species of galectin-3; a phosphorylated form with a pI of 8.2 and non-phosphorylated form with a pI of 8.7 (Cowles et al., 1990). The phosphorylation state, as well as the expression and subcellular distribution of galectin-3 closely correlate with the proliferative state of 3T3 cells in culture (Moutsatsos et al., 1987, Cowles et al., 1990). Both phosphorylated and non-phosphorylated forms are found in the nucleus whereas the cytoplasmic form is almost exclusively phosphorylated. When 3T3 cultures are serum starved to achieve quiescence, galectin-3 is found at low levels, is localized mainly in the cytoplasm and only the phosphorylated form can be detected in the nucleus. Serum stimulation of quiescent cells results in increased amounts of galectin-3. In these cultures, the nucleus now contains both phosphorylated and non-phosphorylated forms and the cytoplasm has increased levels of the phosphorylated form. Proliferating cultures have increased levels of galectin-3. Both phosphorylated and non-phosphorylated forms are found in the nucleus with as much as a 5 fold increase in the non-phosphorylated form (Cowles et al., 1990; Hamann et al., 1991).

Purified galectin-3 isolated after labeling *in vivo* with $^{32}PO_4$ is sensitive to alkaline phosphatase treatment indicating phosphorylation occurred on a Ser, Thr or Tyr residue (Cowles et al., 1990). Recently, a recombinant galectin-3 (rhL-29) has been shown to be phosphorylated at serine^{6} and serine^{12} with 90% of the phosphate at serine^{6} and 10% at serine^{12} (Huflejt et al., 1993). In addition, *in vitro* phosphorylation of galectin-3 can be catalyzed by casein kinase I, incorporating ^{32}P

mainly at serine^{6} and to a much lesser extent at serine^{12} (Huflejt et al., 1993). Further, casein kinase II was unable to phosphorylate galectin-3 (huflejt et al., 1993).

The sequences surrounding Serine^{6} are conserved in all known galectin-3 proteins indicating a possible regulatory function. The dependence of intracellular redistribution and phosphorylation state of galectin-3 on the proliferation status of 3T3 cells suggest phosphorylation may modulate the function of galectin-3. This in turn, may modulate the activity of galectin-3 in splicing of pre-mRNA. These findings led us to determine whether galectin-3 and/or rCBP35 can be phosphorylated under splicing conditions.

MATERIALS AND METHODS

Refer to Chap. 2 materials and methods for preparation of nuclear extract, *in vitro* splicing conditions, SDS-polyacrylamide gel electrophoresis and Western blotting.

***In vitro* phosphorylation.** ^{32}P-labeled MINX pre-mRNA was incubated under splicing conditions in the presence of 30 μCi [γ^{32}P]ATP for 30 min at 30°C in the presence and absence of non-radioactive ATP (1.5 mM) and creatine phosphate (CP) (20 mM) and rCBP35 (0.05 μg). Phosphatase inhibitors NaF (final concentration 10 μM) and $ZnCl_2$ (final concentration 20 μM) were added immediately following *in vitro* phosphorylation. Galectin-3 proteins were purified by LAC-A affinity chromatography (chap. 2, Materials and Methods). The bound material was eluted from the affinity matrix beads with SDS-PAGE sample buffer. Protein samples were resolved on a 12% SDS polyacrylamide gel followed by Western blotting.

Phosphorylated proteins were visualized by autoradiography of transferred proteins. To immunologically detect labeled phosphorylated proteins, after autoradiography the PVDF membranes were analyzed by Western analysis.

Antibody against galactin-3 (anti-Mac-2) was generously provided by Dr. John Wang (Chap 1) and antibody against the hnRNP C proteins (4F4) was generously provided by Dr. Gideon Dreyfuss (Choi and Dreyfuss, 1984).

RESULTS

HeLa nuclear extracts were assayed for their ability to incorporate phosphate into nuclear proteins, galectin-3 and rCBP35. Figure 1A shows the locations of rCBP35 and galectin-3 on the autoradiogram confirmed by immunoblotting with anti-Mac2 antibody against galectin-3/rCBP35 (Fig. 1B, lanes 1-4). There was no detectable ^{32}P incorporation into galactin-3 (Fig. 1A, lanes 1 and 2) or rCBP35 (Fig. 1A, lanes 3 and 4) in the presence (lanes 2 and 4) or absence (lanes 1 and 3) of CP and ATP. However, a number of proteins did incorporate ^{32}P under the phosphorylation conditions used. Previous work has shown that hnRNP C1 and C2 proteins can be phosphorylated *in vitro* by casein kinase II found in HeLa cell hnRNP particle extracts (Holcomb and Friedman, 1984). When 4F4 antibody was used to localize hnRNP C1 and C2 proteins (Fig. 1C), two radioactive bands corresponded to the positions of the immunoblotted proteins in the unbound fractions (Fig. 1A lanes 5-8).

Figure 1A

Proteins phosphorylated during *in vitro* splicing. Splicing reactions using HeLa cell nuclear extract and ^{32}P-labeled MINX pre-mRNA were incubated in the presence of [γ^{32}P]ATP and in the presence (lanes 2, 4, 6 and 8) and absence (lanes 1, 3, 5 and 7) of ATP and CP and in the presence (lanes 3, 4, 7 and 8) and absence (lanes 1, 2, 5 and 6) of 0.05 μg rCBP35. Galectin-3 and rCBP35 were affinity purified by lactose affinity chromatography as described in Materials and Methods (Chap. 2). Bound (lanes 1-4) and unbound (lanes 5-8) proteins were separated by SDS-PAGE then transferred onto PVDF membrane followed by autoradiography. Locations of rCBP35, galectin-3, C1 and C2 are indicated on the right.

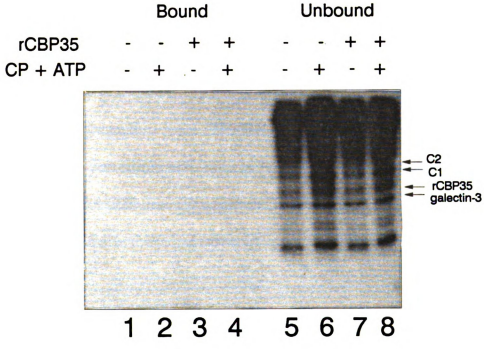

Bound
Unbound
rCBP35 - - + + - - + +
CP + ATP - + - + - + - +
C2
C1
rCBP35
galectin-3
1 2 3 4 5 6 7 8

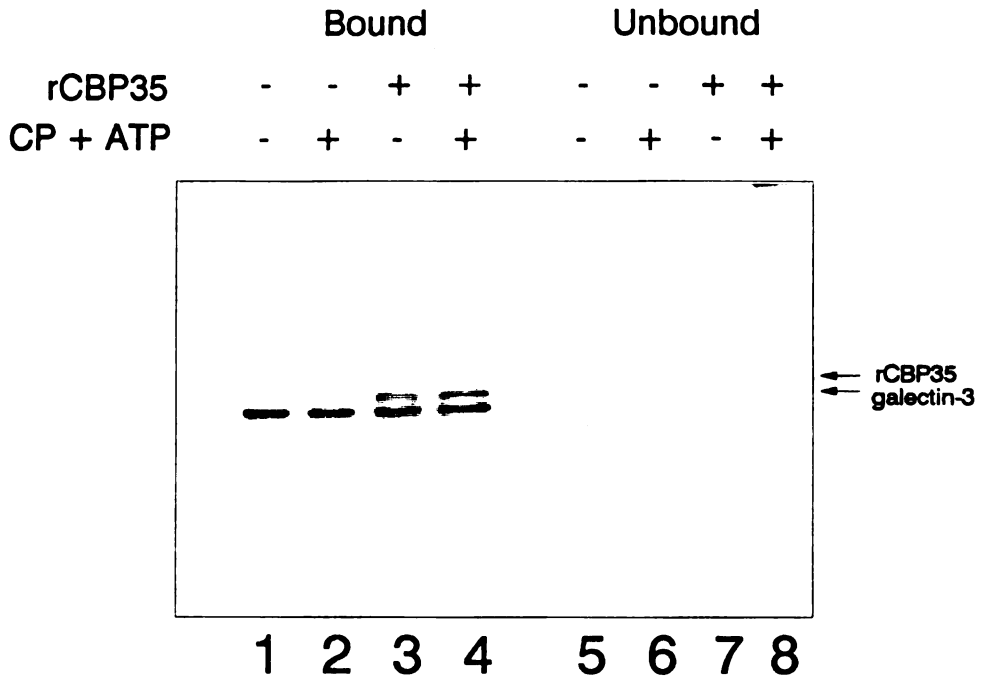
Figure 1A

Figure 1B

Immunoblot analysis of rCBP35 and galectin-3 following *in vitro* phosphorylation of a splicing reaction. The transferred proteins from Fig. 1A were analyzed by Western blotting with anti-Mac-2 monoclonal antibody. Locations of rCBP35 and galectin-3 are indicated on the right.

Figure 1C

Immunoblot analysis of C1 and C2 proteins following *in vitro* phosphorylation of a splicing reaction. The transferred proteins from Fig. 1A were analyzed by Western blotting with 4F4 antibody. Locations of C1 and C2 proteins are indicated on the right.



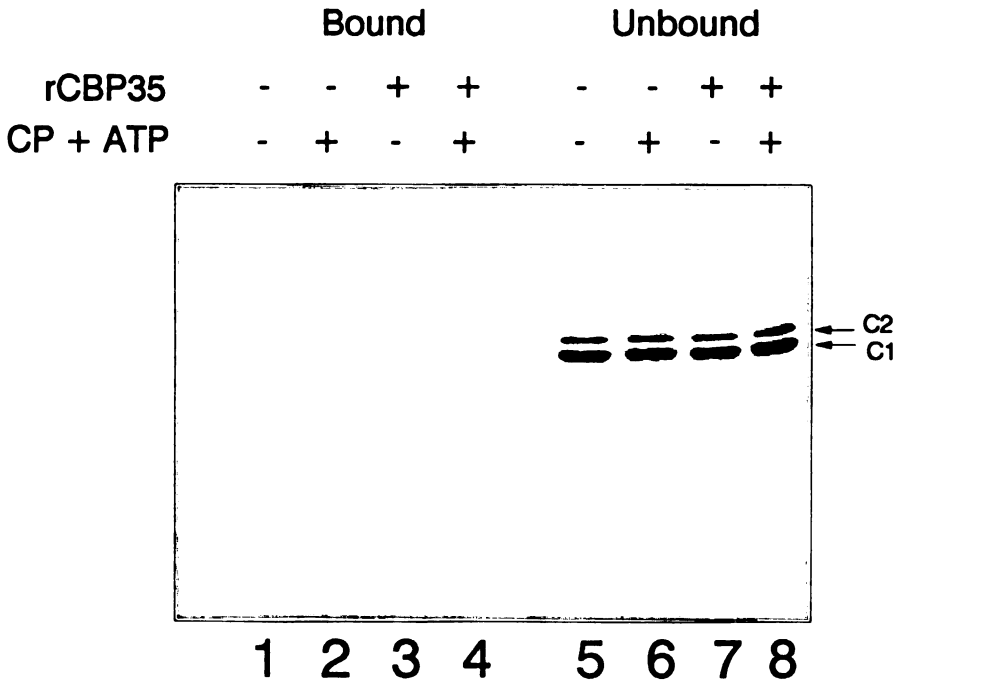


Figure 1B and Figure 1C

DISCUSSION

Digestion with collagenase D has been used to separate the galectin-3 carboxyl-terminal domain (carbohydrate binding domain) from the amino-terminal domain (Raz et al., 1989; Hsu et al., 1992; Agrwal et al., 1993). The carboxyl-terminal domain is insensitive to collagenase and thus remains intact during digestion, however, the amino-terminal domain, which resembles collagen, is extensively digested by collagenase treatment (Raz et al., 1989). Following phosphorylation at Ser^{6}, but not $Serine^{12}$, the Pro^{23}-Gly^{24} bond becomes resistant to collagenase digestion and the Phe^{5}-Ser^{6} bond becomes resistant to α-chymotrypsin digestion, suggesting phosphorylation at Ser^{6} may induce a conformational change at the galectin-3 amino-terminus (Huflejt et al., 1993). The amino-terminal domain has been shown to participate in cooperative binding of galectin-3 to glycoconjugate matrices (Hsu et al., 1992; Massa et al., 1993) indicating phosphorylation could affect interactions between galectin-3 and its ligand (Huflejt et al., 1993).

Galectin-3 appears to be phosphorylated by casein kinase I or a casein kinase I like protein (Milne et al., 1992; Huflejt et al., 1993). Casein kinase I has been demonstrated to affect cellular metabolism and phosphorylation of DNA or RNA binding proteins directly related to regulation of cellular growth (Tuazon and Traugh, 1991). Since galectin-3 appears to be associated with ribonucleoprotein complexes in the nucleus and in the cytoplasm of 3T3 fibroblasts (Laing and Wang, 1988; Wang,1988; Wang et al., 1992), Huflejt et al. (1993) suggest the possibility that the phosphorylation state of galactin-3 dictates its association with RNA. In line with this possibility, it could be postulated that the phosphorylation state of galectin-3

affects its involvement in splicing. If phosphorylation of galectin-3 is required before or during splicing then we would expect to find phosphorylated endogenous galectin-3 and/or phosphorylated rCBP35 following *in vitro* splicing in the presence of [γ-^{32}P]ATP. Neither galectin-3 nor rCBP35 were phosphorylated during *in vitro* splicing in the presence or absence of ATP and CP. It could be argued that the kinase required for phosphorylation of galectin-3/rCBP35 is absent or inactive in our nuclear extract preparation and sufficient phosphorylated endogenous galectin-3 is present to compensate for lack of phosphorylation activity. However, under our assay conditions numerous proteins were phosphorylated including the C1 and C2 hnRNP proteins (Holcomb an Friedman, 1984). These proteins have been shown to be phosphorylated by casein kinase II using an *in vivo* phosphorylation protocol (Huflejt et al., 1993). The hnRNP C proteins, specifically phosphorylated during mitosis (Pinol-Roma and Dreyfuss, 1991), have been implicated in *in vitro* pre-mRNA splicing (Choi et al., 1986). Since the intracellular activity of casein kinase I shows no major changes during the cell cycle (Tauzon and Traugh, 1991), it has been suggested casein kinase I constitutively phosphorylates galectin-3, thereby focusing on dephosphorylation rather then phosphorylation as a mode of regulating the phosphorylation state of galectin-3 (Huflejt et al., 1993). For example, control through phosphorylation/dephosphorylation has been suggested to regulate the transcriptional activity of p53 (Milne et al., 1992). Dephosphorylation regulates cellular processes including transcription, translation, metabolism, initiation of mitosis and progression through the cell cycle (Cohen, 1988). Additionally, Ser/Thr phosphatases appear to play an important role during both catalytic steps in *in vitro*

pre-mRNA splicing (Mermoud et al., 1992). Uncoupling of spliceosome formation from the catalytic steps is observed in the presence of phosphatase inhibitors and can be reversed by the addition of specific phosphatases (Mermoud et al., 1992). Known phosphorylated proteins involved in splicing include the U1 snRNP-specific 70 kDa protein (Woppmann et al. 1990), the U5 specific 52 kDa protein (Behrens and Luhrmann, 1991) and the non-snRNP splicing factor U2AF (Zamore and Green, 1989). Irreversible thiophosphorylation of the U1 specific 70 kDa protein using adenosine phosphorothioates (ATPαS and ATPγS) has been shown to block the first step in splicing (Tazi et al., 1992; Mermoud et al., 1992) adding to the mounting evidence that phosphorylation/dephosphorylation plays a mechanistic role in regulating RNA splicing.

It would be interesting to determine whether the addition of phosphorylated rCBP35 is able to reconstitute splicing activity. If we assume non-phosphorylated galectin-3 is required for splicing activity then successful reconstitution by phosphorylated rCBP35 may imply that rCBP35 would be dephosphorylated before or during the splicing reaction. Analysis by NEPHGE of added phosphorylated rCBP35 for reconstitution of galactin-3 depleted extracts (Cowles et al.,1990) would aid in answering some of these questions.

Photochemical Cross-Linking of Galectin-3 to pre-mRNA During pre-mRNA Splicing

INTRODUCTION

Photochemical (UV) cross-linking has been used to evaluate the numerous protein-RNA interactions that occur during *in vitro* pre-mRNA splicing including interactions with snRNPs and hnRNPs (Economidis and Pederson, 1983; Mayrand and Pederson, 1990; Mayrand et al., 1986; Staknis and Reed, 1994). The method of UV cross-linking relies on the ability of UV light of a specific intensity to generate photoreactive RNA or DNA which react with proteins in direct contact with the RNA or DNA molecule. Proteins in direct contact with DNA or RNA can then be identified when radioactive nucleotides protected by the cross-linked protein following nuclease digestion are essentially transferred to the protein (Smith, 1976).

In this study we have applied the method of UV cross-linking to determine whether galectin-3 binds directly to pre-mRNA during spliceosome formation.

MATERIALS AND METHODS

Refer to Chap. 2 Materials and Methods for preparation of nuclear extracts, *in vitro* pre-mRNA splicing conditions, SDS-PAGE, and Western blotting.

Pre-mRNA synthesis and *in vitro* splicing conditions. [^{32}P]GTP pre-mRNA was synthesized under the transcription conditions specified in Chap. 2. Reaction mixtures containing 35% nuclear extract (final protein concentration, ~12 mg/ml) were incubated at 30°C for 15 min in the presence and absence of 100 mM Lac.

UV cross-linking. To induce covalent cross-linking between ^{32}P-labeled pre-

mRNA and associated proteins, splicing reactions were irradiated on ice with 254-nm UV light for 20 min at a distance of 5 cm from the UV source.

Preparation and analysis of UV cross-linked proteins. ^{32}P-labeled RNA was digested with 50 units RNase T1 and 35 μg of RNase A at 37°C for 30 min. Galectin-3 was affinity purified by Lac affinity chromatography using immobilized Lac (LAC-A) as described in Materials and Methods (Chap. 2). Bound and unbound proteins were separated by SDS-PAGE, transferred onto PVDF membrane and positions of radiolabeled proteins determined by autoradiography. Galectin-3 and hnRNP C1 and C2 proteins were identified by Western blotting (Materials and Methods, Chap 4 under Phosphorylation). Proteins were eluted from LAC-A with SDS-PAGE sample buffer. LAC-A bound and unbound fractions were separated by SDS-PAGE on 15% gels and then transferred to PVDF membrane for Western blotting (Materials and Methods, Chap 4 under Phosphorylation). The positions of radiolabeled proteins were detected by autoradiography of the transferred proteins.

RESULTS

To identify a possible direct interaction between galectin-3 and pre-mRNA during splicing, we analyzed the proteins in a splicing reaction that were ^{32}P-labeled after UV induced cross-linking (label transfer). To improve sensitivity, galectin-3 was affinity selected on LAC-A columns after photo induced cross-linking. The LAC-A bound fraction was analyzed directly for ^{32}P-labeled proteins by autoradiography and then galectin-3 verified by Western analysis using anti-Mac2 (Fig. 2B, lanes 1 and 4). We could not detect ^{32}P-labeled galectin-3 affinity purified over LAC-A in the

absence (fig. 2A, lane 1) or presence (Fig. 2A, lane 2) of Lac. A number of hnRNP proteins including C1 and C2 have been shown to directly bind to pre-mRNA (Garcia-Blanco et al., 1990; Garcia-Blanco et al., 1989; Gil et al., 1991; Patton et al., 1991; Stolow et al., 1991). When cross-linked proteins were analyzed by Western analysis with antibody against hnRNP C proteins, we detected two bands (Fig. 2C, lanes 3 and 4) that could be superimposed onto two radioactive bands (fig. 2A). Thus, direct interactions between hnRNP C1 and C2 proteins and pre-mRNA were demonstrated under our UV cross-linking conditions.

Figure 2A

UV cross-linked proteins following *in vitro* splicing. Splicing reactions using HeLa cell nuclear extract and ^{32}P-labeled MINX pre-mRNA were incubated in the presence (lanes 2 and 4) and absence (lanes 1 an 3) of 100 mM lactose. Galectin-3 was affinity purified by lactose affinity chromatography as described in Materials and Methods (Chap. 2). Bound (lanes 1 and 2) and unbound proteins (lanes 3 and 4) were separated by SDS-PAGE then transferred onto PVDF membrane followed by autoradiography. Locations of galectin-3, C1 and C2 are indicated on the right.

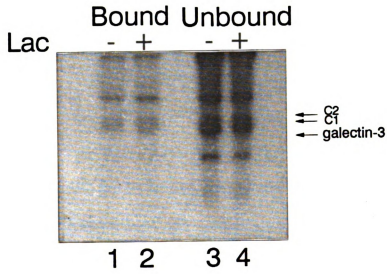


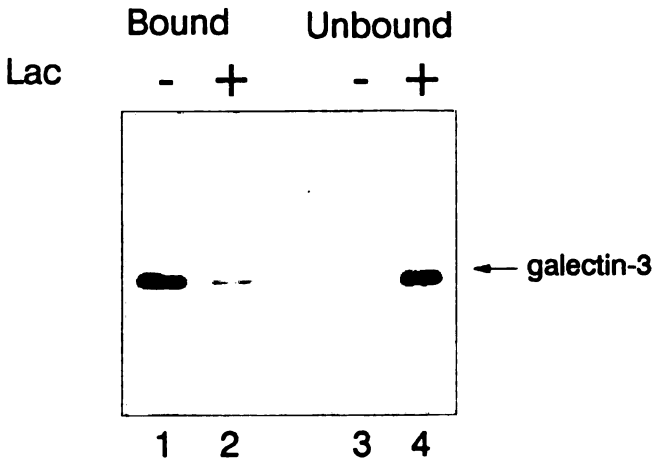
Figure 2A

Figure 2B

Immunoblot analysis of galectin-3 following UV cross-linking of a splicing reaction. The transferred proteins from Fig. 2A were analyzed by Western blotting with anti-Mac-2 monoclonal antibody. The location of galectin-3 is indicated on the right.

Figure 2C

Immunoblot analysis of C1 and C2 proteins following UV cross-linking of a splicing reaction. The transferred proteins from Fig. 2A were analyzed by Western blotting with 4F4 antibody. Locations of C1 and C2 proteins are indicated on the right.



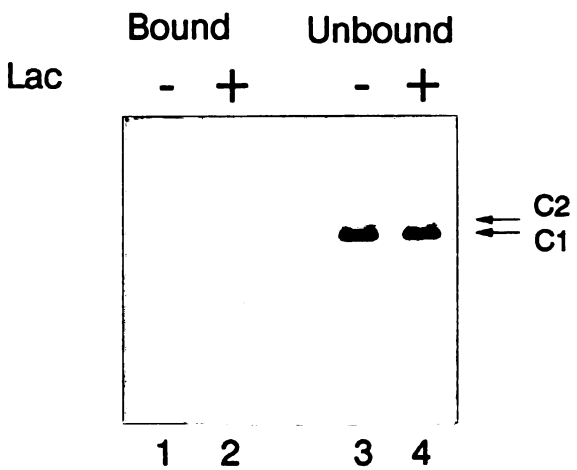


Figure 2B and Figure 2C

DISCUSSION

The results indicate that galectin-3 is not directly associated with pre-mRNA during *in vitro* splicing. Rather, the interaction appears to be indirect such as a protein-protein interaction, possibly through the galectin-3 ligand. It is possible that the UV cross-linking conditions used were not favorable for detecting a direct interaction between galectin-3 and pre-mRNA. Protein cross-linking depends on the pyrimidine richness of the RNA in contact with the protein since pyrimidines are known to be favorable UV-induced RNA-protein cross-linking sites (Smith, 1976). Thus, it is possible for a protein to be tightly associated with RNA and not be UV cross-linkable (Mayrand and Pederson, 1990). We have tested UV cross-linking using pre-mRNA labeled with ^{32}P[CTP] and ^{32}P[UTP] and obtained similar results to this study (data not shown) in which ^{32}P[GTP] was used to label the pre-mRNA.

REFERENCES

Agrwal, N., Sun, Q., Wang, S.-Y., and Wang, J. L. (1992) *J. Biol. Chem.* 268, 14932-14939.

Behrens, S. V., and Luhrmann, R. (1991) *Genes and Dev.* 5, 1439-1452.

Choi, Y. D. and Dreyfuss, G. (1984) *J. Cell Biol.* 99, 1997-2004.

Cohen, P. (1988) *Proc. Roy. Soc. Lond.* 234, 115-144.

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

Economidis, I. V. and Pederson, T. (1983) *Proc. Natl. Acad. Sci. USA* 80, 4296-4300.

Garcia-Blanco, M. A., Jamison, S., and Sharp, P. A. (1989) *Genes Dev.* 3, 1874-1886.

Garcia-Blanco, M. A., Anderson, G. J., Beggs, J. and Sharp, P. A. (1990) *Proc. Natl. Acad. Sci. USA* 87, 3082-3086.

Gil, A., Sharp, P. A., Jamison, S. F., and Garcia-Blanco, M. (1991) *Genes Dev.* 5, 1224-1236.

Hamann, K. K., Cowles, E. A., Wang, J. L., and Anderson, R. L. (1991) *Exp. Cell Res.* 196, 82-91.

Holcomb, E. R. and Friedman, D. L. (1984) *J. of Biol. Chem.* 259, 31-40.

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

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

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

Mayrand, S. H., Pederson, N., and Pederson, T. (1986) *Proc. Natl. Acad. Sci. USA* 83, 3718-3722.

Mayrand, S. H., Pederson, T. (1990) *Nuc. Acids Res.* 18, 3307-3318.

Mermoud, J. E., Cohen, P. and Lamond, A. I. (1992) *Nuc. Acids Res.* 20, 5263-5269.

Milne, D. M., Palmer, R. H., Campbell, D. G., and Meek, D. W. (1992) *Oncogene* 7, 1361-1369.

Patton, J. G., Mayer, S. A., Tempst, P., and Nadal-Ginard, B. (1991) *Genes Dev.* 5, 1237-1251.

Pinol-Roma, S. and Dreyfuss, G. (1991) *Science* 253, 312-314.

Smith, K. C. (1976) *In photochemistry and Photobiology of Nucleic Acids,* ed. Wans, S. Y. (Academic Press, New York) vol. 2, 187-218.

Staknis, D. and Reed, R. (1994) *Mol. and Cell. Biol.* 14, 2994-3005.

Stolow, D. T., and Berget, S. M. (1991) *Proc. Natl. Acad. Sci. USA* 88, 320-324.

Tazi, J., Daugeron, M. C., Cathala, G., Brunel, C., and Jeanteur, P. (1992) *J. Biol. Chem.* 267, 4322-4326.

Wang, J. L., Werner, E. A., Laing, J. G., and Patterson, R. J. (1992) *Biochem. Soc. Trans.* 20, 269-274.

Woppmann, A., Patschinsky, T., Bringmann, P., Godt, F. and Luhrmann, R. (1990) *Nucleic Acids Res.* 18, 4427-4438.

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

CHAPTER V

CONCLUSION

Conclusions and Future Experimentation

The data presented in this thesis identify galectin-3 as a required splicing factor. Two basic assays were used to arrive at this conclusion: (i) perturbation of pre-mRNA splicing by the addition of saccharides with a high affinity for galectin-3 and (ii) depletion - reconstitution experiments in which purified rCBP35 restored splicing activity and spliceosomal complex formation to a galectin-3-depleted NE. Further, galectin-3 is a member of a family of nuclear galectins, each of which fits the definition of a splicing factor. These latter data suggest that nuclear galectins are functionally redundant in RNA splicing.

Many questions remain unanswered. The major question is what is the nuclear ligand for galectin-3? Is it a nuclear glycoconjugate that binds to the CRD of galectin-3 or a structure mimicking Lac? Alternatively, is the ligand a nuclear protein that binds to a region of galectin-3 other then the CRD region? A second important question is what is the role of galectin-3 and other galectins in the splicing pathway? Indeed, product formation is reduced during pre-mRNA splicing in the presence of Lac and derivatives of Lac and not by other sugars. Galectins may act to sequester a pre-mRNA processing inhibitor(s). When the lectin is removed and/or Lac is present, the inhibitor(s) would be free to act. Consequently, a reduction in Lac specific lectins during differentiation and/or serum deprivation may target reduction in processing of certain pre-mRNAs.

The function of the amino-terminus remains a mystery. Could this domain mediate a second functional interaction needed during pre-mRNA processing? The

amino-terminal domain does contain a region that is 55% identical to a region of a human transcription factor, SRF (Norman et al., 1988; Oda et al., 1991). The amino-terminal domain also contains an unusual repeating sequence PGAYPG; seven repeating units in the human (Oda et al., 1991) and nine in mouse (Jia and Wang, 1988; Raz et al., 1989) and rat (Albrandt et al., 1987). So far a GenBank data base search for other proteins containing the PGAY repeat has been unsuccessful, although many proteins contain the solitary sequence (Oda et al., 1991).

Work is currently in progress to elucidate the specific function of galectin-3 in pre-mRNA splicing and to identify the nuclear ligand. However, whatever function galectin-3 may reveal to us in the future it will certainly add new dimensions to the rapidly expanding body of diverse information in so many facets of science.


MICHIGAN STATE UNIV. LIBRARIES
31293010460537