This is to certify that the

thesis entitled

Studies on the nuclear export signal of galectin-3

presented by

Su-Yin Li

has been accepted towards fulfillment
of the requirements for

M.S. degree in Biochemistry

Major professor

Date 08/10/01

STUDIES ON THE NUCLEAR EXPORT SIGNAL OF GALECTIN-3

By

Su-Yin Li

A THESIS

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

MASTER OF SCIENCE

Department of Biochemistry and Molecular Biology
2001

ABSTRACT

STUDIES ON THE NUCLEAR EXPORT SIGNAL OF GALECTIN-3

By

Su-Yin Li

Galectin-3 (Gal3) is a pre-mRNA splicing factor that shuttles between the nucleus and the cytoplasm of cells. To identify the nuclear localization signal (NLS) and the nuclear export signal (NES) in the Gal3 polypeptide, we have engineered a reporter construct expressing a fusion protein containing Gal3 and green fluorescent protein (GFP). Using this system, we found that the carboxyl terminal region of the murine Gal3 polypeptide, spanning leucine 247 through alanine 258, was important for NLS. We also found that the NLS and NES appeared to overlap in this region of the amino acid sequence, precluding us from using the GFP-MalE-Gal3 construct to analyze the NES. Therefore, we turned to an alternative vector, pRev(1.4)-GFP, into which amino acid sequences can be inserted for testing their nuclear export activity. Using this construct, we found the segment of the Gal3 polypeptide, starting at asparagine 240 through leucine 255, exhibited NES activity. The amino acid sequence from leucine 241 through isoleucine 249 corresponds well to a leucine-rich NES and site-directed mutagenesis of leucine 247 and isoleucine 249 to alanine residues affected the nuclear export activity. Consistent with the notion that CRM1, the transport receptor for leucine-rich NESs, is sensitive to inhibition by leptomycin B, the fusion protein containing Rev(1.4)-GFP and Gal3 NES shifts its localization in favor of the nucleus in the presence of the drug. These results suggest that, indeed, the NES of Gal3 appears to be located in the same region of the amino acid sequence that was also important for NLS.

Dedication

To my Mother,

and to my husband, Ren-Song,

my little children, Meg and Albert

Acknowledgements

I owe thanks to Dr. John Wang for the unique educational opportunity he provided for me in his laboratory. I would also like to thank my lab mates, Kyle Openo, Nancy Lin, Patty Voss, Peter Davidson, and Richard Gray, for helpful discussions and caring fellowship.

Most of all, thank you to Ren-Song, who shared with me the care of two young kids and household chores.

TABLE OF CONTENTS

LIST OF TABLES vii

LIST OF FIGURES viii

LIST OF ABBREVIATIONS x

CHAPTER 1: LITERATURE REVIEW 1

INTRODUCTION 2

GALECTIN-3 3

General background 3

Localization 6

A role in pre-mRNA splicing 7

Association with gemin complex 8

Intracellular activities 9

NUCLEAR EXPORT 10

Introduction 10

The Ran GTPase cycle 10

Export of leucine-rich nuclear export signal proteins 12

Recycling of importin-α 13

Export of RNP 13

REFERENCES 16

CHAPTER 2: ANALYSIS OF A PUTATIVE LEUCINE-RICH NUCLEAR EXPORT SIGNAL IN GALECTIN-3 20

INTRODUCTION 21

MATERIALS AND METHODS 23

Cell culture and reagents.....23

Preparation of the pEGFP-C1 vector for the expression of fusion protein GFP-MalE-Gal3 24

The pRev(1.4)-GFP vector and variants 27

Fluorescence microscopy 27

SDS-PAGE and immunoblotting 28

RESULTS 29

A GFP reporter construct for the localization of galectin-3 29

The COOH-terminal region of Gal3 is required for nuclear localization 32

An attempt to identify the NES using the GFP-MalE-Gal3 reporter 36

The Rev(1.4)-GFP vector for the analysis of a functional NES 39

Analysis of the Gal3 NES in the Rev(1.4)-GFP vector 48

The effect of LMB on the fluorescence distribution 48

Site-directed mutagenesis of the Gal3 NES 56

DISCUSSION 58

ACKNOWLEDGMENTS; FOOTNOTES 63

REFERENCES.....64

LIST OF TABLES

CHAPTER 1

Table 1. Export signals and receptors 11

CHAPTER 2

Table 1. The effect of actinomycin D on the localization of GFP fusion proteins 45

Table 2. The effect of leptomycin B on the localization of GFP fusion proteins 52

LIST OF FIGURES

CHAPTER 1

Figure 1. Schematic diagram illustrating the polypeptide architecture of the galectins 4

CHAPTER 2

Figure1. Schematic diagram illustrating the construction of the vector for the expression of the fusion protein GFP-MalE-Gal3 in mammalian cells 25

Figure 2. Summary of the properties characterized for the fusion protein GFP-MalE-Gal3 and variants 30

Figure 3. Representative fluorescence micrographs illustrating the N=C versus C labeling patterns 34

Figure 4. Comparison of the leucine-rich nuclear export signals (NES) identified in protein kinase A inhibitor (PKI) and in HIV-1 Rev protein with the putative NES found in the sequences of Gal3 from various species 37

Figure 5. (A) Schematic diagram of the pRev(1.4)-GFP vector for testing potential nuclear export signals (NES); (B) Schematic diagram illustrating the basis of the assay to determine the nuclear export activity of a potential NES; (C) Summary of the contents of the fusion proteins expressed by various constructs 40

Figure 6. Representative fluorescence micrographs illustrating the classification of GFP localization patterns: N+C, fluorescence observed in the nucleus and in the cytoplasm; N, fluorescence exclusively in the nucleus; and C, fluorescence exclusively in the cytoplasm 43

Figure 7. Display of the data in Table I in the form of histogram distributions of the percent of cells with fluorescence patterns N, N+C, and C 46

Figure 8. Representative fluoresence micrographs showing the nuclear versus cytoplasmic distribution of Rev(1.4)-GFP protein containing the putative NES sequence of Gal3 or site-directed mutants 49

Figure 9. Display of the data in Table II in the form of histogram distributions of the percent of cells with fluorescence patterns N, N+C, and C 52

Figure 10. A comparison of the histogram distributions of the percent of cells with fluorescence patterns N, N+C, and C for Gal3 NES (wt), Gal3 NES (I244A; L247A), and Gal3 NES (L247A; I249A) 54

LIST OF ABBREVIATIONS

ActD	Actinomycin D
CAS	Cellular apoptosis susceptibility protein
CRD	Carbohydrate-recognition domain
CRM1	Chromosomal region maintenance protein
Gal 3	Galectin-3
Gemin4(C50)	The carboxyl-termial 50 residues of Gemin4
GFP	Green fluorescent protein
hnRNP	Heterogeneous nuclear ribonucleoprotein complex
IκB-α	Inhibitior of kappa B-alpha
LMB	Leptomycin B
MalE	Maltose-binding protein
ND	Amino-terminal domain
NES	Nuclear export signal
NLS	Nuclear localization signal
PCR	Polymerase chain reaction
PKI	Protein kinase A inhibitor
RNP	Ribonucleoprotein
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
Sm	Smith antigen of small nuclear ribonucleoprotein complex
SMN	Survival motor neuron protein
snRNP	Small nuclear ribonucleoprotein complex

Chapter 1

Literature Review

INTRODUCTION

With Dr. Eric Arnoys and two undergraduate students in the Wang laboratory, I began a project to identify and characterize the nuclear localization signal and the nuclear export signal in the protein, galectin-3. Together, we engineered a fusion protein containing the green fluorescent protein (GFP) reporter. The original plan was for me to work on the export signal and Dr. Arnoys to work on the import signal, once the GFP fusion construct was developed.

As will become apparent in the experimental chapter, however, the import and export signals appeared to overlap over the same stretch of the galectin-3 amino acid sequence. Therefore, I turned to another GFP fusion system, developed by our collaborator, Dr. Beric Henderson (University of Sydney, Australia). I used this system to study the galectin-3 nuclear export signal.

Hence, the literature review portion of this thesis will cover: (a) the intracellular, particularly the nuclear and cytoplasmic, activity of galectin-3; (b) nuclear export signals and their transport receptors.

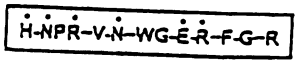
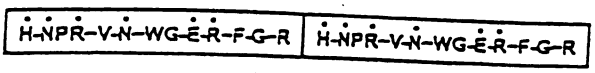
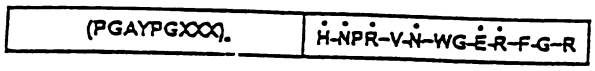
Galectin-3

General Background Galectin-3 was initially isolated from the extracts of cultured mouse Swiss 3T3 fibroblasts and human SL66 fibroblasts by galactose-affinity columns (1, 2). Since then, it has been found in various tissues of many organisms, such as human, rat, mouse, dog and hamsters.

Galectin-3 (M_r~30,000), like other members in the galectin family, is characterized by two major features : (a) the capacity for binding β-galactosides, and (b) a carbohydrate-recognition domain (CRD) with conserved sequence elements (3). The galectin family is grouped into three main subfamilies (Fig 1) : (a) the Prototype group, such as galectins-1, -2, -5, -7, -10, and -11, consists of one CRD, (b) the Tandem Repeat group, including galectin-4, -6, -8, and -9, has two CRDs; and (c) the Chimera group contains an unique proline- and glycine-rich domain in the amino-terminal portion (ND) in addition to a CRD in the carboxyl-terminus, and currently, galectin-3 is the only member identified in this group.

Studies on the localization of galectin-3 indicate it is a predominantly intracellular (cytoplasm and nucleus) protein although it has also been found in the extracellular (cell surface and medium) compartment (4, 5, 6, 7). Therefore, it is reasonable to assume that galectin-3 may have intracellular functions. Indeed, galectin-3 has been shown to be required for the mRNA splicing process in the nucleus (8. 9), using a cell-free assay. Also, it can inhibit apoptosis by interacting with the anti-apoptotic protein Bcl-2 in the cytoplasm. More strikingly, galectin-3 has been recently found associated with two novel proteins, Chrp as well as Gemin4. Chrp is a cytoplasmic protein rich in both cysteine and histidine. It was identified as the interacting partner for galectin-3 in the yeast two-hybrid

Figure 1. Schematic diagram illustrating the polypeptide architecture of the galectins. The Proto Type is composed of a single domain, the CRD. The Tandem Repeat Type has two homologous CRDs. The Chima Type has two parts, a COOH-terminal half containing the CRD and an NH_2-terminal half containing a repeating motif rich in proline and glycine residues. The letter n represents the number of times this proline- and glycine-rich motif is repeated, is species-specific, and varies form 8 to 12. The single letter amino acid code is used. X denotes any amino acid. Conserved amino acid residues are indicated. Dark circles denote residues that interact with carbohydrate by hydrogen bonding.

Subfamily	Member	Domain Organization
Proto Type	1, 2, 5, 7, 10, 11	
Tandem Repeat Type	4, 6, 8, 9	
Chimera Type	3	

system using galectin-3 as the bait to screen a mouse 3T3 cDNA library (10). The other one, Gemin4, is one component of SMN complexes in the nucleus. These data raise the possibility that galectin-3 is a versatile molecule inside the cell. In this review, the discussion is focused on those intracellular activities of galectin-3.

Localization Galectin-3 can be found at the cell surface, in the cytoplasm and in the cell nucleus. Moreover, previous immunofluorescence studies have revealed that the subcellular distrubution of galectin-3 is proliferation-dependent. For example, galectin-3 is primarily cytoplasmic in contact-inhibited or serum-starved quiescent cells but predominantly nuclear in sparse proliferating cells (5).

Subcellular fractionation coupled with biochemical analysis documented that galectin-3 has two isoform: (a) a nonphosphorylated form (pI ~ 8.7), and (b) a phosphorylated derivative (pI ~ 8.2) (11). Although both forms of the polypeptide are found in the nucleus, only the phosphorylated form can exit the nucleus in digitonin-permeabilized cells, where the nucleocytoplasmic transport machinery still remains intact (12). This result implies the phosphorylation may be important in the nuclear export process of galectin-3. The export, moreover, can be blocked by leptomycin B (LMB), a cytotoxin that specifically inhibits the export of proteins containing leucine-rich nuclear export signal (NES) by directly binding to CRM1 export receptors. Indeed, a putative leucine-rich nuclear export signal can be found in the galectin-3 homologs of various species. Based on these obervations, it is likely that galectin-3 is a shuttling protein between the nucleus and cytoplasm, and this means this protein might have a nuclear export signal as well as a nuclear import signal, both of which await further identification.

A role in pre-mRNA splicing Previous studies implicate that galectin-3 is associated with ribonucleoprotein (RNP) structures in the nucleus (13), and is localized in interchromatin spaces as well as the border of condensed chromatin at the ultrastructural level (14). Consequently, the possibility that galectin-3 might play a role in the events of mRNA synthesis was raised since both hnRNPs and snRNPs are closely related to the splicing of pre-mRNA. Furthermore, the border of condensed chromatin is regarded as the site of pre-mRNA processing (15). Indeed, several lines of evidence establish that galectin-3 is a pre-mRNA splicing factor (8, 9). First, a Hela cell nuclear extract (NE) which is able to accomplish pre-mRNA splicing in vitro contains galectin-3. Second, the splicing in vitro is sensitive to saccharide-specific inhibition, which means the splicing reaction is specifically suppressed by lactose or galactose, but not mannose, sucrose or N-acetylglucosamine. Third, depletion of galectin-3 from the NE results in the loss of the splicing activity if galectin-1 is concurrently eliminated by lactose affinity or double antibody adsorption. Depletion of either galectin-1 or galectin-3 fails to remove all of the splicing activity and the remaining activity could still be blocked by saccharides that bind the galectins with high affinity, such as lactose and galactose. Finally, the splicing activity lost from the NE depleted of both galectins can be recovered , at least partially, by addition of either galectin-3 or galectin-1 alone. All of these data strongly indicate that galectin-3 is involved in the pre-mRNA splicing.

Galectin-3 is composed of two domains, a CRD in the carboxyl-terminal half as well as as a proline- and glycine-rich domain in the amino-terminal half (Fig. 1). Interestingly, we found that CRD somehow has positive effects on the splicing reaction because

splicing activity in NEs depleted of both galectin-3 and galectin-1 could be restored by the CRD alone, without the ND of galectin-3 (8). However, the ND has a dominant negative effect on the pre-mRNA splicing because of the observation that the splicing reaction was inhibited by the addition of exogenous ND and the inhibition was dose-dependent (16).

Association with Gemin Complex Recent experiments have identified that the carboxyl-terminal 50 residues of Gemin4 (Gemin4(C50)) interact directly with galectin-3 (16). Gemin4 is one component of a macromolecular complex which includes several other proteins, such as SMN, Sm B/B', Sm D1-3, Gemin2, and Gemin3. This macromolecular complex in the nucleus is crucial for splicing because it provides the H-complex with an essential component, snRNPs, during the process of spliceosome assembly (17). Indeed, galectin-3 is considered as a required factor for the progression from H to higher order complexes in the assembly of splicesomes, based on three lines of experiments. First, when splicing substrates were incubated with galectin-free NE, H complex was found to be the arrested product and no other higher order structure was produced (9). Second, the gel mobility shift assay indicates that galectin-3 is associated with H-complex in the absence of ATP (unpublished data). Third, experimental results have implicated the negative effect of ND on the splicing reaction is the result of its ability to stop spliceosome assembly at the H-complex.

The details of the role of galectin-3 in the spliceosome formation and the fate of the galectin-3 after H-complex converts to active splicesomes are interesting subjects to explore in the future.

Intracellular Activities There are data to suggest that cytosolic galectin-3 can suppress apoptosis by interacting with Bcl-2, a well-known anti-apoptotic protein (18). For example, Yang et al found that human leukemia T cells transfected with galectin-3 cDNA proliferated faster than those without expression of galectin-3. They also discovered that the galectin-3 transfectants were somehow resistant to apoptosis induced by either anti-Fas antibody, directed against a cell surface antigen of the tumor necrosis factor receptor family, or staurosporine, a common inducer of apoptosis. More strikingly, galectin-3 was shown to interact with anti-apoptotic protein Bcl-2. It is thought that this interaction between galectin-3 and Bcl-2 confers resistance to apoptosis. Interestingly, it is shown that Bcl-2 has a proline-, glycine-, and alanine-rich sequence in the amino-terminus, resembling the ND of galectin-3. Also, a highly conserved NWGR element is seen in both Bcl-2 and galectin-3.

Similarly, the expression of galectin-3 can prevent inflammatory cells from undogoing apoptosis, which leads to increased level of inflammatory reactions (19). This phenomenon could be observed in the mice containing the normal galectin-3 gene. Compared to wild-type, mice in which the galectin-3 gene has been inactivated had much fewer macrophages, resulting in less inflammation. Similar observations were made on granulocytes in a second strain of galectin-3 null mice (20).

Nuclear Export

Introduction One of the major differences that distinguish eukaryotic cells from prokaryotes is their nuclear envelopes, which separate RNA biogenesis and DNA replication in the nucleus from the cytoplasmic machinery for protein synthesis. Although the separation provides a means to control gene expression, it requires an additional machinery to govern the nucleocytoplasmic transport process of proteins and some RNPs that need to move between the nucleus and cytoplasm. For instance, various groups of cellular RNA molecules are synthesized in the nucleus, and most of them have to be exported from the nucleus to the cytoplasm, whereas there are lots of proteins that must be imported into the nucleus in order to perform their normal functions. Recently, not only have many factors that mediate nuclear import been uncovered, but also a long list of newly discovered factors (Table 1) implicated in the process of nuclear export have been found (21). These studies have also revealed that a small GTPase Ran is required for the nucleocytoplasmic transport process (22, 23). In this review, the discussion is slanted toward the events of nuclear export, including export of RNPs and the Ran GTPase cycle, which is associated with cargo binding and release.

The Ran GTPase cycle Several lines of research have indicated that the small GTPase Ran plays an important role in the directionality of transport (22, 23). Ran is mainly localized in the nucleus, and, like other small GTPases, it switches between GTP- and GDP-bound states. It has been proposed that nuclear Ran is most likely to be GTP-bound , whereas the small amounts of Ran in the cytoplasm are predominantly

Table 1. Export signals and receptors

Export Cargo	Export Signal	Export Receptor
Proteins with leucine-rich NES	LALKLAGLDI (NES in PKI, the inhibitor of cAMP-dependent protein kinase). Signals are generally leucine-rich.	CRM1
U snRNA	m7G cap binds CBC; CBC or proteins that interact with it contain NES.	CRM1
5S rRNA	Possibly mediated by TFIIIA or ribosomal protein L5.	CRM1
tRNA	Mainly acceptor and TΨC arms. Mature 5' and 3' termini also important.	Exportin-t/los1p
Importin α	Large region (~140 amino acids)	CAS
mRNA	mRNA export factors include:	
	Shuttling hnRNP proteins	N/K
	Gle2p	N/K
	TAP	N/K
	Ddp5p	CRM1

abbreviations: CBC, cap-binding protein complex; N/K, not known.
(This table was modified from the table 2 in reference 21.)

maintained as RanGDP (24, 25, 26). The nucleotide state of Ran is thought to be used for regulating cargo binding and release in nuclear transport. For example, export receptors form stable complexes with their cargoes in the nucleus as a consequence of high concentration of RanGTP. However, after these export receptor-cargo complexes are exported, they dissociate in the cytoplasm where RanGTP is mostly converted to RanGDP. In contrast, import receptors bind their cargos in a RanGTP-independent manner and RanGTP causes disassembly of these complexes in the nucleus. As a result, cargos are easily bound to import receptors in the cytoplasm and released in the RanGTP rich nucleus (27, 28, 29).

Export of leucine-rich nuclear export signal proteins Efforts to unearth how retroviruses export their intron-containing RNA molecules led to the discovery of a leucine-rich NES in the HIV protein Rev, one of the best-characterized NESs. This Rev NES is transferable and consists of a short leucine-rich stretch of amino acids with the consensus: Leu-X_{2-3}-Y-X_{2-3}-Leu-X-Leu/Ile, where X stands for any amino acid and Y represents Leu/Ile/Phe/Val or Met (30, 31). CRM1(chromosomal region maintenance) has been identified as a receptor that recognizes a Rev-like NES and mediate its export pathway (32, 33). An important clue leading to the identification of CRM1 as the NES receptor is from pharmacological experiments using the antibiotic LMB, which inhibits the NES-mediated protein export by directly binding to CRM1. To date, the leucine-rich NES has been found in several proteins, such as PKI (34), IkB (34, 35), MAPKK (36), RanBP1 (37), TFIIIA (30, 38), Gle1P (39), and Mex67p (40).

Recycling of importin-α Importin-α is an adaptor protein that mediates the interaction of classical nuclear localization signal (NLS) with importin β (41, 42). After translocation of the NLS receptor-cargo complex through the pore and release of the cargo in the nucleus, the importin-α has to be returned to the cytoplasm in order to fulfill multiple rounds of nuclear import. The protein charged with the responsibility of recruiting nuclear importin-α to the cytoplasm was identified as the CAS protein (27). In the nuclear compartment, CAS can bind to importin-α in the presence of Ran-GTP and then transport it to the cytoplasm, where importin-α is needed to support the nuclear import machinery. Therefore, the CAS protein plays a crucial role in the recycling process of importin-α.

Export of RNP: U snRNA, rRNA, tRNA, and mRNA

The vast majority of cellular RNA molecules are synthesized in the nucleus and need to exit the nucleus to accomplish their assignments in the cytoplasm. In the nucleus, RNAs are associated with proteins and it is thought that the signal for RNA export resides in these protein components. Indeed, much evidence suggest that the export of all classes of RNAs is selectively saturable, which indicates the existence of class-specific export factors (43, 44, 45). Further, the RNP transport machinery has an interesting feature, which is its ability to distinguish mature from immature RNPs. This property has been extensively studied in the example of tRNA. However, similar principles for confining nonfunctional RNPs to the nucleus could apply to other RNPs. Even though most RNP transports share this general feature, they are different in many other manners.

How each type of RNP selectively gets out of the nucleus is addressed individually below.

i) U snRNA

Its 5' m7G cap structure is essential for export, and some of the proteins involved have been discovered. For instance, a nuclear monomethyl cap-binding protein complex (CBC, consisting of both CBP20 and CBP80) has been found to mediate the export of U snRNAs (46). Also, CRM1 and possibly other factors are shown to get entangled in the event of U snRNA export (33, 44).

ii) 5S rRNA

Two proteins, TFIIIA and ribosomal protein L5, have been shown to have a connection with the export of 5S rRNA, because mutant 5S rRNA unable to bind either protein remains in the nucleus, whereas wild type 5S rRNA that has the capacity to bind either one can be exported (47). Moreover, CRM1 has been implicated as the mediator of the 5S rRNA export because TFIIIA has a Rev-like NES and excess leucine-rich NES conjugates can saturate 5S rRNA export (30, 48).

iii) tRNA

Exportin-t in higher eukaryotes and los1p in S. cerevisiae have been demonstrated to bind directly to tRNA (49, 50, 51). Some features of tRNA export result in the fidelity of export. For example, exportin-t can ignore immature tRNAs and bind exclusively to mature tRNAs (49). Several lines of evidence also disclosed that those charged tRNAs produced by aminoacylation, the last step of forming functional tRNAs, have higher affinity to exportin-t than uncharged ones, which are nonfunctional (49, 52). As a

consequence, those nonfunctional tRNAs are restricted to the nucleus and those tRNAs that can exit the nucleus are functional.

iv) mRNA

The first group of proteins implicated in mRNA export of higher eukaryotes is the hnRNP protein family, and this favors the idea that those proteins involved in the pre-mRNA splicing are somehow responsible for mRNA export . Although most hnRNP proteins are predominantly localized in the nucleus, some of them do shuttle rapidly between the nucleus and cytoplasm (53, 54, 55). Two shuttling hnRNP proteins, hnRNP A1 and hnRNP K, have been found to mediate mRNA export by their bidirectional signals, M9 and KNS, respectively (41, 42, 56, 57).

A large number of other candidate proteins that are involved in exporting mRNA have also been revealed, including Glep2 (58), TAP (40), Dbp5p (59, 60), HIV-Rev proteins (30, 31, 32, 33). All of them hold export signals for mRNA transport. Although nuclear export receptors of Dbp5p and Rev protein are characterized as CRM1, those of the other proteins with the signals for mRNA export are uncharacterized yet. It is the clear expectation that additional receptors for other export pathways will also be found, including for mRNA export.

REFERENCES

1. Roff, C.F. and Rosevear, P.R., Wang, J.L., and Barker, R. (1983). *Biochem. J.* **211**,625.
2. Roff, C.F. and Wang, J.L. (1983). *J. Biol. Chem.* **258**, 10657.
3. Kasai, K. and Hirabayashi, J. (1996) *J. Biochem.* **199**, 1.
4. Lindstedt, R., Apodaca, G., Barondes, S.H., Mostove, and Leffler, H. (1993) *J. Biol. Chem.* **268**, 11750.
5. Moutsatsos, I.K., Wade, M., Schindler, M., and Wang, J.L. (1987). *Proc. Natl. Acad. Sci. USA* **84**, 6452.
6. Vyakarnam, A., Lenneman, A.J., Lakkides, K.M., Patterson, R.J., and Wang, J.L. (1998) *Exp. Cell Res.* **242**, 419.
7. Cowles, E.A., Moutsatsos, I.K., Wang, J.L., and Anderson, R.L. (1989) *Exp. Geront.* **24**, 577.
8. Vyakarnam, A., Dagher, S.F., Wang, J.L., and Patterson, R.J. (1997) *Mol.Cell Biol.* **17**, 4730.
9. Dagher, S.F., Wang, J.L., and Patterson, R.J. (1995). *Proc. Natl. Acad. Sci.USA* **92**, 1213.
10. Menon, R.P., Strom, M., and Hughes, R.C. (2000). *FEBS Lett.* **470**, 227.
11. Cowles, E.A. Agrwal, N., Anderson, R.L., and Wang, J.L. (1990). *J. Biol. Chem* **265**, 17706.
12. Tsay, Y.-G., Lin, N.Y., Voss, P.G., Patterson, R.J., and Wang, J.L. (1999). *Exp. Cell Res.* **252**, 250.
13. Laing, J.G., and Wang, J.L. (1988). *Biochemistry* **27**, 5329.

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

15. Spector, D.L. (1996). *Exp. Cell Res.* **229**, 189.

16. Park, J.W., Voss, P.G., Grabski, S., Wang, J.L., and Patterson, R.J. (2001). *Nucleic Acids Res.*, in press.

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

18. Yang, R.Y., Hsu, D.K., and Liu, F.-T. (1996). *Proc. Natl. Acad. Sci. USA* **93**. 6737.

19. Hsu, D.K., Yang, R.Y., Pan, Z., Yu, L., Salomon, D.R., Fung-Leung, W.-P., and Liu, F.-T. (2000). *Am. J. Pathol.* **156**, 1073.

20. Colnot, C., Ripoche, M.-A., Milon, G., Montagutelli, X., Crocker, P.R., and Poirier, F. (1998). *Immunology* **94**, 290.

21. Nakielny, S., and Dreyfuss, G. (1999) *Cell* **99**, 677.

22. Melchior, F., and Gerace, L. (1998). *Trends Cell Biol.* **8**, 175.

23. Moore, M.S. (1998). *J. Biol. Chem.* **273**, 22857.

24. Mahajan, R., Gerace, L. and Melchior, F. (1998). *J. Cell Biol.* **140**, 259.

25. Matunis, M.J., Wu, J., and Blobel, G. (1998). *J. Cell Biol.* **140**, 499.

26. Ohtsubo, M., Okazaki, H., and Nishimoto, T. (1989). *J. Cell Biol.* **109**, 1389.

27. Kutay, U., Bischoff, F.R.,Kostka, S., Kraft, R., and Gorlich, D. (1997). *Cell* **90**, 1061.

28. Gorlich, D. (1998). *EMBO J.* **17**, 2721.

29. Kehlenbach, R.H., Dickmanns, A., Kehlenbach, A., Guan, T., and Gerace, L. (1999). *J. Cell Biol.* **145**, 645.

30. U. Fischer, Jochen H., Wilbert, C. B., Iain, W. M., and Reinhard, L. (1995). *Cell* **82**, 475.

31. Barbara, E. M., Judy, L. M., and Michael, H. M.. (1996). *J. Virol.* **70**, 2350.

32. Katrin, S., Charleen, S. F., Christine, G., and Karsten, W. (1997). *Cell* **90**. 1041.

33. Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997). *Cell* **90**. 1051.

34. Wen, W., Meinkoth, J. L., Tsien, R.Y., and Taylor, S. S. (1995). *Cell* **82**, 463.

35. Fritz, C.C., and Green, M.R. (1996). *Curr. Biol.* **6**, 848.

36. Makoto, F., Isamu, G., Yukiko, G., and Eisuke, N. (1996). *J. Biol. Chem.* **271**, 20024.

37. Richards, S. A., Lounsbury, K.M., Carey, K.L. and Macara, I.G. (1996). *J. Cell Biol.* **134**, 1157.

38. Fridell, R.A., Fischer, U., Luhrmann, R., Meyer, B.E., Meinkoth, J.L., Malim, M.H., and Cullen, B.R. (1996). *Proc. Natl Acad. Sci. USA* **93**, 2936.

39. Murphy, R., Watkins, J.L., and Wente, S.R. (1996). *Mol. Biol. Cell* **7**, 1921.

40. Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Luhrmann, R., and Hurt, E. (1997). *EMBO J.* **16**, 3256.

41. Gorlich, D., and Mattaj, I.W. (1996). *Science* **271**, 1513.

42. Nigg, E.A. (1997). *Nature* **386**, 779.

43. Bataille, N., Helser, T., and Fried, H.M. (1990). *J. Cell Biol.* **111**, 1571.

44. Jarmolowski, A., Boelens, W.C., Izaurralde, E., and Mattaj, I.W. (1994). *J. Cell Biol.* **124**, 627.

45. Pokrywaka, N.J., and Goldfarb, D.S. (1995). *J. Biol. Chem.* **270**, 3619.

46. Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. (1995). *Nature* **376**, 709.

47. Guddat, U., Bakken, A.H., and Pieler, T. (1990). *Cell* **60**, 619.

48. Mattaj, I.W., and Englmeier, L. (1998). *Annu. Rev. Biochem.* **67**, 265.

49. Arts, G.J., Kuersten, S., Romby, P., Ehresmann, B., and Mattaj, I.W. (1998). *EMBO J.* **17**, 7430.

50. Kutay, U., Lipowsky, G., Izaurralde, E., Bischoff, F.R., Schwarzmaier, P., Hartmann, E., and Gorlich, D. (1998). *Mol. Cell* **1**, 359.

51. Sarkar, S., and Hopper, A.K. (1998). *Mol. Biol. Cell* **9**, 3041.

52. Lund, E., and Dahlberg, J.E. (1998). *Science* **282**, 2082.

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

54. Pinol-Roma, S., and Dreyfuss, G. (1992). *Nature* **355**, 730.

55. Krecic, A.M., and Swanson, M.S. (1999). *Curr. Opin. Cell Biol.* **11**, 363.

56. Lee, M.S., Henry, M., and Silver, P.A. (1996). *Genes Dev.* **10**, 1233.

57. Michael, W.M., Paul, S.E., and Dreyfuss, G. (1997) *EMBO J.* **16**, 3587.

58. Pritchard, C.E., Fornerod, M., Kasper, L.H., and van Deursen, J.M. (1999). *J. Cell Biol.* **145**, 237.

59. Tseng, S.S., Weaver, P.L., Liu, Y., Hitomi, M., Tartakoff, A.M., and Chang, T.H. (1998). *EMBO J.* **17**, 2651.

60. Schmitt, C., von Kobbe, C., Bachi, A., Pante, N., Rodrigues, J.P., Boscheron, C., Rigaut, G., Wilm, M., Seraphin, B., Carmo-Fonseca, M., and Izaurralde, E. (1999). *EMBO J.* **18**, 4332.

Chapter 2

Analysis of a Putative Leucine-rich Nuclear Export Signal in Galectin-3

INTRODUCTION

Galectin-3 (Gal3) is a member of a family of galactose-specific carbohydrate-binding proteins found in a variety of cell types (see reference [1] for a review). It is predominantly an intracellular protein, being found in both the cytoplasm and nucleus of cells [2]. The nuclear localization of Gal3 was sensitive to ribonuclease treatment of permeabilized cells, prior to their fixation for immunofluorescence analysis [3]. Moreover, sedimentation of nucleoplasm over cesium sulfate density gradients identified Gal3 in fractions with densities corresponding to those reported for heterogeneous nuclear ribonucleoprotein complex (hnRNP) and small nuclear RNPs (snRNP). Because these RNPs play important roles in the nuclear processing of pre-mRNA, the possibility was raised that Gal3 was a splicing factor as well. Indeed, using a cell-free assay, depletion and reconstitution experiments documented that Gal3 and another member of the galectin family, galectin-1 (Gal1), were redundant but required factors in the splicing of pre-mRNA [4, 5].

More recently, we have found that Gal3, as well as Gal1, interact with Gemin4, which has been characterized as one of approximately 15 polypeptides of a macromolecular complex, designated as the SMN complex [6]. Co-immunoprecipitation experiments established that Gal1 and Gal3 are *bona fide* members of the SMN complex[1]. The SMN complex is found in both the nucleus and in the cytoplasm. In the cytoplasm, the SMN complex is involved in the biogenesis of snRNPs [7], prior to their entry into the nucleus to function as required components in the splicing of pre-mRNA. In the nucleus, the SMN complex is localized in discrete bodies called Gems [8]. Here,

the role

interme

juncture

galectin

protein

might

both n

mouse

gene

shuttl

(NLS

proce

the role of the SMN complex is to "rejuvenate" the snRNPs and supply them to an intermediate of spliceosome assembly known as the H-complex [9]. This H-complex juncture is also where Gal1 and Gal3 are required, as demonstrated by the effect of galectin depletion on spliceosome assembly [4].

The association of Gal3 with the SMN complex raises the possibility that the protein might perform related functions in both the nucleus and the cytoplasm and that it might shuttle between the two compartments. Indeed, analysis of Gal3 localization in both nuclei of heterodikaryons, derived from fusion of a Gal3 expressing cell (e.g. mouse 3T3 fibroblasts) with a Gal3 null cell (e.g. fibroblasts from mice in which the Gal3 gene has been inactivated), provided definitive evidence for nucleo-cytoplasmic shuttling[2]. The goal of the present study was to identify the nuclear localization signal (NLS) and nuclear export signal (NES) in the Gal3 polypeptide that mediate this shuttling process.

MATERIALS AND METHODS

Cell culture and reagents NIH mouse 3T3 fibroblasts were obtained from the American Type Culture Collection (Rockville, MD). The cells were grown as monolayers in Dulbecco's modified Eagle's medium containing 10 % calf serum, 100 U/ml penicillin, and 100 µg/ml streptomycin at 37 °C in a humidified atmosphere of 10% CO_2. For transfection with vectors expressing fusion proteins containing the Green Fluorescent Protein (GFP) reporter group, cells were seeded in 2-well Lab-Tek Chamber slides (Nalge Nunc International, Naperville, IL) at a density of 2 x 10^4 cells/cm^2 one day before transfection. The cells were grown to ~80% confluency. Transfections were carried out using 1 µg of DNA and 3 µl of lipofectamine (2 mg/ml) following the manufacturer's directions (Life Technologies).

At 9 hours post transfection, either actinomycin D (ActD) and cycloheximide or leptomycin B (LMB) and cycloheximide were added to half of the samples. The other half served as controls. After 5 hours of treatment (14 hours post transfection), the cells were observed under the fluorescence microscope. ActD was purchased from Sigma and was dissolved in H_2O as a 1 mg/ml stock solution and stored at –20 °C. It was added to cultures at a final concentration of 5 µg/ml. Cycloheximide (Boehringer Mannheim) was dissolved directly in culture medium at a concentration of 200 µg/ml and was added to cultures at a final concentration of 10 µg/ml. LMB was a gift of Dr. Minoru Yoshida (University of Tokyo, Japan) (10,11). Purified LMB was dissolved in ethanol as a 10 µg/ml stock solution and stored at -20 °C. It was diluted in culture medium and then added to cultures at a final concentration of 4 ng/ml.

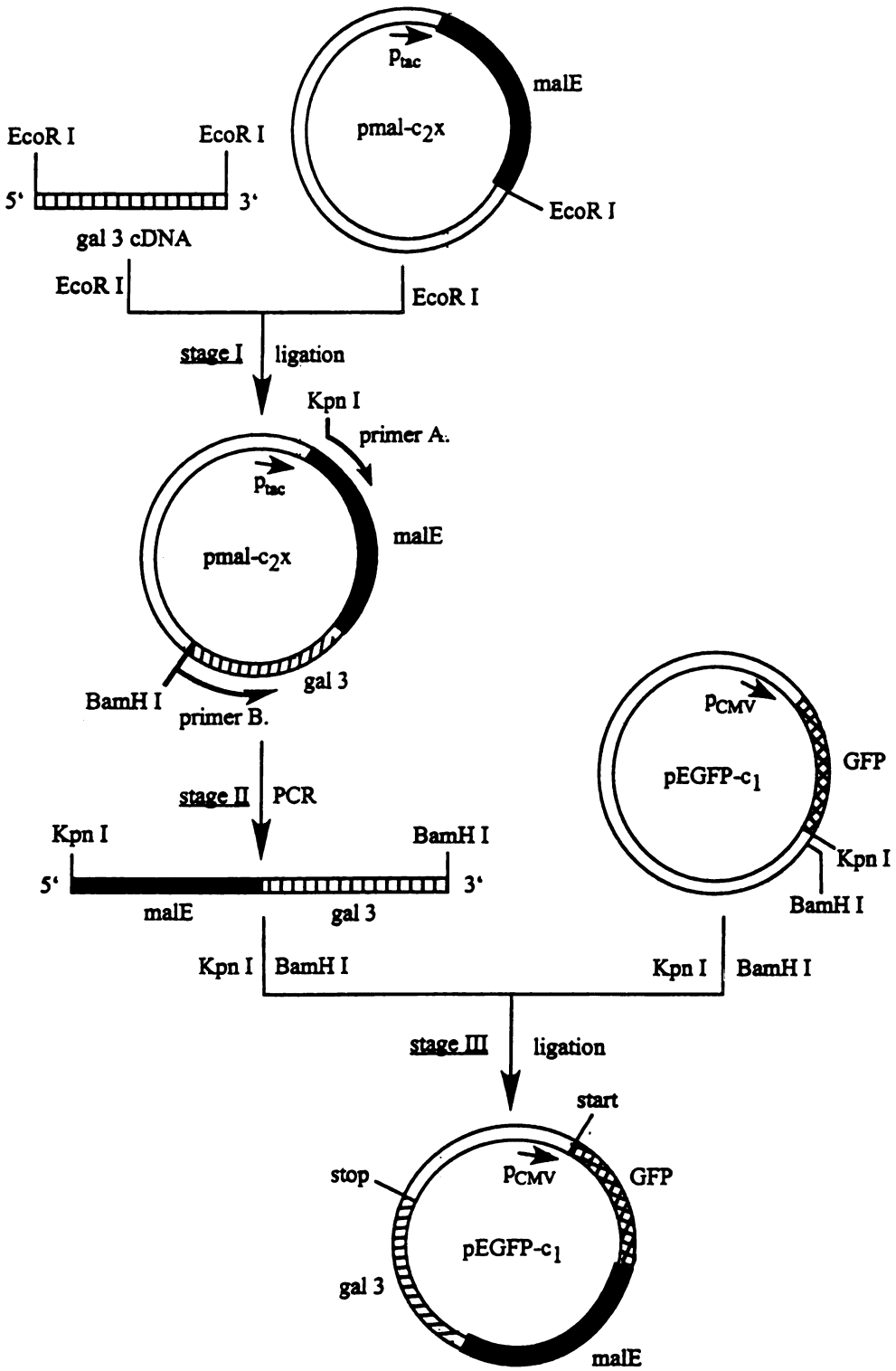
Preparation of the pEGFP-c_1 vector for the expression of fusion protein GFP-MalE-Gal3 The construction of the vector for the expression of the fusion protein GFP-MalE-Gal3 in mammalian cells is summarized in Figure 1. In stage I, the cDNA for Gal3 was liberated from plasmid pWJ31 (12) by EcoRI digestion and then inserted into the corresponding restriction site of the bacterial expression vector pmal-c_2x (New England Biolabs). In stage II, this plasmid encoding the fusion protein MalE-Gal3 was used as the template for polymerase chain reaction (PCR) amplification, using primers:
5'-CGG GGT ACC ATG AAA ATC GAA GAA GGT AAA C-3' (which generates the KpnI restriction site not on the template)
5'- AGG TCG ACT CTA GAG GAT C-3' (which reproduces the BamHI site on the template).

In stage III, this PCR product was ligated into the mammalian expression vector, pEGFP-c_1 (Clontech). The expression of the fusion protein GFP-MalE-Gal3 in transfected cells is driven by a cytomegalovirus promoter (Fig. 1).

The vectors for the production of GFP-Gal3 and GFP-MalE were prepared from the respective cDNAs in a similar fashion, using the same primers and taking advantage of the same restriction sites.

To generate the L247A and I249A mutant of GFP-MalE-Gal3, site-directed mutagenesis was carried out with the QuikChange Site-Directed Mutagenesis Kit (Stratagene), using the vector for GFP-MalE-Gal3 as template and primers:
5'-CGG GAA ATC AGC CAA GCG GGG GCC AGT GGT GAC ATA ACC-3'
5'-GGT TAT GTC ACC ACT GGC CCC CGC TTG GCT GAT TTC CCG-3'

Figure 1: Schematic diagram illustrating the construction of the vector for the expression of the fusion protein GFP-MalE-Gal3 in mammalian cells. In stage I, the cDNA for Gal3 was inserted into the EcoRI restriction site of the bacterial expression vector pmal-c_2x. After ascertaining that this vector expressed the desired fusion protein MalE-Gal3, it was used, in stage II, as template for PCR amplification of the fragment coding for MalE-Gal3, with a KpnI site at the 5'-end and BamHI site at the 3'-end. In stage III, this fragment was ligated into the mammalian expression vector, pEGFP-c_1. The expression of the fusion protein GFP-MalE-Gal3 in mammalian cells is driven by a cytomegalovirus promoter.


EcoR I
EcoR I
5'
3'
gal 3 cDNA
P_{tac}
malE
pmal-c_2x
EcoR I
EcoR I
EcoR I
stage I
ligation
Kpn I
primer A.
P_{tac}
malE
pmal-c_2x
gal 3
BamH I
primer B.
stage II
PCR
Kpn I
BamH I
5'
3'
malE
gal 3
P_{CMV}
GFP
pEGFP-c_1
Kpn I
BamH I
Kpn I
BamH I
Kpn I
BamH I
stage III
ligation
start
stop
P_{CMV}
GFP
pEGFP-c_1
gal 3
malE

Similarly, site-directed mutagenesis was carried out to put stop codons at positions 163, 232, 253, 258, 259, 260, 261, 262, and 263 of the Gal3 sequence.

The pRev(1.4)-GFP vector and variants The pRev(1.4)-GFP vector and its application for testing potential NES sequences were developed by Henderson and Eleftheriou (13). Two previously identified NES sequences were also used in our study as controls: (a) the NES of the inhibitor of cAMP-dependent protein kinase (PKI) was characterized as a strong NES; (b) the NES of IκB-α represented weak NES activity. Each of these sequences was cloned as short fragments between the BamHI and the AgeI sites of pRev(1.4)-GFP, sandwiched between the Rev and the GFP coding sequences. We obtained all three of the above vectors from Dr. Beric Henderson (University of Sydney, Australia).

The vector pRev(1.4)-GFP containing the putative NES sequence of Gal3 (residues 240-255) was derived from the pRev(1.4)-GFP vector containing the NES of PKI by site-directed mutagenesis in five steps, each of which changed multiple amino acids. Finally, two mutants of the Gal3 NES, designated Gal3 NES (I244A; L247A) and Gal3 NES (L247A; I249A) were derived from the wild-type Gal3 NES in the pRev(1.4)-GFP vector by site-directed mutagenesis. All of these experiments used the QuikChange Site-Directed Mutagenesis Kit of Stratagene.

Fluorescence microscopy Transfected cells were examined by fluorescence microscopy directly live in the Lab-Tek Chamber slides, using a Meridian Instruments (Okemos, MI) Insight confocal laser scanning microscope. The cells were counted and

scored for GFP localization: (a) N, fluorescence exclusively in the nucleus; (b) N+C, fluorescence in both the nucleus and cytoplasm; and (c) C, fluorescence exclusively in the cytoplasm. Representative cells were photographed at low (66X) magnification to show a field containing several cells and at high (200X) magnification to show a single cell.

SDS-PAGE and immunoblotting Proteins were resolved on SDS-PAGE (10% acrylamide) as described by Laemmli (14). The procedures for immunoblotting after SDS-PAGE have also been described (15). The antibodies used for immunoblotting and their sources were: (a) polyclonal anti-GFP (Clontech); (b) anti-MalE (New England Biolabs); and (c) polyclonal anti-Gal3 (#32 and #33, see reference [16]).

RESULTS

A GFP reporter construct for the localization of galectin-3 In order to define the NLS and NES of galectin-3, we have developed a reporter construct expressing a fusion protein containing galectin-3 and GFP. This fusion protein also contains bacterial maltose-binding protein (MalE) to serve as a "spacer" that increases the molecular weight of the reporter polypeptide. This was done to insure that the size of the reporter polypeptide would exceed the exclusion limit of nuclear pores (~40 kD), even when the portion of the galectin-3 polypeptide is decreased through deletion mutagenesis.

The cDNA for galectin-3 was digested with EcoRI and ligated into the corresponding restriction site in the prokaryotic expression vector pMAL-c_2x (Fig. 1). The success of this step was indicated by: (a) the same MalE-Gal3 fusion protein (M_r ~ 74,000) could be detected in bacterial lysates by immunoblotting with either anti-MalE or with anti-Gal3; (b) the MalE-Gal3 fusion protein could be isolated on lactose affinity columns. This plasmid then served as the template for PCR amplification of the coding region corresponding to MalE-Gal3 and the product was then ligated into the eukaryotic expression vector, pEGFP-C1 (Fig. 1). This initial construct and its variants were used to transfect mouse 3T3 fibroblasts and the expression of the fusion protein was driven by a cytomegalovirus promoter.

The structure of the wild-type fusion protein, containing the full-length polypeptides of GFP (27 kD), MalE (40 kD), and Gal3 (33 kD), is schematically shown as Construct 4 in Figure 2. Various mutants were derived from this initial construct by site-directed mutagenesis to introduce stop codons to shorten the expressed polypeptide.

Figure 2: Summary of the properties characterized for the fusion protein GFP-MalE-Gal3 and variants. The structure of the polypeptide encoded by each transfected DNA is depicted under the column designated Construct. The numbers below each construct indicate the amino acid residues of the Gal3 polypeptide included in the fusion protein. In Construct 10, site-directed mutagenesis was carried out on residues 247 and 249 of the Gal3 sequence, changing leucine to alanine and isoleucine to alanine, respectively. The DNAs of most constructs were subjected to sequence analysis: OK indicates that the sequence corresponded to the structure depicted for the construct; NT indicates that the DNA was not subjected to sequence analysis (not tested). The expected molecular weight (MW) was calculated on the basis of : GFP (27 kD); MalE (40 kD); and Gal3 (33 kD). The molecular weights of the expressed fusion proteins were determined by immunoblotting of cell extracts with anti-GFP, anti-MalE, and anti-Gal3. Fluorescence patterns of the GFP reporter: N=C denotes the fluorescence was observed in both the nuclear and cytoplasmic compartments, as exemplified in Figure 3, panel A; C denotes the fluorescence observed exclusively in the cytoplasm, as exemplified in Figure 3, panel B.

Construct	DNA Sequence	Expected MW	Western blot MW	Fluorescence Pattern
1. GFP	NT	27 K	27 K	N = C
2. GFP - Gal 3 (1–263)	OK	60 K	NT	N = C
3. GFP - MalE	NT	67 K	NT	N = C
4. GFP - MalE - Gal 3 (1–263)	NT	100 K	100 K	N = C
5. GFP - MalE - Gal 3 (1–258)	OK	99 K	NT	60% C > N 40% N = C
6. GFP - MalE - Gal 3 (1–257)	OK	99 K	NT	C
7. GFP - MalE - Gal 3 (1–252)	OK	99 K	99 K	C
8. GFP - MalE - Gal 3 (1–231)	OK	96 K	96 K	C
9. GFP - MalE - Gal 3 (1–162)	**no mutation**	87 K	100 K	N = C
10. GFP - MalE - Gal 3 (1–263, L247A, I249A)	OK	100 K	NT	C

In each case, the mutations were confirmed by DNA sequence analysis. In addition, several control vectors, GFP, GFP-Gal3 (no MalE), and GFP-MalE (no Gal3) were prepared (Constructs 1-3 in Figure 2). Unless specifically noted as NT (not tested), parallel cultures were transfected and processed for: (a) immunoblotting (by anti-GFP, anti-MalE, and anti-Gal3) to check the molecular weight and integrity of the reporter polypeptide; and (b) fluorescence microscopy to record the nuclear versus cytoplasmic distribution.

The COOH-terminal region of Gal3 is required for nuclear localization

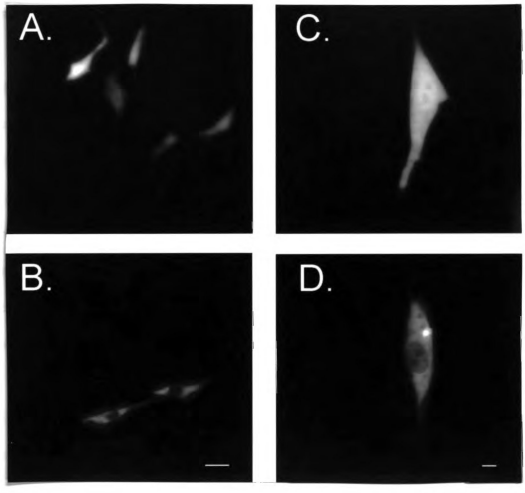
Transfection of 3T3 cells with GFP (Construct 1 in Fig. 2) resulted in the expression of a 27 kD polypeptide and fluorescence in both the nuclear and cytoplasmic compartments. Examples of what we have designated as the N=C fluorescence pattern are shown in Figure 3 (panels A and C). Because the molecular weight of GFP is below the exclusion limit of the nuclear pores, the observed nuclear and cytoplasmic localization is rationalized in terms of its ability to diffuse in and out of the nucleus (17). More surprising, however, was the observation that a fusion protein containing GFP and MalE, but no Gal3, also showed N=C fluorescence (Construct 3 in Fig. 2). The size of the fusion protein (67 kD) should have restricted it to the cytoplasm, in the absence of a NLS. This represented one complication in our analysis.

Transfections with Construct 2 (GFP-Gal3; 60 kD) and Construct 4 (GFP-MalE-Gal3; 100 kD) both yielded N=C fluorescence (Fig. 2). Because of the above complication due to the N=C localization observed with GFP-MalE, however, we could not conclude, from Constructs 2 and 4, that the Gal3 polypeptide contained a NLS. The

presence of a NLS and its approximate location in the Gal3 polypeptide came out of analysis of truncation mutants, in which a stop codon was put at specific residues to terminate the polypeptide. Transfection with Construct 8 (GFP-MalE-Gal3 (1-231); 96 kD) yielded an exclusively cytoplasmic (C) fluorescence pattern, as illustrated in Fig. 3 (panels B and D). This suggested that the carboxyl-terminal 30 amino acids of the Gal3 polypeptide contained information which allowed GFP-MalE-Gal3 (1-263) to localize to the nuclear compartment. Starting at residue 263, we put stop codons in individual mutants (263, 262, 261, 260, 259, 258, 257, etc). These mutants have narrowed the putative NLS-containing boundary to residue 258 of the polypeptide; deletion of six or more amino acids from the carboxy terminus results in an exclusively cytoplasmic localization (Constructs 5 and 6, Fig. 2).

In the course of these analyses, we had intended to engineer Construct 9 (GFP-MalE-Gal3 (1-162)), by site-directed mutagenesis of residue 163 into a stop codon. When this construct was transfected into target cells, we observed, quite surprisingly, the N=C fluorescence pattern (Fig. 2). In light of the cytoplasmic fluorescence patterns obtained with Constructs 6-8, this result of Construct 9 seemed confounding, if the putative NLS indeed resided in the carboxy terminus of the Gal3 polypeptide. Immunoblotting of extracts derived from transfected cultures yielded the reporter polypeptide at 100 kD, rather than the expected 87 kD. These apparent discrepancies were finally resolved when DNA sequence analysis showed that our site-directed mutagenesis had failed to insert the stop codon at residue 163 (highlighted in bold, Fig. 2) and, as a result, the wild-type reporter polypeptide (corresponding to Construct 4) was actually expressed. Although we did not accomplish the original intent of Construct 9,

Figure 3: Representative fluorescence micrographs illustrating the N=C versus C labeling patterns. The N=C fluorescence patterns shown here are derived from 3T3 fibroblasts transfected with Construct 1, GFP: panel A, low magnification image showing a field containing several cells; panel C, high magnification image showing a single cell. The C fluorescence patterns shown here are derived from 3T3 cells transfected with Construct 8, GFP-MalE-Gal3 (1-231): panel B, low magnification image showing a field containing several cells (bar=50 μm); panel D, high magnification image showing a single cell (bar=10 μm).


A.
C.
B.
D.

this experiment did emphasize the importance of checking the DNA sequence of the mutant constructs and the size of the expected reporter polypeptides. Moreover, this experience also lent a considerable amount of confidence in our assignment of the fluorescence patterns (N=C versus C). [I scored this N=C because it was truly N=C, rather than the expected pattern of C.]

An attempt to identify the NES using the GFP-MalE-Gal3 reporter Previous studies had documented that LMB inhibits the export of Gal3 from the nucleus, concentrating the protein in the nuclear compartment [15, 18]. This suggested that nuclear export of Gal3 was mediated by the CRM1 exportin, which recognizes leucine-rich NES [19]. Indeed, a putative leucine-rich NES, with the requisite spacing of leucine/isoleucine, can be identified between residues 241 and 249 of the murine Gal3 sequence. Moreover, the putative NES motif appears to be conserved in the Gal3 homologs of various species (Fig. 4).

On this basis, we wanted to test the effect of mutagenizing two key residues in the putative NES: leucine 247 to alanine (L247A) and isoleucine 249 to alanine (I249A) (Construct 10, Fig. 2). These two residues were chosen for mutagenesis because they occupy corresponding positions that had been shown to be critical for the functioning of the leucine-rich NES in PKI [20]. If the putative NES was indeed functional in CRM1-mediated nuclear export, we would expect that the fusion protein expressed by Construct 10 to exhibit a nuclear localization. Transfection of 3T3 cells with Construct 10 resulted, however, in an exclusively cytoplasmic fluorescence pattern. DNA sequence analysis confirmed that the mutations had been correctly carried out. We interpret the results to

Figure 4: Comparison of the leucine-rich nuclear export signals (NES) identified in in protein kinase A inhibitor (PKI) and in HIV-1 Rev protein with the putative NES found in the sequences of Gal3 from various species.

Nuclear Export Signals

Protein Kinase A Inhibitor	L	A	L	K	L	A	G	L	D	I
Rev (HIV-1)	L	P		P	L	E	R	L	T	L

Putative NES in Galectin-3

Mouse	L	R		E	I	S	Q	L	G	I
Rat	L	R		E	I	S	Q	L	G	I
Human	L	N		E	I	S	K	L	G	I
Hamster	L	R		E	I	N	Q	M	E	I
Rabbit	L	K		E	I	N	K	L	G	I

indicate that this stretch of the Gal3 sequence was also important for a functional NLS and our mutagenesis on residues 247 and 249 inactivated the nuclear import signal. This notion is consistent with the results obtained with Constructs 6-8, which implicated the carboxyl terminal 30 amino acids as necessary for nuclear import.

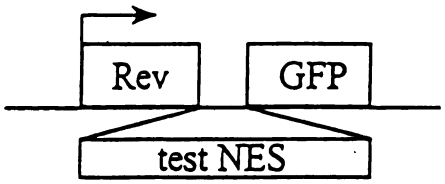
The Rev(1.4)-GFP vector for the analysis of a functional NES The apparent overlap of the NLS and NES in the Gal3 polypeptide precluded us from using the GFP-MalE-Gal3 reporter construct to define the NES. Mutations intended to inactivate the putative NES also inactivated the NLS and therefore, the properties of the NES cannot be studied in a protein that fails to enter the nucleus. To circumvent these difficulties, we have taken advantage of the pRev(1.4)-GFP vector developed by Henderson and Eleftheriou [13]. Although the HIV-1 Rev protein normally contains both a NLS as well as a NES, the Rev(1.4) variant is NES-deficient. Instead, test sequences representing a putative NES can be cloned in frame between the Rev(1.4) segment and the GFP reporter (see schematic in Fig. 5A). The fusion protein expressed by this vector contains a NLS of Rev, which could be "inactivated" by treatment of cells with ActD [21, 22]. This allows the activity of very weak NESs to be detected. Thus, the relative activity of different NESs can be distinguished by their ability to shift the fusion protein to the cytoplasm, both in the presence or absence of active nuclear import (i.e., in the absence or presence of ActD, respectively). If the test NES is recognized by the CRM1 exportin, then nuclear export is expected to be sensitive to LMB inhibition [10, 11]. On this basis, ActD and LMB will play critical roles in our dissection of the NLS-based nuclear import and the CRM1-mediated nuclear export (Fig. 5B).

Figure 5: (A) Schematic diagram of the pRev(1.4)-GFP vector for testing potential nuclear export signals (NES). A potential NES is cloned into this vector for the expression of a fusion protein containing Rev(1.4)-test sequence-GFP. Normal Rev has both a nuclear localization signal (NLS) and a NES; the Rev(1.4) variant is NES deficient. The expressed fusion protein uses the NLS of Rev to import the reporter into the nucleus. The activity of the potential NES to export the reporter is determined by the nuclear and/or cytoplasmic distribution of fluorescence due to GFP.

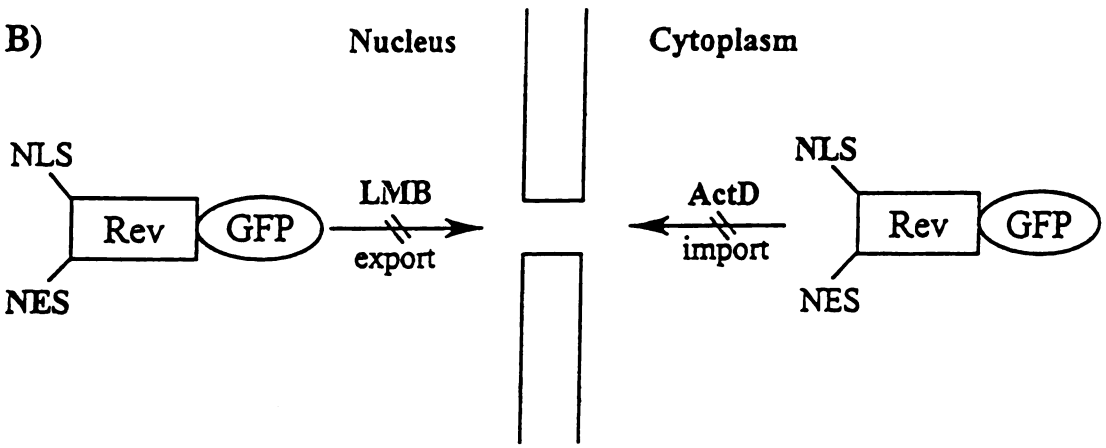
(B) Schematic diagram illustrating the basis of the assay to determine the nuclear export activity of a potential NES. The fusion protein contains a NLS derived from Rev, a test sequence for NES activity, and the GFP reporter. The NLS of the fusion protein is responsible for nuclear import and is highlighted by bold letters in the cytoplasm; this NLS-mediated import is specifically blocked by treatment of cells with actinomycin D (ActD). The NES of the fusion protein is responsible for nuclear export and is highlighted by bold letters in the nucleus; nuclear export mediated by leucine-rich NES is specifically blocked by leptomycin B (LMB).

(C) Summary of the contents of the fusion proteins expressed by various constructs. The construct designated as GFP is simply the mammalian expression vector for the production of GFP. The construct designated as Rev1.4 expresses a NES-deficient variant of Rev as a fusion protein with GFP. The construct designated as PKI NES expresses the Rev(1.4)-GFP fusion protein containing the specific NES sequence shown, derived from PKI. Similarly, the other constructs express the Rev(1.4)-GFP fusion protein containing the specific test NES sequence shown.

A) pRev(1.4)-GFP



B)



C)

Construct	Test NES sequence	comment
GFP		GFP-N1(No NLS; No NES)
Rev1.4	NONE	NLS but No NES
PKI NES	33-N S N E L A L K L A G L D I N K T E	residues 33-50 of PKI
IκB-α NES	261-P S T R I Q Q Q L G Q L T L E N L Q	residues 261-278 of IκB-α
Gal3 NES (wt)	240-N L R E I S Q L G I S G D I T L	residues 240-255 of Gal3
Gal3 NES (I244A; L247A)	A A	residue 244: mutate I to A residue 247: mutate L to A
Gal3 NES (L247A; I249A)	A A	residue 247: mutate L to A residue 249: mutate I to A

The GFP protein contains neither NLS nor NES (Fig. 5C, line 1). As observed previously (Fig. 3, panels A and C), transfection of 3T3 fibroblasts yielded fluorescence in both the nuclear and cytoplasmic compartments; this N+C labeling pattern is illustrated in Figure 6 (panel A). For the Rev(1.4)-GFP fusion protein (M_r ~40,000), we quantitated the number of cells showing an exclusively nuclear (N) fluorescence (Fig. 6, panel B) versus a labeling pattern in which there was fluoresence in both the nuclear and cytoplasmic compartments (N+C). This quantitation is documented in Table I and the data are displayed in the form of histograms in Figure 7. In 80% of the transfected cells, the Rev(1.4)-GFP polypeptide exhibited the N labeling pattern in the absence of ActD (Table I and Fig. 7). Addition of ActD increased the percentage of cells with the N+C labeling pattern, at the expense of the N labeling pattern (Table I and Fig. 7). This is consistent with the notion that the Rev(1.4)-GFP polypeptide contains a NLS that could be inactivated by ActD (Fig. 5C, line 2).

We also tested the effect of inserting test NES sequences whose strengths had been previously characterized (Henderson, 2000). The NES of PKI (Fig. 5C, line 3) was strong. In 100% of the transfected cells, the fluorescence labeling pattern was exclusively cytoplasmic (C) (Fig. 6, panel C). This was true irrespective whether ActD was included in the cultures (Table I and Fig. 7). Thus, the PKI NES was sufficiently strong to overcome an active NLS. The NES of IκB-α could neutralize an active NLS, resulting in all of the cells exhibiting the N+C labeling pattern in the absence of ActD (Fig. 6, panel D; Table I and Fig. 7). In the presence of ActD, some 20% of the cells yielded the exclusively C pattern (Fig. 6, panel E; Table I and Fig. 7), since the NLS has been inactivated.

Figure 6: Representative fluorescence micrographs illustrating the classification of GFP localization patterns: N+C, fluorescence observed in the nucleus and in the cytoplasm; N, fluorescence exclusively in the nucleus; and C, fluorescence exclusively in the cytoplasm. (A): cells expressing the GFP protein, which yielded the N+C fluorescence pattern. (B): cells expressing Rev(1.4)-GFP protein, which yielded the N fluorescence pattern. (C): cells expressing Rev(1.4)-GFP containing the NES of PKI, which yielded the C fluorescence pattern. (D) : cells expressing Rev(1.4)-GFP containing the NES of IκB-α, which yielded the N+C fluorescence pattern. (E): cells, in the presence of actinomycin D (ActD), expressing Rev(1.4)-GFP containing the NES of IκB-α, which yielded the C fluorescence pattern. (F): cells, in the presence of leptomycin B (LMB), expressing Rev(1.4)-GFP containing the NES of IκB-α, which yielded the N fluorescence pattern. Bar=50 μm.

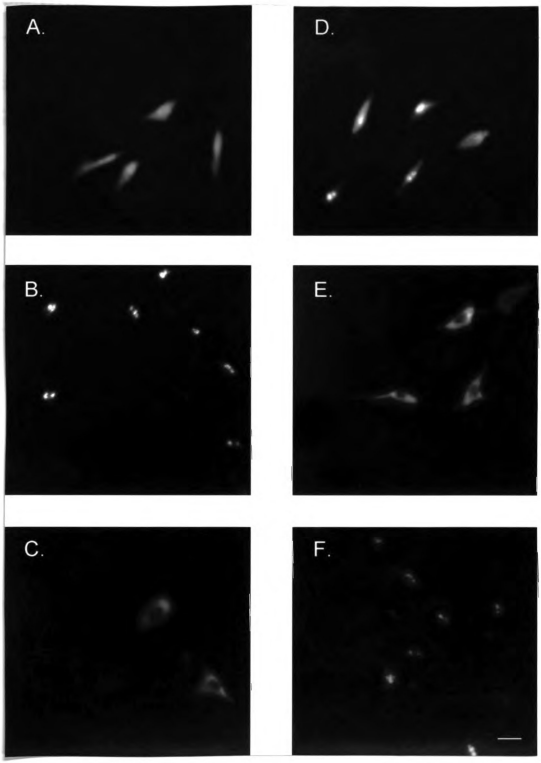
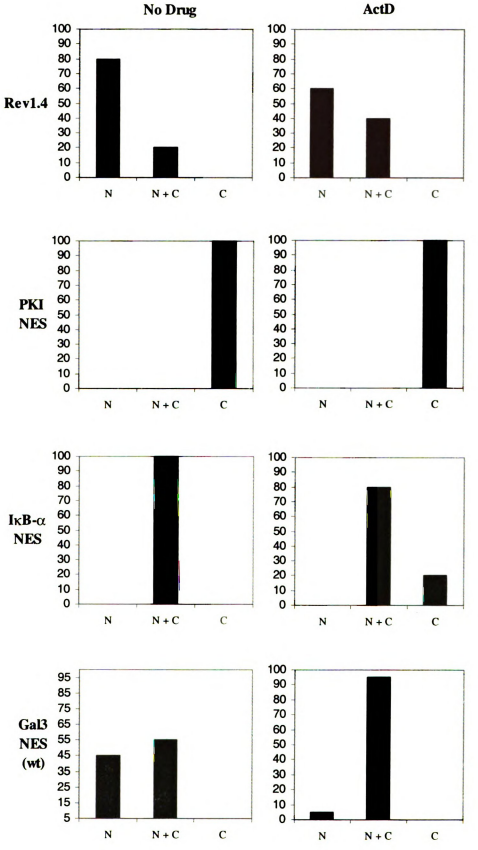

A.
D.
B.
E.
C.
F.

Table 1.The effect of actinomycin D on the localization of GFP fusion proteins

Construct	ActD	Cellular localization of GFP(%) N	N+C	C	Total cells counted
Rev1.4	-	80	20	0	180
	+	60	40	0	177
PKI NES	-	0	0	100	210
	+	0	0	100	198
IκB-α	-	0	100	0	122
	+	0	80	20	110
Gal3 NES(wt)	-	45	55	0	144
	+	5	95	0	120

Figure 7: Display of the data in Table I in the form of histogram distributions of the percent of cells with fluorescence patterns N, N+C, and C.


No Drug
ActD
Rev1.4
PKI NES
IκB-α NES
Gal3 NES (wt)
N
N + C
C

Analysis of the Gal3 NES in the Rev(1.4)-GFP vector When the putative NES of Gal3 was inserted into the Rev(1.4)-GFP vector as the test sequence, about 45% of the transfected cells showed the N labeling pattern while 55% of the cells showed the N+C fluorescence pattern (Fig. 8, panel A). This approximately 40%-60% distribution between the N versus the N+C labeling patterns was reproducible from experiment to experiment (see, for example, Gal3 NES (wt) in Table II and Fig. 9). This distribution should be compared to the corresponding distribution obtained in the transfection with Rev(1.4)-GFP vector (Table I and Fig. 7). There was a higher percentage of cells showing the N+C labeling pattern with the Gal3 NES than with no NES. Nevertheless, the activity of the Gal3 NES appeared weaker than that of the IκB-α NES, neutralizing the effect of active nuclear import in only half of the cells (Table I and Fig. 7).

When nuclear import was inhibited by ActD, virtually all cells transfected with Gal3 NES yielded the N+C fluorescence pattern (Fig. 8, panel B) (Table I and Fig. 7). In some experiments, ~3% of the transfected cells showed a shift of the GFP fluorescence completely to the cytoplasm (see, for example, Fig. 10, Gal3 NES (wt), ActD column). Thus, the NES activity of the Gal3 sequence becomes more apparent when nuclear import is inactivated.

The effect of LMB on the fluorescence distribution The Gal3 NES activity, as reported by the pRev(1.4)-GFP vector, should be sensitive to LMB inhibition, as had been documented for endogenous Gal3 of 3T3 cells [15]. Indeed, incubation with LMB shifts the distribution in favor of exclusively nuclear (N) pattern (Fig. 8, panel C). Not all the cells showed exclusively nuclear fluorescence (Table II and Fig. 9). When 3T3 cells

Figure 8: Representative fluorescence micrographs showing the nuclear versus cytoplasmic distribution of Rev(1.4)-GFP protein containing the putative NES sequence of Gal3 or site-directed mutants. (A): cells expressing Gal3 NES (wt), which yielded the N+C fluorescence pattern. (B): cells, in the presence of actinomycin D (ActD), expressing Gal3 NES (wt), which yielded the N+C fluorescence pattern. (C): cells, in the presence of leptomycin B (LMB), expressing Gal3 NES (wt), which yielded the N fluorescence pattern. (D): cells expressing Gal3 NES (I244A; L247A), which yielded the N fluorescence pattern. (E): cells, in the presence of ActD, expressing Gal3 NES (I244A; L247A), which yielded the N+C fluorescence pattern. (F): cells, in the presence of LMB, expressing Gal3 NES (L247A; I249A), which yielded the N fluorescence pattern. Bar=50 μm.

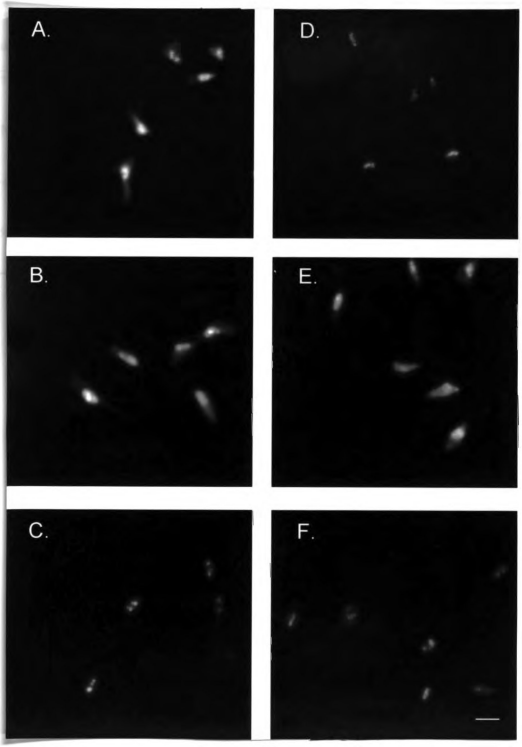

A.
D.
B.
E.
C.
F.

Table 2. The effect of leptomycin B on the localization of GFP fusion proteins

Construct	LMB	Cellular localization of GFP(%) N	N+C	C	Total cells counted
Rev1.4	-	86	14	0	201
	+	71	29	0	126
PKI NES	-	0	0	100	210
	+	68	30	2	198
IκB–α NES	-	2	98	0	230
	+	44	55	1	110
Gal3 NES(wt)	-	36	64	0	211
	+	54	46	0	197

Figure 9: Display of the data in Table II in the form of histogram distributions of the percent of cells with fluorescence patterns N, N+C, and C.

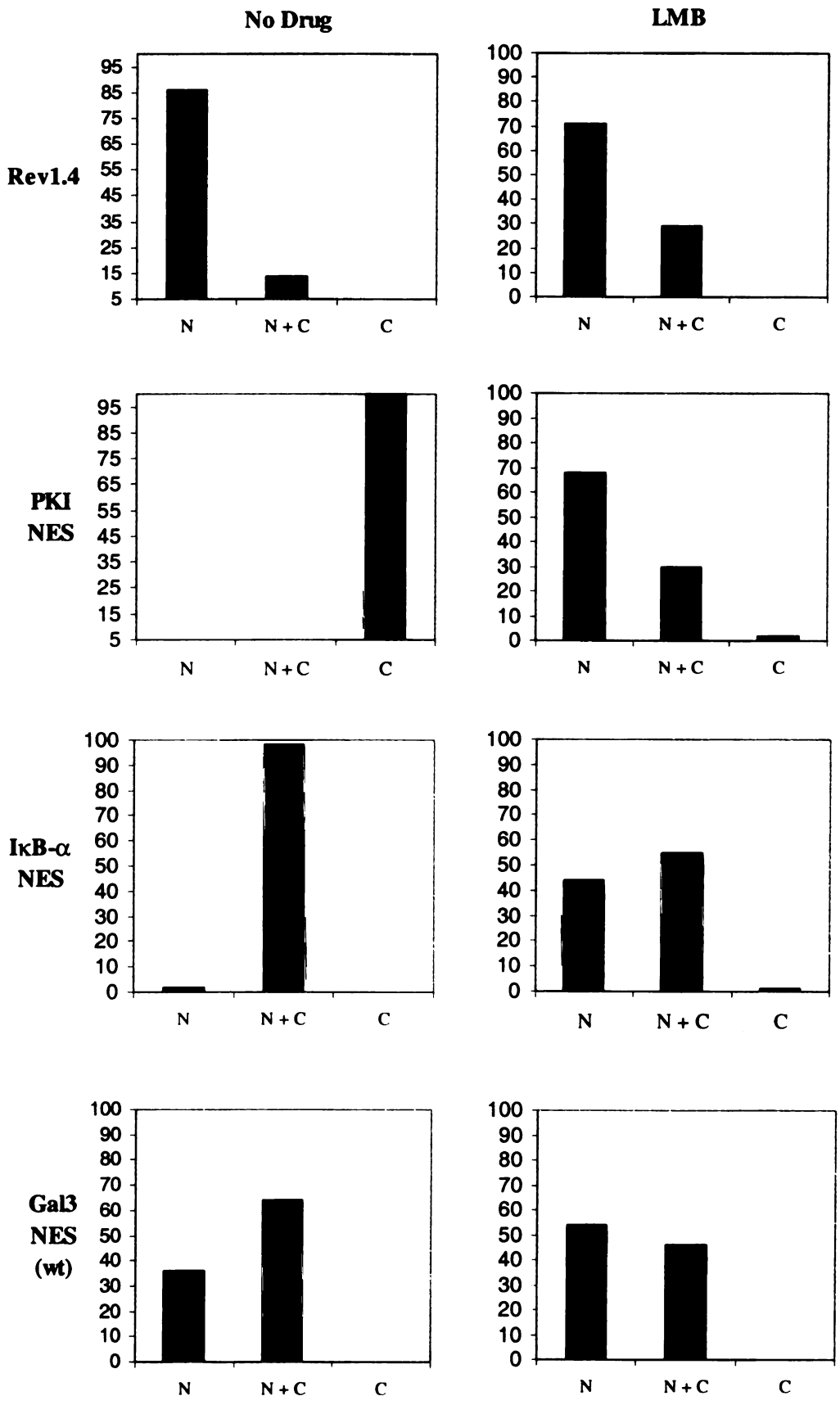
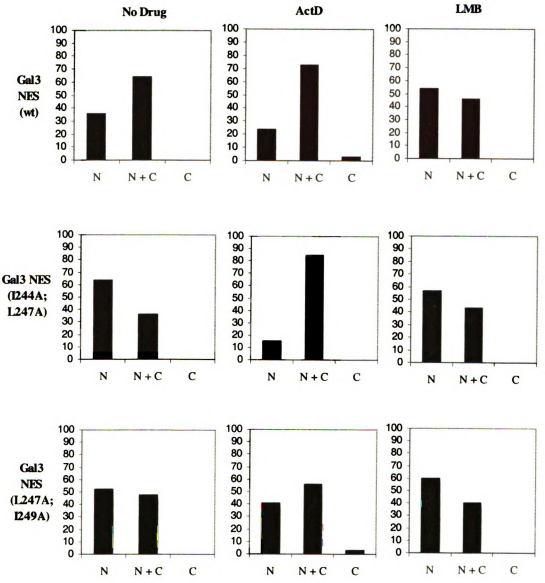

No Drug
LMB
Rev1.4
PKI NES
IκB-α NES
Gal3 NES (wt)
N
N+C
C
0
10
20
30
40
50
60
70
80
90
100
5
15
25
35
45
55
65
75
85
95

Figure 10: A comparison of the histogram distributions of the percent of cells with fluorescence patterns N, N+C, and C for Gal3 NES (wt), Gal3 NES (I244A; L247A), and Gal3 NES (L247A; I249A). In the left column, the transfected cells were cultured in the absence of any drugs. In the middle column, the transfected cells were treated with actinomycin D (ActD). In the right column, the transfected cells were treated with leptomycin B (LMB).


No Drug
ActD
LMB
Gal3 NES (wt)
Gal3 NES (I244A; L247A)
Gal3 NES (L247A; I249A)
100
90
80
70
60
50
40
30
20
10
0
N
N + C
C

were incubated with LMB, they accumulated endogenous Gal3 in the nucleus, as reflected by an accentuation of the nuclear staining [15]; however, there was always some cytoplasmic fluorescence in these LMB treated cells.

LMB also affected the fluorescence of other test NES sequences, shifting the distribution in favor of the exclusively nuclear (N) labeling pattern. In the presence of LMB, about 70% of the PKI NES showed the N labeling pattern while some 30% were N+C. This should be compared with the exclusively cytoplasmic (C) pattern obtained in the absence of the export inhibitor (Table II and Fig. 9). Similarly, LMB shifted the fluorescence distribution from 98% N+C to 44% N (Fig. 6, panel F) and 55% N+C for the IκB-α test sequence (Table II and Fig. 9). In both cases, the effect of LMB was partial; not all of the cells showed an exclusively N labeling pattern.

Finally, the Rev(1.4)-GFP construct contains no NES; therefore, it should not be sensitive to LMB. Consistent with this notion, LMB did not shift the fluorescence distribution of the Rev(1.4)-GFP fusion protein in favor of the N labeling pattern. If anything, the presence of LMB resulted in a shift in the opposite direction, decreasing the percentage of cells showing he exclusively nuclear (N) labeling pattern (Table II and Fig. 9).

Site-directed mutagenesis of the Gal3 NES Site-directed mutagenesis was carried out to generate Gal3 NES (L247A; I249A), the two positions corresponding to critical residues in the leucine-rich NES of PKI [20]. In parallel, the Gal3 NES (I244A; L247A) mutant was also generated. The fluorescence labeling pattern of Gal3 NES (I244A; L247A) and Gal3 NES (wt) were very similar, particularly in terms of the effect of ActD

and LMB. ActD shifted the histogram distributions in favor of N+C, at the expense of exclusively nuclear (N) pattern (Fig. 10). In the presence of LMB, the predominant labeling pattern was exclusively nuclear (N). On the other hand, the histograms of Gal3 NES (L247A; I249A) showed little change upon addition of either ActD or LMB (Fig. 10). If the NES activity were destroyed in this latter mutant, inactivation of nuclear import would not be expected to result in a major shift of fluorescence from the nuclear staining to the cytoplasm. The effect of ActD on Gal3 NES (L247A; I249A) was comparable to the effect of the drug on Rev(1.4)-GFP, which does not carry a NES (compare Fig. 10 with Fig. 7). Similarly, if the leucine-rich NES were destroyed, one would also not expect a dramatic shift toward more nuclear staining upon LMB inhibition of CRM1.

DISCUSSION

The key findings of this study include: (a) The carboxyl terminal region of the murine Gal3 polypeptide is important for nuclear localization, as assayed using a GFP-MalE-Gal3 reporter system. The critical residues appear to lie, at the least, within a region spanning leucine 247 and alanine 258. Truncation of the polypeptide at serine 257 or mutagenesis of leucine 247 and isoleucine 249 results in an exclusively cytoplasmic localization. (b) The segment of the polypeptide, starting at asparagine 240 through leucine 255 exhibits weak CRM1-mediated (LMB-sensitive) nuclear export activity, as assayed in the Rev(1.4)-GFP fusion protein system. The amino acid sequence from leucine 241 through isoleucine 249 corresponds well to a leucine-rich NES and site-directed mutagenesis of leucine 247 and isoleucine 249 affected the nuclear export activity. (c) Inasmuch as mutagenesis of leucine 247 and isoleucine 249 to alanine residues affected both nuclear import and export, the NLS and NES of murine Gal3 appear to overlap in this stretch of amino acids.

The definition of an NLS for Gal3 has been problematic. Two other laboratories reported different regions of the Gal3 polypeptide as being required for its nuclear localization. Gong *et al.* [23] reported that deletion of the first 11 amino acids of Gal3 resulted in a mutant exhibiting cytoplasmic (and no nuclear) localization. Moreover, when the first 11 amino acids were fused to GFP, a predominantly nuclear distribution of the reporter was observed. In contrast to these results, Gaudin *et al.* [24] transfected Cos-7 cells with cDNAs encoding mutants of hamster Gal3 containing amino-terminal or internal deletions and showed that nuclear localization does not require the first 103

amino acid residues. Further deletion of residues 104-110 drastically reduced nuclear localization but the specific sequences between residue 104-110 were not obligatory (these residues could be substituted by unrelated sequences). One caveat is that in both of these studies, the demonstration of the minimal region required for nuclear localization was achieved with polypeptides well below the exclusion limit of nuclear pores, thus allowing diffusion to cloud the issue.

Even more confounding is the fact that our own results point to still a third region of the molecule, at the carboxy end, as being important for nuclear localization. It remains for us to document that a short stretch of the Gal3 sequence (e.g. leucine 247 through alanine 258) can serve as an NLS for a polypeptide above the exclusion limit of nuclear pores. Unfortunately, the use of the GFP-MalE reporter system for such an experiment is complicated by the observation that GFP-MalE yields a nuclear and cytoplasmic localization, despite the fact that its molecular weight (67 kD) should have precluded its passing through nuclear pores in the absence of an NLS. We propose to link longer and shorter variations of this stretch of 12 amino acids (leucine 247 through alanine 258) to a pyruvate kinase reporter [25] to carry out this critical test. If these 12 amino acids fail to target the reporter to the nucleus, it would indicate that this segment of the Gal3 polypeptide is necessary but insufficient for nuclear localization. In this case, our results can at least be reconciled with those of Gaudin *et al.* [24], which have pointed to the importance of residues 104-110, as well as the rest of the carboxyl-terminal domain, for nuclear localization. Finally, it should be noted that the amino acid sequence from leucine 247 through alanine 258 does not match any of the groups of NLSs that have been defined (the basic NLS of SV40 T-antigen, the bipartitie NLS, the M9 domain,

the KNS sequence, etc. [19]). It is also known, however, that although cyclin E does not exhibit any sequences recognizable as one of the defined NLSs, it binds to importin-α and is imported into the nucleus via the importin-β mediated pathway [26].

Henderson and Eleftherious [13] developed the Rev(1.4)-GFP reporter system for testing potential NES sequences. This fusion protein carries an NLS of the HIV-1 Rev protein which could be inactivated by ActD [21, 22]. Thus, each putative NES sequence is challenged to overcome the active NLS in the absence of ActD, resulting in an exclusively cytoplasmic localization of the GFP reporter. Such an NES, classified as a "strong," was found in proteins such as PKI, the mitogen-activated protein kinase kinase (MAPKK), and the c-ABL oncogene [13]. Some test NES sequences display "weak" nuclear export activity. These can partially neutralize the NLS of the Rev(1.4)-GFP reporter, resulting in nuclear and cytoplasmic localization in the absence of ActD. In the presence of ActD, the GFP fusion protein is shifted partially to the cytoplasm in the majority of the cells. "Very weak" NESs cannot normally overcome the rate of Rev NLS-mediated nuclear import in the absence of ActD but are able to shift the GFP fluorescence partially to the cytoplasm in 20-50% of the cells when import is blocked by ActD. The tumor suppressor p53 and its hdm2 regulator each has an NES that fit this latter category.

By the criteria established in the development of the Rev(1.4)-GFP test vector [13], the NES of Gal3 (residues 240-255 tested) would fall in the "weak" category. This weak NES activity may be important for the nuclear function of the protein. A strong NES might result in futile shuttling of Gal3 between the nucleus and cytoplasm while a weak NES would allow longer residence in the nucleus so that the protein can accumulate

to sufficient concentrations to assemble into the SMN complex for pre-mRNA splicing. This notion was first advanced to explain the very low affinity observed between IκB-α and the CRM1 exportin [27], which appears to be consistent with our own observation that its NES exhibits "weak" nuclear export activity in the Rev(1.4)-GFP assay system.

In this connection, it may be useful to note that, in the study of Lee and Hannink [27], the addition of LMB only shifted the cytoplasmic localization of IκB-α to a nuclear and cytoplasmic (N+C) pattern, rather than the exclusively nuclear (N) pattern. This corresponds well with our present results, in which there were appreciable percentages of cells showing the N+C fluorescence pattern in the presence of LMB for all NESs tested in the Rev(1.4)-GFP fusion constructs (Table 2). Similarly, the nuclear accumulation of GFP-IκB-α [28] and of endogenous Gal3 [15] was increased by LMB addition but there were still appreciable levels of the respective proteins remaining in the cytoplasm.

We had generated the double mutant, L247A and I249A, in the GFP-MalE-GFP system to test whether the putative leucine-rich NES of Gal3 (leucine 241 through isoleucine 249) was functional. The exclusively cytoplasmic localization of GFP-MalE-Gal3 (L247A; I249A) indicated, however, that the mutations may have disrupted the NLS. This, in turn, implies that the NLS and NES overlap in this segment of the Gal3 polypeptide. The M9 and KNS sequences represent two other examples of overlapping signals, in which the same stretch of amino acid sequence is capable of mediating both nuclear import and nuclear export [29, 30]. The M9 signal, a stretch of 38 amino acids with critical glycine and proline residues, was identified on hnRNP A1 protein, responsible for its shuttling property between the nucleus and the cytoplasm. The 39-amino acid KNS shuttling signal was identified on hnRNP K protein. For nuclear export,

the critical residues include negatively charged acidic amino acids. The possible involvement of these amino acids in KNS-mediated nuclear import has not been investigated.

The polypeptide of Gal3 contains two domains: (a) the proline- and glycine-rich amino-terminal domain; and (b) the carbohydrate-binding carboxyl-terminal domain. Physico-chemical studies carried out on the mouse [16] and hamster [31] homologs of Gal3 have indicated that the two domains are structurally, as well as functionally, distinct and independent. The three-dimensional structure of the carboxyl-terminal carbohydrate-binding domain has been elucidated by X-ray crystallography [32]. In this structure, the region of overlap between the NLS and the NES lies in the middle of a β-sandwich. It seems somewhat puzzling, therefore, how the transport receptors can gain access to this portion of the sequence. Alternatively, the possibility is raised that this stretch of the amino acid sequence simply provides a scaffold for folding, allowing other surface residues to interact with transport chaperones or components of a macromolecular complex, such as the SMN complex, that move in and out of the nucleus as an ensemble.

ACKNOWLEDGMENTS

We thank Peter Davidson for his help in the preparation of this manuscript. This work was supported by grants GM-38740 from the National Institutes of Health and MCB 97-23615 from the National Science Foundation.

FOOTNOTES

[1]Park, J., Voss, P.G., Grabski, S., Wang, J.L., and Patterson, R.J. (2001) Association of galectin-1 and galectin-3 with Gemin4 in complexes containing the SMN protein. Nucleic Acids. Res., in press.

[2]Davidson, P.J., Davis, M.J., Patterson, R.J., Ripoche, M.-A., Poirier, F., and Wang, J.L. (2001) Shuttling of galectin-3 between the nucleus and cytoplasm. Submitted for publication.

REFERENCES

1. Leffler, H. (1997). Introduction to galectins. *Trends Glycosci. Glycotechnol.* **9**, 9-19.
2. Moutsatsos, I.K., Davis, J.M., and Wang, J.L. (1986). Endogenous lectins from cultured cells: subcellular localization of carbohydrate-binding protein 35 in 3T3 fibroblasts. *J. Cell Biol.* **102**, 477-483.
3. Laing, J.G. and Wang, J.L. (1988). Identification of carbohydrate-binding protein 35 in heterogeneous nuclear ribonucleoprotein complex. *Biochemistry* **27**, 5329-5334.
4. Dagher, S.F., Wang, J.L., and Patterson, R.J. (1995). Identification of galectin-3 as a factor in pre-mRNA splicing. *Proc. Natl. Acad. Sci. USA* **92**, 1213-1217.
5. Vyakarnam, A., Dagher, S.F., Wang, J.L., and Patterson, R.J. (1997). Evidence for a role for galectin-1 in pre-mRNA splicing. *Mol. Cell. Biol.* **17**, 4730-4737.
6. Charroux, B., Pellizzoni, L., Perkinson, R.A., Yong, J., Shevchenko, A., Mann, M., and Dreyfuss, G. (2000). Gemin4: a novel component of the SMN complex that is found in both gems and nucleoli. *J. Cell Biol.* **148**, 1177-1186.
7. Fisher, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**, 1023-1029.
8. Liu, Q. and Dreyfuss, G. (1997). A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* **15**, 3555-3565.
9. Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell* **95**, 615-624.
10. Nishi, K., Yoshida, M., Fujiwara, D., Nishikawa, M., Horinouchi, S., and Beppu, T.

(1994). Leptomycin B targets a regulatory cascade of crm1, a fission yeast nuclear protein, involved in control of higher order chromosome strucutre and gene expression. *J. Biol. Chem.* **269**, 6320-6324.

11. Kudo, N., Wolff, B. Sekimoto, T., Schreiner, E. P., Yoneda, Y., Yanagida, M., Horinouchi, S., and Yoshida, M. (1998). Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1. *Exp. Cell Res.* **242**, 540-547.

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

13. Henderson, B.R. and Eleftheriou, A. (2000). A comparison of the activity, sequence specificity, and CRM1-dependence of different nuclear export signals. *Exp. Cell Res.* **256**, 213-224.

14. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680-685.

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

16. Agrwal, N., Sun, Q., Wang, S.-Y., and Wang, J.L. (1993). Carbohydrate binding protein 35. I. Properties of the recombinant polypeptide and the individuality of the domains. *J. Biol. Chem.* **268**, 14931-14939.

17. Fujii, G., Tsuchiya, R., Ezoe, E., and Hirohashi, S. (1999). Analysis of nuclear localization signals using a green fluorescent protein-fusion protein library. *Exp. Cell Res.* **251**, 299-306.

18.Openo, K.P., Kadrofske, M.M., Patterson, R.J., and Wang, J.L. (2000). Galectin-3 expression and subcellular localization in senescent human fibroblasts. *Exp. Cell Res.* **255**, 278-290.

19.Nakielny, S. and Dreyfuss, G. (1999). Transport of proteins and RNAs in and out of the nucleus. *Cell* **99**, 677-690.

20.Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463-473.

21.Meyer, B.E. and Malim, M.H. (1994). The HIV-1 rev transactivator shuttles between the nucleus and the cytoplasm. *Genes Dev.* **8**, 1538-1547.

22.Kalland, K.H., Szilvay, A.M., Brokstad, K.A., Seatrevik, W., and Haukenes, G. (1994). The human immunodeficiency virus type 1 (HIV-1) Rev protein shuttles between the cytoplasm and nuclear compartments. *Mol. Cell. Biol.* **14**, 7436-7444.

23.Gong, H.C., Honjo, Y., Nangia-Makker, P., Hogan, V., Mazurak, N., Bresalier, R.S., and Raz, A. (1999). The NH_2-terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells. *Cancer Res.* **59**, 6239-6245.

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

25.Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499-509.

26.Moore, J.D., Yang, J., Truant, R., and Kornbluth, S. (1999). Nuclear import of Cdk/cyclin complexes: identification of distinct mechanisms for import of Cdk2/cyclin E and Cdc2/cyclin B1. *J. Cell Biol.* **144**, 213-224.

27.Lee, S.-H. and Hannink, M. (2001). The N-terminal nuclear export sequence of IκBα

is required for RanGTP-dependent binding to CRM1. *J. Biol. Chem.* **276**, 23599-23606.

28. Huang, T.T., Kudo, N., Yoshida, M., and Miyamoto, S. (2000). A nuclear export signal in the N-terminal regulatory domain of IκBα controls cytoplasmic localization of inactive NFκB/IκBα complexes. *Proc. Natl. Acad. Sci. USA* **97**, 1014-1019.

29. Michael, W.M., Choi, M., and Dreyfuss, G. (1995). A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* **83**, 415-422.

30. Michael, W.M., Eder, P.S., and Dreyfuss, G. (1997). The K nuclear shuttling domain: A novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J* **16**, 3587-3598.

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

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