





This is to certify that the  
dissertation entitled  
Carbohydrate Binding Protein 35:  
In Vivo and In Vitro Expression Properties  
of the Polypeptide  
presented by  
Neera Agrwal

has been accepted towards fulfillment  
of the requirements for  
Ph. D degree in Biochemistry

  
Major professor

Date August 25, 1992

**CARBOHYDRATE BINDING PROTEIN 35: *IN VIVO* and *IN VITRO*  
EXPRESSION PROPERTIES OF THE POLYPEPTIDE**

By

Neera Agrwal

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1992

## **ABSTRACT**

### **CARBOHYDRATE BINDING PROTEIN 35: *IN VIVO* AND *IN VITRO* EXPRESSION PROPERTIES OF THE POLYPEPTIDE**

By

Neera Agrwal

Previous studies had shown that the amount and subcellular localization of the endogenous lectin Carbohydrate Binding Protein 35 (CBP35) in murine 3T3 fibroblasts is regulated by the growth state of the cell. In proliferating cells, the protein levels are higher than in quiescent cells and largely localized in the nucleus. Using the cDNA clone for CBP35 as a probe, we have examined the expression of the CBP35 gene in quiescent and serum-stimulated cells. The main conclusions of these studies include: a) higher levels of accumulated CBP35 mRNA are found in proliferating cells than in quiescent cells; b) the rise in mRNA levels is detected within 30 minutes of serum stimulation, and increases until ~20 hours after serum addition; c) the mRNA is "superinduced" in the presence of cycloheximide; d) the transcriptional rate for the CBP35 gene increases within 3 hours of serum stimulation, and reaches a maximum level at 10 hours following serum addition; and e) the transcription of the CBP35 gene occurs even in the presence of cycloheximide, indicating that this is a primary event in the response to serum growth factors by the cell.

Using the cDNA clone for CBP35, the full length recombinant protein and the NH<sub>2</sub>- and COOH-terminal domains have been expressed in *E. coli* cells. Analyses of the expressed proteins have demonstrated that the galactose binding activity of CBP35 is contained entirely within the COOH-terminus. Differential scanning calorimetry with the recombinant polypeptides has shown that the two domains are folded independently.

The recombinant CBP35 (rCBP35) has also been used to examine the uptake of the protein by nuclei from 3T3 cells. The results indicate that the protein uses a pathway distinct from that used by the synthetic substrate, human serum albumin bearing the nuclear localization signal of the SV40 large T antigen. Cytosolic factors which are sufficient for the entry of the synthetic substrate into the nucleus cannot support the nuclear import of rCBP35. This suggests that the control of nucleo-cytoplasmic distribution of rCBP35 may be mediated by the presence of a cytoplasmic anchor, or the availability of a factor which serves as a carrier for rCBP35 in nuclear import.

To David  
and  
To my parents  
for their love and faith in me

## ACKNOWLEDGEMENT

I thank Dr. John Wang for his infinite patience in guiding me through the time that I have spent in his laboratory. He is an extraordinary teacher, and I am fortunate to have done my graduate work under his care.

I thank my committee members for their insights into this research project: Dr. Zachary Burton, Dr. Susan Conrad, Dr. Jerry Dodgson, and Dr. Bill Smith. I would also like to thank Dr. Richard Anderson for his thoughts and views.

Much of the best work in science arises as a result of collaborations between investigators. We could not have been successful without the assistance of the following: Dr. Ron Patterson, Dr. Betty Werner and Sue Dagher for their scientific input on the nuclear import assays; Dr. John Wilson for the DSC analysis; Dr. Richard Wozniak and Dr. Günter Blobel for the use of the microinjection facilities; and Dr. Natasha Raikhel for the use of the fluorescence microscope. I would also like to thank Dr. David Lerner for his support throughout the years, and Dave Schwab for the discussions regarding *in vitro* protein production.

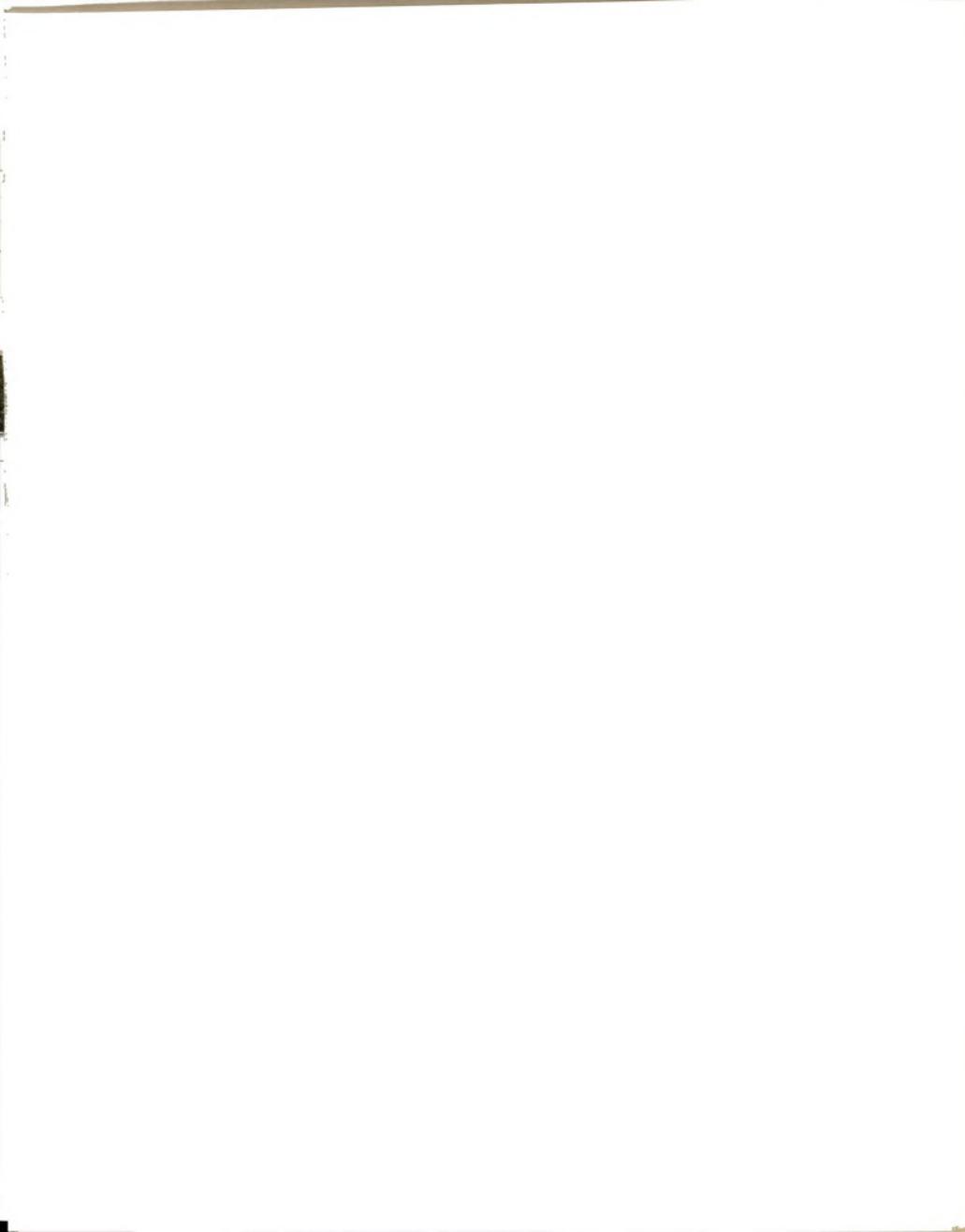
The entire staff in the Biochemistry Department has always done their utmost to facilitate administrative matters. I especially thank Linda Lang for all of her help. For the photographic work, I thank Kurt Stepnitz, who always found a way to get the work done "as soon as possible".

Finally, I wish to thank my colleagues in the Wang and Schindler laboratories for their advice and helpful criticisms: Dr. Elizabeth Cowles, Dr. James Laing, Patty Voss, Dr. John Ho, Dr. Shizhe Jia, Dr. Quan Sun, Sung Yuan Wang, Kim Hamann, John Loh, Anandita, and Mark Kadrofske. I would like to especially acknowledge the help given to me by Liz Cowles, Jamie Laing, and Patty Voss during my first few years in the lab.



## TABLE OF CONTENTS

	Page
<b>LIST OF TABLES</b> .....	x
<b>LIST OF FIGURES</b> .....	xi
<b>LIST OF ABBREVIATIONS</b> .....	xiv
<b>INTRODUCTION</b> .....	1
<b>CHAPTER I: LITERATURE REVIEW</b> .....	3
<b>INTRODUCTION TO LECTINS</b> .....	3
Classification of Animal Lectins .....	4
The C-type lectins .....	6
The S-type lectins .....	10
The L-14 family of S-type lectins .....	11
The L-30 family S-type lectins .....	18
Saccharide binding characteristics of the S-type lectins .....	19
Subcellular localization of the L-14 and L-30 lectins .....	20
The bifunctional nature of lectins .....	25
a) Selectins .....	25
b) The S-type L-14 lectins .....	27
<b>CARBOHYDRATE BINDING PROTEIN 35</b> .....	27
Structure of the CBP35 gene and protein .....	28
Ligands for CBP35/Mac-2 .....	30
Proliferation dependent localization and expression of CBP35 .....	30
<b>THE REGULATION OF EXPRESSION OF C-FOS</b> .....	32
Introduction to Primary Response Genes .....	32
Kinetics of the expression of <i>c-fos</i> .....	34
Superinduction of the <i>c-fos</i> mRNA .....	34



The <i>c-fos</i> promoter .....	35
Factors controlling the induction and function of <i>fos</i> .....	37
Regulation of transcription by the PRG .....	38
NUCLEAR TRANSPORT .....	39
General characteristics of nuclear transport .....	39
Nuclear localization signal .....	40
The nuclear pore complex .....	41
Cytosolic factors are required for nuclear transport .....	43
NLS binding proteins .....	45
Alternative methods of nuclear transport .....	46
Control of nuclear transport .....	48
LITERATURE CITED .....	50
<b>CHAPTER II: CARBOHYDRATE BINDING PROTEIN 35:</b>	
<b>Levels of Transcription and mRNA Accumulation in</b>	
<b>Quiescent and Proliferating Cells .....</b>	<b>63</b>
FOOTNOTES .....	64
ABSTRACT .....	65
INTRODUCTION .....	66
MATERIALS AND METHODS .....	68
Cell Culture .....	68
RNA Isolation and Northern Blot Analysis .....	68
Nuclear Run-off Transcription Assays .....	71
Plasmids and Preparation of Probes .....	73
Indirect Immunofluorescence .....	74
RESULTS .....	75
Kinetics of the Accumulation of CBP35	
mRNA upon Mitogenic Stimulation .....	75
Analysis of the Transcription Rate of the	
CBP35 Gene after Stimulation .....	78
Effect of Cycloheximide on Transcription Rate	
and mRNA Accumulation following Mitogen Addition .....	83



Comparison of the Expression of the CBP35 Gene in Normal and Transformed Cells .....	86
Comparison of the Expression of the CBP35 Gene in Sparse and Confluent Cultures of 3T3 Cells .....	96
DISCUSSION .....	97
ACKNOWLEDGEMENTS .....	101
REFERENCES .....	102
<b>CHAPTER III: CARBOHYDRATE BINDING PROTEIN 35:</b>	
<b>Properties of the Recombinant Polypeptide     and the Individuality of the Domains .....</b>	<b>104</b>
FOOTNOTES .....	105
SUMMARY .....	106
INTRODUCTION .....	107
MATERIALS AND METHODS .....	109
Construction of prCBP35s and Purification of rCBP35 .....	109
Site-Directed Mutagenesis and Generation of Recombinant Clones .....	110
Preparation of the N- and C-domains .....	111
Antibodies .....	112
Analytical Techniques .....	113
Differential Scanning Calorimetry (DSC) .....	114
RESULTS .....	116
Expression and Purification of Recombinant CBP35 .....	116
Generation of the NH <sub>2</sub> - and COOH-terminal Domains of rCBP35 .....	122
Development and Characterization of Antisera Against rCBP35 .....	130
DSC Analysis .....	133
DISCUSSION .....	144
ACKNOWLEDGEMENTS .....	147



REFERENCES .....	148
<b>CHAPTER IV CARBOHYDRATE BINDING PROTEIN 35:</b>	
<b>Preliminary Studies on the Transport of the</b> <b>Recombinant Polypeptide into the Nucleus .....</b>	150
FOOTNOTES .....	151
SUMMARY .....	152
INTRODUCTION .....	153
MATERIALS AND METHODS .....	155
Cell Cultures .....	155
Preparation of Import Substrates .....	155
Nuclear Import Assays .....	157
Fluorescence Microscopy .....	158
RESULTS .....	160
Nuclear Transport of Rh-HSA-NLS Assayed with Permeabilized Cells .....	160
Behavior of Rh-rCBP35 in the Permeabilized Cell Assay System .....	163
Nuclear Transport in Intact Cells .....	166
Nuclear Transport of Rh-rCBP35 in Quiescent and Proliferating Cells .....	174
Microinjection of NH <sub>2</sub> - and COOH- Terminal Domains of rCBP35 .....	175
DISCUSSION .....	180
ACKNOWLEDGEMENT .....	185
REFERENCES .....	185
<b>CHAPTER V CONCLUDING STATEMENT .....</b>	187
REFERENCES .....	188

## LIST OF TABLES

	Page
<b>CHAPTER I</b>	
TABLE I. Classes of Animal Carbohydrate Binding Proteins .....	5
TABLE II. The C-Type Lectins .....	7
TABLE III. L-14 Group of S-Type Lectins .....	12
TABLE IV. L-30 Group of S-Type Lectins .....	13
TABLE V. S-Type Lectins of M <sub>r</sub> 16000-22000 .....	14
<b>CHAPTER III</b>	
TABLE I. Parameters Affecting Yield of rCBP35 .....	121
<b>CHAPTER IV</b>	
TABLE I. Nuclear Localization in Serum-Starved and Serum-Stimulated 3T3 Cells .....	176

## LIST OF FIGURES

	<b>Page</b>
<b>CHAPTER I</b>	
Figure 1. Summary of the structural features of the C-type lectins .....	9
Figure 2. Summary of the structural features of the S-type lectins .....	16
<b>CHAPTER II</b>	
Figure 1. Kinetics of the accumulation of mRNA for CBP35 after serum stimulation of growth-arrested 3T3 cells .....	77
Figure 2. Gene transcription rates after serum stimulation of quiescent 3T3 cells in the absence (A) and presence (B) of cycloheximide (10 $\mu$ g/ml) .....	80
Figure 3. Kinetics of the changes in relative transcription rates of the CBP35 gene following serum stimulation of quiescent 3T3 fibroblasts .....	82
Figure 4. Effect of cycloheximide on the levels of accumulated mRNA for CBP35 .....	85
Figure 5. Immunofluorescence analyses comparing CBP35 in 3T3 and 3T3-KiMSV cells .....	88
Figure 6. Northern blot analysis of CBP35 mRNA in 3T3 and 3T3-KiMSV cells .....	91
Figure 7. Autoradiogram of nuclear run-off assays comparing the transcription rate of the CBP35 gene in 3T3 and 3T3-KiMSV cells .....	93

Figure 8.	Northern blot analysis to compare half-lives of CBP35 mRNA in 3T3 and 3T3-KiMSV cells . . . . .	95
-----------	---	----

### CHAPTER III

Figure 1.	Schematic diagram of the construction of the recombinant expression vector prCBP35s and site-directed mutagenesis to obtain the amino terminal and carboxyl terminal domains . . . . .	118
Figure 2.	Purification of rCBP35 from <i>E. coli</i> transformed with the expression vector prCBP35s . . . . .	120
Figure 3.	Purification of the N-domain from <i>E. coli</i> transformed with expression vector containing a fragment of mutagenized cDNA . . . . .	125
Figure 4.	Digestion of rCBP35 with collagenase D and affinity purification of C-domain on Asialofetuin—Affi-gel . . . . .	129
Figure 5.	Characterization of anti-CBP35 antisera using immunoblotting . . . . .	132
Figure 6.	Representative thermograms from DSC analysis of rCBP35 . . . . .	135
Figure 7.	Deconvolution of the thermogram of rCBP35 in pH 10 buffer . . . . .	137
Figure 8.	Effect of heat denaturation of the N-domain on collagenase digestion of rCBP35 . . . . .	140
Figure 9.	Representative thermograms from DSC analysis of N- and C-domains . . . . .	143

### CHAPTER IV

Figure 1.	<i>In vitro</i> nuclear import assay comparing the translocation of Rh-HSA-NLS with Rh-HSA . . . . .	162
Figure 2.	<i>In vitro</i> nuclear import assay for Rh-rCBP35 . . . . .	165

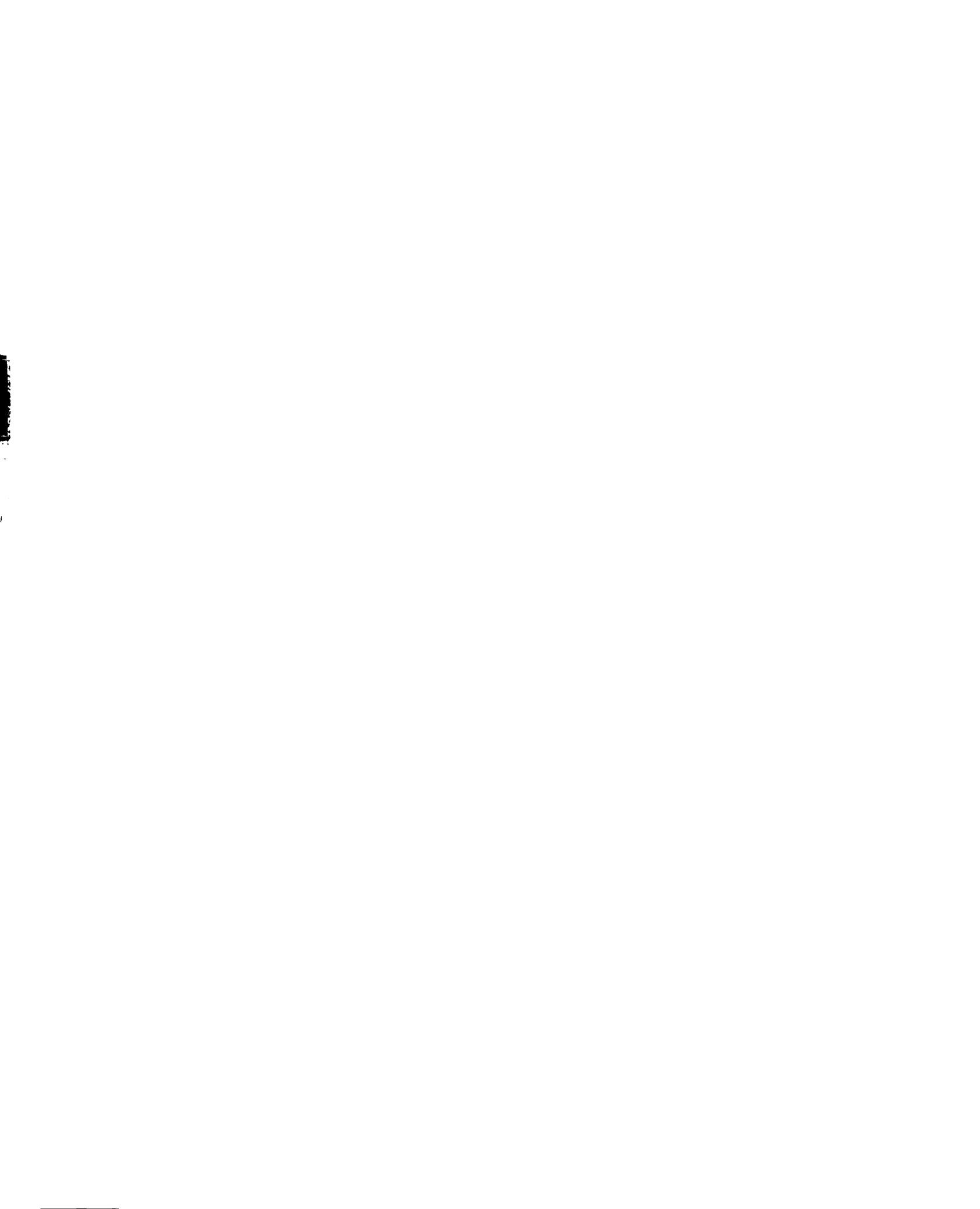
Figure 3.	<i>In vivo</i> nuclear import assay for Rh-HSA-NLS in the absence and presence of WGA . . . . .	168
Figure 4.	<i>In vivo</i> nuclear import assay comparing the translocation of Rh-rCBP35 with Rh-HSA . . . . .	171
Figure 5.	<i>In vivo</i> nuclear import assay for Rh-rCBP35 showing the nuclear and non-nuclear localization of the protein . . . . .	173
Figure 6.	<i>In vivo</i> nuclear import assay for the NH <sub>2</sub> - and COOH-terminal portions of rCBP35 . . . . .	179



## LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
BHK	baby hamster kidney
BSA	bovine serum albumin
CBP	carbohydrate binding protein
CBP35	carbohydrate binding protein 35
cDNA	complementary deoxyribonucleic acid
CRD	carbohydrate recognition domain
CRE	cAMP response element
DME	Dulbecco's modified Eagle Medium
DNA	deoxyribonucleic acid
DSC	differential scanning calorimetry
$\epsilon$ BP	IgE binding protein
EGF	epidermal growth factor
ER	endoplasmic reticulum
FBS	fetal bovine serum
FGF	fibroblast growth factor
Fuc	fucose
Gal	galactose
GalNAc	N-acetylgalactosamine
GlcNAc	N-acetylglucosamine
HEPES	<i>N</i> -2-hydroxyethylpiperazine- <i>N'</i> -2-ethanesulfonic acid
hnRNP	heterogeneous nuclear ribonucleoprotein
HPLC	high performance liquid chromatography
HSA	human serum albumin
hsc70	heat shock cognate 70
hsp70	heat shock protein 70
IE	immediate early gene
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
kD, kDa	kilodaltons
Lac	lactose
LBP	laminin binding protein
Man	mannose
2-ME	2-mercaptoethanol
M6P	mannose-6-phosphate

MEF	mouse embryonic fibroblast
mGBP	mouse galactoside binding protein
mRNA	messenger ribonucleic acid
NBP	nuclear localization signal binding protein
NEM	N-ethylmaleimide
NGF	nerve growth factor
NLS	nuclear localization signal
NPC	nuclear pore complex
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PDGF	platelet-derived growth factor
PRG	primary response gene
rCBP35	recombinant CBP35
rCBPCD	carboxyl terminus of rCBP35
rCBPND	amino terminus of rCBP35
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
snRNP	small nuclear ribonucleoprotein
SRE	serum response element
SRF	serum response factor
TGF	transforming growth factor
TMG	trimethylguanosine
TRE	TPA response element
WGA	wheat germ agglutinin



## INTRODUCTION

Carbohydrate Binding Protein 35 (CBP35;  $M_r$  35,000) was initially purified from extracts of murine Swiss 3T3 fibroblasts on the basis of its affinity for  $\beta$ -galactosyl-containing glycoconjugates. It has since been isolated and studied from other species, as well as many tissue types. Our interest in CBP35 stems from three key properties of the protein: a) its structure; b) its subcellular localization; and c) its proliferation dependent expression.

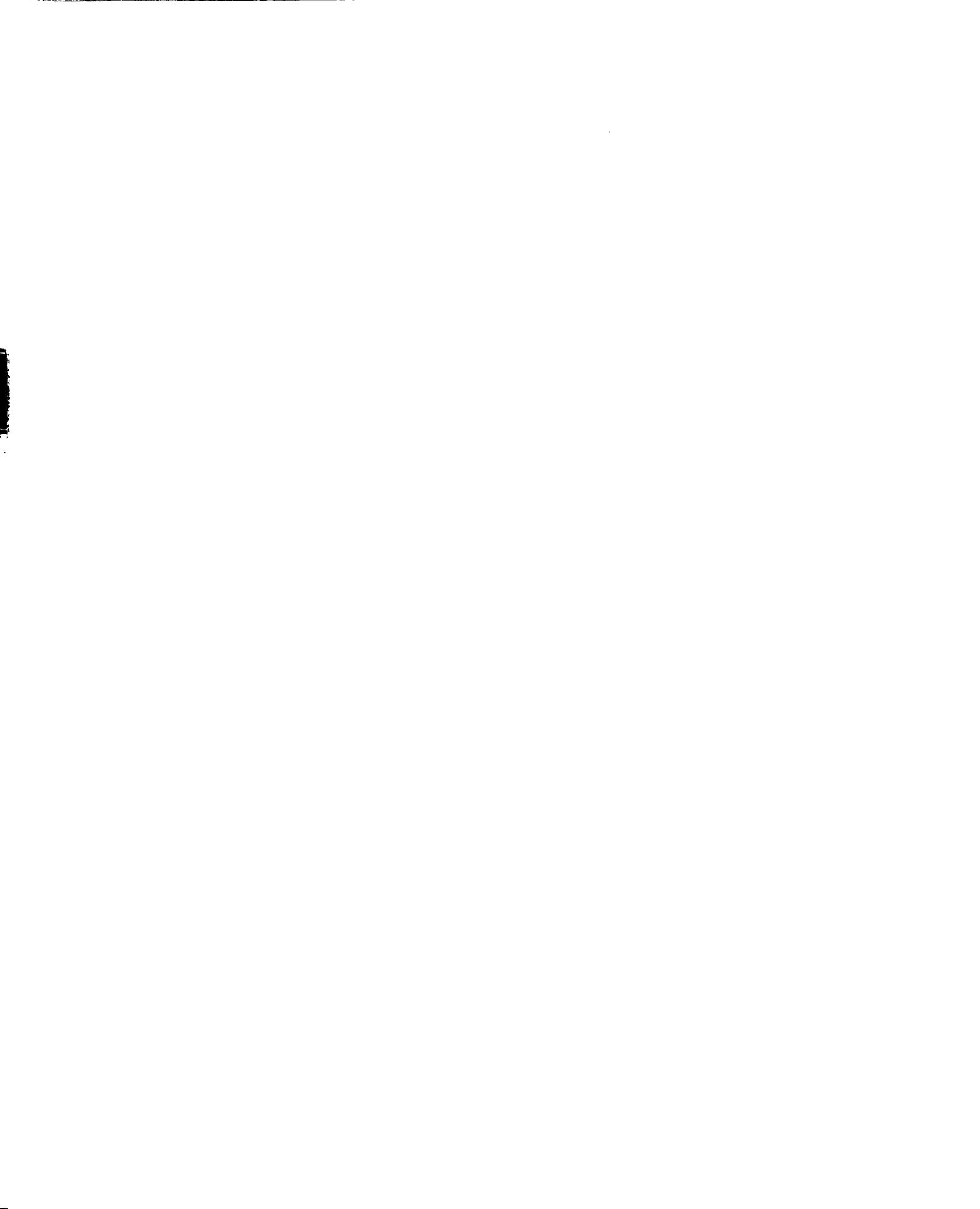
First, the amino acid sequence of the polypeptide, deduced from the nucleotide sequence of the cDNA clone, showed that the protein contains two distinct domains. The amino terminal half contains eight contiguous repeats of the 9-amino acid sequence PGAYPGXXX. The carboxyl terminal half contains 13 amino acids that are invariant in CBP35 and other  $\beta$ -galactoside specific lectins and, therefore, is believed to constitute the carbohydrate recognition domain of the lectin. These features of the primary sequence prompted us to ask questions concerning the physico-chemical properties of the individual domains, and their roles in the biological activity of the protein.

Second, studies on the subcellular localization of CBP35 showed that the majority of the lectin is intracellular, although a small percentage is also found at the cell surface, despite the lack of an obvious signal sequence to target it through

the secretory pathway. Moreover, the intracellular lectin can be found in the nucleus and the cytoplasm, depending on the proliferative state of the cell.

Finally, when quiescent 3T3 cells were stimulated by the addition of serum, the levels of CBP35 rose dramatically, and the protein was translocated into the nucleus. These observations formed the basis for our studies on the transcriptional regulation of CBP35 expression, as well as factors controlling its nuclear localization.

Thus, this literature review will discuss the following topics: a) the structure and subcellular localization of carbohydrate binding proteins; b) the proliferation dependent expression of proteins; and c) mechanisms by which proteins translocate into the nucleus.



## CHAPTER I

### LITERATURE REVIEW

#### INTRODUCTION TO LECTINS

Specific recognition is a key event in every biological process. For example, the specificity of recognition is seminal in cell-cell interaction, a step which is necessary in such diverse functions as immunomodulatory activity, fertilization, development, and infection (1). The discovery that all cells carry carbohydrates on their surface led to the notion that these glycoproteins may play a fundamental role in recognition processes. This idea was furthered by the finding that there exist molecules which specifically recognize these sugar structures. These molecules were termed lectins (2). Lectins are classified as non-enzymatic, non-immune carbohydrate-binding molecules (3). The traditional definition of a lectin has implied that it is at least a bivalent molecule with respect to saccharide binding. The terms carbohydrate binding protein and lectin will be used interchangeably in the discussion to follow.

Lectins were first discovered in plants as hemagglutinating substances. They have since been found in almost all organisms, in all tissues and cell types, and both intracellularly and extracellularly. Their exquisite ability to differentiate



between different monosaccharides and oligosaccharides, as well as to recognize subtle variations in sugar structures and linkages has made the lectins a potentially central figure in the cellular recognition system.

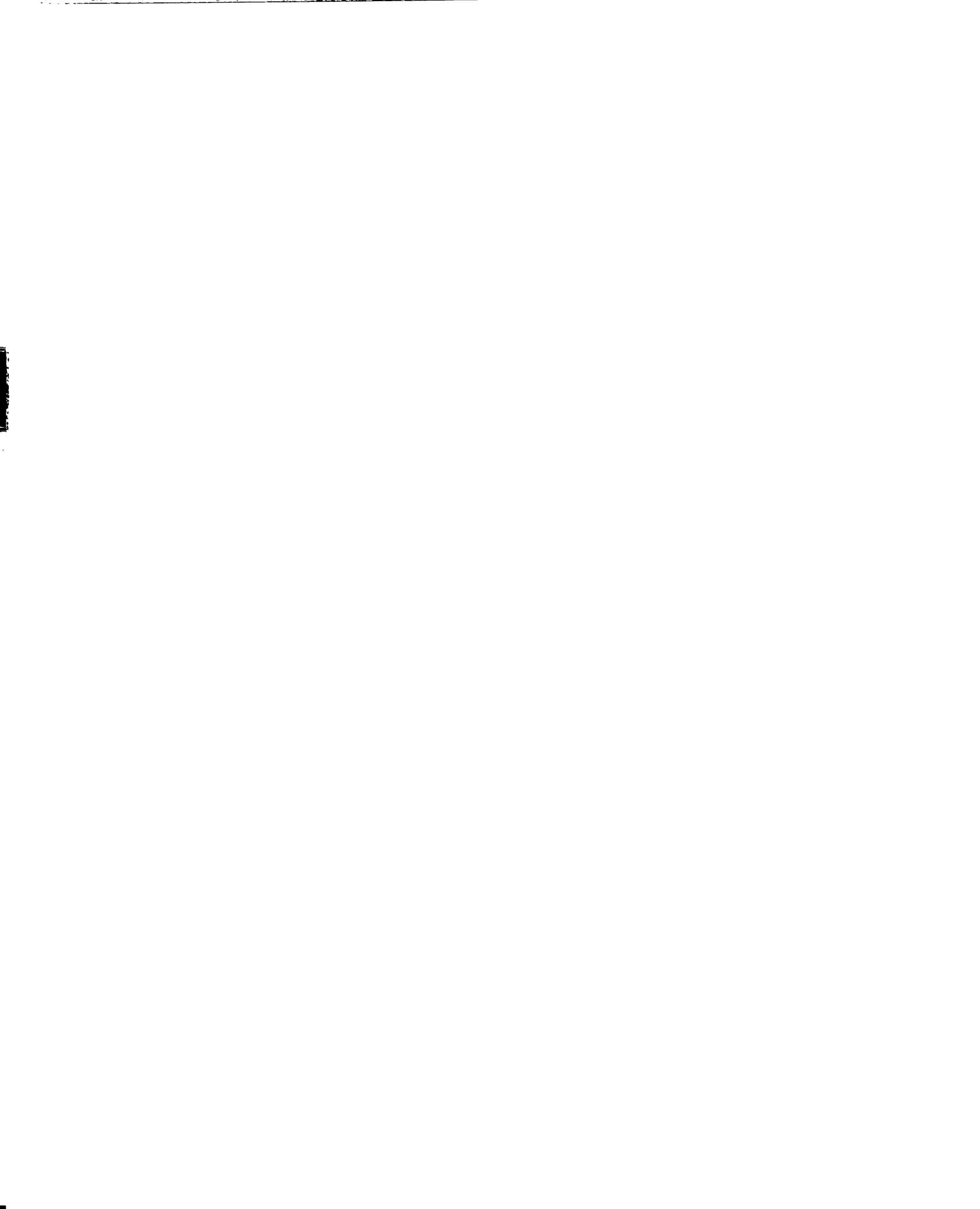
In this chapter, I will concentrate on the animal lectins. The plant lectins have been reviewed elsewhere (4,5).

### **Classification of Animal Lectins**

The region of the polypeptide which has the ability to bind to carbohydrates is designated as the Carbohydrate Recognition Domain (CRD). Animal lectins can be grouped into several categories based on sequence similarities within the CRD and/or saccharide binding characteristics (Table I). By these criteria, the most well studied lectins fall into the C-type and S-type categories (7). The C-type lectins are those which require the bivalent cation  $\text{Ca}^{2+}$  for sugar binding activity. These proteins may be membrane bound or soluble, and may also be glycosylated. The S-type lectins were originally defined to be proteins which require reducing agents for sugar binding. They have only been found as soluble, non-glycosylated entities.

The heparin binding lectins that have been isolated have no similarities with the other lectin families. They may or may not exhibit hemagglutinating activity, and may require divalent cations for binding to heparin. Their binding to heparin is inhibitable by competing sugars (8,9).

There are two other groups of carbohydrate binding proteins which bear no common structural or functional features with the above mentioned lectins.



**TABLE I. CLASSES OF ANIMAL CARBOHYDRATE BINDING PROTEINS\***

Family Type	Examples	Ca <sup>2+</sup> <sup>b</sup>	Cysteines	Location
C-type	receptors mediating glycoprotein endocytosis	yes	disulfide	plasma membrane
	mannose binding proteins	yes	disulfide	serum, liver
	proteoglycan core lectins	yes	disulfide	extracellular matrix
	selectins	yes		cell surface
	pulmonary surfactant	yes	disulfide	lung, fluid
S-type	L-14 and L-30 $\beta$ -galactoside binding lectins	no	sulfhydryl	extracellular, cell surface, nucleus
Mannose-6-P receptor	250 kDa M-6-P receptor	no	disulfide	plasma membrane, extracellular
	46 kDa M-6-P receptor	yes	disulfide	plasma membrane, extracellular
Pentraxin	serum amyloid protein	yes	disulfide	serum
Heparin binding lectins	placental lectin, p33, p41	?		?

\* adapted from reference 6

<sup>b</sup> Ca<sup>2+</sup> requirement for sugar binding

These include the serum amyloid protein, and the mannose-6-phosphate receptors.

I will focus on the C-type and S-type lectins in this review.

### The C-type Lectins

The C-type CRDs are found in a variety of proteins with a diversity of structure and function (Figure 1 and Table II). These include the following subgroups: a) the selectins, such as  $gp90^{MEL}$ , ELAM-1, and GMP-140 (10,11,12); b) transmembrane receptors, as exemplified by the asialoglycoprotein receptors (13), the Kupffer cell receptor (14), and the hepatic lectins (13); c) the macrophage receptors which are involved in the phagocytosis of pathogens (15); d) the soluble mannose binding proteins, found in liver and serum (16); e) the soluble pulmonary surfactant apoproteins (17); and f) the proteoglycan core proteins (18). Some C-type lectins have also been found in such divergent species as *Dictyostelium discoideum* (19), the fly (20), barnacle (21), sea urchin (22), and the tunicate *Polyandrocarpa misakiensis* (23).

The initial studies on the C-type lectins denoted the fact that these proteins required  $Ca^{2+}$  for binding activity. Subsequent cloning and sequence analyses have shown that the CRD for this type of lectin can be recognized by a sequence motif which contains approximately 30 conserved residues over a  $\sim 120$  amino acid range (7). The alignment of the CRD regions of 22 distinct proteins shows that there are 14 invariant residues, with an additional 18 residues that are conserved in character. These residues fall into three classes: a) 4 cysteines, involved in disulfide bond formation; b) a "WIGL" sequence which indicates a region packed

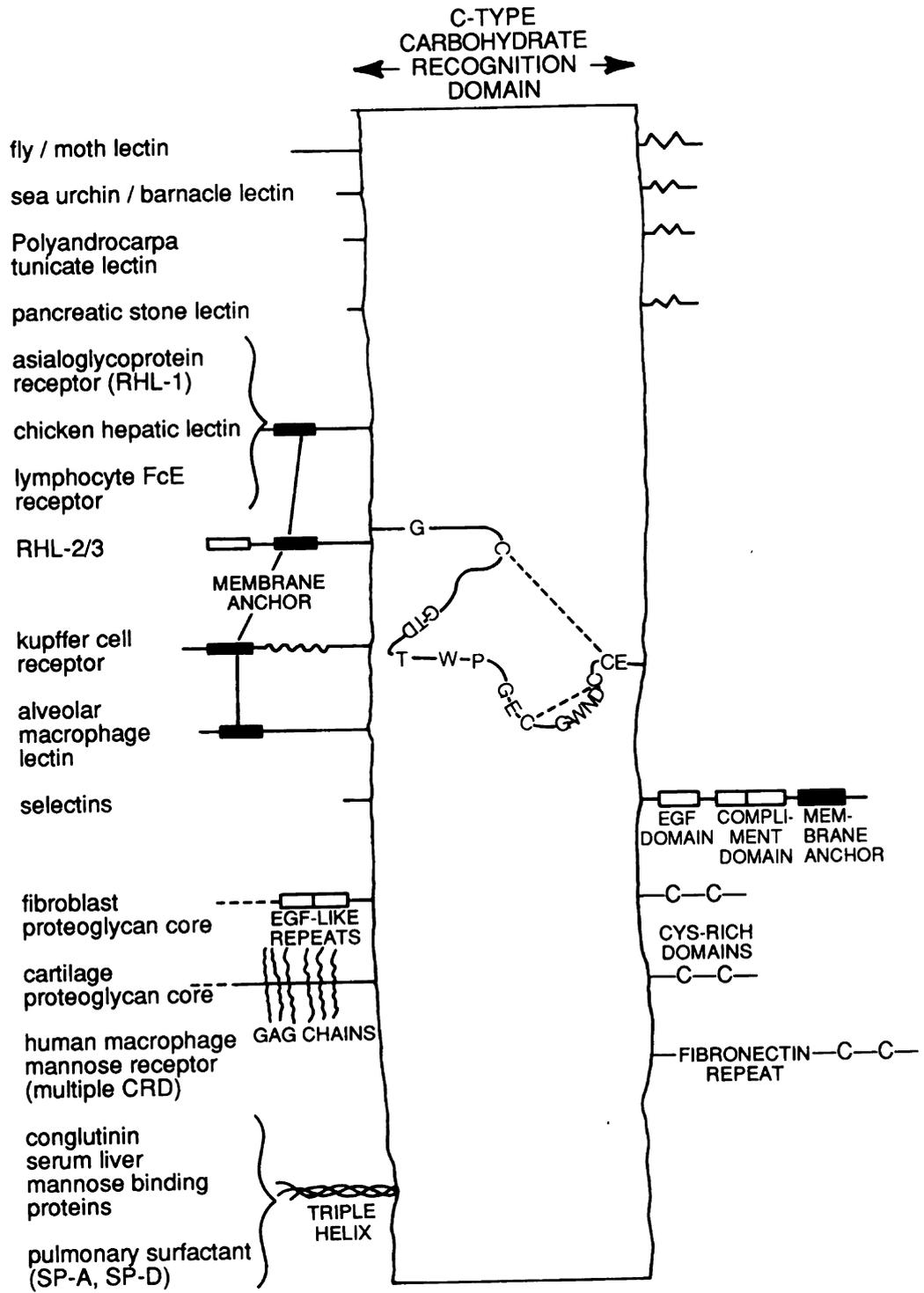
TABLE II. THE C-TYPE LECTINS\*

Species/Source	Lectin	Ligand <sup>b</sup>	Reference
flesh fly hemolymph	fly lectin	Gal	20
sea urchin	sea urchin lectin	Gal	22
<i>Polyandrocarpa misakiensis</i>	tunicate lectin	Gal	23
avian liver	chicken hepatic lectin	GlcNAc	13
sheep, goat, buffalo liver	hepatic lectin	Gal	27
rat liver	RHL-1, RHL-2/3	Gal	28,29
rat liver	Kupffer cell receptor	Fuc	14
mouse spleen	gp190 <sup>MEL</sup>	M6P	10
mouse macrophage	macrophage lectin	Gal	30
human lung	pulmonary surfactant	Man	17
human macrophage	mannose receptor	Man	15
human pancreas	pancreatic stone lectin	Gal	31
human fibroblast	fibroblast proteoglycan core	Gal	18

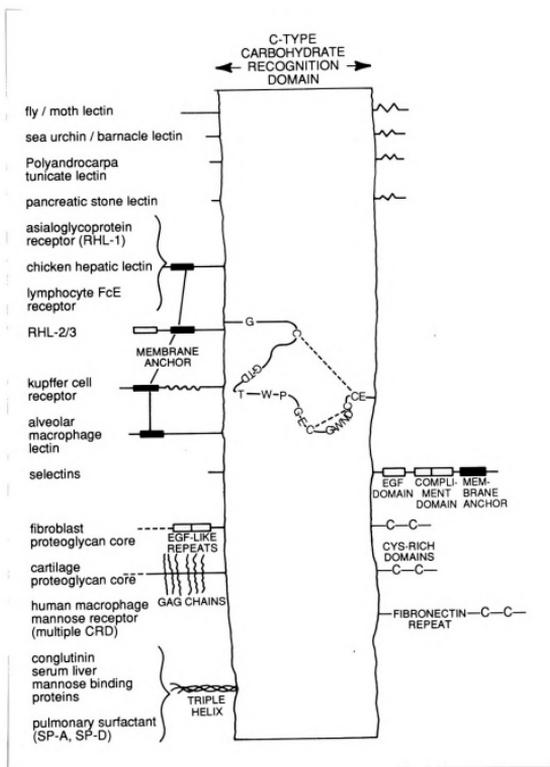
\* A representative group of the C-type lectins

<sup>b</sup> Gal, galactose; GlcNAc, N-acetylglucosamine; Fuc, fucose; M6P, mannose-6-phosphate; Man, mannose

**Figure 1. Summary of the structural features of the C-type lectins.** The invariant residues in the C-type carbohydrate recognition domain are shown, combined with the effector domains (if any) found in the various members of the C-type lectin family. GAG, glycosaminoglycan; EGF, epidermal growth factor. Adapted from reference 6.







into hydrophobic cores; and c) residues forming the ligands for  $\text{Ca}^{2+}$  (24).

Analysis of the gene structure for several C-type CRDs has shown that they fall into three distinct types (6,25). In the first type, the single CRD is encoded by three separate exons. This can be found in the rat asialoglycoprotein receptor (RHL-1), the Kupffer cell receptor, and the chicken proteoglycan core protein. On the other hand, the (single) CRDs for the mannose binding proteins, the pulmonary surfactant SP-A, and the murine lymphocyte homing receptor are encoded in a single exon. Finally, the cell surface mannose receptor of macrophages and hepatic sinusoidal cells has eight tandemly placed CRDs which comprise the carbohydrate recognition domain of this protein (26). Studies to determine the role of each of the CRDs in this protein have revealed that while CRD4 can function independently of the other CRDs, it does not exhibit the high affinity of binding of the whole molecule. Thus, it is probable that the multiple CRDs are required for tight binding to multivalent ligands.

### **The S-type lectins**

The S-type differ from the C-type lectins in that they: a) do not depend on cations for carbohydrate binding activity; b) have an affinity exclusively for Gal/Lac containing structures; c) have a specific consensus sequence, differing from that in the C-type lectins, in the CRD domain (7) (figure 2); d) are isolated invariably (thus far) as soluble proteins; and e) are susceptible to oxidation inactivation of sugar binding activity in the absence of reducing agents. The nature of the oxidation may lie in the cysteine or tryptophan residues in these

proteins (32). These lectins do not appear to be glycosylated.

The S-type lectins (also called the S-lac lectins) can be subdivided into several groups, based on their subunit molecular weight. The two best characterized groups are the L-14 and L-30 lectins. The L-14 group of lectins ranges in size from  $M_r$  12,000-14,500 daltons (Table III). The L-30 group of lectins consists of proteins of  $M_r$  29,000-35,000 daltons (Table IV). There is another group of S-type lectins which are of  $M_r$  16,000-22,000 (Table V). This group of lectins has been isolated from rat intestinal tissue (56,57), mouse lungs (45,58), *Xenopus laevis* skin tissue (59), and the nematode *C. elegans* (60). Although these lectins share many amino acid and structural similarities with the L-14 group, they appear to be novel members of the S-lac family, encoded by a separate gene.

The 67 kDa component (the large component) of the elastin receptor has also been classified as an S-type lectin based on the following data: a) the 67 kDa protein can be eluted from an asialofetuin affinity column by the addition of  $\beta$ -galactoside sugars; b) the asialofetuin-purified protein can be eluted from a column derivatized with elastin peptides by lactose; and c) antibody raised against the rat lung L-14 lectin reacts with the 67 kDa protein (63). However, the S-type CRD consensus sequence has not been found in the 67 kDa lectin.

#### **The L-14 family of S-type lectins**

The most abundant of the S-type lectins are the L-14 proteins. They are isolated as dimers, and have one or more cysteine residues. Evidence obtained

TABLE III. L-14 GROUP OF S-TYPE LECTINS<sup>a</sup>

Species	Lectin	Tissue Source	M <sub>r</sub> <sup>b</sup>	References
Conger eel	congerin	skin mucus	15000	33
<i>C. elegans</i>	GBP32	whole worm	32000	237
<i>B. arenarum</i>	L-15	ovary	15000	34
Chicken	CLL-II	intestine, liver	12000	35,36
	CHL	heart	13000	37
	C-14	skin	14000	38
	CLL-1	muscle, liver	15000	35,39
Marmoset	L-15	neonate	15000	40
Rabbit	Galaptin	bone marrow	13000	41
Bovine	BHL	heart, lung, spleen, thymus	12000	42,43
Porcine	PHL	heart	14700	44
Mouse	CBP13.5	lung, fibroblast	13500	45
	L-14.5	fibrosarcoma, melanoma	14500	46
	mGBP	embryonic fibroblast	14735	47
Rat	RL-14.5	muscle, lung, brain, neurons, intestine	14500	48,49,50
Human	HLBP14	melanoma	14000	235
	H14	placenta, hepatoma	14000	51
	HL-14	lung	14000	52,53
	HBL	brain	14500	54
	L-14-II	hepatoma	14650	55

<sup>a</sup> A representative group of the L-14 lectins  
Some of these lectins may be identical

<sup>b</sup> Subunit M<sub>r</sub>, as determined by reducing SDS-PAGE

TABLE IV. L-30 GROUP OF S-TYPE LECTINS<sup>a</sup>

Species	Lectin	Tissue/Cell Source	M <sub>r</sub> <sup>b</sup>	References
Mouse	CBP35	lung, fibroblast	35000	45,58 75,76
	LBP	macrophages	35000	74
	L-34	fibrosarcoma, melanoma	34000	46,72
	Mac-2	macrophages	32000	77,78
Hamster	hamster lectin	BHK cells	30000	79
Rat	RL-29	lung, brain	29000	48,56 49,50
	εBP	basophilic leukemia cells	31000	80
Human	HL-29	lung, brain	29000	62,81
	εBP	basophilic leukemia cells	31000	82
	Mac-2	macrophages	32000	83
	CBP35	lung, fibroblast	35000	58,84

<sup>a</sup> Some of these lectins may be identical

<sup>b</sup> Subunit M<sub>r</sub>, as determined by reducing SDS-PAGE

TABLE V. S-TYPE LECTINS OF M<sub>1</sub> 16000-22000<sup>a</sup>

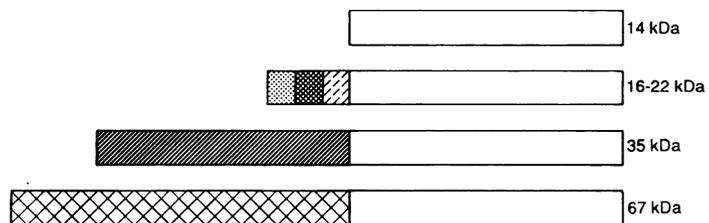
Species	Lectin	Cell/Tissue Source	M <sub>1</sub> <sup>b</sup>	References
Electric eel	electrolectin	electric organ	16000	32
<i>C. elegans</i>	L-16	whole worm	16000	60
<i>Xenopus laevis</i>	skin lectin	skin	16000	59
Chicken	C-16	liver	16000	38
Mouse	CBP16	lung, fibroblast	16000	45,58
Rat	L-17	intestinal mucosa	17000	56,57
	L-19	intestinal mucosa	19000	56,57
	L-21.5	intestinal mucosa	21500	57
	RL-22	lung	22000	48
	IgE binding protein	intestine	17500	61
Human	CBP16	lung, fibroblast	16000	58
	HL-22	lung	22000	62

<sup>a</sup> A representative group of the S-type lectins of M<sub>1</sub> 16-22 kDa  
Some of these lectins may be identical

<sup>b</sup> Subunit M<sub>1</sub>, as determined by reducing SDS-PAGE

**Figure 2. Summary of the structural features of the S-type lectins.**

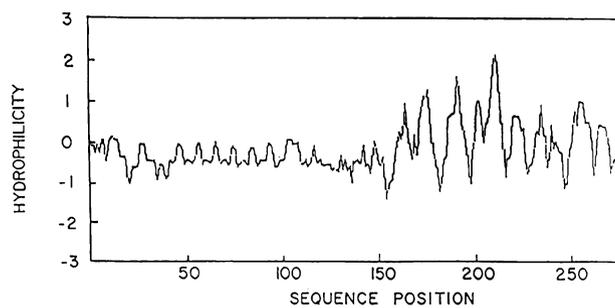
**A.** The multidomain structure of the S-type lectins. The carboxyl terminal domain (white box) contains the carbohydrate recognition domain. The shaded boxes represent various different effector domains, whose function is unknown. **B.** The features of the domains of the L-14 and L-30 lectins. The 13 invariant amino acid residues that occur in a 39-residue sequence in the CRD are shown. Also shown is the 9-amino acid sequence that is repeated in the amino-terminal domain of the L-30 lectins. The letter n, designating the number of repeats, ranges from 5 in the human Mac-2 sequence to 10 in the rat  $\epsilon$ BP sequence. A single residue between invariant residues is denoted by hyphen (-). Sequences of two or more residues are denoted by the symbol ( $\sim$ ). **C.** Hydropathy plot of murine CBP35, as determined by the deduced amino acid sequence. The distinctive hydropathy plot pattern of the two domains of an L-30 lectins is illustrated. Positive values indicate hydrophilicity and negative values indicate hydrophobicity. Adapted from references 67 and 88.

**A****B****L-14**

~H-NPRF~V-N-WG-E-R~F-G~

**L-30**~(PGAYPG...)<sub>n</sub>~

~H-NPRF~V-N-WG-E-R~F-G~

PROLINE, GLYCINE-  
RICH DOMAINCARBOHYDRATE RECOGNITION  
DOMAIN**C**

from analysis of the cDNA clones indicates that the L-14 lectins are encoded by a multi-gene family. At least three distinct L-14 lectins have been hinted at by the cDNA clones. The L-14-I lectin is representative of the first type of L-14 lectin isolated from vertebrate tissues. These lectins exhibit greater than 85% conservation at the amino acid level (55,64,65,66). The L-14-II lectin, isolated from a human hepatoma library, shares 43% amino acid identity with L-14-I (55). The avian L-14 cDNA appears to be more divergent showing 50% amino acid identity with the L-14-I lectin (65,66). The gene for the L-14-I type of protein contains 4 exons, with the third exon containing the CRD. The upstream region of the gene indicates the presence of a possible heat shock element, a putative steroid binding site, a putative metal regulatory element, and a sequence related to the Y box of the histocompatibility genes. In addition, there are intronic Alu sequences, and a G/T cluster downstream of the polyadenylation signal (53). The upstream region of L-14-II gene differs from that of L-14-I. There are two tandem TATA boxes, an Sp1-binding site, and a putative site for regulation by AP-1 (55).

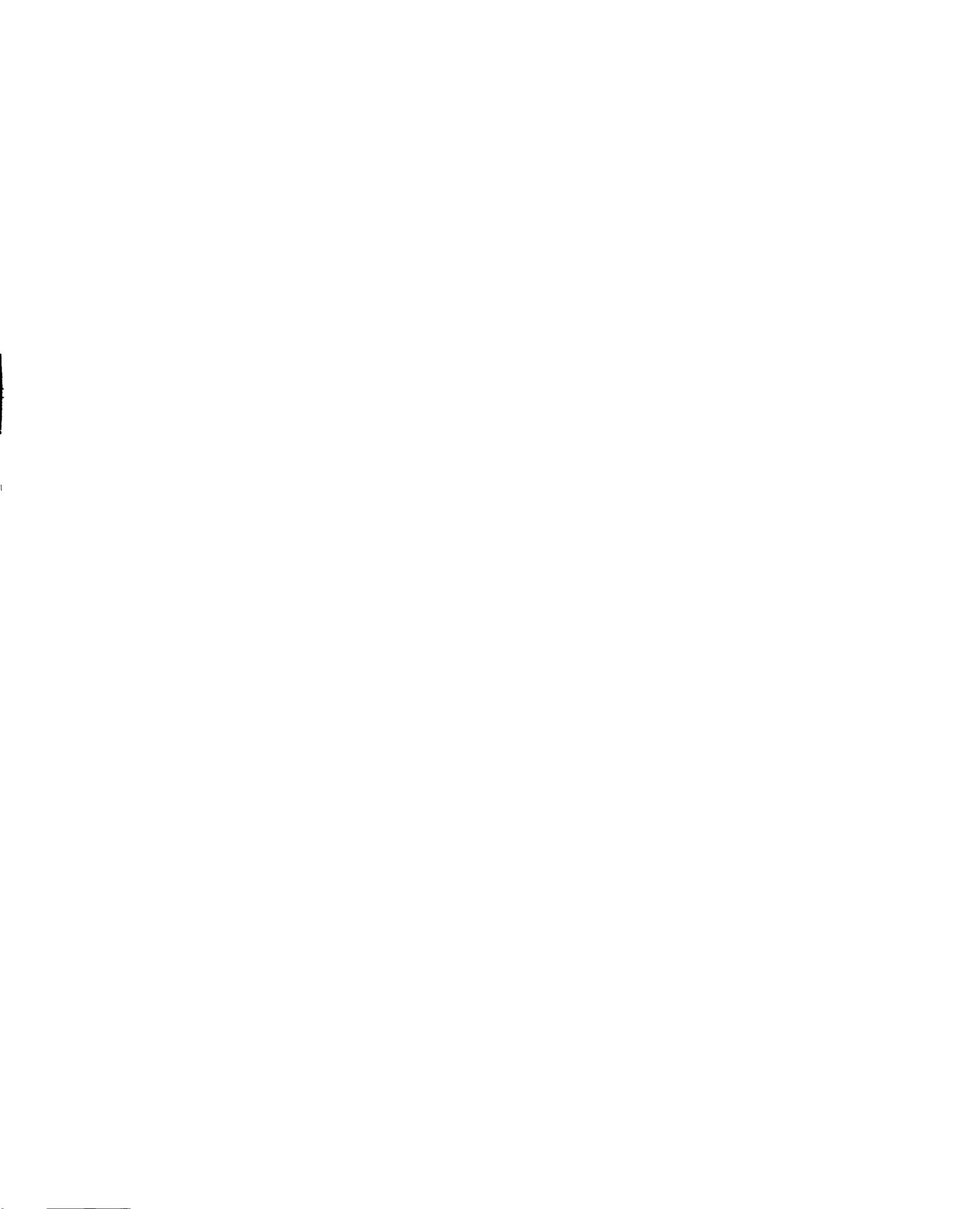
A novel type of L-14 lectin has been isolated and cloned from the nematode *C. elegans* (60,237). The lectin has a  $M_r$  of 32 kDa, and hence has been named 32 kDa GBP. Analysis of the cDNA clone for the protein shows that it has two tandem repeats of the L-14 unit. Each of the units show approximately 30% amino acid conservation between themselves, as well as the other S-type lectins. The S-type consensus sequence is well conserved, with an exact match of 9 amino acids. However, unlike the other L-14 lectins, there appear to be no

cysteines in the 32 kDa GBP. The significance of the repeated domains is unknown, especially since the hemagglutination activity is weak.

Histochemical and biochemical studies have shown that the L-14 lectin is ubiquitously distributed in tissues. It is found in skeletal muscle (68,69), nervous tissue (49,50,54), connective tissue (70,71), and tumor tissue (46,72,235). Perhaps, the richest sources of the L-14 lectin are the fetal lung and uterine tissues. This is further corroborated by experiments which were performed to isolate genes which are selectively expressed during embryogenesis. One of the cDNA clones isolated was that for the L-14 lectin (73).

### **The L-30 family S-type lectins**

The members of the L-30 family (Table IV) are all identical to each other or highly homologous ( $\geq 85\%$ ), as determined by cDNA sequence analysis. Despite their similarities in structure, a number of different functions have been ascribed to these proteins. The lectins RL-29, HL-29, BHK lectin, L-34, and CBP35 were isolated as  $\beta$ -galactoside binding proteins (45,48,62,72,79).  $\epsilon$ BP was isolated as a protein that bound to Immunoglobulin E (80). Mac-2 was initially identified as a cell surface antigen for thioglycolate-elicited peritoneal macrophages (77). LBP was identified as the major non-integrin laminin binding protein in macrophages, and subsequent sequencing showed its identity to CBP35/Mac-2 (and hence, the other members of the L-30 group) (74). Southern blotting analysis of genomic DNA suggests that there is a single gene which encodes these lectins (Jia, S. and Wang, J.L., unpublished results). Northern



blotting analysis has identified a major mRNA transcript of approximately 1.1-1.3 kB (75,80,96).

Sequence analysis of the cDNA clones corresponding to these lectins has demonstrated that they are composed of two distinct domains (Figure 2). The carboxyl terminal domain houses the CRD (7). This domain contains the 13 invariant residues found in the CRD regions of all the known S-type lectins. The L-14 lectins consist only of the CRD; however, the L-30 lectins contain a second effector domain. This domain is highly proline and glycine-rich, since it has eight contiguous repeats of the nine amino acid sequence PGAYPGXXX (76). The two domain structure of these lectins is further borne out by hydropathy analysis of the cDNA clones. The carboxyl terminal domain exhibits both hydrophobic and hydrophilic stretches, as characteristic of most globular proteins, whereas the amino terminus does not contain these characteristics.

### **Saccharide binding characteristics of the S-type lectins**

All the S-type lectins share a higher affinity of binding for Lac than for Gal. The critical determinants within the disaccharide are the hydroxyls at position 4 and 6 of Gal and position 3 of Glc, since substitution at any of these positions greatly reduces binding. The addition of an acetamido group at position 2 of the glucose in the Lac molecule (i.e. to yield N-acetyllactosamine) increases the binding affinity of the lectins for these sugars.

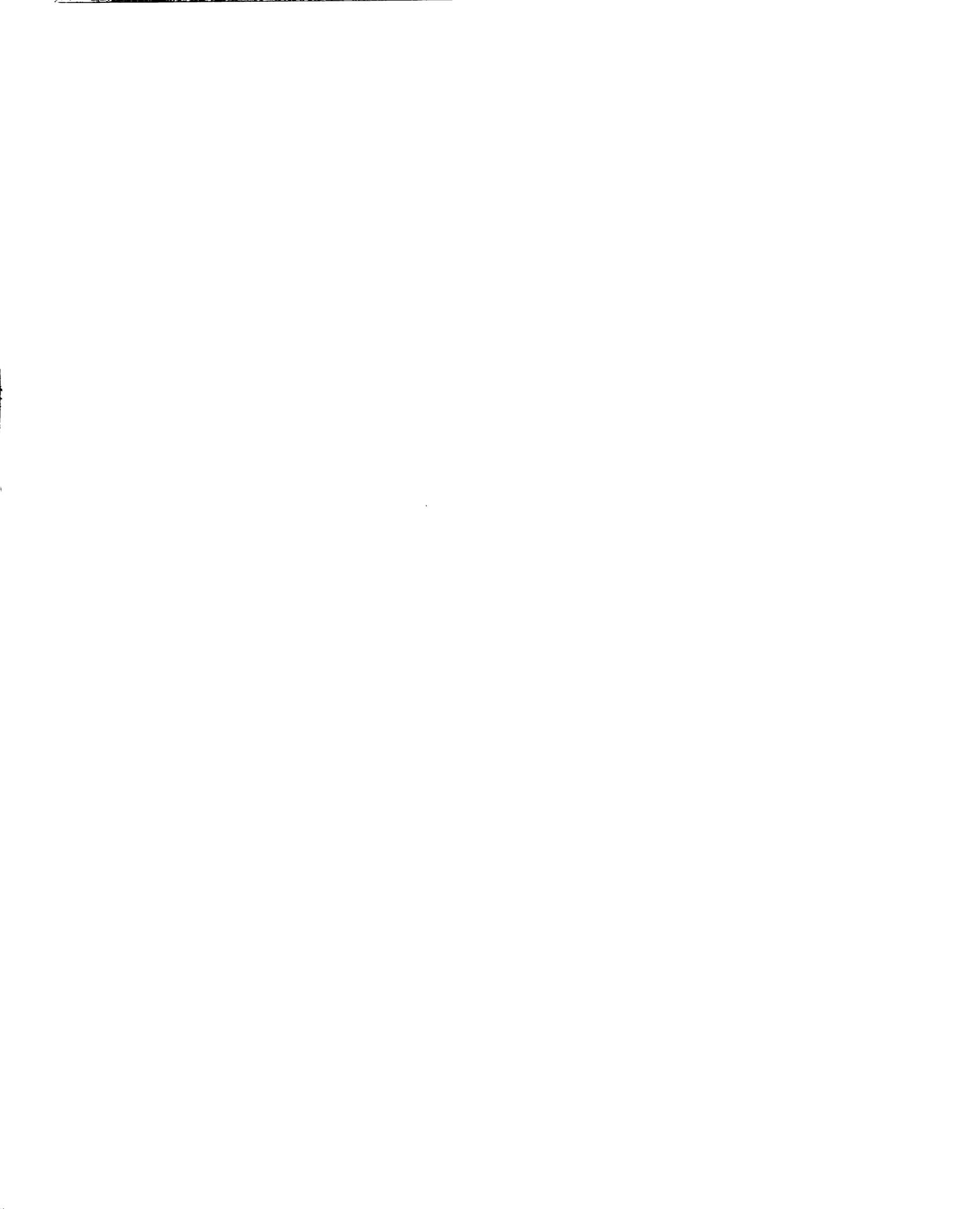
Site directed mutagenesis has been carried out on the bovine and human 14 kDa lectin to determine the critical residues for saccharide binding.

Hirabayashi and Kasai showed that, for the human lectin, neither the cysteine residues nor the conserved tryptophan residue are essential for saccharide binding (85). In contrast, Abbott and Feizi's results indicate that changing the tryptophan or cysteines either greatly reduces or eliminates the binding of the bovine lectin to Lac. In addition, deletion mutation analyses predict that almost the entire bovine lectin polypeptide chain is necessary for binding activity (86). The varying results obtained by the two groups point to the possible inefficacy of these techniques in analyzing structure-function relationships for the lectins.

The L-30 lectins can be distinguished from the L-14 lectins by their preferential binding (approximately 100-fold greater) for the blood group A tetrasaccharide (62). This results from the substitution of the GalNAc $\alpha$ 1- at the 3 position of the Gal of the lactose moiety. Within the L-30 family, it seems that the hamster lectin displays a slightly different binding specificity than do RL-29 and HL-29 (87). Whereas substitution at the C6 of the terminal galactose of reactive saccharides eliminates binding of the 29 kDa lectins, it appears to have little effect on the hamster lectin. Whether the differences in binding characteristics of these lectins is due to experimental conditions, species diversity, or protein isoforms is open to question.

#### **Subcellular localization of the L-14 and L-30 lectins**

The L-14 and L-30 lectins are found in the nuclear, cytosolic, and extracellular compartments (see references 88 and 240 for a review). The extracellular localization is difficult to explain since the proteins do not contain a



signal sequence for secretion. I will discuss the implications of their dual localization below.

The predominant portion of the L-14 proteins is found intracellularly, although reports vary regarding the nuclear and cytoplasmic distribution of the lectins. Using antibodies directed against CLL-1 and BHL-1 to label cryostat sections, staining is noted both in the nucleus and the cytoplasm (39,89). Immunoelectron microscopy has localized the 14 kDa lectin in the nucleus of the epidermal cells of the intermediate layer of chick embryonic skin (90). However, the lectin was not found in the nuclei of the basal cells of chick embryonic skin (90). Immunolocalization of the L-14 lectin in murine myoblasts shows that the lectin is only in the cytoplasm, and not in the nucleus (91). Finally, in neuronal cells, RL-14.5 is found in both the cytoplasm and the nucleus (50).

In addition to being found intracellularly, there is considerable evidence for the extracellular localization of the L-14 lectins. It has also been demonstrated that the extracellular localization can arise as a result of a specific stimulus. As an example, the L-14 lectin in *Xenopus laevis* skin tissue is found in the cytoplasm of granular and mucous gland cells. Upon the injection of epinephrine, the lectin is externalized using a novel secretory method (92). Similarly, the 14 kDa lectin in chick embryonic muscle is found intracellularly, but upon maturation of the organism, is exported from polynucleated myotubules (68). In an identical situation, the 14 kDa lectin in mouse cultured myoblasts is found both intracellularly and extracellularly. However, as the cells fuse to form multinucleate myotubules, the lectin is less abundant in the cytoplasm, and is found in vesicles in

the extracellular milieu (91). The export of both of these lectins is proposed to occur by a novel secretory mechanism, whereby the protein is packed into vesicles which "bud off". The process has been termed ectocytosis (69). An L-14 lectin (mGBP) has been purified from mouse embryonic fibroblast conditioned media and identified as a growth inhibitory substance (47), providing another example of a secreted S-type lectin.

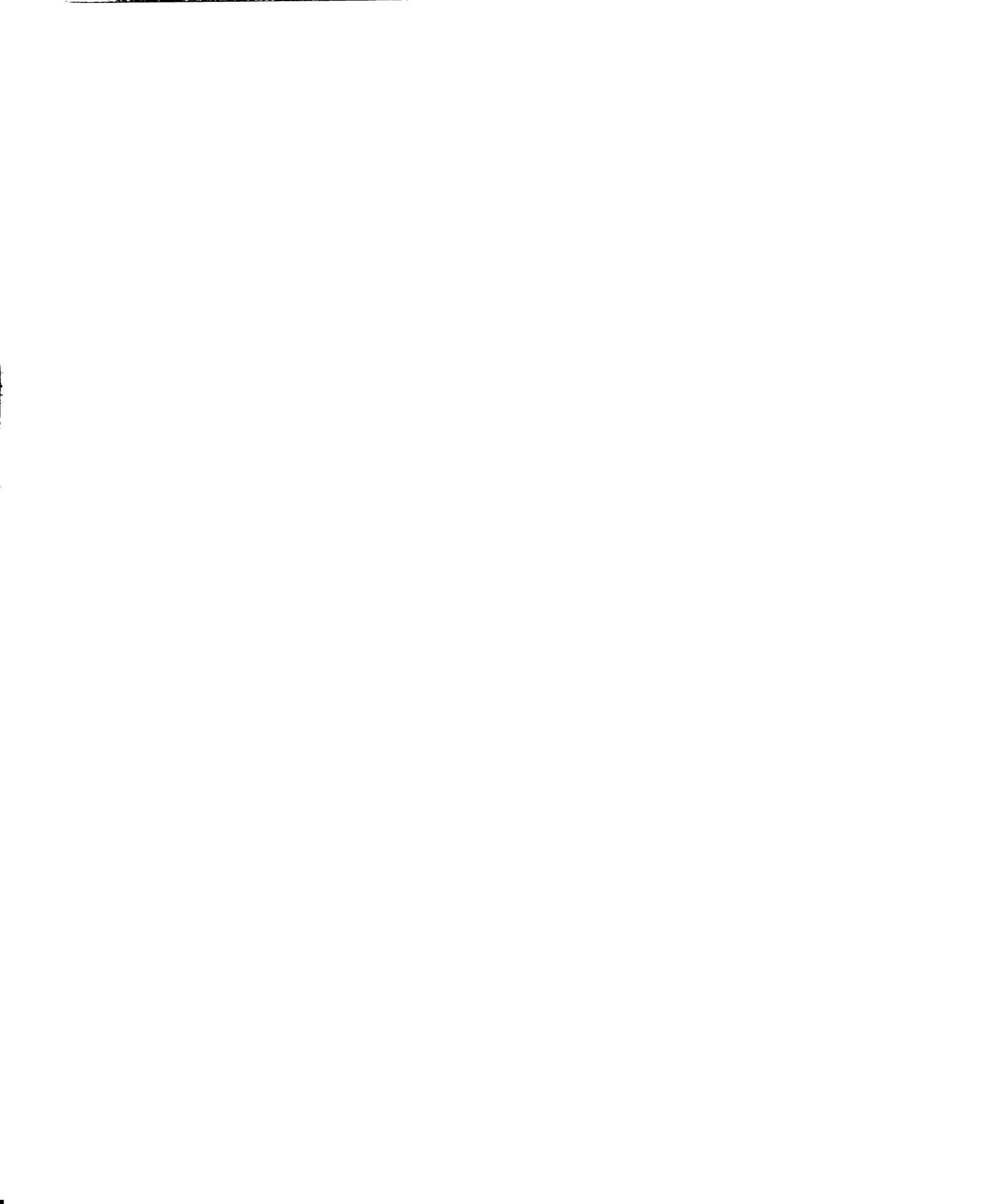
Similarly, the L-30 lectins are found on the cell surface and inside the cell. Since the proteins are identical or homologous to each other, presumably information pertaining to one member can also be applied to the others. CBP35 has been found mostly intracellularly, although a minor fraction is cell surface localized (93). Anti-CBP35 antibody staining of the proliferating cell shows a prominent staining of the nucleus and variable staining of the cytoplasm (94). Within the nucleus, there is a punctate staining pattern, which can be eliminated by prior treatment with RNase (95). Subcellular fractionation studies with  $\epsilon$ BP have also indicated that the majority of the protein is found in the cytoplasm and nucleus (96). Immunocytochemical studies have shown that RL-29 can be detected in both the cytoplasm and the nucleus (49,50).

The proteins Mac-2 and LBP were both identified by virtue of their cell surface localization. LBP can be isolated by cell surface iodination of murine macrophages, followed by laminin-Sepharose affinity chromatography (74). The Mac-2 antigen was isolated as a cell surface protein on thioglycollate-elicited macrophages (77). Subsequent experiments with the anti-Mac-2 monoclonal antibody indicate that the protein is also found in the nucleus of the P388D<sub>1</sub>,

macrophage cell line (J.L. Wang, unpublished observations). The increased cell surface expression of the L-34 lectin has been proposed to be involved in transformation and metastases of cells (46,72,97). Cells that exhibit the greatest metastatic potential have the highest levels of L-34 on the surface.

The S-type lectins fall into a growing group of proteins which are localized in two distinct milieus, the intracellular and extracellular compartments. These include: a) proteins with known nuclear function, such as SV40 large T antigen (98), adenovirus E1A gene product (99), probasin (100), and the La RNP identified by autoimmune anti-nuclear antibodies (101); b) members of the growth factor families as exemplified by the heparin-binding growth factors (102), and platelet-derived endothelial cell growth factor (103); and c) other proteins, including interleukin  $1\alpha$  and  $1\beta$  (104), yeast mating  $\alpha$ -factor (105), and CAP-50 (106), which is a member of the annexins.

These proteins can be divided into two groups, those that have a signal sequence for extracellular transport, and those that do not. Within the former group are probasin, platelet-derived growth factor (PDGF), and the product of the mouse *int-2* gene, which is a member of the fibroblast growth factor (FGF) family. Probasin is a rat prostatic protein which is found in secretions and in the nucleus of prostatic epithelial cells. The dual localization of probasin occurs as a result of alternative AUG-codon usage during translation, with the protein derived from the upstream AUG-codon containing a signal sequence (100). In the case of the Int-2 oncoprotein, the N-terminally extended protein initiated at a CUG-codon is nuclear, while the downstream AUG-initiated product is found in the secretory



pathway (107). Alternative splicing of the transcript for the PDGF A-chain determines whether the protein will be localized to the nucleus, or contain a signal sequence for secretion (108).

However, the basic fibroblast growth factor (bFGF) and the interleukins  $1\alpha$  and  $1\beta$  have not been shown to contain a signal sequence. It has been postulated that these proteins can be externalized via a mechanism of exocytosis that is independent of the ER-Golgi endomembrane secretory pathway (104,109). This conclusion is based on results obtained from studies using drug inhibitors which are specific for the ER-Golgi pathway. Their nuclear localization is mediated through nuclear targeting signals (see section below on nuclear transport). The significance of the cell surface localization of the FGF has been questioned. Acidic FGF mutant molecules, lacking the nuclear targeting signal, failed to induce DNA synthesis and cell proliferation in target cells, even though they could initiate membrane events such as tyrosine-phosphorylation (110). Thus, it has been suggested that acidic FGF may ultimately act as an intracellular, nuclear-translocated polypeptide mitogen, therefore obviating the need for a signal sequence.

A possible explanation for the export of proteins independently of the classical secretory pathway may lie in the discovery of the ATP-dependent translocators (111). In yeast *S. cerevisiae*, the product of the *STE6* gene has been shown to be a key factor in the export of the mating  $\alpha$ -factor (105). Mutants lacking the STE6 protein fail to export  $\alpha$ -factor, while the mutants of the normal secretory pathway (the *sec* mutants) have no effect. The STE6 protein displays

significant homology with the P-glycoprotein, which is involved in multi-drug resistance (reviewed in reference 112), as well as with bacterial permeases (105). The STE6 protein shares about 60% amino acid identity with the mammalian *MDR1* gene product. *MDR1* is homologous to bacterial permeases and contains an ATP-binding domain, suggesting that *MDR1* functions as an ATP-driven pump. The similarity of STE6 with *MDR1* implies that it, too, may act as a pump for the  $\alpha$ -factor. There may be similar translocation systems for the export of proteins such as the interleukins and the FGFs.

### **The bifunctional nature of lectins**

It has become increasingly obvious that the physiological role of many lectins extends further than just the binding of carbohydrates. This has been demonstrated for the C-type lectins (Figure 1), as well as for the S-type lectins (Figure 2). I will discuss two examples below.

#### **a) Selectins**

One of the most exciting discoveries in the field of lectins has been that of the family of cell-adhesion proteins called selectins. Their nomenclature is as follows: a) L-selectin (peripheral lymph node homing receptor), also known as gp190<sup>MEL</sup>, LAM-1, LECAM-1, LEC.CAM-1, DREG.56, TQ-1, and Leu-8; b) E-selectin, also known as ELAM-1; and c) P-selectin, also known as PADGEM, GMP-140, and CD62. These proteins were initially isolated as important mediators of adhesion of leukocytes to the blood vascular compartment. The presence of a Ca<sup>2+</sup> dependent CRD within these molecules was discerned only

after they were cloned and sequenced (10,11,12). Structurally, these proteins are very similar, with an identical arrangement of the C-type lectin domain attached to an EGF-like and a complement binding domain. In addition, there is a signal sequence, a transmembrane domain, and a cytoplasmic tail. It has been proposed that the selectin gene arrangement arose from exon shuffling mechanisms (113).

L-selectin is a surface antigen on lymphocytes which facilitates their binding specifically to lymph node endothelium during lymphocyte circulation. It is, hence, commonly called the lymphocyte homing receptor since it allows for selective trafficking of particular lymphocyte populations to specific sites. P-selectin is a glycoprotein found on the surface of platelets and endothelial cells after stimulation by thrombogenic agents, thereby allowing these cells to bind to neutrophils and monocytes at areas of tissue injury. The E-selectin is generated by endothelial cells as a result of inflammatory agents, and promotes adhesion of neutrophils, monocytes, and a subpopulation of lymphocytes to the endothelial cells.

Tentative carbohydrate ligands have been identified for these selectins (see references 114,115 for reviews). The SSEA-1/Lewis<sup>x</sup>/CD15 antigen has been implicated as the binding determinant for P-selectin. E-selectin seems to prefer a sialylated Lewis<sup>x</sup> structure. The endothelial ligand for L-selectin appeared to be more elusive, until a recent discovery by Lasky *et al.* (116). They have cloned a cDNA for a sulfated 50 kDa glycoprotein, Sgp50, which has a mucin-like domain. The protein backbone on this molecule may serve as a scaffold on which to present the carbohydrates.

**b) The S-type L-14 lectins**

Although there is a paucity of information on the functional ligand for the L-14 lectins, there have been several reports implicating these lectins in growth regulation. Yamaoka *et al.* have reported that the overexpression of a rat 14 kDa galactose-binding-protein (GBP) causes transformation of murine fibroblasts (117). In fact, they present information suggesting that the GBP protein is identical to the growth regulatory factor TGF $\gamma$ 2. Furthermore, the growth stimulatory activity of TGF $\gamma$ 2/GBP is not inhibited by the addition of  $\beta$ -galactoside, implying that the two activities are distinct. The idea that GBP may be a transforming growth factor, or act like one, is strengthened by the observation by Wells and Malluci that a galactoside binding protein (mGBP) secreted by mouse embryonic fibroblasts (MEF) is a cyostatic growth factor (47). mGBP, when added to MEF *in vitro*, inhibited their growth in the G<sub>0</sub> phase of the cell cycle. The inhibitory activity was not reversed by the addition of lactose. This may be evidence that L-14 is a multifunctional molecule, much like transforming growth factor which can both stimulate and inhibit cell growth.

**CARBOHYDRATE BINDING PROTEIN 35**

CBP35 was originally isolated from mouse lung tissue as a monomeric protein of M<sub>r</sub> 35,000 by its affinity for galactose containing glycoconjugates (45). Crittenden *et al.* demonstrated that the protein is widely expressed in a variety of species and tissues (58). It was also shown that the protein is more highly expressed in embryonic tissue than in adult tissue. CBP35 has been shown to be

identical or homologous to all of the other members of the L-30 family (Table IV) (88).

Immunolocalization data have shown that the majority of CBP35 is found intracellularly, although a small fraction is also found on the cell surface (93). The control of the intracellular localization of CBP35 will be discussed further below.

CBP35 is speculated to be a component of the heterogenous nuclear ribonucleoprotein complex (hnRNP). This is based on the following observations (76,95): a) CBP35 is released from permeabilized nuclei by treatment with RNase A, but not by treatment with DNase I; b) CBP35 is found in the same density fractions as the hnRNP proteins on cesium sulfate gradients ( $\sim 1.30$  g/ml); c) CBP35 co-isolates with hnRNP proteins by sucrose gradient centrifugation (40 S); d) fractionation of nucleoplasm on a galactose affinity column yields CBP35 and a set of polypeptides whose molecular weights match those for the hnRNP proteins; and e) sequence analysis of the cDNA clone for CBP35 suggests a homology between CBP35 and some hnRNP proteins. The hnRNP proteins are thought to aid in the processing and transport of mRNA. Further support for the involvement of CBP35 in this process comes from the evidence that antibodies directed against CBP35 and galactose-containing saccharides perturb the *in vitro* splicing of pre-mRNA (Patterson, *et al.*, unpublished data).

### **Structure of the CBP35 gene and protein**

There is a single gene for CBP35 in the mouse genome. The gene encompasses approximately 9 kb of genomic DNA, and is comprised of five exons

and four introns (Jia, S., and Wang, J.L., unpublished observations). The upstream promoter region contains the TATA and CCAAT sequences, as well as the serum response element sequence about 200 nucleotides upstream of the transcription start site. A polyadenylation signal has been delineated in the 3' untranslated region.

Since the CBP35 protein has been only observed as a monomeric species, its hemagglutination property is difficult to explain. The stoichiometry of saccharide binding by the CRD was recently determined by equilibrium dialysis, using [<sup>14</sup>C]-lactose and recombinant L-30 expressed in and purified from *E. coli*. The results indicate one Lac binding site per M<sub>r</sub> 30000 of protein (Knibbs, R., manuscript in preparation, and ref.118). Woo *et al.* suggest that CBP35/Mac-2 forms intermolecular dimers using the single cysteine residue (119). However, since their study used a non-reducing SDS-PAGE system, the question arises whether the dimer that they detect is actually an artifact of the analysis process (120). The IgE binding protein (εBP) has also been shown to form oligomers using crosslinking reagents (118). Hsu *et al.* suggest that the N-terminal domain of the εBP may contribute to the multivalency of the molecule, possibly by engaging in protein-protein interactions (118). These results are refuted, however, by the observation that CBP35 does not form oligomers under non-denaturing conditions as determined by HPLC gel filtration columns and equilibrium sedimentation centrifugation (Anandita, *et al.*, manuscript in preparation).

**Ligands for CBP35/Mac-2**

Mac-2, the CBP35 homolog, has been shown to bind to two intestinal epithelial glycoproteins, M2BP-1 ( $M_r$  98 kDa) and M2BP-2 ( $M_r$  70 kDa). These proteins were isolated from a human adenocarcinoma cell line by virtue of their association with Mac-2 (121). The interaction between these proteins and Mac-2 is mediated through the carbohydrate binding portion of Mac-2. The M2BP-1 protein is secreted into the media, leading the authors to speculate that it is an extracellular ligand for the surface antigen Mac-2. However, no glycosylated nuclear ligand has yet been found for CBP35/Mac-2. It remains to be seen if the intracellular pool of CBP35 mediates its function through its sugar binding activity.

**Proliferation dependent localization and expression of CBP35**

Initial studies using indirect immunofluorescence techniques showed that quiescent or serum starved 3T3 cells exhibited very low levels of CBP35. The protein was found only in the cytoplasm of these cells. On the other hand, proliferating or serum fed cells demonstrated a significant increase in the levels of CBP35, and in this case the protein was localized largely in the nucleus (94). Upon serum stimulation, the protein levels increase in the  $G_1$  phase of the cell cycle, before the onset of DNA synthesis. These observations have been corroborated by immunoblotting. The results at the protein level prompted us to examine the transcriptional regulation of the CBP35 gene. The conclusions are presented in Chapter II of this thesis. One of the reasons why CBP35 may be stimulated by the addition of serum is the presence of a serum response element



(SRE) in the upstream regulatory region of the gene. The SRE has been shown to operate in the serum-mediated activation of the *c-fos* and  $\beta$ -actin genes (reviewed in references 122,123).

Two dimensional gel electrophoresis analyses of quiescent and proliferating populations of cells revealed that CBP35 exists as two isoelectric variants in the cell. The unmodified polypeptide has an isoelectric point (pI) of 8.7, whereas the singly phosphorylated derivative has a pI of 8.2. The pI 8.2 form is found both in the cytoplasm and the nucleus. The pI 8.7 form is restricted to the nucleus. More interestingly, quiescent cells only express the 8.2 variant. But, when cells are proliferating, the levels of the 8.7 variant increase dramatically and it is entirely nuclearly localized (124).

A similar situation is encountered with SL66 normal human fibroblasts. Using immunofluorescence, immunoblotting, and two dimensional electrophoresis analyses with these cells, it has been revealed that the expression of CBP35 is dramatically different in young passage (passage 11) as compared to old passage (passage 31-35) cells. When young SL66 cells are serum stimulated, there is a rise in the levels of CBP35. Both isoelectric variants are expressed in these cells and the pI 8.7 form dramatically increases upon serum addition. On the other hand, old cells did not exhibit any increase in CBP35 levels, and they also did not contain any pI 8.7 form (84). The absence of any CBP35 induction in old cells raises the intriguing possibility that certain induction mechanisms are abrogated in senescence. It has been suggested that a similar fate may occur for the *c-fos* protein (see below).

The significance of the two isoelectric variants remains to be elucidated. However, the compartmentalization of the two forms suggests that the phosphorylation may serve as a partitioning signal. I will discuss this aspect further below in the section on nuclear transport of proteins.

## THE REGULATION OF EXPRESSION OF C-FOS

### Introduction to Primary Response Genes

It is generally acknowledged that control of vertebrate cell proliferation is exerted largely in the  $G_1$  phase of the cell cycle. The complex cellular process of proliferation is initiated by an interaction between extracellular factors and specific cell surface receptors (125). The cytoplasmic activation signals thus engendered cross the nuclear membrane and alter the expression of a set of genes known as immediate early (IE) or primary response genes (PRG). The PRGs are characterized by the following: a) they do not require *de novo* protein synthesis for induction; and b) their rapid and transient induction occurs in a wide variety of cell types (reviewed in references 126,127).

Within the ranks of the PRGs are genes which encode structural proteins, transcription factors, and proto-oncogenes (125). Many cDNA clones for PRGs have also been identified by subtractive hybridization screening, although their products have not yet been characterized (128,129,130). Several of the proto-oncogenes have been well studied, although their exact function is unknown. The product of the *c-fos* gene has been identified as a transcription factor.

The proto-oncogene *c-fos* is a paradigm for the regulation of proto-oncogenes. In addition, the CBP35 gene and protein share several similarities with the *c-fos* gene and FOS protein: a) they both satisfy the criteria for a PRG; b) CBP35 contains a DNA promoter element which fits the consensus sequence for the serum response element, also identified in *c-fos*; c) both are nuclear proteins; and d) quiescent and senescent cells express negligible levels of FOS and CBP35. I shall, therefore, discuss the expression of *c-fos* with respect to the cell cycle.

There are four members of the *fos* family: *c-fos*, *fosB*, *fra-1*, and *fra-2* (125). The FOS proteins are found associated with the JUN proteins. Proteins within the *jun* family include *c-jun*, *junB*, and *junD* (125). The FOS and JUN proteins form heterodimers, through leucine zippers, in all possible combinations (131,132). The resulting dimer then forms the AP-1 transcription factor (131,132). The FOS and JUN proteins can also dimerize with proteins from the CREB/ATF family using the leucine zipper (133). DNA binding by the AP-1 factor occurs by a region adjacent to the leucine zipper which is rich in basic amino acids. Those proteins containing the dual motifs of the basic amino acids and the leucine zipper have been clumped into the "bzip" family of transcription factors (134). Although the FOS/JUN proteins are generally regarded as transcription site AP-1 binding proteins, they can also bind to the TRE (TPA response element) and the CRE (cAMP response element) sites (133,135).



**Kinetics of the expression of *c-fos***

The *c-fos* gene product is a M<sub>r</sub> 55,000 nuclear phosphoprotein (p55<sup>*c-fos*</sup>) (136,137). The mRNA transcript has a size of 2.2 kb. Studies with the expression of *c-fos* show that the levels of the mRNA and protein are transiently high in proliferating cells (136,137). There is increased expression in placental and extraembryonal tissues (138). Experiments with cell cultures have established that *c-fos* mRNA is detectable within 15 minutes of serum addition to quiescent cells, thus making this the earliest PRG identified so far (136,137). At least part of the increase in the mRNA can be attributed to an increase in the transcription rate (139). The peak of expression of the mRNA at about 60 minutes is followed by a rapid decay (136). The FOS protein has been reported to have a  $t_{1/2} \sim 2$  hours (136). The presence of multiple AUUUA sequence motifs in the 3' noncoding region of the *c-fos* mRNA probably accounts, in part, for its instability (140). Removal of these sequences has been shown to cause transformation by *c-fos*. Senescent fibroblasts exhibit very low levels of the *c-fos* mRNA, even when serum stimulated (141). The results obtained by Seshadri and Campisi indicate that the *c-fos* gene is subject to transcriptional repression in these cells.

**Superinduction of the *c-fos* mRNA**

The *c-fos* mRNA exhibits a phenomenon, common to all the PRGs, known as superinduction. Although normally very transient, the level of these transcripts is elevated when cells are incubated with protein synthesis inhibitors. Several reasons have been postulated to explain this occurrence. First, it is possible that

the labile transcripts are stabilized. Indeed, the *c-fos* mRNA half life increases from 9 minutes to several hours (140). This may be accomplished by the loss of labile RNAses (142), by the shielding of mRNAs which stay trapped on polysomes (128), or because the concurrent translation of *fos* is necessary for mRNA degradation (142). Second, it is possible that the continued synthesis of the FOS protein is required for transcriptional shut-off of the gene. This implies that the FOS protein has an autorepressor function. It has been proposed that the autorepression may occur through the serum response element (143). Finally, a third explanation is that labile repressors normally keep the *c-fos* gene inactive in quiescent cells. When protein synthesis ceases, the gene is available for transcription. Although no such repressors have been identified, there are hints that such an effect may be mediated through the serum response element-serum response factor complex (144). In a recent study by Edwards and Mahadevan on the possible mechanism of superinduction of *c-fos*, the notion of the labile repressor has been questioned. Using several protein synthesis inhibitors, they find that anisomycin and cycloheximide cause the transcriptional induction (as determined by nuclear run-on assays) of *c-fos* mRNA (145). Thus, in addition to inhibiting protein synthesis, these compounds also act as nuclear signalling agonists.

### **The *c-fos* promoter**

Analysis of the upstream promoter region of the *c-fos* gene reveals multiple regulatory sites. The serum response element (SRE), located at -300 bp, has been

hypothesized to be a target of many growth factors (reviewed in reference 123). There is a cAMP response element (CRE) at -60 bp which is regulated by elevated levels of cAMP or  $\text{Ca}^{2+}$  (146). A TPA response element (TRE) is located at -295 bp (147). A region at -345 bp is thought to be involved in induction by platelet derived growth factor (PDGF), although no protein binding factor for the site has been seen (148). A site called the SRE-2, at -276 bp, is involved in induction by nerve growth factor (NGF); the proteins binding to this region are different from those binding to the SRE (149). In addition to the positive regulatory elements, the retinoblastoma protein down-regulates *c-fos* expression by binding to a site at -90 bp (150). There is also evidence that cooperation between the SRE and a fosATF/AP1 sequence downstream from the SRE leads to either repression or activation from the *c-fos* promoter, depending on the growth state of the cell (151). Finally, an intragenic sequence at the exon1-intron1 boundary is involved in the blocking of transcription elongation (152). Interestingly, many of the regulatory sites to which the FOS protein binds are present in the *c-fos* gene.

The SRE in the *fos* promoter has been studied in great detail. It is a 20-bp region of dyad symmetry which is the site for the binding of the serum response factor (SRF). The 14-bp inner core element of the SRE,  $\text{CC(A/T)}_6\text{GG}$ , is sufficient for the binding of the SRF and the induction of serum-stimulated transcription (153). The *zif268* gene has 4 SREs, none of which show symmetry outside of the core, but each of the four confer serum inducibility on reporter genes (154). The outer palindromic arms of the SRE may enhance the binding of

the SRF. Alternatively, these arms may also serve as binding sites for other regulatory factors, which may act in concert with the SRF.

The SRF itself is also a member of the PRG group (155). SRF is a 62-67 kDa protein, depending on its state of phosphorylation, which binds the SRE as a dimer. The phosphorylated SRF has a higher binding affinity for DNA (reviewed in reference 238). The protein is also modified by an O-linked N-acetylglucosamine moiety (156). The p62/ternary complex factor (p62/TCF) interacts with the SRF when it is bound to the SRE, thus enhancing the binding between the SRE and the SRF (156). In yeast, another protein called the SRF accessory protein-1 (SAP-1) is recruited to the SRE-SRF complex (157). Although SAP-1 does not seem to be identical to p62/TCF, it does appear to be structurally and functionally related.

### **Factors controlling the induction and function of *fos***

As indicated by analysis of the promoter, *c-fos* is subject to both negative and positive regulation. It has been demonstrated that in quiescent cells the transcription of the *c-fos* gene is induced by the addition of excess copies of the *fos* promoter element, thus suggesting that some negative regulatory factor can be competed out. Likewise, in proliferating cells, addition of the same element reduces *c-fos* transcription, presumably by competing out a positive regulatory factor (158).

The activation of *c-fos* through the SRE can occur by at least two distinct intracellular pathways, one that is protein kinase C (pkC) dependent, and another

that is pkC-independent. There is also a third cAMP-dependent pathway which seems to act independently of the SRE (159). Therefore, the SRE is necessary and sufficient for activation by protein kinase-C, but not for activation by cAMP.

A specific inhibitor of FOS/JUN proteins has been identified from nuclear and cytoplasmic extracts (160). IP-1 associates with FOS and JUN and prevents the proteins from binding to DNA. It is unclear how IP-1 exerts its control over FOS/JUN activity. Phosphorylation of IP-1, possibly by protein kinase A, causes the inhibitor to dissociate from the proteins. This can be seen by an increase in AP-1 activity. The presence of a cytoplasmic inhibitor of FOS implies that cells always contain a basal level of the protein. Although this has not been rigorously determined, Bravo *et al.* have reported that *c-fos* is inducible at low levels throughout the cell cycle (161). This suggests that IP-1 may have a specific role in sequestering FOS until the appropriate time. Finally, another mode of regulation of FOS may occur by the formation of disulfide bonds within the basic residues of a FOS-JUN dimer. This prevents it from binding to the TRE sequence (162,163).

### **Regulation of transcription by the PRG**

Many genes are known to be either activated or repressed by FOS, FOS/JUN, or the FRA proteins (reviewed in ref. 125). It is of great interest to examine the regulatory activity of FOS on other genes as well as on itself. The repressor activity of FOS does not require it to be a heterodimer. However, the repressor activity is dependent upon the presence of the SRE sequence in the target promoter, and the phosphorylation of the serine residues in the carboxyl

terminus of FOS (164,165,238). Therefore, the PRGs not only regulate a wide variety of target genes, but also themselves. This indicates the complexity of the cross talk that occurs at molecular levels in transcriptional regulation.

## **NUCLEAR TRANSPORT**

### **General characteristics of nuclear transport**

Proteins and RNA enter and exit the nucleus in a specific and regulated manner. In this review I will specifically discuss the import of macromolecules into the nucleus.

The nuclear envelope, which renders the nuclear compartment distinct from the cytoplasm, is a complex assembly consisting of inner and outer nuclear membranes, nuclear pore complexes, and the nuclear lamina (reviewed in reference 166). The outer membrane of the envelope is contiguous with the endoplasmic reticulum (ER), and the perinuclear space created by the outer and inner membranes is continuous with the ER. The inner and outer membranes are joined at the nuclear pore complex.

It has been suggested that nuclear pores can accommodate the passive entry of proteins  $\leq M_r$  20-40 kDa. Larger proteins enter the nucleus by a facilitated process (reviewed in references 167,168,169). Facilitated nuclear transport has the following characteristics: a) it requires the presence of a nuclear localization signal (NLS) which is both necessary and sufficient for transport; b) presumably requires a receptor; c) requires ATP; d) it is temperature dependent;

and e) proteins enter in a folded state.

The diffusion of proteins through the pore complex is an open issue. There are data that suggest that nuclear proteins, regardless of size, enter the pore complex in a specific manner. As an example, histone H1 ( $M_r$  21 kDa) and histone H2B ( $M_r$  13.8 kDa), despite their small size, enter the nucleus using mechanisms which are distinct from simple diffusion (170,171). It has been also been shown that some proteins can enter the nucleus by a "piggyback method" where they are co-transported with an NLS-containing protein (172).

Nuclear import can be divided into two discrete steps: a) a relatively rapid binding of proteins to the nuclear pore; and b) a slower translocation step into the nucleus (173). The translocation step is sensitive to the lack of ATP, and is inhibitable by the inclusion of the N-acetylglucosamine-specific lectin wheat germ agglutinin (WGA) in the transport reaction (174).

### **Nuclear localization signal**

The nuclear localization signals of many nuclear proteins have been compiled and compared in several review papers (168,175). There appears to be no consensus sequence. However, there are certain characteristics which are obviously specific enough to ensure that only nuclear proteins enter the nucleus. These are as follows: a) they are short sequences, usually no more than 8-10 amino acids; b) the NLS may be contained within one sequence or may be divided into a bipartite signal; c) they contain a high proportion of basic amino acids, usually lysine and arginine; d) they can occur at any site in the polypeptide; e)

NLSs are retained following transport; and f) a protein may contain more than one NLS. The well-characterized NLS of the SV40 large T antigen, PKKKRKV, has been regarded as the classic unipartite signal. The lys<sup>128</sup> residue in the SV40 NLS is the critical residue in determining whether the NLS is sufficient for transport (176,177). On the other hand, nucleoplasmin, a major nuclear protein of *Xenopus* oocytes and embryos, requires a bipartite signal for import (178,179). The bipartite signal is located at the terminus of a 16 amino acid stretch in the carboxyl end of the protein. The recent elucidation of an NLS for an agrobacterium protein shows that it is homologous to that of nucleoplasmin (180). This indicates that the NLS structure has been conserved in evolution.

### **The nuclear pore complex**

The nuclear pore complex (NPC) is an aqueous channel which acts as a molecular sieve that spans the nuclear envelope. The functional diameter of the pore is 7-10 nm, thus allowing for the diffusion of ions and small molecules (166). However, particles of sizes up to 280 Å have been shown to traverse through the NPC. Studies by Feldherr and Akin have suggested that the permeability of the NPC may be linked to the physiological state of the cell (181,182). Using nucleoplasmin-coated colloidal gold particles, they have demonstrated that proliferating cells can transport particles of 230-250 Å, whereas growth arrested cells can only transport particles of 160-200 Å. It has been estimated that the nuclear envelope of a eucaryotic cell contains approximately 2000-4000 pores (183).

The architecture of the pore complex consists of three prominent substructures: nuclear and cytoplasmic rings or annuli, central spokes, and a central plug (reviewed in reference 166). These structures are complexed with proteins to form the pore complex. The mass of the nuclear pore complex has been estimated at 120 mDa.

Two major classes of pore proteins have been identified. These are: a) the integral membrane protein gp210; and b) the nucleoporins. The integral membrane protein gp210 contains N-linked high mannose sugar modification. Primary structure and antibody epitope mapping studies suggest that gp210 spans the nuclear membrane once. A short region in the carboxyl terminus protrudes towards the pore, with the rest of the protein being found in the perinuclear space (184). The microinjection of anti-gp210 antibody into the ER, which is continuous with the perinuclear space, drastically reduced nuclear import (185). This has raised the question whether the NPC can be regulated via the ER.

Nucleoporins are proteins containing the O-linked N-acetylglucosamine (GlcNAc) sugar which are found on both the cytoplasmic and nuclear faces of the NPC (reviewed in reference 186). Three nucleoporins, p62 (187), the 110 kDa yeast NSP1 (188), and the 130 kDa yeast NUP1 (189) have been studied in some detail. It is uncertain whether the yeast nucleoporins are modified with O-GlcNAc (189). A monoclonal antibody developed against mammalian p62 cross-reacts with the yeast nucleoporins (189,190). Since there is no primary sequence homology between p62 and NUP1, it is assumed that the anti-p62 antibodies recognize some secondary structural feature.

Perturbation of transport by anti-nucleoporin antibodies and the GlcNAc-specific lectin WGA has implicated nucleoporins in the transport process. The addition of monoclonal antibodies against nucleoporins blocked nuclear import and RNA export (191). WGA inhibits the facilitated import of proteins into the nucleus, although diffusion through the pores is unaffected (192). Finlay and Forbes have demonstrated that WGA can deplete the nuclear transport capabilities of an extract, but that reconstitution with the WGA-bound fraction restores nuclear transport (193). The analysis of the WGA-bound fraction from rat cytosol yields three proteins of M<sub>r</sub> 62 kDa, 58 kDa, and 57 kDa, which seem to be a part of a ~600 kDa complex (194).

#### **Cytosolic factors are required for nuclear transport**

The nuclear pore complex does not contain all the components necessary for transport, as evidenced by the inability of isolated nuclei to support nuclear import with fidelity (195,196). Using an *in vitro* system involving digitonin-permeabilized cells, Adam *et al.* have shown that transport requires soluble cytosolic factors (196). These factors are found in both nucleate and anucleate cell extracts, and cannot be pelleted by centrifugation at 100000 x *g*. No RNA component is involved in the cytosolic factor, since the extract can be treated with micrococcal nuclease with no adverse effects. Treatment of the cytosol with the sulfhydryl alkylating agent N-ethylmaleimide (NEM) leads to an inhibition of transport. Also, cytosolic extracts from one species can support *in vitro* nuclear transport with another cell system, indicating that the factors are not species-

specific.

Newmeyer and Forbes have isolated two factors from *Xenopus* oocyte extract which no longer support nuclear transport after NEM treatment (197). The factors have been named NIF-1 and NIF-2. The factor NIF-1 is required for the binding of the nuclear protein to the pore, and thus may act as a cytoplasmic carrier protein that binds to the NLS. Using the nuclear pore glycoproteins as an affinity matrix, Sterne-Marr *et al.* have depleted cytosol of transport factors such that nuclear transport was reduced  $\sim 80\%$  (198). The factors did not exhibit binding to WGA, and also, were not sensitive to NEM inactivation. These factors may also act as docking proteins with the nuclear pore complex.

In agreement with the two step model of nuclear transport, Moore and Blobel have fractionated oocyte transport extract on a DE-52 anion exchange column and isolated two fractions, A and B, both of which are required for efficient transport (199). Fraction A seems to be involved in the NLS recognition; addition of fraction A alone leads to accumulation of the protein around the nuclear periphery. Fraction B is necessary for translocation into the pore, and cannot by itself support nuclear transport. Fraction A is NEM-sensitive, whereas fraction B is not affected by NEM. Thus, the data from the various laboratories indicate that there are several different cytosolic components which are involved in the transport of proteins into the nucleus.

Finally, hsp70 (heat shock protein 70) and its cognate hsc70 have been shown to be required for the transport of proteins into the nucleus (200). The hsp70 is not NEM sensitive.

Hsp70 has been found to be necessary for the translocation of proteins into mitochondria and the endoplasmic reticulum (201,202). In addition, transport into the ER and mitochondria requires signal sequences and receptors. It has also been documented that NEM inactivates transport into the Golgi (203). Together, these requirements suggest that there are certain basic principles for the translocation of proteins within the cell (for a commentary, see reference 236).

### **NLS binding proteins**

The existence of the NLS binding proteins (NBPs) has been inferred from the demonstration that nuclear transport is a saturable event (204,205). The presence of several of these adaptor molecules or NBPs has been substantiated by their purification using the SV40 T antigen NLS as an affinity matrix (204,206,207). The NBPs may either stay fixed to the nuclear pore, or alternatively, they may act as shuttling molecules between the cytoplasm and nucleus (167,169,208).

A mammalian 55 kDa protein has been found to stimulate *in vitro* transport while demonstrating a sensitivity to N-ethylmaleimide (204,207). It is unclear whether the 55 kDa protein is related to the cytosolic factor NIF-1. A 70 kDa protein, isolated from yeast, has been localized to the nuclear envelope, and it has been suggested to interact with the nucleoporin NSP1 (169,209,210). Proteins which are immunologically cross-reactive with the yeast 70 kDa protein are also found in *Drosophila*, HeLa cells, and *Z. mays*. Stochaj and Silver report that these eucaryotic NBPs are phosphorylated, and this modification is necessary

for binding by the NLS (211).

Lee and Mélése have identified a 67 kDa protein from yeast, encoded by the *NSR1* gene, which not only binds to the SV40 NLS, but also is localized to the nucleus and the nucleolus (212,213). The NSR1 protein also contains two RNA recognition motifs, the significance of which is unclear. Meier and Blobel have reported the identification of a 140 kDa protein, Nopp140, which binds to NLSs and is also found in the nucleolus (214). The binding of Nopp140 to NLSs requires that the protein be phosphorylated (239). The dual localization of the p67 and p140 proteins has led to the debate of whether non-nucleolar proteins use the nucleolar/ribosomal protein transport pathway. This could be a demonstration of transport efficiency, since the transport of ribosomal proteins constitutes the majority of all trafficking into the nucleus/nucleolus. Following a general transport of all nuclear proteins into the nucleolus, there could be a re-routing of the proteins into the appropriate nuclear compartment.

#### **Alternative methods of nuclear transport**

Studies with transport of the spliceosomal U small nuclear ribonucleoproteins (snRNPs) have revealed the existence of other methods of nuclear import. The transport of these molecules into the nucleus cannot be competed by the SV40 T antigen NLS, thus indicating that they use a distinct pathway (215,216). The assembly pathway of the U snRNPs involves the export of the newly synthesized RNAs into the cytoplasm where they are complexed with proteins, either of the Sm class or the U-specific proteins. The initial export of

the snRNA from the nucleus into the cytoplasm may be due, in part, to the presence of a 7-methylguanosine ( $m^7GpppG$ ) cap at the 5' end (217). After the snRNA is complexed with proteins in the cytoplasm, the  $m^7GpppG$  structure is hypermethylated to form the 2,2,7-trimethylguanosine ( $m_3GpppG$ ) cap. All U snRNA species, except U6 snRNA, carry the trimethylguanosine (TMG) cap; the U6 snRNA has a  $\gamma$ -monomethylphosphate at its 5' end. The Sm proteins and the TMG cap have been shown to make up a bipartite NLS for some of the U snRNPs (218,219).

Using microinjection of the U snRNAs in oocyte nuclei, Michaud and Goldfarb have hypothesized that these snRNPs are imported into nuclei by three different pathways (220): a) the karyophilic pathway employed by SV40 large T antigen; b) the TMG dependent pathway; and c) a pathway distinct from the other two. The import of U1, U2, U4, and U5 snRNPs can be inhibited by the injection of excess TMG cap, but not by excess SV40 T antigen NLS (220). On the other hand the U6 snRNP import is inhibited by the SV40 T antigen NLS, but not by the TMG cap (220). The U6 snRNP does not contain a consensus Sm-protein binding site, but is bound by U6-specific proteins (221). The U3 snRNP also does not contain Sm proteins, but does contain a TMG cap. Its import is inhibited neither by free TMG cap, nor by excess SV40 T antigen NLS (220).

Despite the various pathways used in the nuclear import of these U snRNPs, they all cross into the nucleus through the nuclear pore complex. Evidence for this notion resides in the observation that anti-nucleoporin antibody and WGA both affect the import of the snRNPs (220). The WGA inhibition of

transport varies, with the U6 snRNP being highly sensitive (216,220). On the other hand, it has been reported that the U1 and U5 snRNPs are not inhibited by the same concentrations of WGA as used to block U6 entry (216,220).

### **Control of nuclear transport**

Two general mechanisms, not exclusive of each other, for regulating the nuclear-cytoplasmic distribution of proteins are phosphorylation-dephosphorylation and anchor proteins (reviewed in references 222,238). Posttranslational mechanisms, such as phosphorylation, may change a protein's conformation or mask/unmask the NLS, thus rendering the protein translocation competent. Alternatively, anchor proteins may act as scaffolds which secure the protein in the cytoplasm until the appropriate signal causes its release. The yeast transcription factor SW15 is cytoplasmic in the S, G<sub>2</sub>, and M phases of the cell cycle. The phosphorylation of three serine residues near the NLS by the CDC28 kinase is instrumental in preventing nuclear entry. Dephosphorylation of these residues at the end of M phase, when the CDC28 kinase is inactive, causes SW15 to enter the nucleus (223). Studies with the *in vitro* nuclear uptake of SV40 T antigen- $\beta$ -galactoside fusion proteins have shown that phosphorylation at residues away from the NLS by casein kinase II accelerates transport (224). Phosphorylation within the SV40 NLS by cdc2 kinase reduces the nuclear accumulation of the fusion proteins (225). Finally, *Xenopus* oocytes contain large amounts of the protein c-MYC in the cytoplasm. The fertilization process triggers the phosphorylation of c-MYC which then leads to its nuclear localization (169,226).

Studies with the transcription factor NF- $\kappa$ B have shown that in quiescent cells the protein is associated with its anchor I $\kappa$ B in the cytoplasm (reviewed in reference 227). The stimulation of cells leads to the phosphorylation of I $\kappa$ B through a protein kinase C dependent pathway, thus resulting in the dissociation of the anchor and the entry of NF- $\kappa$ B into the nucleus (228). A similar situation is seen with the gene product of the *dorsal* gene which controls the asymmetric expression of genes in *Drosophila*. The *dorsal* product is sequestered in the cytoplasm of the embryonic cells by the *cactus* gene product (229,230,231). The subsequent entry of the *dorsal* protein into the nucleus is influenced by many other gene products (reviewed in reference 232). The c-FOS protein has been reported to be cytoplasmic in serum-starved cells but nuclear in serum-stimulated cells. The regulation of c-FOS nuclear transport has been suggested to depend on two factors: a labile control protein, and mediation by the cAMP dependent protein kinase (233). Although the labile inhibitor of transport has not been identified by Roux *et al.*, it could be the IP-1 inhibitor protein of FOS/JUN (160).

Positive regulation of nuclear transport has been reported for the pancreas-specific transcription factor PTF1. There are two forms of PTF1,  $\alpha$  and  $\beta$ , which are distinguished by their subcellular localization. The only difference between the two forms is the association of a protein, p75, with the  $\alpha$  form. PTF1 $\alpha$  is found in both the nucleus and the cytoplasm, while PTF1 $\beta$  is found exclusively in the cytoplasm. Based on these localization data, it has been hypothesized that the stable association of p75 with PTF1 $\alpha$  allows the protein to enter the nucleus (234).

## LITERATURE CITED

1. Sharon N & Lis H (1989) *Science* **246**: 227-234.
2. Liener IE, Sharon N & Goldstein IJ, eds (1986) *The Lectins: Properties, Functions and Applications in Biology and Medicine*. Academic Press, Orlando.
3. Goldstein IJ, Hughes RC, Monsigny A, Osawa T & Sharon N (1980) *Nature* **285**: 66-68.
4. Etzler ME (1985) *Ann Rev Plant Physiol* **36**: 209-234.
5. Sharon N & Lis H (1989) in *Lectins*. Chapman and Hall, London.
6. Drickamer K (1989) in *Carbohydrate Recognition in Cellular Function*. Bock G & Harnett S, eds. Wiley & Sons, Chichester, UK.
7. Drickamer K (1988) *J Biol Chem* **263**: 9557-9560.
8. Kohnke-Godt B & Gabius H-J (1989) *Biochemistry* **28**: 6531-6538.
9. Kojima K, Ogawa HK, Seno N & Matsumoto I (1992) *J Chromatog* **597**: 323-330.
10. Lasky LA, Singer MS, Yednock TA, Dowbenko D, Fennie C, Rodriguez H, Nguyen T, Stachel S & Rosen SD (1989) *Cell* **56**: 1045-1055.
11. Bevilacqua MP, Stengelin S, Gimbrone Jr. MA & Seed B (1989) *Science* **243**: 1160-1172.
12. Johnston GI, Cook RG & McEver RP (1989) *Cell* **56**: 1033-1044.
13. Ashwell G & Harford J (1982) *Ann Rev Biochem* **51**: 531-554.
14. Hoyle GW & Hill RL (1988) *J Biol Chem* **263**: 7487-7492.
15. Lennartz MR, Wileman TE & Stahl PD (1987) *Biochem J* **245**: 705-711.
16. Taylor ME & Summerfield JA (1986) *Clin Sci* **70**: 539-546.
17. Rust K, Grosso L, Zhang V, Chang D, Persson A, Longmore W, Cai G-Z & Crouch E (1991) *Arch Biochem Biophys* **290**: 116-126.

18. Halberg DF, Proulx G, Doege K, Yamada Y & Drickamer K (1988) *J Biol Chem* **263**: 9486-9490.
19. Cooper DN, Lee S-C & Barondes SH (1983) *J Biol Chem* **258**: 4698-4701.
20. Komano H, Mizuno D & Natori S (1987) *J Biol Chem* **255**: 2919-2924.
21. Muramoto K & Kamuja H (1986) *Biochim Biophys Acta* **874**: 285-295.
22. Giga Y, Ikai A & Takahasi K (1987) *J Biol Chem* **262**: 6197-6203.
23. Suzuki T, Takagi T, Furukohri T, Kawamura K & Nakauchi M (1990) *J Biol Chem* **265**: 1274-1281.
24. Weis WI, Kahn R, Fourme R, Drickamer K & Hendrickson WA (1991) *Science* **254**: 1608-1615.
25. Bezouska K, Crichlow GV, Rose JM, Taylor ME & Drickamer K (1991) *J Biol Chem* **266**: 11604-11609.
26. Taylor ME, Bezouska K & Drickamer K (1992) *J Biol Chem* **267**: 1719-1726.
27. Ali N & Salahuddin A (1989) *Biochim Biophys Acta* **992**: 30-34.
28. Drickamer K, Mamon JF, Binns G & Leung JO (1984) *J Biol Chem* **259**: 770-778.
29. Halberg DF, Wager RE, Farrell DC, Hildreth IV J, Quesenberry MS, Loeb JA, Holland EC & Drickamer K (1987) *J Biol Chem* **262**: 9828-9838.
30. Oda S, Sato M, Toshima S & Osawa T (1988) *J Biochem* **104**: 600-605.
31. Patthy L (1988) *Biochem J* **253**: 309-311.
32. Levi G & Teichberg VI (1981) *J Biol Chem* **256**: 5735-5740.
33. Muramoto K & Kamiya H (1992) *Biochim Biophys Acta* **1116**: 129-132.
34. Fink de Cabutti NE, Caron M, Joubert R, Elola MT, Bladier D & Herkovitz J (1987) *FEBS lett* **223**: 330-334.
35. Beyer EC & Barondes SH (1982) *J Cell Biol* **92**: 23-27.
36. Beyer EC & Barondes SH (1982) *J Cell Biol* **92**: 28-33.

37. Barak-Briles E, Gregory W, Fletcher P & Kornfeld S (1979) *J Cell Biol* **81**: 528-537.
38. Sakakura Y, Hirabayashi J, Oda Y, Ohyama Y & Kasai K (1990) *J Biol Chem* **265**: 21573-21579.
39. Beyer EC & Barondes SH (1980) *J Supramol Struct* **13**: 219-227.
40. Oshawa F, Hirano F & Natori S (1990) *J Biochem* **105**: 431-434.
41. Fitzgerald JE, Catt JW & Harrison FL (1988) *Eur J Biochem* **140**: 137-141.
42. Abbott WM, Mellor A, Edwards Y & Feizi T (1989) *Biochem J* **259**: 283-290.
43. Carding SR, Thorpe SJ, Thorpe R & Feizi T (1985) *Biochem Biophys Res Comm* **127**: 680-686.
44. Merkle RA, Zhou Q, Schultz TK, Harper WB & Cummings RD (1989) *Arch Biochem Biophys* **274**: 404-416.
45. Roff CF & Wang JL (1983) *J Biol Chem* **258**: 10657-10663.
46. Raz A, Carmi P & Pazerini G (1988) *Cancer Res* **48**: 645-649.
47. Wells V & Malucci L (1991) *Cell* **64**: 91-97.
48. Leffler H & Barondes SH (1986) *J Biol Chem* **261**: 10119-10126.
49. Regan LJ, Dodd J, Barondes SH & Jessell TM (1986) *Proc Natl Acad Sci USA* **83**: 2248-2252.
50. Hynes MA, Gitt M, Barondes SH, Jessell TM & Buck LB (1990) *J Neurosci* **10**: 1004-1013.
51. Hirabayashi J, Kawasaki H, Suzuki K & Kasai K (1987) *J Biochem* **101**: 775-787.
52. Gitt MA & Barondes SH (1986) *Proc Natl Acad Sci USA* **83**: 7603-7607.
53. Gitt MA & Barondes SH (1991) *Biochemistry* **30**: 82-89.
54. Bladier D, Le Caer JP, Joubert R, Caron M & Rossier J (1991) *Neurochem Int* **18**: 275-281.

55. Gitt MA, Massa SM, Leffler H & Barondes SH (1992) *J Biol Chem* **267**: 10601-10606.
56. Leffler H, Masiarz FR & Barondes SH (1989) *Biochemistry* **28**: 9222-9229.
57. Ruggiero-Lopez D, Louisot P & Martin A (1992) *Biochem Biophys Res Comm* **185**: 617-623.
58. Crittenden SL, Roff CF & Wang JL (1984) *Mol Cell Biol* **4**: 1252-1259.
59. Marschal P, Herrmann J, Leffler H, Barondes SH & Cooper DNW (1992) *J Biol Chem* **267**: 12942-12949.
60. Hirabayashi J, Satoh M, Ohyama Y & Kasai K (1992) *J Biochem* **111**: 553-555.
61. Brassart D, Kolodziejczyk E, Granato D, Woltz A, Pavillard M, Perotti F, Frigeri LG, Liu F-T, Borel Y & Neeser J-R (1992) *Eur J Biochem* **203**: 393-399.
62. Sparrow CP, Leffler H & Barondes SH (1987) *J Biol Chem* **262**: 7383-7390.
63. Hinek A, Wrenn DS, Mecham RP & Barondes SH (1988) *Science* **239**: 1539-1541.
64. Abbott WM & Feizi T (1989) *Biochem J* **259**: 291-294.
65. Greer Wilson TJ, Firth MN, Powell JT & Harrison FL (1989) *Biochem J* **261**: 847-852.
66. Harrison L (1991) In *Lectin Reviews Vol I*. Kilpatrick DC, Van Driessche E & Bøg-Hansen TC, Eds. Sigma Chemical Co, St Louis.
67. Powell JT & Harrison FL (1991) *Am J Physiol* **261**: L236-L239.
68. Barondes SH & Haywood-Reid PL (1981) *J Cell Biol* **91**: 568-572.
69. Cooper DN & Barondes SH (1990) *J Cell Biol* **110**: 1681-1691.
70. Catt JW & Harrison FL (1985) *J Cell Sci* **73**: 347-359.
71. Wasano K, Hirakawa Y & Yamamoto T (1990) *Cell Tissue Res* **259**: 43-49.
72. Raz A, Meromsky L, Zvibel I & Lotan R (1987) *Int J Cancer* **39**: 353-360.

73. Poirier F, Timmons PM, Chan C-TJ, Guenet J-L & Rigby PWJ (1992) *Develop* **115**: 143-155.
74. Woo H-J, Shaw LM, Messier JM & Mercurio AM (1990) *J Biol Chem* **265**: 7097-7099.
75. Jia S, Mee RP, Morford G, Agrwal N, Voss PG, Moutsatsos IK & Wang JL (1987) *Gene* **60**: 197-204.
76. Jia S & Wang JL (1988) *J Biol Chem* **263**: 6009-6011.
77. Ho MK & Springer TA (1982) *J Immunol* **128**: 1221-1228.
78. Cherayil BJ, Weiner SJ & Pillai S (1989) *J Exp Med* **170**: 1959-1972.
79. Foddy L, Stamatoglou SC & Hughes RC (1990) *J Cell Sci* **97**: 139-148.
80. Laing JG, Robertson MW, Gritzmacher CA, Wang JL & Liu F-T (1989) *J Biol Chem* **264**: 1907-1910.
81. Oda Y, Leffler H, Sakakura Y, Kasai K & Barondes SH (1990) *Gene* **99**: 279-283.
82. Robertson MW, Albrandt K, Keller D & Liu F-T (1990) *Biochemistry* **28**: 8093-8100.
83. Cherayil BJ, Chaitovitz S, Wong C & Pillai S (1990) *Proc Natl Acad Sci* **87**: 7324-7328.
84. Hamann KK, Cowles EA, Wang JL & Anderson RL (1991) *Exp Cell Res* **196**: 82-91.
85. Hirabayashi J & Kasai K (1991) *J Biol Chem* **266**: 23648-23653.
86. Abbott WM & Feizi T (1991) *J Biol Chem* **266**: 5552-5557.
87. Sato S & Hughes RC (1992) *J Biol Chem* **267**: 6983-6990.
88. Wang JL, Laing JG & Anderson RL (1991) *Glycobiology* **1**: 243-252.
89. Childs RA & Feizi T (1980) *Cell Biol Int Reports* **4**: 775.
90. Akimoto Y, Kawakami H, Oda Y, Obinata A, Endo H, Kasai K & Hirano H (1992) *Exp Cell Res* **199**: 297-304.

91. Harrison FL & Greer Wilson TJ (1992) *J Cell Sci* **101**: 635-646.
92. Bols NC, Roberson MM, Haywood-Reid PL, Cerra RF & Barondes SH (1986) *J Cell Biol* **102**: 492-499.
93. Moutsatsos IK, Davis JM & Wang JL (1986) *J Cell Biol* **102**: 477-483.
94. Moutsatsos IK, Wade M, Schindler M & Wang JL (1987) *Proc Natl Acad Sci USA* **84**: 6452-6456.
95. Laing JG & Wang JL (1988) *Biochemistry* **27**: 5329-5334.
96. Gritzmacher CA, Robertson MW & Liu F-T (1988) *J Immunol* **141**: 2801-2806.
97. Raz A, Zhu D, Hogan V, Shah N, Raz T, Karkash R, Pazerini G & Carmi P (1990) *Int J Cancer* **46**: 871-877.
98. Butel JS & Jarvis DL (1986) *Biochim Biophys Acta* **865**: 171-195.
99. Bellgrau D, Walker TA & Cook JL (1988) *J Virol* **62**: 1513-1519.
100. Spence AM, Sheppard PC, Davie JR, Matuo Y, Nishi N, McKeehan WL, Dodd JG & Matusik RJ (1989) *Proc Natl Acad Sci USA* **86**: 7843-7847.
101. Bachmann M, Chang S-H, Slor H, Kukulies J & Müller WEG (1990) *Exp Cell Res* **191**: 171-180.
102. Sano H, Forough R, Maier JAM, Case JP, Kackson A, Engleka K, Maciag T & Wilder RL (1990) *J Cell Biol* **110**: 1417-1426.
103. Ishikawa F, Miyazono K, Hellman U, Drexler H, Wernstedt C, Hagiwara K, Usuki K, Takaku F, Risau W & Helden C-H (1989) *Nature* **338**: 557-562.
104. Rubartelli A, Cozzolino F, Talio M & Sitia R (1990) *EMBO J* **9**: 1503-1510.
105. Kuchler K, Sterne RE & Thorner J (1989) *EMBO J* **8**: 3973-3984.
106. Mizutani A, Usuda N, Tokumitsu H, Minami H, Yasui K, Kobayashi R & Hidaka H (1992) *J Biol Chem* **267**: 13498-13504.
107. Acland P, Dixon M, Peters G & Dixon C (1990) *Nature* **343**: 662-665.

108. Maher DW, Lee BA & Donoghue DJ (1989) *Mol Cell Biol* **9**: 2251-2253.
109. Mignatti P, Morimoto T & Rifkin DB (1992) *J Cell Physiol* **151**: 81-93.
110. Imamura T, Engleka K, Zhan X, Tokita Y, Forough R, Roeder D, Jackson A, Maier JAM, Hla T & Maciag T (1990) *Science* **249**: 1567-1570.
111. Featherstone C (1990) *Trends Biochem Sci* **15**: 169-170.
112. Juranka PF, Zastawny RL & Ling V (1989) *FASEB J* **3**: 2583-2592.
113. Dowbenko DJ, Diep A, Taylor BA, Lusic AJ & Lasky LA (1991) *Genomics* **9**: 270-277.
114. Feizi T (1991) *Trends Biochem Sci* **16**: 84-86.
115. Springer TA & Lasky LA (1991) *Nature* **349**: 196-197.
116. Lasky LA, Singer MS, Dowbenko D, Imai Y, Henzel WJ, Grimley C, Fennie C, Gillett N, Watson SR & Rosen SD (1992) *Cell* **69**: 927-938.
117. Yamaoka K, Ohno S, Kawasaki H & Suzuki K (1991) *Biochem Biophys Res Comm* **179**: 272-279.
118. Hsu DK, Zuberu RI & Liu F-T (1992) *J Biol Chem* **267**: 14167-14174.
119. Woo H-J, Lotz MM, Jung JU & Mercurio AM (1991) *J Biol Chem* **266**: 18419-18422.
120. Kumar AM & Davidson VL (1992) *Biotechniques* **12**: 198-202.
121. Rosenberg I, Cherayil BJ, Isselbacher KJ & Pillai S (1991) *J Biol Chem* **266**: 18731-18736.
122. Rivera VM & Greenberg ME (1990) *New Biol* **2**: 751-758.
123. Treisman R (1990) *Cancer Biol* **1**: 47-58.
124. Cowles EA, Agrwal N, Anderson RL & Wang JL (1990) *J Biol Chem* **265**: 17706-17712.
125. McMahon SB & Monroe JG (1992) *FASEB J* **6**: 2707-2715.
126. Denhardt DT, Edwards DR & Parfett CLJ (1986) *Biochim Biophys Acta* **865**: 83-125.

127. Herschman HR (1991) *Ann Rev Biochem* **60**: 281-319.
128. Cochran BH, Reffel AC & Stiles CD (1983) *Cell* **33**: 939-947.
129. Lau LF & Nathans D (1987) *Proc Natl Acad Sci USA* **84**: 1182-1186.
130. Mohn KL, Laz TM, Hsu JC, Melby AE, Bravo R & Taub R (1991) *Mol Cell Biol* **11**: 381-390.
131. Nakabeppu Y, Ryder K & Nathans D (1988) *Cell* **55**: 907-915.
132. Halazonetis TD, Georgopoulos K, Greenberg ME & Leder P (1988) *Cell* **55**: 917-924.
133. Hai T & Curran T (1991) *Proc Natl Acad Sci USA* **88**: 3720-3724.
134. Vinson CR, Sigler PB & McKnight SL (1989) *Science* **246**: 911-922.
135. Rauscher III FJ, Voulalas PJ, Franza Jr BR & Curran T (1989) *Genes Dev.* **2**: 1687-1699.
136. Kruijer W, Cooper JA, Hunter T & Verma IM (1984) *Nature* **312**: 711-716.
137. Müller R, Bravo R, Burckhardt J & Curran T (1984) *Nature* **312**: 716-720.
138. Müller R, Verma IM & Adamson ED (1983) *EMBO J* **2**: 679-684.
139. Greenberg ME & Ziff EB (1984) *Nature* **311**: 433-438.
140. Wilson T & Treisman R (1988) *Nature* **336**: 396-399.
141. Seshadri T & Campisi J (1990) *Science* **247**: 205-209.
142. Pontecorvi A, Tata JR, Phyllaier M & Robbins J (1988) *EMBO J* **7**: 1489-1495.
143. Sassone-Corsi P, Sisson JC & Verma IM (1988) *Nature* **334**: 314-319.
144. Subramaniam M, Schmidt LJ, Crutchfield CE & Getz MJ (1989) *Nature* **340**: 64-66.
145. Edwards DR & Mahadevan LC (1992) *EMBO J* **11**: 2415-2424.
146. Sheng M, Dougan ST, McFadden G & Greenberg ME (1988) *Mol Cell Biol* **8**: 2787-2796.

147. Fisch TM, Prywes R & Roeder RG (1989) *Mol Cell Biol* **9**: 1327-1331.
148. Hayes TE, Kitchen AM & Cochran BH (1987) *Proc Natl Acad Sci USA* **84**: 1272-1276.
149. Visvader J, Sassone-Corsi P & Verma IM (1988) *Proc Natl Acad Sci USA* **85**: 9474-9478.
150. Robbins PD, Horowitz JM & Mulligan RC (1990) *Nature* **346**: 668-671.
151. Morgan IM & Birnie GD (1992) *Cell Prolif* **25**: 205-215.
152. Mechti N, Piechaczyk M, Blanchard J-M, Jeanteur P & Lebleu B (1991) *Mol Cell Biol* **11**: 2832-2841.
153. Rivera VM, Sheng M & Greenberg ME (1990) *Genes Dev* **4**: 255-268.
154. Christy BA & Nathans D (1989) *Mol Cell Biol* **9**: 4889-4895.
155. Norman C, Runswick M, Pollock R & Treisman R (1988) *Cell* **55**: 989-1003.
156. Schroter H, Mueller CGF, Meese K & Nordheim A (1990) *EMBO J* **9**: 1123-1130.
157. Dalton S & Treisman R (1992) *Cell* **68**: 597-612.
158. Sassone-Corsi P & Verma IM (1987) *Nature* **326**: 507-510.
159. Gilman MZ (1988) *Genes Dev* **2**: 394-402.
160. Auwerx J & Sassone-Corsi P (1991) *Cell* **64**: 983-993.
161. Bravo R, Burckhardt J, Curran T & Muller R (1986) *EMBO J* **5**: 695-700.
162. Bannister AJ, Cook A & Kouzarides T (1991) *Oncogene* **6**: 1243-1250.
163. Abate CL, Patel L, Rauscher III FJ & Curran T (1990) *Science* **249**: 1157-1161.
164. Gius D, Cao X, Rauscher III FJ, Cohen DR, Curran T & Sukhatme VP (1990) *Mol Cell Biol* **10**: 4243-4255.
165. Konig H, Ponta H, Rahmsdorf U, Buscher M, Schonthal A, Rahmsdorf HJ & Herrlich P (1989) *EMBO J* **8**: 2559-2566.

166. Gerace L & Burke B (1988) *Ann Rev Cell Biol* **4**: 335-374.
167. Silver P (1991) *Cell* **64**: 489-497.
168. Garcia-Bustos J, Heitman J & Hall MN (1991) *Biochim Biophys Acta* **1071**: 83-101.
169. Nigg EA, Baeuerle PA & Lührmann R (1991) *Cell* **66**: 15-22.
170. Moreland RB, Langevin GL, Singer RH, Garcea RL & Hereford LM (1987) *Mol Cell Biol* **7**: 4048-4057.
171. Breeuwer M & Goldfarb DS (1990) *Cell* **60**: 999-1008.
172. Zhao L-J & Padmanabhan R (1988) *Cell* **55**: 1005-1015.
173. Richardson WD, Mills AD, Dilworth SM, Laskey RA & Dingwall C (1988) *Cell* **52**: 655-664.
174. Newmeyer DD & Forbes DJ (1988) *Cell* **52**: 641-653.
175. Dingwall C & Laskey RA (1991) *Trends Biochem Sci* **16**: 478-481.
176. Kalderon D, Richardson WD, Markham AF & Smith AE (1984) *Nature* **311**: 33-38.
177. Kalderon D, Roberts BL, Richardson WD & Smith AE (1984) *Cell* **39**: 499-509.
178. Dingwall C, Robbins J, Dilworth SM, Roberts B & Richardson WD (1988) *J Cell Biol* **107**: 841-849.
179. Robbins J, Dilworth SM, Laskey RA & Dingwall C (1991) *Cell* **64**: 615-623.
180. Citovsky V, Zupan J, Warnick D & Zambryski P (1992) *Science* **256**: 1802-1805.
181. Feldherr CM & Akin D (1990) *J Cell Biol* **111**: 43-53.
182. Feldherr CM & Akin D (1991) *J Cell Biol* **115**: 933-939.
183. Maul GG (1977) *Int Rev Cytol Suppl* **6**: 75-186.
184. Greber UF, Senior A & Gerace L (1990) *EMBO J* **9**: 1495-1502.

185. Greber UF & Gerace L (1992) *J Cell Biol* **116**: 15-30.
186. Hart GW, Haltiwanger RS, Holt GD & Kelly WG (1989) *Ann Rev Biochem* **58**: 841-874.
187. Starr CM, D'Onofrio M, Park MK & Hanover JA (1990) *J Cell Biol* **110**: 1861-1871.
188. Nehrbass U, Kern H, Mutvei A, Horstmann H, Marshallsay B & Hurt EC (1990) *Cell* **61**: 979-989.
189. Davis LI & Fink GR (1990) *Cell* **61**: 965-978.
190. Aris JP & Blobel G (1989) *J Cell Biol* **108**: 2059-2067.
191. Featherstone C, Darby MK & Gerace L (1988) *J Cell Biol* **107**: 1289-1297.
192. Finlay DR, Newmeyer DD, Price TM & Forbes DJ (1987) *J Cell Biol* **104**: 189-200.
193. Finlay DR & Forbes DJ (1990) *Cell* **60**: 17-29.
194. Finlay DR, Meier E, Bradley P, Horecka J & Forbes DJ (1991) *J Cell Biol* **114**: 169-183.
195. Newmeyer DD, Finlay DR & Forbes DJ (1986) *J Cell Biol* **103**: 2091-2102.
196. Adam SA, Marr RS & Gerace L (1990) *J Cell Biol* **111**: 807-816.
197. Newmeyer DD & Forbes DJ (1990) *J Cell Biol* **110**: 547-557.
198. Sterne-Marr R, Blevitt JM & Gerace L (1992) *J Cell Biol* **116**: 271-280.
199. Moore MS & Blobel G (1992) *Cell* **69**: 939-950.
200. Shi Y & Thomas JO (1992) *Mol Cell Biol* **12**: 2186-2192.
201. Chirico WJ, Waters MG & Blobel G (1988) *Nature* **332**: 805-810.
202. Deshaies RJ, Koch BD, Werner-Washburne M, Craig EA & Schekman R (1988) *Nature* **332**: 800-805.
203. Block MR, Glick BS, Wilcox CA, Wieland FT & Rothman JE (1988) *Proc Natl Acad Sci USA* **85**: 7852-7856.

204. Adam SA, Lobl TJ, Mitchell MA & Gerace L (1989) *Nature* **337**: 276-279.
205. Garcia-Bustos JF, Wagner P & Hall MN (1991) *J Biol Chem* **266**: 22303-22306.
206. Yamasaki L, Kanda P & Lanford RE (1989) *Mol Cell Biol* **9**: 3028-3036.
207. Adam SA & Gerace L (1991) *Cell* **66**: 837-847.
208. Borer RA, Lehner CF, Eppenberger HM & Nigg EA (1989) *Cell* **56**: 379-390.
209. Silver PA, Sadler I & Osborne MA (1989) *J Cell Biol* **109**: 983-989.
210. Stochaj U, Osborne M, Kurihara T & Silver PA (1991) *J Cell Biol* **113**: 1243-1254.
211. Stochaj U & Silver PA (1992) *J Cell Biol* **117**: 473-482.
212. Lee W-C & Mèlèse (1989) *Proc Natl Acad Sci USA* **86**: 8808-8812.
213. Lee W-C, Xue Z & Mèlèse T (1991) *J Cell Biol* **113**: 1-12.
214. Meier UT & Blobel G (1990) *J Cell Biol* **111**: 2235-2245.
215. Michaud N & Goldfarb DS (1991) *J Cell Biol* **112**: 215-223.
216. Fischer U, Darzynkiewicz E, Tahara SM, Dathan NA, Lührmann R & Mattaj IW (1991) *J Cell Biol* **113**: 705-714.
217. Hamm J & Mattaj IW (1990) *Cell* **63**: 109-118.
218. Hamm J, Darzynkiewicz E, Tahara SM & Mattaj IW (1990) *Cell* **62**: 569-577.
219. Fischer U & Lührmann R (1990) *Science* **249**: 786-790.
220. Michaud N & Goldfarb D (1992) *J Cell Biol* **116**: 851-861.
221. Hamm J & Mattaj IW (1989) *EMBO J* **8**: 4179-4187.
222. Hunt T (1989) *Cell* **59**: 949-951.
223. Moll T, Tebb G, Surana U, Robitsch H & Nasmyth K (1991) *Cell* **66**: 743-758.

224. Rihs H-P, Jans DA, Fan H & Peters R (1991) *EMBO J* **10**: 633-639.
225. Jans DA, Ackermann MJ, Bischoff JR, Beach DH & Reiners P (1991) *J Cell Biol* **115**: 1203-1212.
226. Gusse M, Ghysdael J, Evan G, Soussi T & Méchali M (1989) *Mol Cell Biol* **9**: 5395-5403.
227. Baeuerle PA (1991) *Biochim Biophys Acta* **1072**: 63-80.
228. Ghosh S & Baltimore D (1990) *Nature* **344**: 678-682.
229. Rushlow CA, Han K, Manley JL & Levine M (1989) *Cell* **59**: 1165-1177.
230. Steward R (1989) *Cell* **59**: 1179-1188.
231. Roth S, Stein D & Nüsslein-Volhard C (1989) *Cell* **59**: 1189-1202.
232. Govind S & Steward R (1991) *Trends Genet* **7**: 119-125.
233. Roux P, Blanchard J-M, Fernandez A, Lamb N, Jeanteur P & Piechaczyk M (1990) *Cell* **63**: 341-351.
234. Sommer L, Hagenbüchle O, Wellnauer PK & Strubin M (1991) *Cell* **67**: 987-994.
235. Castronovo V, Luylen F, van den Brûle F & Sobel ME (1992) *Arch Biochem Biophys* **297**: 132-138.
236. Goldfarb DS (1992) *Cell* **70**: 185-188.
237. Hirabayashi J, Satoh M & Kasai K (1992) *J Biol Chem* **267**: 15485-15490.
238. Hunter T & Karin M (1992) *Cell* **70**: 375-387.
239. Meier UT & Blobel G (1992) *Cell* **70**: 127-138.
240. Harrison FL (1991) *J Cell Sci* **100**: 9-14.

## **CHAPTER II**

### **CARBOHYDRATE BINDING PROTEIN 35: Levels of Transcription and mRNA Accumulation in Quiescent and Proliferating Cells**

Neera Agrwal, John L. Wang, and Patricia G. Voss

Department of Biochemistry  
Michigan State University  
East Lansing MI 48824

## FOOTNOTES

- \* This work was supported by Grants GM-38740 and GM-27203 from the National Institutes of Health.
  
- <sup>1</sup> N. Agrwal was supported by a Patricia Roberts Harris Fellowship.
  
- <sup>2</sup> The abbreviations used are: CBP, carbohydrate-binding protein; hnRNP, heterogeneous nuclear ribonucleoprotein complex; SDS, sodium dodecyl sulfate; BSA, bovine serum albumin; HEPES, *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid.
  
- <sup>3</sup> This work was previously published in *Journal of Biological Chemistry* (1989) 264: 17236-17242 and is reproduced with permission from the publisher.  
Copyright 1989 American Society for Biochemistry and Molecular Biology.

**ABSTRACT**

In previous studies, we observed proliferation-dependent expression and nuclear localization of the lectin, designated carbohydrate-binding protein 35 (CBP35), in 3T3 fibroblasts at the polypeptide level by Western blot and immunofluorescence analysis. In the present study, we have compared the expression of the CBP35 gene in quiescent and proliferative 3T3 cells at the levels of (a) accumulated mRNA by Northern blotting and (b) nuclear transcription by run-off assays. When serum-starved, quiescent cultures of 3T3 cells were stimulated by the addition of serum, there was an increase in the nuclear transcription of the CBP35 gene and in the accumulation of its mRNA early (1-3 h) in the activation process, well before the first wave of DNA synthesis. These increases were not dependent on *de novo* protein synthesis inasmuch as they occurred even in the presence of cycloheximide. There was also an elevated transcription rate and mRNA level in transformed cells when compared to their normal counterparts. Finally, the expression of CBP35 was compared between sparse, proliferating cultures of 3T3 cells and density-inhibited confluent monolayers of the same cells. Although the rate of transcription of the CBP35 gene was approximately the same in the two cultures, there was a higher level of CBP35 mRNA in the dense cells. Thus, it appears that post-transcriptional mechanisms may be involved in the accumulation of mRNA.

## INTRODUCTION

Carbohydrate-binding protein (CBP) 35 ( $M_r$  35,000) was initially purified from extracts of mouse 3T3 fibroblasts on the basis of its binding to galactose-containing glycoconjugates (1). More recent studies have suggested that CBP35 is a component of the heterogeneous nuclear ribonucleoprotein complex (hnRNP). This conclusion was based on the following observations (2): (a) CBP35 was released from permeabilized nuclei by treatment with ribonuclease A but not by similar treatment with deoxyribonuclease I; (b) when nucleoplasm was fractionated on a cesium sulfate gradient, CBP35 was found in fractions with the same densities as those reported for hnRNP ( $\approx 1.30$  g/ml); (c) CBP35 co-isolates with hnRNP by sucrose gradient centrifugation (40 S); and (d) fractionation of nucleoplasm on a column derivatized with  $N\epsilon$ -amino-caproylgalactosamine yielded a bound fraction containing CBP35, as well as a set of polypeptides whose molecular weights matched those reported for the proteins in hnRNP.

By using antibodies specifically directed against CBP35 to screen a  $\lambda$ gt 11 expression library derived from the mRNA of 3T3 cells, we have identified and characterized a cDNA clone for CBP35 (3). The fusion protein expressed by this clone, upon V8 protease digestion, yielded a  $M_r$  30,000 polypeptide that exhibited carbohydrate binding activity and immunoblotting pattern with anti-CBP35 identical to those of CBP35 itself. Moreover, sequence analysis of this cDNA clone revealed two distinct domains within the polypeptide. The carboxyl-terminal domain showed sequence homology with other  $\beta$ -galactoside-binding lectins,

whereas the amino-terminal portion showed homology with certain hnRNP proteins (4).

Previous studies had shown that, when quiescent 3T3 fibroblasts were stimulated by the addition of serum, there was increase in the overall level of the CBP35, as well as a dramatic rise in the amount of the polypeptide in the nucleus (5). This increase in the level of CBP35 occurred well before the onset of the first S phase after serum stimulation of 3T3 cells. We had also demonstrated that, in a comparison of the levels of CBP35 in normal fibroblasts and their virally transformed counterparts, there was always more CBP35 in the transformed cells (5, 6). The availability of the cDNA probe provided the opportunity to analyze the expression of the CBP35 gene under these conditions. In the present article, we document the levels of nuclear transcription and mRNA accumulation in quiescent and proliferating cells.

## MATERIALS AND METHODS

**Cell Culture.** Swiss 3T3 fibroblasts and 3T3 cells transformed by Kirsten murine sarcoma virus cells (3T3-KiMSV cells) were obtained from the American Type Culture Collection (CCL92 and CCL163.3, respectively). They were cultured in Dulbecco's modified Eagle's Medium (GIBCO) supplemented with 10% calf serum (GIBCO) at 37°C in 10% CO<sub>2</sub>. Proliferating cultures of 3T3 cells were seeded at a density of 1 X 10<sup>4</sup> cells/cm<sup>2</sup> in culture medium containing 10% serum. These cultures were synchronized by incubation in Dulbecco's medium containing 0.2% serum for 48 h and were stimulated to proceed through the cell cycle by the addition of 10% serum. At various times after stimulation, cells were harvested for the isolation of polyadenylated (poly(A)<sup>+</sup>) RNA. Alternatively, nuclei were isolated for run-off transcription assays. Cells cultured at a density of 5 X 10<sup>4</sup> cells/cm<sup>2</sup> formed a quiescent monolayer. In some experiments, cycloheximide was added to a final concentration of 10 µg/ml (36 µM) at the same time as the addition of serum to quiescent cells.

**RNA Isolation and Northern Blot Analysis.** Poly(A)<sup>+</sup> RNA isolation and Northern blot analysis were carried out as described by Stuart *et al.* (7). The fibroblasts were washed with calcium and magnesium-free phosphate-buffered saline (PBS; 0.14 M NaCl, 2.7 mM KCl, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, 4.3 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.4) and lysed with 2 ml of lysis buffer (0.5 M NaCl, 10 mM Tris-HCl, 1 mM EDTA, 1% sodium dodecyl sulfate (SDS), 100 µg of proteinase K/ml, pH 7.5) per

flask (150 cm<sup>2</sup>). The lysed cells were removed from the flask, cellular DNA was sheared by passage through a 22-gauge needle, and the lysate was incubated with additional proteinase K (100 µg/ml) for 1 h at 37°C. The lysate was incubated with 30 mg of oligodeoxythymidylate-cellulose (Sigma) for 1 h at room temperature to absorb the poly(A)<sup>+</sup> RNA. This material was packed into an Isolab Quick-Sep microcolumn ( $V_T = 2$  ml) and washed with 4 ml of wash buffer no. 1 (0.5 M LiCl, 10 mM Tris-HCl, 1 mM EDTA, 0.2% SDS, pH 7.5) followed by 6 ml of wash buffer no. 2 (0.1 M NaCl, 10 mM Tris-HCl, 1 mM EDTA, 0.2% SDS, pH 7.5). The poly(A)<sup>+</sup> RNA was eluted with 10 mM Tris-HCl, 1 mM EDTA, pH 7.5. After the addition of the carrier yeast tRNA (10 µg) and sodium acetate (0.3 M final concentration), the eluted material was ethanol-precipitated.

Precipitated RNA was suspended in gel sample buffer (50% formamide, 1 X running buffer, 2.2 M formaldehyde), heated to 65°C, and electrophoresed on a 1.2% agarose gel containing 20 mM morpholinepropanesulfonic acid, 2.2 M formaldehyde. Running buffer was 20 mM morpholinepropanesulfonic acid, 1 mM EDTA, 5 mM sodium acetate, 2.2 M formaldehyde, pH 7.0. After electrophoresis, gels were washed in water and then in 20 X SSC (1 X SSC = 0.15 M NaCl, 15 mM trisodium citrate), and the separated RNA was transferred to nitrocellulose filters.

The filters were incubated in prehybridization solution (50% formamide, 5 X SSC, 5 X Denhardt's solution, 25 mM sodium phosphate, 0.5 mg/ml denatured sheared salmon sperm DNA) for 2-4 h in plastic bags at 42°C. The composition of the 5 X Denhardt's solution was 0.1% bovine serum albumin, 0.1%

polyvinylpyrrolidone, and 0.1% Ficoll in H<sub>2</sub>O. Filters were then incubated overnight at 42°C in hybridization solution (50% formamide, 3 X SSC, 1 X Denhardt's solution, 10 mM sodium phosphate, 0.2 mg/ml denatured sheared salmon sperm DNA) containing 10<sup>6</sup> cpm/ml of DNA probe that was labeled with [ $\alpha$ -<sup>32</sup>P]dCTP using random oligodeoxynucleotide primer labeling. Following hybridization with the radiolabeled probes, filters were washed for 15 min in 2 X SSC, 0.1% SDS at room temperature and washed twice for 45 min in 2 X SSC, 0.1% SDS at 65°C. The washed filters were exposed to Kodak X-Omat AR film with an intensifying screen where indicated. Relative intensities of the bands were determined by scanning densitometric analysis (Gelman ACD-18 automatic computing densitometer).

In experiments to determine the half-life of mRNA, actinomycin D was added to cell cultures at a final concentration of 8  $\mu$ g/ml for the indicated times prior to RNA isolation as described above. In experiments comparing poly(A)<sup>+</sup> RNA levels in 3T3 and 3T3-KiMSV cells, the method used was essentially that of Maniatis *et al.* (8). Cells were scraped into ice-cold PBS and pelleted by centrifugation, followed by incubation in cold lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.5% Nonidet P-40). The nuclei were removed for nuclear run-off experiments while the cytoplasmic layer was incubated with 2X PK buffer (0.2 M Tris-HCl, pH 7.5, 25 mM EDTA, 0.3 M NaCl, 2% SDS, 400  $\mu$ g/ml proteinase K), after which the RNA was ethanol-precipitated. The RNA was quantitated, and equal amounts were passaged over oligo(dT)-cellulose columns as described above to isolate poly(A)<sup>+</sup> RNA.

**Nuclear Run-off Transcription Assays.** Run-off transcriptions were performed essentially as described by Linial *et al.* (9) and Stewart *et al.* (10). The cells were washed in cold hypotonic buffer (20 mM Tris-HCl, 5 mM MgCl<sub>2</sub>, 6 mM CaCl<sub>2</sub>, 0.5 mM dithiothreitol, pH 8.0) and scraped off the dishes in this buffer with a rubber policeman. The scraped cells were collected by centrifugation and lysed in 0.6 M sucrose with 0.2% Nonidet P-40 and 0.5 mM dithiothreitol. The nuclei were suspended in freezing buffer (40% glycerol, 50 mM Tris-HCl, 5 mM MgCl<sub>2</sub>, 0.1 M EDTA, pH 8.3), frozen, and stored in liquid nitrogen until they were used. In experiments comparing 3T3 and 3T3-KiMSV cells, nuclei were isolated as described under "RNA Isolation and Northern Blot Analysis."

A typical transcription assay (total volume of 200  $\mu$ l) consisted of approximately 5-10 X 10<sup>6</sup> nuclei in a volume of 130  $\mu$ l; 40  $\mu$ l of 5 X run-off buffer (12 mM magnesium acetate, 25 mM Tris-HCl, 12.5 mM MgCl<sub>2</sub>, 750 mM KCl, 0.5 mM EDTA, pH 8.0), 20  $\mu$ l of 10 X nucleotides (4.0 mM CTP, 4.0 mM GTP, 10 mM ATP, 5 mM S-adenosylmethionine), and 100  $\mu$ Ci of either [ $\alpha$ -<sup>32</sup>P]UTP or [ $\alpha$ -<sup>35</sup>S]UTP (800 Ci/mmol or 1500 Ci/mmol, respectively). The reaction was carried out at 25°C for 30 min, after which human placental ribonuclease inhibitor (RNA-guard, Pharmacia LKB Biotechnology Inc.) and ribonuclease-free deoxyribonuclease I (Boehringer Mannheim) were added. The suspension was then mixed with an equal volume of 2 X SETY (2% SDS, 10 mM EDTA, 20 mM Tris-HCl, pH 7.5, 200  $\mu$ g/ml yeast tRNA) with 200  $\mu$ g/ml proteinase K, incubated at 37°C for 20-45 min, extracted with an equal volume of phenol:chloroform (phenol:chloroform:isoamyl alcohol (25:24:1) (v/v) saturated with 10 mM Tris-

HCl, 1 mM EDTA, pH 7.5), and precipitated overnight at  $-20^{\circ}\text{C}$  with ammonium acetate (final concentration 2.3 M) and 2 volumes of ethanol. The RNA was pelleted by centrifugation for 30 min at 12,000 X *g*. Incorporated radioactive label in the RNA was verified by precipitating a small amount of the RNA onto glass fiber filters using trichloroacetic acid and counting the filters in a scintillation counter. NaOH was added to the resuspended RNA pellet (to a final concentration of 0.2 N) for 10 min on ice, and then HEPES was added to a final concentration of 0.24 M to neutralize the base. The RNA was then ethanol-precipitated. Equal amounts of radioactivity were added to each hybridization solution.

Complementary CBP35 RNA transcripts and DNA plasmids were bound to nitrocellulose fibers by using a slot blot manifold (Bethesda Research Laboratories). For the RNA probes, approximately 1-2  $\mu\text{g}$  of RNA probe/slot was suspended in  $\text{H}_2\text{O}$ . Formaldehyde (5  $\mu\text{l}$ ) was added, and the RNA was heated to  $65^{\circ}\text{C}$  for 10 min. Then, 130  $\mu\text{l}$  of 20 X SSC was added, and the samples were immediately applied to the individual slots. The plasmid DNAs were linearized by restriction enzyme digestion, extracted with an equal volume of phenol:chloroform and precipitated with ethanol at  $-20^{\circ}\text{C}$ . Approximately 1-2  $\mu\text{g}$  of DNA/slot was suspended in 25  $\mu\text{l}$  of  $\text{H}_2\text{O}$  and mixed with 5  $\mu\text{l}$  of freshly made 2 M  $\text{NH}_4\text{OH}$ . The sample was heated for 3 min at  $90^{\circ}\text{C}$ , quenched in ice, and 20  $\mu\text{l}$  of cold 5 M NaCl was added prior to sample application to the nitrocellulose filters. The filters were washed with 2 X SSC and baked under vacuum at  $80^{\circ}\text{C}$ . The filters were prehybridized in 50% formamide, 5 X SSC, 50 mM sodium phosphate, 1 X

Denhardt's solution, 250  $\mu\text{g/ml}$  denatured sheared salmon sperm DNA for 2-24 h at 42°C. Then, hybridization solution (1 part of 50% dextran sulfate added to 4 parts of prehybridization solution) containing the denatured radioactive RNA transcripts was added, and hybridization was carried out at 42°C for 72 h. The filters were washed as directed for Northern blot analysis. The filters were allowed to air dry and were exposed to Kodak X-Omat AR film with an intensifying screen where appropriate.

**Plasmids and Preparation of Probes.** The plasmid pWJ31 containing the cDNA insert for CBP35 was constructed by excising the 883-base pair *EcoRI* fragment out of the clone identified in the  $\lambda\text{gt}11$  library and subcloning into the pUC-13 vector (3). The probe for murine  $\beta_2$ -microglobulin was an 8-kilobase pair *XhoI* fragment subcloned into the pKc7 vector (11).

Complementary RNA transcripts for CPB35 were made using insert fragments subcloned into the pSp65 vector (Promega Biotec.). The *in vitro* transcription reactions were performed as described by Promega Biotec, and the amount of RNA transcribed was quantitated by comparing the intensity of ethidium bromide staining of the RNA transcripts against known amounts of control RNA after electrophoresis on agarose-formaldehyde gels.

DNA probes were linearized by restriction enzyme digestion and extracted with phenol:chloroform. The insert cDNA for CBP35 was purified by gel electrophoresis. DNA labeled with [ $\alpha$ - $^{32}\text{P}$ ]dCTP was prepared using a random oligodeoxynucleotide primer labeling kit (Pharmacia LKB Biotechnology Inc.).

Typical reactions yielded probe with a specific activity of  $1 \times 10^9$  cpm/ $\mu$ g. Usually  $1 \times 10^6$  cpm/ml of hybridization solution was used.

**Indirect Immunofluorescence.** 3T3 and 3T3-KiMSV cells were seeded on coverslips at a subconfluent density ( $1 \times 10^4$  cells/cm<sup>2</sup>) in Dulbecco's modified Eagle's medium containing 10% calf serum. 3T3 cells were synchronized by serum starvation and then stimulated to enter the cell cycle by serum addition following the procedure described (5). The coverslips were washed in PBS, fixed in 3.7% formaldehyde for 15 min at 4°C, washed in PBS and finally permeabilized in PBS containing 0.2% Triton X-100 for 30 min at 4°C (12). The cells were washed with 20 mM Tris-HCl, 0.5 M NaCl, 2.5% bovine serum albumin, pH 7.5. Each coverslip was then inoculated for 1 h in 100  $\mu$ l of a 1:10 dilution of rabbit antiserum raised against CBP35 (1) in 20 mM Tris-HCl, 0.5 M NaCl, 2.5% bovine serum albumin pH 7.5. The cells were washed in the same Tris buffer containing bovine serum albumin and then incubated in 100  $\mu$ l of a 1:30 dilution of rhodamine-conjugated goat anti-rabbit immunoglobulin (ICN Biomedical) in the same buffer. The coverslips were then rinsed and mounted in 70% glycerol, PBS containing the anti-bleaching agent *n*-propylgallate (5%; Sigma). The slides were viewed with a Leitz epiphase fluorescence microscope using a 25 X objective lens.

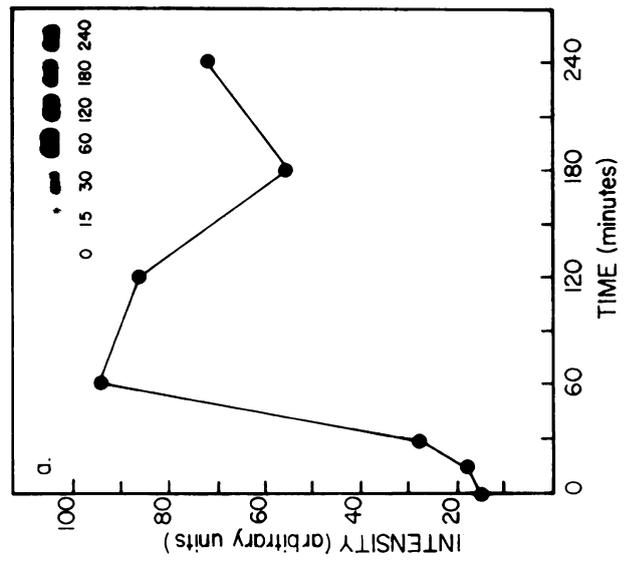
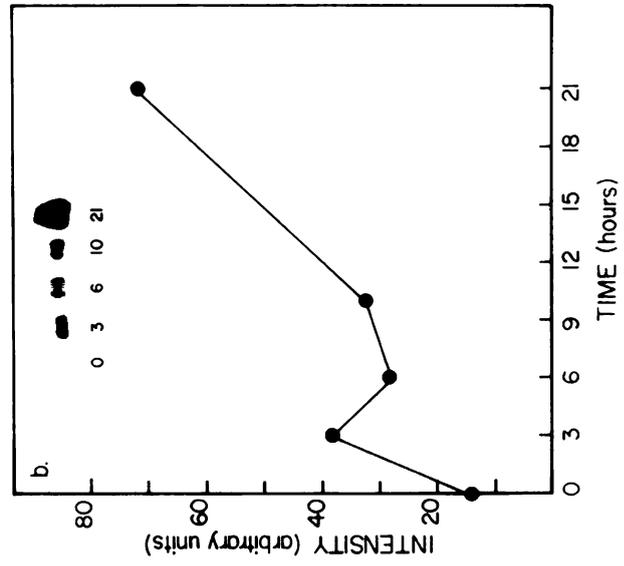
## RESULTS

**Kinetics of the Accumulation of CBP35 mRNA upon Mitogenic Stimulation.** In previous studies, we observed that when serum-starved quiescent 3T3 fibroblasts were activated by the addition of serum, there was an increase in the overall level of CBP35 as detected by Western blotting of total cell extracts (5). This increase in the amount of CBP35 polypeptide accounts, at least in part, for an increase in the percentage of cells that are fluorescently labeled by rabbit antibodies directed against CBP35. The increase in the amount of CBP35 at the protein level, as detected by immunoblotting and by immunofluorescence, occurred as early as 5-8 h after the addition of serum.

Serum-starved 3T3 cells were stimulated by the addition of serum; at various times thereafter, poly(A)<sup>+</sup> RNA was isolated and subjected to Northern analysis using the cDNA probe for CBP35. A single mRNA species, ~1.3 kilobases, was revealed by this probe (3). This mRNA was virtually absent in quiescent cells but could be detected within 30 min after the serum activation (Fig. 1a). The level reached a peak value ~1-2 h after stimulation and then decreased. A more extended study revealed that the early rise (<3 h) was followed by a slight decline before the mRNA level for CBP35 increased some 6-fold at 21 h (Fig. 1b). The generation time and the length of S phase for the 3T3 cells used in the present study have been estimated to be about 24 and 9 h, respectively (5,13). Therefore, the early time points (Fig. 1a) cover the G<sub>0</sub>/early G<sub>1</sub> transition, whereas the 9-h time point (Fig. 1b) is in the middle of the G<sub>1</sub> period. DNA synthesis, as



**Figure 1. Kinetics of the accumulation of mRNA for CBP35 after serum stimulation of growth-arrested 3T3 cells.** Poly(A)<sup>+</sup> mRNA was isolated from equal numbers of cells ( $3 \times 10^6$ ) at the beginning of the experiment and at various time points after serum addition. CBP35 mRNA was detected using a <sup>32</sup>P-labeled cDNA probe. CBP35 mRNA levels were quantitated from densitometric scans of autoradiograms of the Northern blots (insets). *a*, time course study of CBP35 mRNA levels immediately following serum addition; *b*, extended time course study of CBP35 mRNA levels covering the duration of one cell cycle.



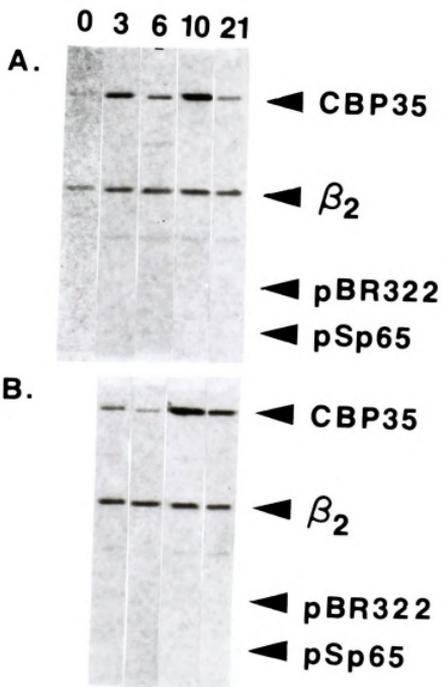
assayed by the incorporation of radioactive thymidine, starts about 15 h after serum addition (5). Thus, the 21-h time point (Fig. 1b) is at late S phase, close to the G<sub>2</sub> period.

In both kinetic studies, it was apparent that the level of mRNA for CBP35 showed an early increase, well before the onset of DNA synthesis in the synchronized population (5). The same blot was subjected to probing with  $\beta_2$ -microglobulin, whose mRNA level has been reported to remain relatively constant during the cell cycle (14). This served to ascertain that approximately the same amount of RNA was electrophoresed in the samples representing the various time points.

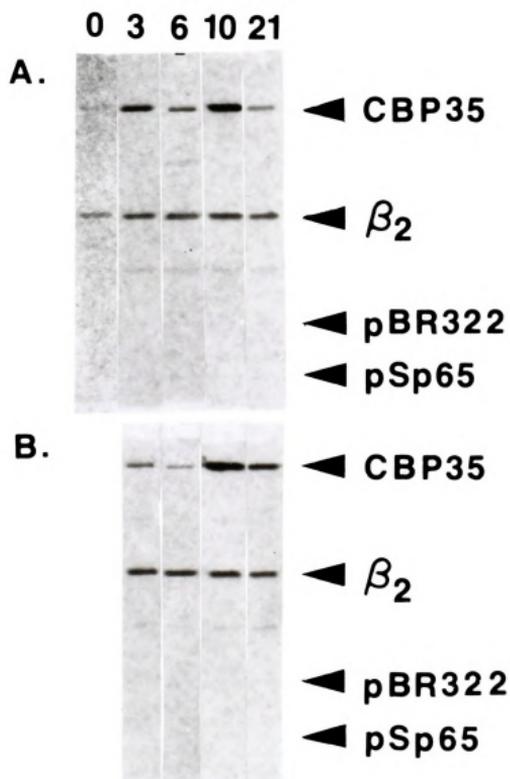
**Analysis of the Transcription Rate of the CBP35 Gene after Stimulation.** Nuclear run-off transcription assays were performed on nuclei derived from cells stimulated for various times after serum addition. The autoradiogram for the slot blots, representing hybridization with various probes, are presented in Fig. 2A; in addition, the data for CBP35 were quantitated and presented in the form of a kinetic profile in Fig. 3. The rate of transcription of the CBP35 gene increased within 3 h after stimulation (Figs. 2A and 3). This may account, at least in part, for the observed rise in the level of accumulated mRNA for CBP35 (Fig. 1). The rate of transcription increased to a maximal level at ~10 h and then decreased by ~21 h (Fig. 3).

The rate of transcription for  $\beta_2$ -microglobulin did not show much variation throughout this time course. This is consistent with previous reports that the

**Figure 2. Gene transcription rates after serum stimulation of quiescent 3T3 cells in the absence (A) and presence (B) of cycloheximide (10  $\mu\text{g/ml}$ ).** Nuclei were isolated at 0 h and at various times after serum addition. Run-off assays were performed, at  $^{35}\text{S}$ -labeled RNA transcripts were hybridized to slot blots containing CBP35 antisense RNA, mouse  $\beta_2$ -microglobulin ( $\beta_2$ ) cDNA, and vectors pBR322 and pSp65. The blots were subjected to autoradiography.

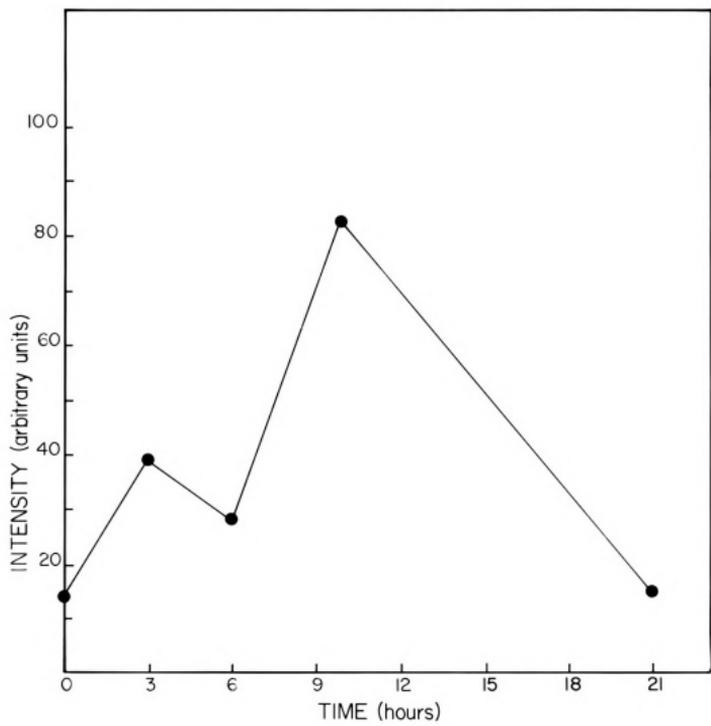


**Figure 2. Gene transcription rates after serum stimulation of quiescent 3T3 cells in the absence (A) and presence (B) of cycloheximide (10  $\mu\text{g/ml}$ ).** Nuclei were isolated at 0 h and at various times after serum addition. Run-off assays were performed, at  $^{35}\text{S}$ -labeled RNA transcripts were hybridized to slot blots containing CBP35 antisense RNA, mouse  $\beta_2$ -microglobulin ( $\beta_2$ ) cDNA, and vectors pBR322 and pSp65. The blots were subjected to autoradiography.

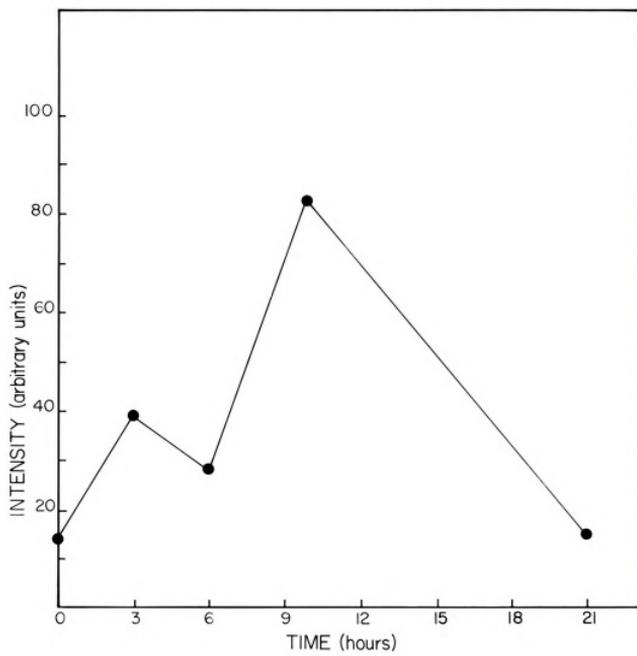




**Figure 3. Kinetics of the changes in relative transcription rates of the CBP35 gene following serum stimulation of quiescent 3T3 fibroblasts.** Densitometric scans of the autoradiograms shown in Fig. 2A were performed to quantitate the relative transcription rates.



**Figure 3. Kinetics of the changes in relative transcription rates of the CBP35 gene following serum stimulation of quiescent 3T3 fibroblasts.** Densitometric scans of the autoradiograms shown in Fig. 2A were performed to quantitate the relative transcription rates.





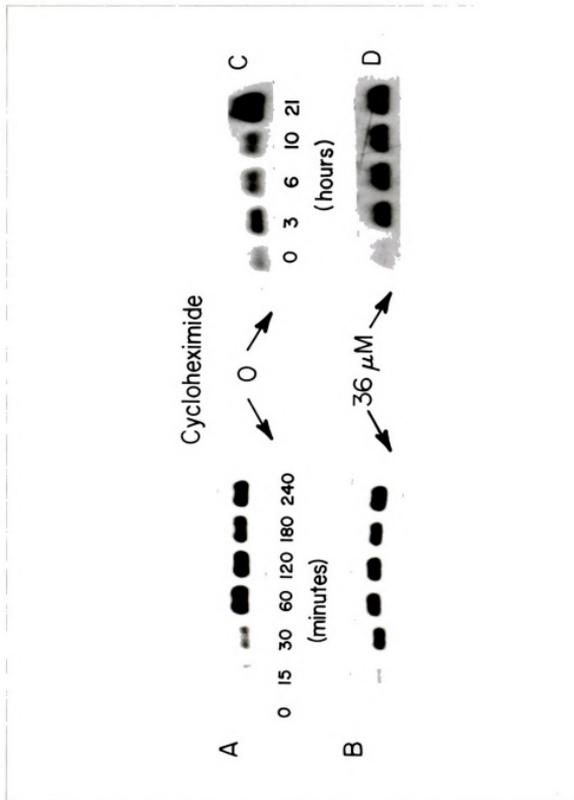
transcription rate of the  $\beta_2$ -microglobulin gene is relatively constant throughout the cell cycle (10). There was no detectable hybridization between the radioactive mRNA derived from the cell-free transcription and the vectors pBR322 and pSp65 (Fig. 2A). These results provide information on the specificity of our nuclear run-off assay and the kinetics of the increase in transcription rate of the CBP35 gene.

#### Effect of Cycloheximide on Transcription Rate and mRNA Accumulation

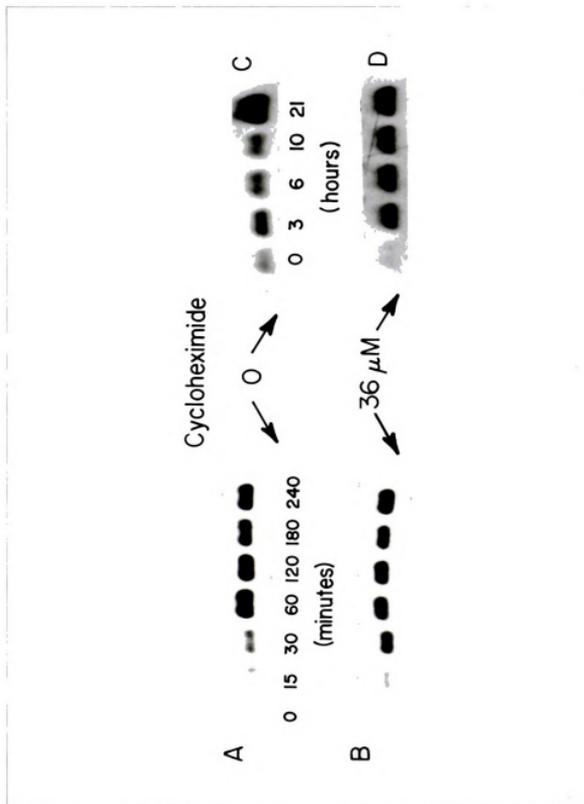
following Mitogen Addition. The elevated transcription of the CBP35 gene could be a primary event, directly resulting from the signals transduced from the binding of growth factors to their cell surface receptors. Alternatively, this increased expression of the CBP35 gene could be a secondary phenomenon, in which the mitogenic signal first induces the transcription of other genes and synthesis of the corresponding gene products, which in turn enter the nucleus to activate the CBP35 gene. To distinguish these possibilities, the 3T3 cell cultures were stimulated by serum in the presence of cycloheximide (10  $\mu$ g/ml). In our 3T3 cell system, this concentration of cycloheximide inhibited >95% of the incorporation of  $^{14}$ C-labeled amino acids into trichloroacetic acid-precipitable material, in agreement with the result of previous studies on the effect of the drug on protein synthesis in 3T3 cells (15,16). In the absence of *de novo* protein synthesis, the transcription of the CBP35 gene was nevertheless elevated, increasing at least 5-fold over a course of 10 h (Fig. 2B). Over the same time course, the transcription rate of the gene for  $\beta_2$ -microglobulin remained constant. Thus, it appears that the gene for CBP35 can be induced directly as a result of mitogen addition, without



**Figure 4. Effect of cycloheximide on the levels of accumulated mRNA for CBP35.** Poly(A)<sup>+</sup> mRNA was isolated from quiescent and serum-stimulated 3T3 cells. Identical isolations were performed from quiescent and serum-stimulated cells that had been cultured in the presence of cycloheximide (10  $\mu$ g/ml). <sup>32</sup>P-labeled CBP35 cDNA was used as a probe. The Northern blots were subjected to autoradiography. *A*, early time course analysis of CBP35 mRNA levels; *B*, early time course levels CBP35 mRNA levels in cultures stimulated in the presence of cycloheximide; *C*, analysis of CBP35 mRNA levels covering the time course of cell cycle following serum stimulation; and *D*, extended time course analysis of CBP35 mRNA levels in cultures stimulated in the presence of cycloheximide. *A* and *C* represent data from Fig. 1; the experiments in *B* and *D* were performed in parallel with those in *A* and *C*, respectively.



**Figure 4. Effect of cycloheximide on the levels of accumulated mRNA for CBP35.** Poly(A)<sup>+</sup> mRNA was isolated from quiescent and serum-stimulated 3T3 cells. Identical isolations were performed from quiescent and serum-stimulated cells that had been cultured in the presence of cycloheximide (10  $\mu$ g/ml). <sup>32</sup>P-labeled CBP35 cDNA was used as a probe. The Northern blots were subjected to autoradiography. *A*, early time course analysis of CBP35 mRNA levels; *B*, early time course levels CBP35 mRNA levels in cultures stimulated in the presence of cycloheximide; *C*, analysis of CBP35 mRNA levels covering the time course of cell cycle following serum stimulation; and *D*, extended time course analysis of CBP35 mRNA levels in cultures stimulated in the presence of cycloheximide. *A* and *C* represent data from Fig. 1.; the experiments in *B* and *D* were performed in parallel with those in *A* and *C*, respectively.





the requirement of synthesis of other proteins.

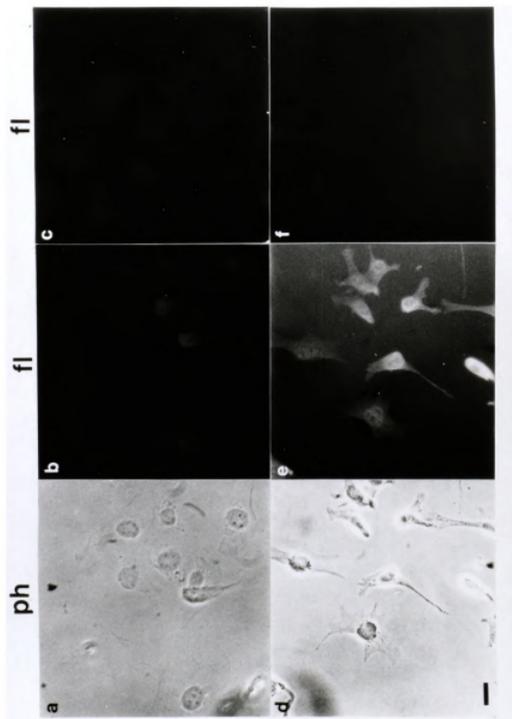
Northern blotting analysis showed that, in the presence of cycloheximide (10  $\mu\text{g/ml}$ ), there was also increased accumulation of the mRNA for CBP35 when quiescent 3T3 cells are stimulated with serum (Fig. 4). Moreover, the level of mRNA for CBP35 was found to be higher in the presence of cycloheximide than in its absence. For example, during the first 30 min after the addition of serum, there was a higher level of CBP35 mRNA in cultures with cycloheximide (Fig. 4B) than in cultures without the drug (Fig. 4A). Similarly, over the long term time course, there was always more CBP35 mRNA in cultures treated with cycloheximide (Fig. 4D) than in cultures devoid of the inhibitor (Fig. 4C). This phenomenon of superinduction, in which the increase in mRNA level for a given gene is higher in the presence of cycloheximide than in its absence, had been documented for a number of genes activated early during the process of mitogenic stimulation, including the oncogenes *c-myc* and *c-fos* (17,18).

#### **Comparison of the Expression of the CBP35 Gene in Normal and Transformed**

**Cells.** By using Western blotting techniques, we previously documented that the level of CBP35 was considerably higher in transformed cells than in their normal counterparts (6). For example, chicken embryo fibroblasts transformed by Rous sarcoma virus showed 5-10 times more CBP35 than secondary cultures of the same fibroblasts. This was also true when mouse 3T3 fibroblasts were compared to 3T3 cells transformed with Kirsten murine sarcoma virus (3T3-KiMSV cells).

When these latter two cell types were subjected to immunofluorescence analysis

**Figure 5. Immunofluorescence analyses comparing CBP35 in 3T3 and 3T3-K1MSV cells.** 3T3 and 3T3-K1MSV cells were prepared for immunofluorescence (*f*) as described under "Materials and Methods" and incubated with rabbit anti-CBP35 (*panels b* and *e*, respectively) or preimmune rabbit serum (*panels c* and *f*, respectively). The binding of the rabbit antibodies was detected with rhodamine-conjugated goat anti-rabbit immunoglobulin. *Panels a* and *d* are the corresponding phase contrast photographs (*ph*). *Bar* = 45  $\mu$ m.



with antibodies directed against CBP35, the 3T3-KiMSV cells showed much higher intensity of fluorescence than their normal counterparts (Fig. 5). At least part of the reason for this drastic difference in staining intensity must be ascribed to a difference in the level of the CBP35 polypeptide in the two cell types (6). In both 3T3 and 3T3-KiMSV cells, the staining pattern reflected the nuclear localization of the lectin.

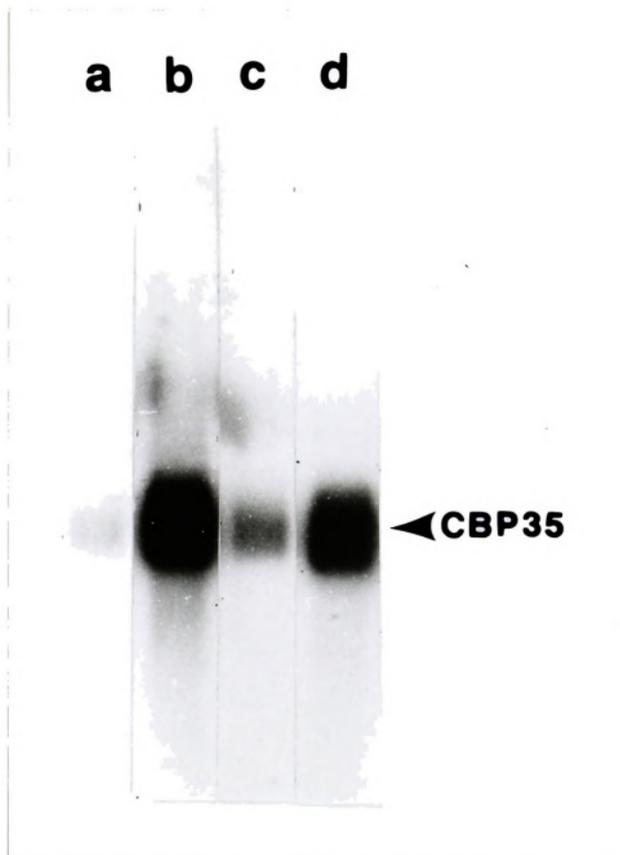
We have corroborated these differences between 3T3 and 3T3-KiMSV cells in the expression of CBP35 at the level of accumulated mRNA, as well as at the level of nuclear transcription rates. By using nonsynchronized cell populations, poly(A)<sup>+</sup> RNA was prepared from an equal number of cells derived from sparse and confluent cultures and subjected to Northern blot analysis. In both cases, much higher amounts (10-15-fold) of CBP35 mRNA were detected in 3T3-KiMSV cells than in normal 3T3 cells (Fig. 6). The RNA species in 3T3-KiMSV cells is of the same size (~1.3 kilobases) as that in 3T3 cells.

To determine whether the differences seen in the levels of CBP35 mRNA between normal and virally transformed cells are due to transcription rate differences, nuclear run-off transcription assays were performed. Under both sparse and confluent culture conditions, 3T3-KiMSV cells had a higher transcription rate for the CBP35 gene than 3T3 cells (Fig. 7). This conclusion takes into consideration the fact that equal amounts of radioactivity from the different run-off experiments were hybridized in the slot blots shown in Fig. 7.

Experiments were also carried out to compare the half-lives of the CBP35 mRNA in normal and transformed 3T3 cells. Cells were cultured in the presence

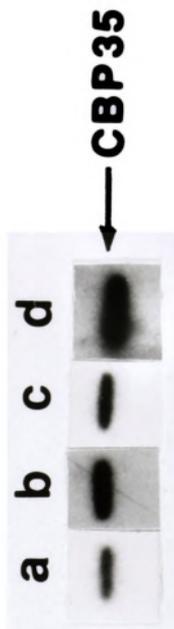


**Figure 6. Northern blot analysis of CBP35 mRNA in 3T3 and 3T3-KiMSV cells.** *Lane a*, 3T3 cells seeded at a density of  $1 \times 10^4$  cells/cm<sup>2</sup>; *lane b*, 3T3-KiMSV cells seeded at  $1 \times 10^4$  cells/cm<sup>2</sup>; *lane c*, 3T3 cells seeded at  $5 \times 10^4$  cells/cm<sup>2</sup>; *lane d*, 3T3-KiMSV cells seeded at  $5 \times 10^4$  cells/cm<sup>2</sup>. Equal amounts of total RNA (120  $\mu$ g) from each culture were used to isolate poly(A)<sup>+</sup> RNA, which was then electrophoresed and hybridized with a <sup>32</sup>P-labeled cDNA probe for CBP35. The autoradiograms were exposed for 72 h without an intensifying screen.



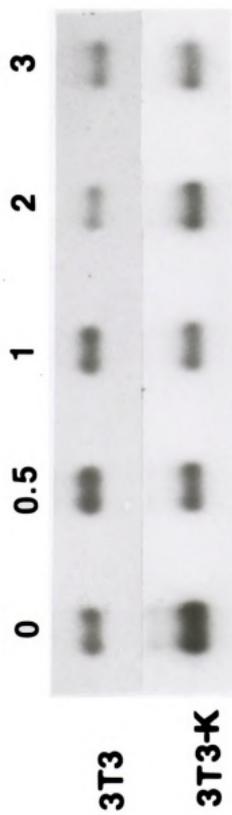


**Figure 7. Autoradiogram of nuclear run-off assays comparing the transcription rate of the CBP35 gene in 3T3 and 3T3-KiMSV cells.** *Lane a*, 3T3 cells seeded at an initial density of  $1 \times 10^4$  cells/cm<sup>2</sup>; *lane b*, 3T3-KiMSV cells seeded at  $1 \times 10^4$  cells/cm<sup>2</sup>; *lane c*, 3T3 cells seeded at  $5 \times 10^4$  cells/cm<sup>2</sup>; *lane d*, 3T3-KiMSV cells seeded at  $5 \times 10^4$  cells/cm<sup>2</sup>. <sup>32</sup>P-Labeled RNA transcripts from run-off assays were subjected to scintillation counting, and equal amounts of radioactivity were hybridized to slot blots containing CBP35 antisense RNA. The autoradiogram was exposed for 7 days with an intensifying screen.

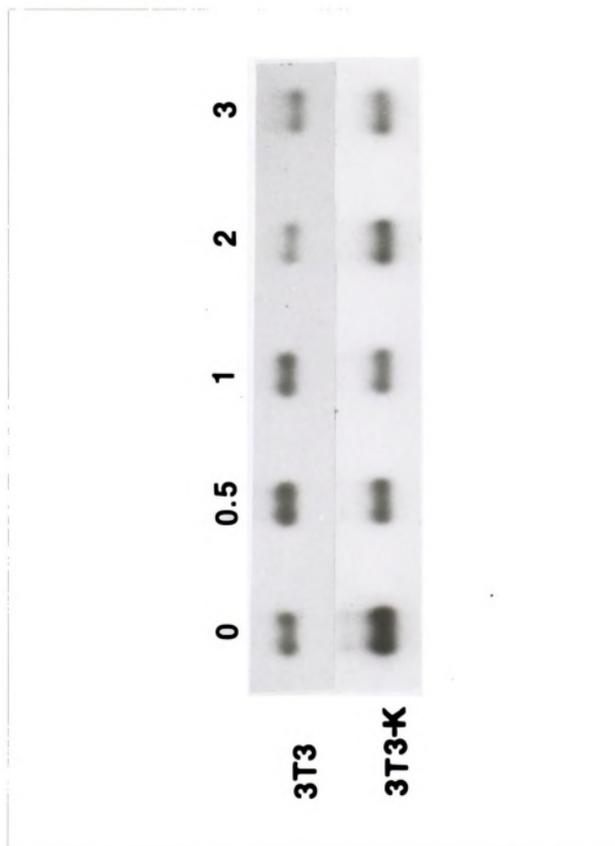




**Figure 8. Northern blot analysis to compare half-lives of CBP35 mRNA in 3T3 and 3T3-KiMSV cells.** The poly(A)<sup>+</sup> RNA was isolated after incubation of the cells in actinomycin D (8 μg/ml) for the indicated times (h). The autoradiogram for 3T3 cells was exposed for 17 h with an intensifying screen, and the autoradiogram for the 3T3-KiMSV cells was exposed for 27 h without an intensifying screen. Half-lives were determined by performing scanning densitometric analysis of the band intensities. The less intense band immediately above the major band seen in blots of 3T3-KiMSV cells corresponds to poly(A)<sup>+</sup> RNA that is incompletely processed. It is observed only in RNA preparations derived from whole cells (including nuclear RNA), but is not observed in cytoplasmic poly(A)<sup>+</sup> RNA samples.



**Figure 8. Northern blot analysis to compare half-lives of CBP35 mRNA in 3T3 and 3T3-KiMSV cells.** The poly(A)<sup>+</sup> RNA was isolated after incubation of the cells in actinomycin D (8 µg/ml) for the indicated times (h). The autoradiogram for 3T3 cells was exposed for 17 h with an intensifying screen, and the autoradiogram for the 3T3-KiMSV cells was exposed for 27 h without an intensifying screen. Half-lives were determined by performing scanning densitometric analysis of the band intensities. The less intense band immediately above the major band seen in blots of 3T3-KiMSV cells corresponds to poly(A)<sup>+</sup> RNA that is incompletely processed. It is observed only in RNA preparations derived from whole cells (including nuclear RNA), but is not observed in cytoplasmic poly(A)<sup>+</sup> RNA samples.





of actinomycin D for various times; RNA from these treated cells was isolated and subjected to Northern blotting with the cDNA probe for CBP35. The autoradiogram showed that the 3T3-KiMSV cells displayed an enhanced stability of the CBP35 mRNA (Fig. 8). The levels of the mRNA for CBP35 were quantitated from the autoradiogram by densitometric scanning. CBP35 mRNA in 3T3 cells had a half-life of 120 min, compared to that of 200 min in 3T3-KiMSV cells. It appears, therefore, that increased stability of the mRNA allows for its accumulation in the transformed cells.

#### **Comparison of the Expression of the CBP35 Gene in Sparse and Confluent**

**Cultures of 3T3 Cells.** In previous studies, we observed that sparse cultures of 3T3 cells showed intense immunofluorescence staining for CBP35, predominantly within the nuclei, whereas in confluent monolayers of the same cells, the staining intensity decreased with a dramatic loss of CBP35 from the nuclei (5,12). In the present study, it was quite to our surprise, therefore, to find that the transcription rate, as well as level of accumulated mRNA, for CBP35 remained appreciable in the confluent cultures. As quantitated by densitometric analysis, the transcription rate for the CBP35 gene in sparse cultures of 3T3 cells was approximately the same as that for confluent cultures (Fig. 7). However, the levels of CBP35 mRNA, as detected by Northern blot analysis, were higher for confluent cultures of 3T3 cells than for sparse cells (Fig. 6).

## DISCUSSION

We have compared the rate of nuclear transcription and the level of accumulated mRNA for CBP35 in cultures of 3T3 cells representing quiescent and proliferative states: (a) serum-starved *versus* serum-stimulated cells, (b) confluent monolayers *versus* sparse cultures, and (c) normal fibroblasts *versus* cells transformed by KiMSV. The basis for these comparisons is the observation that at the polypeptide level, as detected by Western blotting and/or immunofluorescence, the expression of CBP35 was correlated with the proliferating state of the cell (5,6). Our present studies showed that, in general, high levels of CBP35 protein in proliferating 3T3 cells reflect higher levels of mRNA as well as elevated nuclear transcription rates of the gene. In one instance, however, there was an exception: although confluent monolayers of 3T3 cells exhibited much lower levels of CBP35 protein than sparse proliferating cells, their mRNA levels for CBP35 were higher when compared on the basis of equal amount of RNA.

When serum-starved 3T3 cells were stimulated by the addition of serum, there was an early rise in the transcription rate of the CBP35 gene. Within 0.5 h, there was a detectable increase in the level of CBP35 mRNA, which exhibited a peak at about 3 h. This was followed by a transient decrease and then a second rise, leading up to a ~6-fold elevation over approximately the next 10 h. Such a bimodal kinetic pattern for the level of mRNA as a function of time after mitogen activation has also been seen when the mRNA corresponding to the *c-myc* gene

was monitored after stimulation of quiescent 3T3 cells by fetal calf serum (19).

The effect of serum growth factors on the regulation of CBP35 and *c-myc* genes is similar in another respect: when the stimulation of the cells was carried out in the presence of cycloheximide, superinduction of mRNA accumulation for these genes was observed. This phenomenon may be rationalized in terms of the inhibitory effect of cycloheximide on the production of shortlived nucleases that normally could affect the stability of mRNA. Alternatively, the results may be interpreted to indicate that the increased expression of the CBP35 gene is a direct result of signals transduced by the binding of growth factors to their plasma membrane receptors, without the requirement of prior synthesis of other gene products. The observation that a similar pattern of transcription rate variations occurred in the presence and absence of cycloheximide lends support to the latter hypothesis. This pattern of changes includes an early increase, a slight decline at 6 h, and an approximately 5-fold elevation of CBP35 mRNA under both sets of conditions.

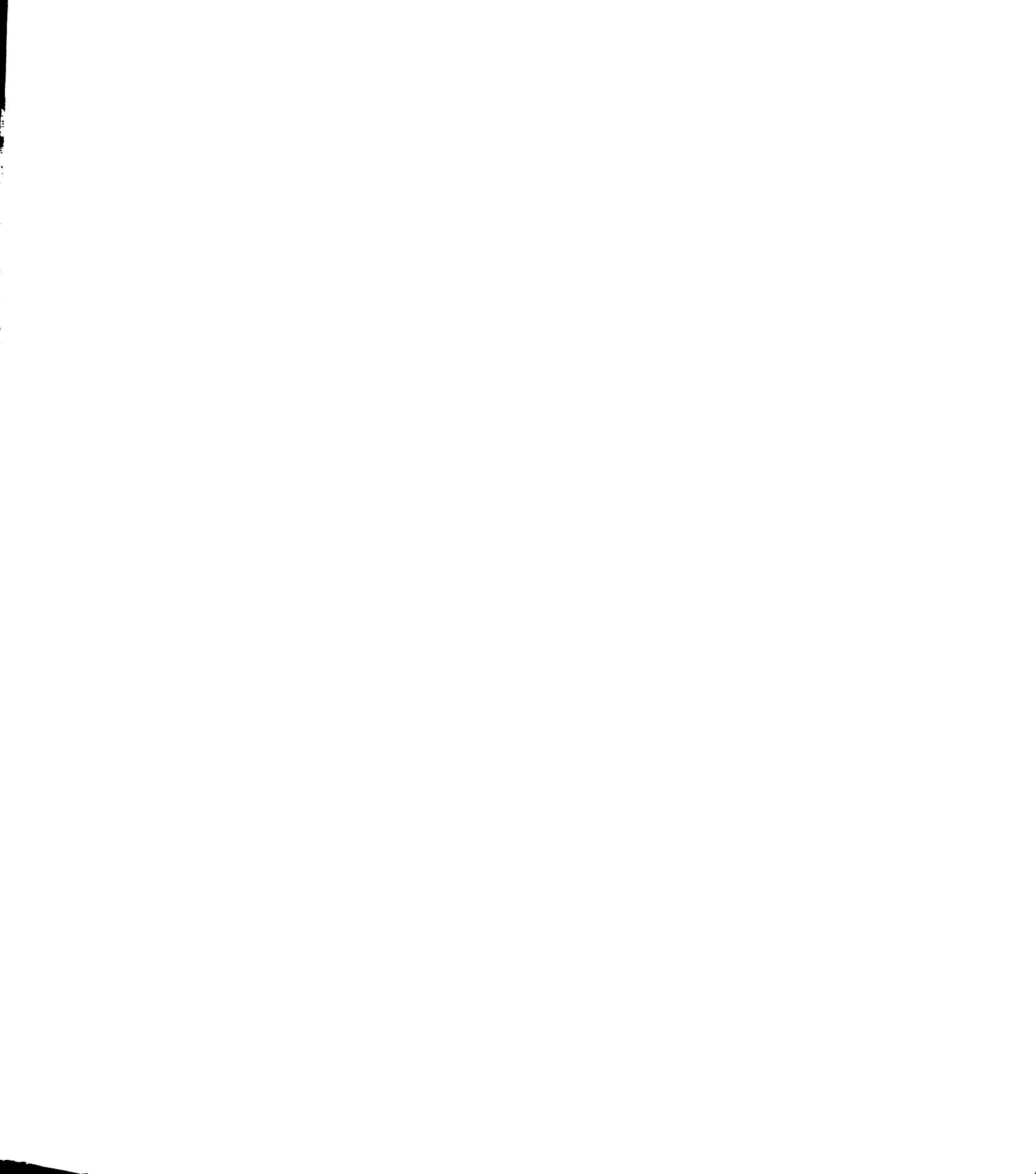
Thus, CBP35 appears to be regulated in manner comparable to other mitogen activated genes, including the oncogenes *c-fos* and *c-myc* (17,18). Recent evidence indicates that inhibition of the expression of these oncogene products (*e.g.* by antisense RNA) results in a blockage of the progression of a stimulated cell into DNA synthesis and cell division (20,21). This suggests that oncogene products are required for normal cell cycle progression during cell activation. It would be of interest, therefore, to test if inhibition of the expression of CBP35 protein might also result in growth inhibition.

Like the oncogenes, the amounts of CBP35 are elevated in "deregulated" cell populations, *i.e.* transformed cells. The data suggest that the expression of the CBP35 gene is directly affected by transformation at the transcriptional level. In addition, our results also indicate that there is greater stability of the CBP35 mRNA in 3T3-KiMSV cells than in their normal counterparts. Both of these effects contribute to a marked elevation in the accumulated mRNA for CBP35 and the protein product in transformed cells. In the transformed murine fibrosarcoma cell line, UV-2237-IP3, there is a selectively elevated expression of one lectin, L-34 (which corresponds to CBP35), but not of another, L-14 (which corresponds to CBP13.5) (22). It was suggested that increased amounts of L-34 may reflect increased malignancy of the cells (23).

By using differential screening and cDNA cloning techniques, a set of some 10 mRNAs that appears in 3T3 cells soon after growth stimulation by serum has been identified (17,18). These have been termed growth-related immediate early genes. The kinetics of the induction of expression of the CBP35 gene are similar to these immediate early genes (3CH61, 3CH77, 3CH268, etc.) in that (a) they are all elevated within 10-30 min after serum addition, and (b) they are superinduced in the presence of cycloheximide. The mRNAs corresponding to most of these immediate early genes reached peak levels between 40 and 120 min after serum addition and rapidly decayed thereafter (17,18), whereas that for CBP35 showed bimodal kinetics, with a long term increase over a period of approximately 20 h. Recent nucleotide sequence analyses have revealed that one of the immediate early genes (3CH268) contains three tandem "zinc finger" sequences typical of a

class of eukaryotic transcription factors (24) and another (3CH77) encodes a member of the superfamily of ligand-binding transcription factors that includes the steroid and thyroid hormone receptors (25). We have recently provided evidence that CBP35 is a component of hnRNP on the basis of immunochemical localization of the lectin in hnRNP fractions (2) and on the basis of amino acid homology of the CBP35 polypeptide to other hnRNP proteins (4). Our previous observations of the proliferation-dependent expression of the CBP35 polypeptide (5,6) and our present documentation of the correlation of increased nuclear transcription and mRNA accumulation of the CBP35 gene with data indicating that several other hnRNP polypeptides are also regulated upon quiescent cell to proliferative cell transition (26,27). Although the function of CBP35 is not known, its identification as a component of hnRNP suggests that it may play a role in the processing, packaging, or transport of mRNA from the nucleus into the cytoplasm. Glycosylated components (28) and CBP35 (12) are found in both the nucleus and the cytoplasm, suggesting the possibility that they may shuttle between the two subcellular compartments in a transport role. Indeed, preliminary experiments using a cell-free assay for RNA transport from isolated nuclei indicate that CBP35 is cotransported with mRNA in the form of a ribonucleoprotein complex (J. G. Laing, E. A. Werner, and R. J. Patterson, unpublished observations).

Finally, it is noteworthy to discuss the one exception to the correlation between CBP35 protein levels and CBP35 mRNA levels and gene transcription rates. Although dense cultures of 3T3 cells exhibit approximately the same transcription rate for the CBP35 gene as sparse cultures, there are higher levels of



detectable mRNA in the dense monolayers. Thus, posttranscriptional mechanisms, such as mRNA stability, must be involved in the accumulation of CBP35 RNA in 3T3 cells. Such mechanisms have been reported in the regulation of histone mRNA (29) and *c-myc* RNA (30). More surprisingly, despite having higher amounts of mRNA for CBP35, dense 3T3 cells exhibit less of the protein than sparse 3T3 cells (5). Therefore, additional mechanisms for this difference at the translation level (*e.g.* ribosome binding) or posttranslational level (*e.g.* protein stability) may be involved. Clearly, there is need for caution in interpreting correlative data from different sets of experiments; simultaneously Northern and Western blot determinations on CBP35 mRNA and protein, respectively, from the same cell population are required.

#### **ACKNOWLEDGEMENTS**

We thank Mrs. Linda Lang for her help in the preparation of the manuscript.

## REFERENCES

1. Roff CF & Wang JL (1983) *J Biol Chem* **258**: 10657-10663.
2. Laing JG & Wang JL (1988) *Biochemistry* **27**: 5329-5334.
3. Jia S, Mee RP, Morford G, Agrwal N, Voss PG, Moutsatsos IK & Wang JL (1987) *Gene (Amst.)* **60**: 197-204.
4. Jia S & Wang JL (1988) *J Biol Chem* **263**: 6009-6011.
5. Moutsatsos IK, Wade M, Schindler M & Wang JL (1987) *Proc Natl Acad Sci USA* **84**: 6452-6456.
6. Crittenden SL, Roff CF & Wang JL (1984) *Mol Cell Biol* **4**: 1252-1259.
7. Stuart P, Ito M, Stewart C & Conrad SE (1985) *Mol Cell Biol* **5**: 1490-1497.
8. Maniatis T, Fritsch EF & Sambrook J (1982) *Molecular Cloning: A Laboratory Manual* pp. 191-193, Cold Spring Harbor Laboratory, Cold Spring Harbor NY.
9. Linial M, Gunderson N & Groudine M (1985) *Science* **230**: 1126-1132.
10. Stewart CJ, Ito M & Conrad SE (1987) *Mol Cell Biol* **7**: 1156-1163.
11. Parnes JR & Seidman JG (1982) *Cell* **29**: 661-669.
12. Moutsatsos IK, Davis JM & Wang JL (1986) *J Cell Biol* **102**: 477-483.
13. Steck PA, Voss PG & Wang JL (1979) *J Cell Biol* **83**: 562-575.
14. Kelly K, Cochran BH, Stiles CD & Leder P (1983) *Cell* **35**: 603-610.
15. Brooks RF (1977) *Cell* **12**: 311-317.
16. Campisi J, Medrano E, Morreo G & Pardee AB (1982) *Proc Natl Acad Sci USA* **79**: 436-440.
17. Lau LF & Nathans D (1985) *EMBO J.* **4**: 3145-3151.
18. Lau LF & Nathans D (1987) *Proc Natl Acad Sci USA* **84**: 1182-1186.

19. Müller R, Bravo R, Burckhardt J & Curran T (1984) *Nature* **312**: 716-720.
20. Nishikura K, & Murray JM (1987) *Mol Cell Biol* **7**: 639-649.
21. Freytag SO (1988) *Mol Cell Biol* **8**: 1614-1624.
22. Raz A, Carmi P & Pazerini G (1988) *Cancer Res* **48**: 645-659.
23. Raz A, Meromsky L, Zvibel I & Lotan R (1987) *Int J Cancer* **39**: 353-360.
24. Christy BA, Lau LF & Nathans D (1988) *Proc Natl Acad Sci USA* **85**: 7857-7861.
25. Hazel TG, Nathans D & Lau LF (1988) *Proc Natl Acad Sci USA* **85**: 8444-8448.
26. Celis JE, Bravo R, Arenstorf HP & LeStourgeon WM (1986) *FEBS Lett* **194**: 101-109.
27. Leser GP & Martin TE (1987) *J Cell Biol* **105**: 2083-2094.
28. Hart GW, Holt GD & Haltiwanger RS (1988) *Trends Biochem Sci* **13**: 380-384.
29. Marzluff WF & Pandey PB (1988) *Trends Biochem Sci* **13**: 49-51.
30. Piechaczyk M, Yang J-Q, Blanchard J-M, Jeanteur P & Marcu KB (1985) *Cell* **42**: 589-597.

**CHAPTER III**

**CARBOHYDRATE BINDING PROTEIN 35:**

**Properties of the Recombinant Polypeptide  
and the Individuality of the Domains\***

Neera Agrwal, Quan Sun, Sung-Yuan Wang, and John L. Wang

Department of Biochemistry  
Michigan State University  
East Lansing, MI 48824



**FOOTNOTES**

- This work was supported by grants GM-38740 and GM-27203 from the National Institutes of Health.
  
- <sup>1</sup> The abbreviations used are: rCBP35, recombinant carbohydrate binding protein 35; N- and C-domains, NH<sub>2</sub>-terminal and COOH-terminal portions, respectively, of the CBP35 polypeptide; IPTG, isopropyl- $\beta$ -D-thiogalactopyranoside; 2-ME, 2-mercaptoethanol; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; DSC, differential scanning calorimetry.

**SUMMARY**

The cDNA clone for Carbohydrate Binding Protein 35 (CBP35) was engineered into the bacterial expression vector pIN III *ompA2*, which directs the secretion of the expressed protein into the periplasmic space. Recombinant CBP35 was purified from this system, at a level of  $\sim 50$  mg per liter of bacterial culture. Digestion of recombinant CBP35 with collagenase D, followed by purification using saccharide-specific affinity chromatography yielded a  $M_r \sim 16,000$  polypeptide, corresponding to the COOH-terminal domain (residues 118-264) of the CBP35 polypeptide. This indicates that the COOH-terminal half of CBP35 contains the carbohydrate recognition domain, consistent with its sequence homology to other S-type lectins. The NH<sub>2</sub>-terminal domain (residues 1-137) was derived by site-directed mutagenesis of the cDNA, in which stop codons are inserted in place of Gly 138 and Gly 139, and expression of the mutant cDNA in the same pIN III *ompA2* system. The purified NH<sub>2</sub>-terminal domain failed to bind to saccharide-specific affinity resins. Differential scanning calorimetry of rCBP35 and its individual domains yielded transition temperatures of  $\sim 39^\circ\text{C}$  and  $\sim 56^\circ\text{C}$  for the NH<sub>2</sub>- and COOH-terminal domains, respectively. Lactose binding by the COOH-terminal domain shifted the transition temperature to  $65^\circ\text{C}$ , whereas sucrose failed to yield the same effect. These results suggest that the individual domains of the CBP35 polypeptide are folded independently.



## INTRODUCTION

Carbohydrate Binding Protein 35 (CBP35,  $M_r$  35,000) is a galactose specific lectin which was initially purified from murine fibroblasts (1). CBP35 or its homologs have since been found in a variety of tissues and species (2). Moreover, the available data indicate that the same protein has been studied under various different names (for a review, see (3)): (a) L-34, a lectin from murine tumor cells (4,5); (b) Mac-2, a surface antigen of mouse macrophages (6,7); (c) IgE-binding protein from rat basophilic leukemia cells (8,9); and (d) RL-29 (10) and HL-29 (11), lectins from rat and human lungs, respectively. Using indirect immunofluorescence and cell fractionation studies, we have localized CBP35 predominantly to the cytoplasm and nucleus of mouse 3T3 fibroblasts (12). CBP35 gene expression, both at the transcriptional level and protein level, is elevated upon mitogenic stimulation of quiescent cultures of 3T3 cells (13,14).

Sequence and hydropathy analysis of the cDNA clone for CBP35 showed that it has two distinct structural domains (3,15). The carboxyl terminal half contains significant homology to other  $\beta$ -galactoside binding lectins, and thus may contain the carbohydrate recognition domain. The amino terminal half has eight contiguous nine amino acid repeats with the sequence Pro-Gly-Ala-Tyr-Pro-Gly-X-X-X, which consequently makes this domain highly proline- and glycine- rich. Thus, CBP35 has been classified as an S-type lectin with a bifunctional motif (16,17).

In this paper, we have taken advantage of an improved prokaryotic expression vector to produce milligram amounts of recombinant CBP35 (rCBP35) and its corresponding NH<sub>2</sub>-terminal and COOH-terminal halves (designated as N-domain and C-domain, respectively). Comparison of the physico-chemical properties of the N- and C-domains with the parent rCBP35 polypeptide suggest independent folding of the individual domains.

## MATERIALS AND METHODS

### Construction of prCBP35s and Purification of rCBP35

The CBP35 cDNA was excised out of the plasmid pWJ31 using EcoRI restriction. The cDNA was then ligated into the prokaryotic expression vector pIN III *ompA2* (18) using standard subcloning techniques (19). The resulting recombinant clone, containing the plasmid for the production of rCBP35 in a soluble form (prCBP35s) (Fig. 1A), was then transformed into the *E. coli* JA221 strain.

prCBP35s was freshly transformed into JA221 cells prior to each purification. An overnight culture of prCBP35s was grown in LB broth containing 100  $\mu\text{g/ml}$  ampicillin at 37°C. One liter of Terrific Broth (2.31 g  $\text{KH}_2\text{PO}_4$ , 12.54 g  $\text{K}_2\text{HPO}_4$ , 12 g bacto-tryptone, 24 g bacto-yeast extract, and 4 ml glycerol in 1 liter (ref. 19)) containing 100  $\mu\text{g/ml}$  ampicillin was then inoculated with the overnight culture, and then allowed to grow at room temperature for 2 hours, after which isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG; Research Organics) was added to a final concentration of 50  $\mu\text{M}$ , and the culture allowed to grow at room temperature for 16-24 hours. After harvest, the cell pellet was washed with ice cold phosphate-buffered saline (0.13 M NaCl, 5 mM sodium phosphate, pH 7.5), and then resuspended in 25 ml ice cold 1 M Tris-HCl (pH 7.0) containing 10 mM 2-mercaptoethanol (2-ME) (Pierce Chemical), 1 mM phenylmethanesulfonyl fluoride (BMB), 0.5  $\mu\text{g/ml}$  leupeptin (BMB), 30 KIU/ml aprotinin (BMB), and .01 mg/ml soybean trypsin inhibitor (Sigma). The resuspended pellet was kept on

ice for 30 minutes, after which it was centrifuged at 90,000 x g for 30 minutes at 4°C.

The supernatant, representing the periplasmic fraction, was then fractionated by ammonium sulfate (65% of saturation) at 4°C. The precipitated protein was recovered by centrifugation at 12,000 x g, and resuspended in loading buffer (75 mM Tris-HCl pH 7.5, 75 mM NaCl, 2 mM EDTA, 1 mM 2-ME, 1 mM phenylmethanesulfonyl fluoride). The resuspended protein was then combined with 10 ml of asialofetuin—Affi-gel 15 (BioRad), and the slurry dialyzed overnight at 4°C (2 changes of 2 liters each) against loading buffer. The slurry was packed into a column, washed with 20 column volumes of loading buffer (25 ml/hour), and eluted with 0.4 M lactose in loading buffer. The eluted protein was concentrated by ultrafiltration, and the buffer exchanged to 10 mM Tris-HCl pH 8.5, 1 mM EDTA, 1 mM 2-ME, using a PM10 filter (Amicon) in an Amicon stirred cell. Where noted, rCBP35 was purified entirely without the addition of 2-ME.

### **Site-Directed Mutagenesis and Generation of Recombinant Clones**

The CBP35 cDNA was subcloned into the phagemid pUC119 (20), and the recombinant vector, designated pWJ1131, was transformed into the *E. coli* strain CJ236 (21). After uracil containing single stranded DNA was obtained, site-directed mutagenesis was carried out using the BioRad Muta-Gene Phagemid *in vitro* Mutagenesis kit (BioRad Laboratories, Richmond CA). All oligonucleotides were synthesized in the Michigan State University Macromolecular Structure Facility.

In order to obtain the amino terminal half of CBP35, glycines 138 and 139 (the numbering system follows that of murine L-34 (4) and Mac-2 (6), since these cDNAs identified the translation initiation methionine (residue 1)) were converted to stop codons by using the oligonucleotide 5'-GATCAGCATGCGAAGCTTGA CTCATCAAGGCAACGGCAGGTC-3' (Fig. 1B). The above oligonucleotide also inserted a HindIII restriction site downstream of the stop codons. The fragment covering the 5'-end of the cDNA through the HindIII site was subcloned into pIN III *ompA2*. The construct was transformed into JA221 cells.

HindIII digestion of the mutagenized cDNA also allowed us to subclone the fragment corresponding to the carboxyl terminal half of the polypeptide (Fig. 1B). The carboxyl terminal half was excised out of pWJ1131 with HindIII-BamHI restriction, and then ligated into the vector pIN III *ompA1*. The recombinant expression plasmid was then transformed into *E. coli* JA221.

The periplasmic fractions of *E. coli* transformed with the mutant recombinant clones were harvested in the same manner as that for rCBP35.

### **Preparation of the N- and C-domains**

For the purification of N-domain, the ammonium sulfate (65% of saturation) precipitate of the periplasmic fraction was dialyzed against 10 mM sodium phosphate buffer (pH 7.2). The dialyzed material was mixed with 20 ml of hydroxylapatite (Biogel HT, BioRad) and rocked at 4°C for 2 hours. The suspension was then centrifuged (1,500 x g; 10 minutes). The supernatant represented the unbound fraction. The beads were washed four times with 10



mM sodium phosphate (25 ml each), and the supernatant of each wash was combined with the unbound fraction. Material bound to hydroxylapatite was eluted by incubating the beads with 0.4 M sodium phosphate (pH 7.2). The unbound and bound fractions were concentrated by pressure dialysis in Amicon filters and analyzed by SDS-PAGE. The unbound fraction (containing the N-domain) was dialyzed against 10 mM Tris (pH 8.0) and then fractionated over a column (1.8 x 14 cm) of DEAE-cellulose (DE-52, Whatman). A linear gradient (0 - 0.2 M KCl in 0.01 M Tris, pH 8.0) in a total volume of 150 ml was used to develop the column and purified N-domain was found in fractions eluted by 0.05 - 0.1 M KCl. For sequence analysis, the N-domain containing fractions were pooled and subjected to high pressure liquid chromatography over a Poly LC (Columbus, MD) hydroxyethyl A column (4.6 x 250 mm) in 50 mM formic acid. The flow rate was 250  $\mu$ l/minute and the absorbance of the effluent was monitored at 214 nm.

For the preparation of C-domain, rCBP35 was dialyzed against 75 mM Tris-HCl pH 7.0, 75 mM NaCl, 10 mM CaCl<sub>2</sub> overnight at 4°C, followed by digestion with an equal weight of collagenase D (BMB) at 37°C for 1.5 hours. The digestion mixture was then dialyzed overnight against loading buffer, and then fractionated over an asialofetuin—Affi-gel 15 column as described above for rCBP35.

### **Antibodies**

CBP35 was purified from mouse lung (2) and was used to immunize a female New Zealand White rabbit via the popliteal lymph node (1). The

specificity of this antiserum, which serves as a reference in the present study, has been characterized in terms of immunoprecipitation of CBP35 from a nondenatured protein mixture and in terms of immunoblotting the polypeptide after SDS-PAGE (12). This antiserum is hereafter referred to as #24. Two new antisera, designated #32 and #33, have been generated using rCBP35 as the immunogen. Flemish Giant rabbits were immunized subcutaneously with 50  $\mu$ g of rCBP35 in complete Freund's adjuvant; they were boosted two weeks later with 50  $\mu$ g of rCBP35 in incomplete Freund's adjuvant. Antisera were collected 7-10 days after each subsequent booster injection.

Rat monoclonal antibody directed against the Mac-2 antigen (22) was isolated from the hybridoma line M3/38.1.2.8.HL.2 (American Type Culture Collection T1B166). The hybridoma line was cultured in serum-free medium (RPMI 1640 containing Nutridoma SP (BMB)). After centrifugation to pellet the cells, supernatants from the cultures were pooled, subjected to ammonium sulfate precipitation (45% of saturation), dialyzed against phosphate-buffered saline exhaustively, and stored in aliquots (25  $\mu$ g/ml).

### **Analytical Techniques**

Protein concentrations were determined by the Bradford assay (23). Amino acid sequence analysis was carried out in the Michigan State University Macromolecular Structure Facility. Protein samples were electrophoresed on 12.5% or 15% SDS-PAGE as described by Laemmli (24). The gels were stained for protein by silver (25, 26) or Coomassie blue, or transferred onto Immobilon-P

membrane (Millipore), using semidry blotting and the buffer system of Bjerrum and Schafer-Nielsen (27).

The immunoblots were blocked for several hours in 2% gelatin (BioRad) Tris-buffered saline (20 mM Tris-HCl, pH 7.5, 0.5 M NaCl), and subsequently incubated in polyclonal antiserum raised against CBP35 or monoclonal anti-Mac-2 (in Tris-buffered saline containing 1% gelatin) for 2 hours at room temperature. The blots were washed in Tris-buffered saline containing 0.05% Tween-20 extensively, prior to the addition of secondary antibody conjugated to either horseradish peroxidase (BioRad) or alkaline phosphatase (BMB). The blots were colorimetrically developed using either 3,3'-diaminobenzidine/H<sub>2</sub>O<sub>2</sub> for horseradish peroxidase or nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate for alkaline phosphatase.

### **Differential Scanning Calorimetry (DSC)**

Samples were dialyzed at 4°C against either 10 mM Tris (pH 7), 1 mM EDTA, 10 mM 2-ME (pH 7 buffer) or the same buffer adjusted to pH 10 (pH 10 buffer). Both samples and buffers were deaerated with stirring for 10 minutes under vacuum. Aliquots were removed for the determination of protein concentration.

The DSC experiments were carried out in a Microcal MC-2 scanning calorimeter (Microcal Inc., Amherst, MA), with a Model 150 B (Keithley Instruments, Cleveland, OH) microvolt ammeter added to increase sensitivity. The calorimeter was interfaced with an IBM XT. Once the temperatures of the

sample and reference buffer were equilibrated, a computer-controlled scan was initiated, at a rate of 90°/hour. The data were analyzed using the "F1" option of the DA-2 program, included in the software provided by Microcal. Deconvolution analysis was performed using the "F1" option of the deconvolution subroutine.

To test the effect of prior heat denaturation of the N-domain on collagenase digestion of rCBP35, parallel samples (20  $\mu$ g rCBP35 in 400  $\mu$ l of 75 mM Tris, 75 mM NaCl, 10 mM CaCl<sub>2</sub>, pH 10) were incubated at 30°C and 40°C for 15 minutes. After reequilibration at 30°C, 5  $\mu$ g of collagenase D was added in 20  $\mu$ l of the same Tris buffer. Aliquots were taken from the two digestion mixtures every 3 minutes, quenched with SDS-PAGE buffer and analyzed by gels.

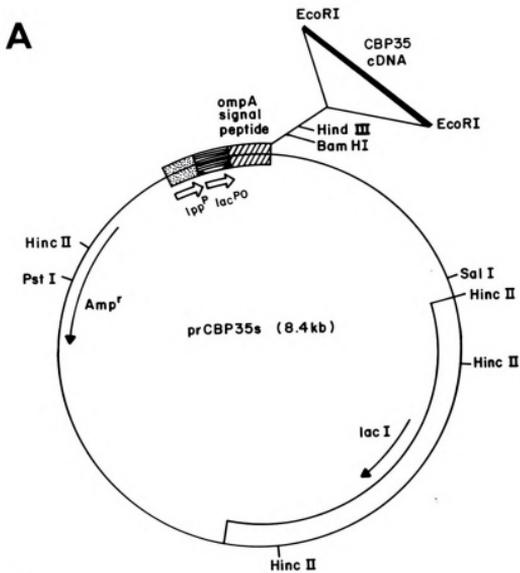
## RESULTS

### Expression and Purification of Recombinant CBP35

The cDNA clone for CBP35 was inserted into the pIN III *ompA2* secretion vector in *E. coli* for the production of rCBP35 (Fig. 1). This vector uses the signal sequence of *ompA*, an *E. coli* outer membrane protein, to direct expressed proteins into the periplasmic space (18). Using the *lpp-lac* fusion promoter, the protein can be induced by the addition of IPTG (Fig. 1). The soluble periplasmic fraction of *E. coli* JA221 cells transformed with prCBP35s was collected 16-24 hours after induction, subjected to ammonium sulfate precipitation and then affinity chromatography on a column of asialofetuin—Affi-gel. The material bound to the column and eluted upon lactose addition yielded a single band upon SDS-PAGE, as revealed by both silver-staining (Fig. 2A, lane 4) and immunoblotting (Fig. 2B, lane 4). The position of migration of this polypeptide corresponded to that observed with mouse CBP35 ( $M_r \sim 35,000$ ). This material is designated as rCBP35.

The initiator methionine of the protein product is provided by the ATG of the signal sequence. Due to the location of the cleavage site of the signal sequence, three extra amino acids are added on to the amino-terminus of the mature protein. Thus, the amino terminal end of rCBP35 has the following sequence: Ala-Glu-Phe-Arg-Asp-Ser-, with the fourth residue (Arg) being the first amino acid of the cDNA clone for CBP35 (15). The yield of the isolated protein is highly dependent on the type of culture broth used, the concentration of IPTG,

**Figure 1. Schematic diagram of the construction of the recombinant expression vector prCBP35s and site-directed mutagenesis to obtain the amino terminal and carboxyl terminal domains.** A) The cDNA clone for CBP35 was cloned into the EcoRI site of the pIN III *ompA2* vector. The cDNA is downstream of the IPTG inducible promoter and the *ompA* signal sequence. B) Using oligonucleotide-directed mutagenesis, glycines 138-139 were changed to translation stop codons, and a HindIII restriction site was introduced at amino acid positions 141-142. The mutant cDNA was digested with HindIII and the fragment corresponding to the NH<sub>2</sub>-terminal portion of the CBP35 polypeptide was subcloned and inserted into the EcoRI-HindIII sites of pIN III *ompA2* vector as above to express the NH<sub>2</sub>-terminal domain.

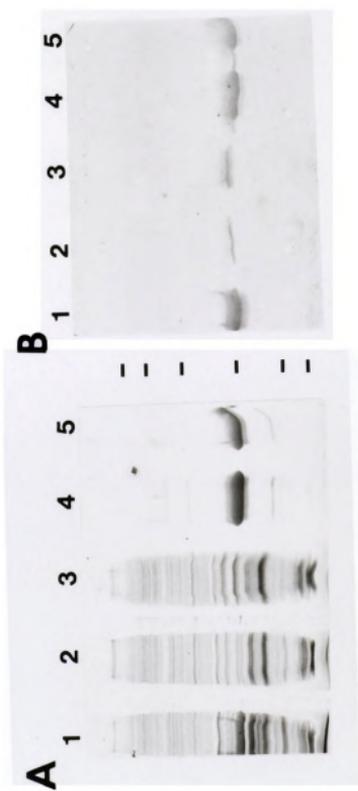


**B**

	137			140				145		
wild type	{	Pro	Gly	Gly	Val	Met	Pro	Arg	Met	Leu
		CCT	GGA	GGA	GTC	ATG	CCC	CGC	ATG	CTG
mutant	{	CCT	<u>IGA</u>	<u>IGA</u>	GTC	<u>AAG</u>	<u>CTT</u>	CGC	ATG	CTG
		Pro	<u>Stop</u>	<u>Stop</u>	Val	<u>Lys</u>	<u>Leu</u>	Arg	Met	Leu



**Figure 2. Purification of rCBP35 from *E. coli* transformed with the expression vector prCBP35s.** Approximately 10  $\mu$ g of protein was electrophoresed in each lane of a reducing SDS-PAGE (12.5% acrylamide) and analyzed by A) silver-staining (reference 25) or by B) immunoblotting using rabbit anti-CBP35 antiserum (#33). The immunoblot was revealed by using goat anti-rabbit IgG-horseradish peroxidase conjugate and the substrate 3,3'-diaminobenzidine. Lanes: 1) total contents of periplasmic supernatant; 2) material not bound to the asialofetuin—Affi-gel column; 3) material washed off the column; 4) material eluted by lactose; 5) rCBP35 isolated in the absence of 2-ME. The molecular weight standards indicated by bars are: 14.4 kDa, 21.5 kDa, 31 kDa, 42.7 kDa, 66.2 kDa, and 97.4 kDa.





**TABLE I. Parameters Affecting Yield of rCBP35**

<b>Media</b>	<b>[IPTG]</b>	<b>Temp</b>	<b>Yield<sup>a,b</sup></b>
LB	2-4 mM	37°C	~ 10-20 $\mu$ g
M9 + 20% CA.	2-4 mM	37°C	1-5 $\mu$ g
LB	2-4 mM	30°C	2 mg
LB	0.1 mM	22-25°C	20 mg
TB	0.05 mM	22-25°C	50 mg
LB	0.05 mM	22-25°C	20 mg

<sup>a</sup> All preparations were carried out with freshly transformed cells.

<sup>b</sup> Amount is based on yield from 1 liter cultures.

and the temperature at which the culture is grown (Table I) (28,29). Under optimal growth conditions (Terrific Broth medium, 0.05 mM IPTG, 22-25°C), approximately 50 mg of rCBP35 could be isolated per liter of culture. This represented ~ 35% of the total protein in the soluble periplasmic fraction.

It had been thought that the definition of CBP35 as an S-type lectin (16,17) meant that reducing agents were necessary for retaining saccharide binding activity. The initial isolation of CBP35 from mouse 3T3 fibroblasts in our laboratory, using extraction buffers that contained Triton X-100, reported the requirement of 2-ME to maintain carbohydrate-binding activity (1). However, recent data showing that CBP35 can be purified from 3T3 in the total absence of reducing agents indicate that such is not necessarily the case (30). In the present study, we found that approximately equal amounts of rCBP35 can be isolated from the periplasmic fraction in the presence and absence of 2-ME (Fig. 2A and 2B, lanes 4 and 5). Thus, the previously supposed requirement for reducing agents may be due to an artifact of the purification procedure, most probably involving the use of impure Triton-X-100 (1).

### **Generation of the NH<sub>2</sub>- and COOH-terminal Domains of rCBP35**

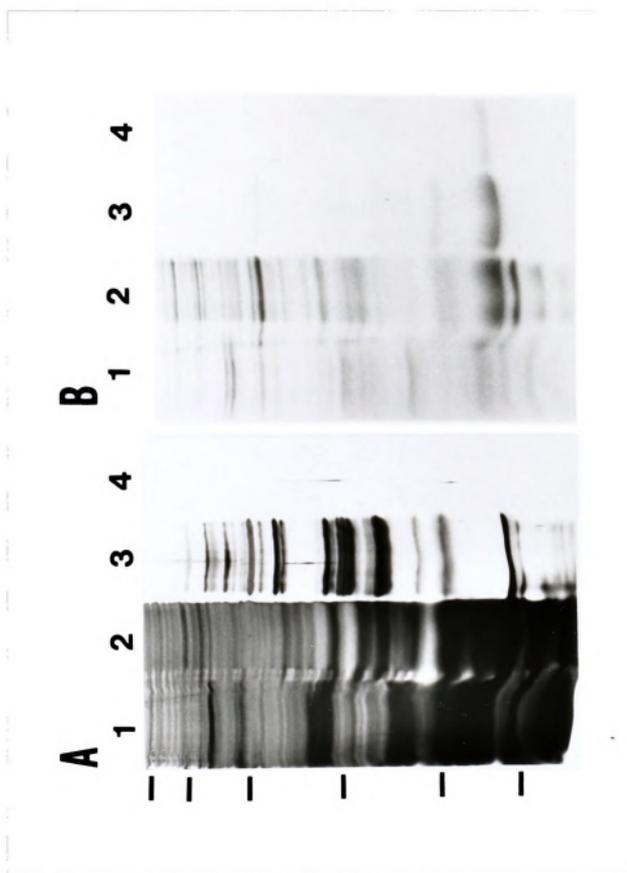
The strategy for obtaining the N-domain was to carry out site-directed mutagenesis on the CBP35 cDNA. In this scheme (Fig. 1B), two stop codons were introduced at Gly 138 and Gly 139. Thus, the translation open reading frame stops at Pro 137, giving an N-domain covering residues 1-137. In this step, a HindIII restriction site is simultaneously introduced at positions occupied by Met



141 and Pro 142 (Fig. 1B). Digestion of the mutant cDNA allowed for the subcloning of the fragment covering the 5'-end through the HindIII site and its subsequent insertion into the pIN III *ompA2* secretion vector. *E. coli* JA221 cells transformed with this plasmid expressed polypeptide(s) in the periplasmic fraction that were immunoblotted with antibodies against CBP35. The mobility of the predominant band was consistent with a molecular weight of  $\sim 15,000$ ; this would be expected if the N-domain spanned residues 1-137.

The N-domain was purified using two ion-exchange resins. During each of these steps, fractions containing the N-domain were identified by immunoblotting, while the purity of the fractions was assessed by silver and Coomassie blue staining. Both of these staining procedures were used because the N-domain polypeptide does not stain well with the silver reagent. Thus, Coomassie blue staining was used to reveal the N-domain polypeptide while the more sensitive silver staining was used to monitor the presence of contaminants. The ammonium sulfate precipitate of the periplasmic fraction was first subjected to adsorption onto hydroxylapatite beads. This resulted in significant purification, inasmuch as a major portion of the periplasmic proteins bound to the hydroxylapatite, whereas the N-domain was found in the unbound fraction (Fig. 3A and 3B, lanes 1-3). The unbound fraction was chromatographed over a DEAE-cellulose column and eluted with a linear 0 - 0.2 M KCl gradient. The N-domain was found in fractions eluted by 0.05 - 0.1 M KCl. These fractions yielded one predominant band ( $M_r \sim 15,000$ ) on SDS-PAGE, as revealed by Coomassie blue staining (Fig. 3B, lane 4).

**Figure 3. Purification of the N-domain from *E. coli* transformed with expression vector containing a fragment of mutagenized cDNA.** Approximately 10  $\mu$ g of protein was electrophoresed in each lane of a reducing SDS-PAGE (12.5% acrylamide) and analyzed by A) silver staining (reference 26) or by B) Coomassie blue staining. Lanes: 1) total contents of periplasmic supernatant; 2) ammonium sulfate precipitate of periplasmic contents; 3) unbound fraction of the hydroxylapatite adsorption step; 4) material pooled from fractions of the DEAE-cellulose column (0.05 - 0.1 M KCl) that contained the N-domain. The molecular weight standards are identical to those in Figure 2 and are indicated by the bars.



This material was devoid of any contaminants, as monitored by the silver stain (Fig. 3A, lane 4). The identity and purity of the isolated N-domain was further checked by amino acid sequence analysis. Upon high pressure liquid chromatography on a hydroxyethyl A column, two peaks were observed when the effluent was monitored at 214 nm. The minor peak yielded neither a band on SDS-PAGE nor any identifiable amino acids through five cycles of the Edman reaction during sequence analysis. In contrast, the predominant peak exhibited a single band ( $M_r \sim 15,000$ ) on SDS-PAGE, which can be immunoblotted by anti-CBP35 (see below) and yielded the sequence Ala-Glu-Phe-Arg-Asp-Ser-Phe-Ser-Leu-Asn, exactly as expected for the  $\text{NH}_2$ -terminus of the polypeptide. This material is, therefore, designated as the N-domain. Upon affinity chromatography on an asialofetuin—Affi-gel column, all of the N-domain was found in the flow-through fraction, indicating that it does not exhibit any carbohydrate-binding activity. In addition, the isolated N-domain was sensitive to digestion by collagenase, similar to the behavior of this domain in the intact polypeptide (see below).

In generating the mutant cDNA described above for the N-domain, the HindIII restriction site was designed also with the production of the C-domain in mind (Fig. 1B). A double digestion with HindIII and EcoRI of the mutant DNA would result in a fragment containing the C-domain (residues 145-264), which can be directionally cloned into the pIN III *ompA1* vector. However, the recombinant clones that were generated failed to yield any proteins immunoreactive with any of

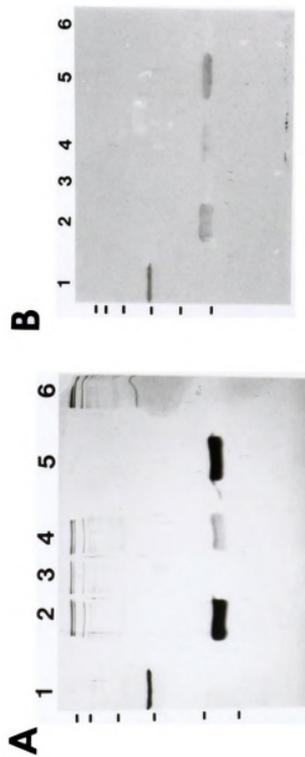


the anti-CBP35 antibodies. Thus, an alternative strategy was used to obtain the C-domain.

Raz and co-workers had reported that the NH<sub>2</sub>-terminal half of L-34 (which is identical to CBP35) bears striking resemblance to collagen and that treatment with collagenase reduced the molecular weight of the L-34 polypeptide (4). More recently, the same approach was used to prepare the C-domain of recombinant human IgE-binding lectin, the human homolog of rCBP35 (31). On this basis, rCBP35 was digested with collagenase D. The digestion mixture contained the major degradation product derived from rCBP35 (at M<sub>r</sub> ~ 16,000), revealed by silver staining and immunoblotting (Fig. 4A and 4B, lane 2). Collagenase D, in amounts corresponding to those included in the digestion mixture, yielded high molecular weight bands upon silver-staining (Fig. 4A, lane 6), but did not react with anti-CBP35 (Fig. 4B, lane 6). Affinity chromatography of this digestion mixture on asialofetuin—Affi-gel resulted in a bound fraction that was eluted upon lactose addition. The bound polypeptide exhibited a mobility corresponding to M<sub>r</sub> ~ 16,000 (Fig. 4A and 4B, lane 5). Amino acid sequence analysis yielded Ala-Gly-Pro-Tyr-Gly-Val, suggesting that the M<sub>r</sub> ~ 16,000 polypeptide represented the C-domain of CBP35, spanning residues 118-264. Thus, the C-domain of CBP35 accounts for the carbohydrate-binding activity of the polypeptide, consistent with the homology of this portion of the protein with the sequences of other galactose-specific lectins (3,15) and with the results obtained with the C-domain of recombinant human IgE-binding lectin (31).



**Figure 4. Digestion of rCBP35 with collagenase D and affinity purification of C-domain on Asialofetuin—Affi-gel.** rCBP35 (20 mg) was digested with collagenase D (20 mg) and fractionated over an affinity column containing asialofetuin—Affi-gel. The samples were analyzed by reducing SDS-PAGE (15% acrylamide) which was either A) silver-stained (reference 25) or B) immunoblotted using rabbit anti-CBP35 antiserum (#32). Approximately equal amounts of protein were loaded in each of the lanes. Lanes: 1) undigested substrate, rCBP35; 2) digestion mixture containing rCBP35 and collagenase D prior to affinity purification; 3) material not bound to asialofetuin—Affi-gel; 4) material washed off the column; 5) bound material eluted from the affinity column using lactose; 6) collagenase D. The immunoblot was developed using goat anti-rabbit IgG-horseradish peroxidase conjugate and the substrate 3,3',-diaminobenzidine. The molecular weight standards are identical to those in Figure 2 and are indicated by the bars.





Finally, the isolated C-domain was not sensitive to treatment with collagenase, also similar to the behavior of this domain in the intact polypeptide.

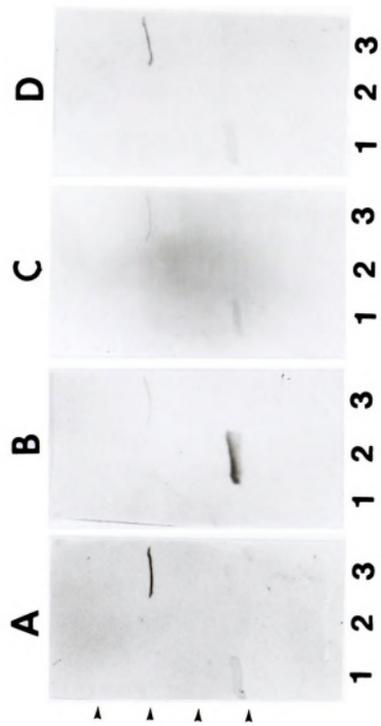
### **Development and Characterization of Antisera Against rCBP35**

Purified rCBP35 was used as an immunogen to generate polyclonal antiserum in two rabbits, designated as #32 and #33. Their antisera were compared against: (a) an antiserum derived from a rabbit (designated #24) immunized with CBP35 purified from mouse lung (2); and (b) the immunoglobulin fraction of a rat monoclonal antibody directed against mouse Mac-2 (22), whose cDNA sequence had indicated it to be identical to CBP35 (6). All four antibody preparations recognized CBP35 in Western blots of the immunogen, rCBP35 (Fig. 5A-D, lane 3), as well as the extracts of mouse 3T3 cells (data not shown). Preimmune serum from each of the rabbits, when used at the same dilutions, failed to yield any immunoreactive bands on control Western blots. The four antibody preparations also immunoblotted CBP35 in extracts of human HeLa cells (data not shown). In this case, the molecular weight of the immunoreactive band was slightly lower than that of mouse CBP35, consistent with the length of the polypeptides predicted from the nucleotide sequences of the mouse and human cDNAs (5,7,9,11).

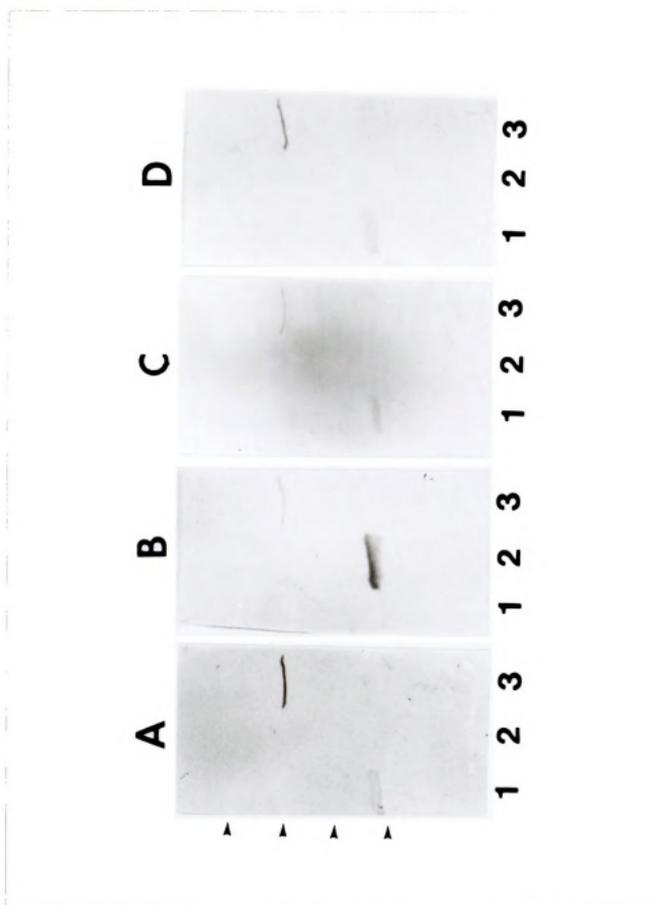
Immunoblotting of the N- and C-domains was carried out with the four preparations of anti-CBP35. Antisera #24 and #33 immunoblotted the N-domain (Fig. 5A and 5C, lane 1), indicating that the principal epitopes recognized by these antibodies were in this portion of the polypeptide. Neither antisera reacted with



**Figure 5. Characterization of anti-CBP35 antisera using immunoblotting.** Equal amounts of N-domain (lane 1), C-domain (lane 2) and rCBP35 (lane 3) were analyzed by reducing SDS-PAGE (15% acrylamide), and immunoblotting. The antibodies against CBP35 were: A) rabbit antiserum #24 (1:150 dilution); B) rabbit antiserum #32 (1:1000 dilution); C) rabbit antiserum #33 (1:1000 dilution); D) rat anti-Mac-2 (25 ng/ml). The polypeptide bands were either revealed by goat anti-rabbit IgG-alkaline phosphatase or goat anti-rat IgG-alkaline phosphatase followed by the substrates nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate. The molecular weight standards are identical to those in Figure 2 and are indicated by bars.



**Figure 5. Characterization of anti-CBP35 antisera using immunoblotting.** Equal amounts of N-domain (lane 1), C-domain (lane 2) and rCBP35 (lane 3) were analyzed by reducing SDS-PAGE (15% acrylamide), and immunoblotting. The antibodies against CBP35 were: A) rabbit antiserum #24 (1:150 dilution); B) rabbit antiserum #32 (1:1000 dilution); C) rabbit antiserum #33 (1:1000 dilution); D) rat anti-Mac-2 (25 ng/ml). The polypeptide bands were either revealed by goat anti-rabbit IgG-alkaline phosphatase or goat anti-rat IgG-alkaline phosphatase followed by the substrates nitroblue tetrazolium/5-bromo-4-chloro-3-indolyl phosphate. The molecular weight standards are identical to those in Figure 2 and are indicated by bars.





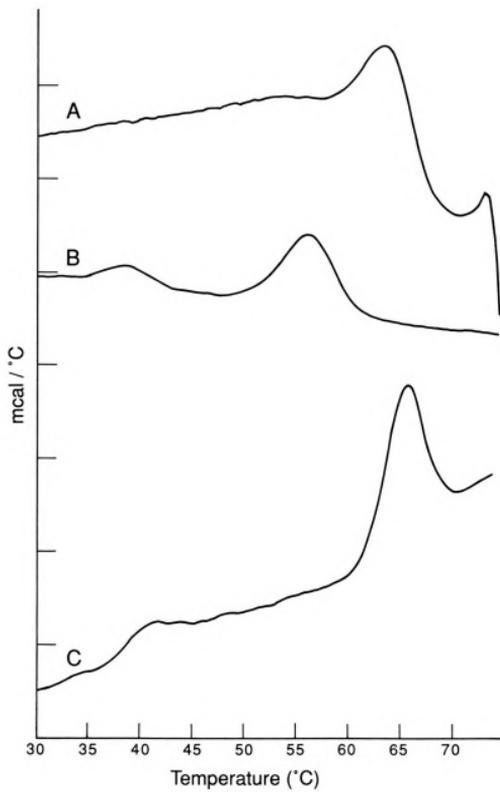
the C-domain (Fig. 5A and 5C, lane 2). Similarly, rat anti-Mac-2 was also N-domain specific (Fig. 5D, lanes 1 and 2). In this case, the epitope of the monoclonal antibody must be restricted to the NH<sub>2</sub>-terminal half of the CBP35 polypeptide. In contrast, antiserum #32 immunoblotted with the C-domain but not with the N-domain (Fig. 5B, lanes 1 and 2). At the level of immunoblotting, therefore, it appears that we have generated domain-specific antisera in our repertoire.

### **DSC Analysis**

In pH 7 Tris buffer, rCBP35 yielded the thermogram shown in Figure 6A. There was a transition at  $\sim 64^{\circ}\text{C}$ , followed by precipitation of the protein between 65 and 70°C. This made the definition of a baseline for deconvolution analysis difficult. Therefore, different conditions were sought in order to perform the DSC studies. We found that rCBP35 yielded a thermogram with stable baselines around transition regions when the Tris buffer was adjusted to pH 10 (Fig. 6B). At this pH, rCBP35 retained its saccharide-binding activity, on the basis of a comparison of the recovery of the protein, at pH 10 versus pH 7, in the fraction bound to asialofetuin—Affi-gel and eluted by lactose. The thermal denaturation of rCBP35 as seen in Figure 6B was irreversible; cooling and rescanning the sample over the same temperature range yielded no observable deviations from the baseline.

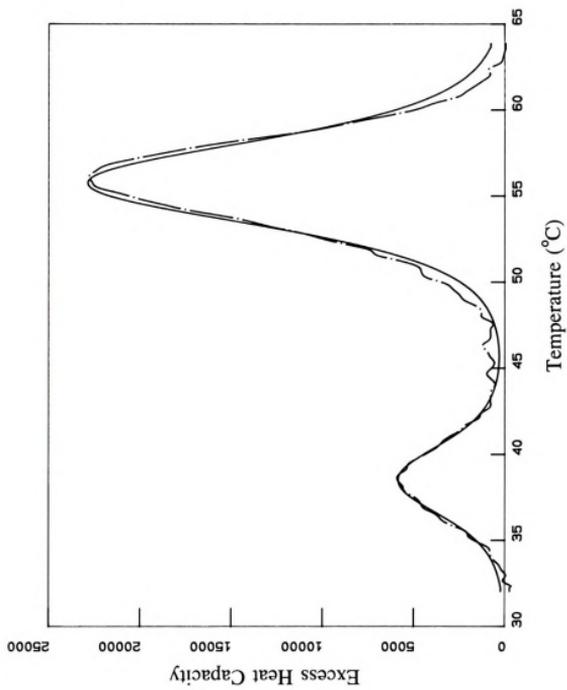


**Figure 6. Representative thermograms from DSC analysis of rCBP35.** (A) 0.77 mg/ml in pH 7 buffer; (B) 0.77 mg/ml in pH 10 buffer; (C) 0.75 mg/ml in pH 10 buffer containing 0.1 M lactose. Distance between tick marks on the y axis corresponds to 1 mcal/°C.

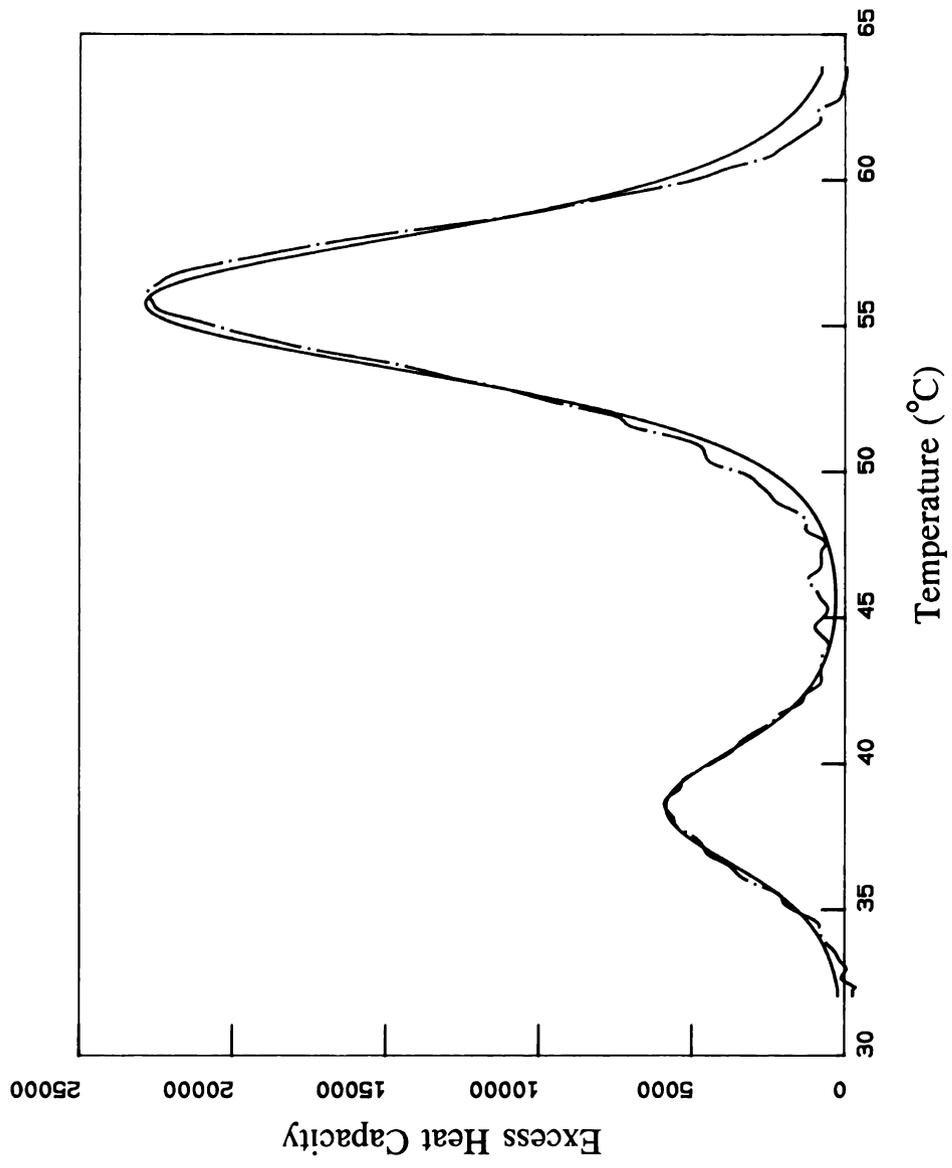




**Figure 7. Deconvolution of the thermogram of rCBP35 in pH 10 buffer.** The scan shown in curve B of Figure 6 was subjected to deconvolution using the Microcal DA-2 program. The dashed curve represents the actual experimental data, while the solid curve represents the transitions obtained by deconvolution. Distance between tick marks on the y axis represents 5000 cal/mole/°C.



**Figure 7. Deconvolution of the thermogram of rCBP35 in pH 10 buffer.** The scan shown in curve B of Figure 6 was subjected to deconvolution using the Microcal DA-2 program. The dashed curve represents the actual experimental data, while the solid curve represents the transitions obtained by deconvolution. Distance between tick marks on the y axis represents 5000 cal/mole/°C.



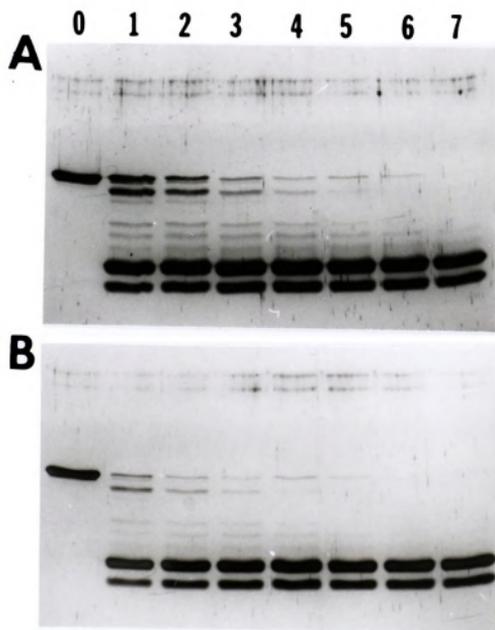


Deconvolution analysis was performed on the scan shown in Figure 6B. There was an excellent fit between the calculated and experimental results (Fig.7). This analysis showed the two transitions in the thermal denaturation of rCBP35 have  $T_m$  and enthalpy values of: (a) first transition: 39°C, 32 kcal/mole; and (b) second transition: 56°C, 149.5 kcal/mole. In light of the fact that hydrophathy analyses of the amino acid sequence had predicted two distinct structural domains, it was of interest to assign each of the observed transitions to a particular domain of the polypeptide. The most clear-cut hint on this issue is derived from the fact that in the presence of the saccharide ligand lactose, the second transition is shifted to  $\sim 65^\circ\text{C}$  (Fig. 6C). Sucrose, which does not bind to CBP35, failed to yield this effect. Amino acid sequence homologies predict that the C-domain is responsible for carbohydrate binding; we have indeed documented this by the purification of the C-domain using saccharide-specific affinity chromatography (Fig. 4). Therefore, the second transition ( $T_m \sim 56^\circ\text{C}$ ) most likely represents the thermal denaturation of the C-domain in the rCBP35 polypeptide.

The N-domain, containing the proline- and glycine-rich repeats (15), exhibits certain features similar to collagen (e.g., sensitivity to collagenase digestion (4,31)). Inasmuch as collagen has reported  $T_m$  values of 38-39°C (32,33), it is likely that the first transition ( $T_m \sim 39^\circ\text{C}$ ) represents the thermal denaturation of the N-domain. Consistent with this notion was the observation that collagenase digestion of the N-domain in rCBP35 proceeded faster in samples that had been previously denatured at 40°C. In this experiment, parallel samples of rCBP35 were incubated at 30°C and 40°C for 15 minutes. The samples were then



**Figure 8. Effect of heat denaturation of the N-domain on collagenase digestion of rCBP35.** Parallel samples (20  $\mu$ g rCBP35 in 400  $\mu$ l of 75 mM Tris, 75 mM NaCl, 10 mM  $\text{CaCl}_2$ , pH 10) were incubated: (A) 30°C for 15 minutes; or (B) 40°C for 15 minutes. Samples were then digested at 30°C with 5  $\mu$ g collagenase D. Aliquots were taken from the two digestion mixtures every 3 minutes, subjected to reducing SDS-PAGE (15% acrylamide) and silver staining (reference 26). For each sample, the lane marked 0 represents the undigested material and the lanes labeled 1-7 represent time points (3 minutes/time point).





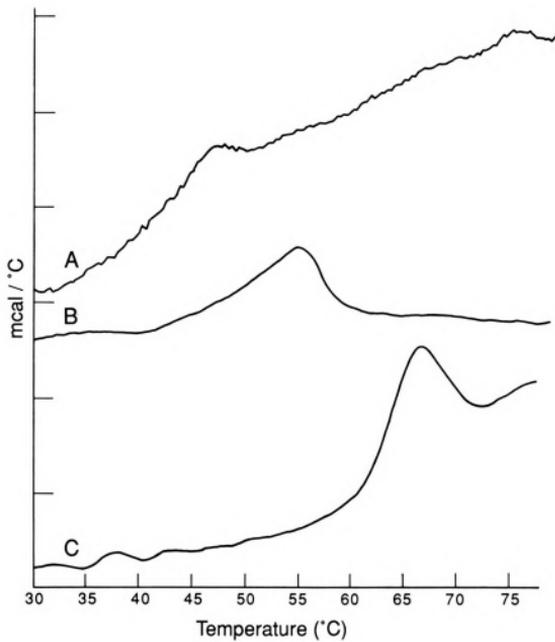
subjected to digestion by collagenase at 30°C. Aliquots of the digestion mixtures were taken every 3 minutes and subjected to SDS-PAGE. The rCBP35 sample that had been denatured at 40°C prior to collagenase addition quickly lost the  $M_r \sim 35,000$  band, as well as most of the intermediate bands (Fig. 8B). In contrast, the rCBP35 sample that had not previously been denatured showed a slower loss of the  $M_r \sim 35,000$  band and many of the intermediate bands (Fig. 8A).

The availability of purified N- and C-domains provided the opportunity to test directly these assignments of  $T_m$  values to the individual domains. DSC analysis on the isolated N-domain yielded a transition at  $\sim 48^\circ\text{C}$  (Fig. 9A), somewhat higher in temperature than the  $T_m$  of the first transition in the intact polypeptide. The features of this transition in the thermogram suggest that it corresponds to the low enthalpy first  $T_m$  of rCBP35. Moreover, this transition was not affected by the addition of lactose. On the other hand, DSC analysis of the C-domain yielded a single transition ( $\sim 55^\circ\text{C}$ ) (Fig. 9B). Lactose shifted this transition to  $\sim 66^\circ\text{C}$  (Fig. 9C), whereas sucrose did not. These results provide strong evidence that the first and second transitions in the thermogram of rCBP35 represent, respectively, the denaturation of the N- and C-domains of the polypeptide.

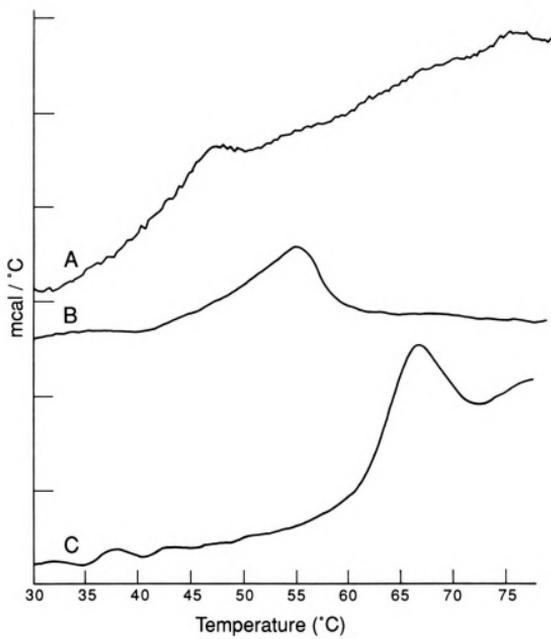




**Figure 9. Representative thermograms from DSC analysis of N- and C-domains.**  
(A) N-domain (2.5 mg/ml) pH 10 buffer; (B) C-domain (0.77 mg/ml) pH 10 buffer; (C) C-domain (0.77 mg/ml) pH 10 buffer containing 0.1 M lactose.  
Distance between tick marks on the y axis represents /mcal/°C.



**Figure 9. Representative thermograms from DSC analysis of N- and C-domains.**  
(A) N-domain (2.5 mg/ml) pH 10 buffer; (B) C-domain (0.77 mg/ml) pH 10 buffer; (C) C-domain (0.77 mg/ml) pH 10 buffer containing 0.1 M lactose.  
Distance between tick marks on the y axis represents  $\text{J/mcal}/^\circ\text{C}$ .





## DISCUSSION

An *E. coli* expression system, which secreted the expressed protein into the periplasmic space, was optimized for the production of rCBP35. Consistent with similar findings by other groups, our studies suggest that the secretion vector system may be susceptible to deleterious effects of "normal" growth conditions such as high concentrations of IPTG and 37°C growth temperature. In fact, low concentrations of IPTG (0.05 mM) and low growth temperature (22-25°C) resulted in the highest yield of rCBP35 (~ 50 mg per liter of bacterial culture). Thus, the yield of our expression system is about 25 times higher than that reported for the expression system of the human homolog of CBP35 (IgE-binding lectin), which was isolated at a level of ~ 2 mg per liter of *E. coli* culture (31).

Taking advantage of the availability of large amounts of protein, we have compared the properties of the whole polypeptide and the individual N- and C-domains. The following key conclusions can be derived: (a) the N-domain exhibits no apparent binding to asialofetuin—Affi-gel affinity resins; (b) the C-domain is sufficient for saccharide-binding and, therefore, contains the carbohydrate recognition domain; (c) rCBP35 is not dependent on thiol reagents for retention of carbohydrate-binding activity; (d) DSC analysis of rCBP35 yielded two independent transitions, with  $T_m$  values of 39°C and 56°C representing, respectively, the thermal denaturation of the N- and C-domains; and (e) binding of lactose to rCBP35 and to the C-domain results in a stabilization of the folded polypeptide to thermal denaturation ( $T_m$  rises from 56°C to 65°C).

The most striking of these findings is that the rCBP35 polypeptide appears to be folded into two distinct domains. Hydropathy analysis of the amino acid sequence had clearly delineated the polypeptide into two parts (3). The carboxyl terminal portion (residues 127-264) contains both hydrophilic and hydrophobic regions, typical of globular proteins. In contrast, the amino terminal portion (residues 1-126) exhibits neither a highly hydrophilic nor a hydrophobic nature.

The N-domain is characterized by a highly conserved repetitive sequence (4-11,15) rich in proline and glycine residues. Although it does not have the exact (Gly-X-Y)<sub>n</sub> repeat seen in collagen, its folding might be expected to be stabilized by many reinforcing bonds, each of which is relatively weak. Consistent with this prediction, the  $T_m$  value for thermal denaturation is rather low ( $\sim 39^\circ\text{C}$ ), similar to the reported  $T_m$  values of collagen (32,33). Thus, in addition to sensitivity to digestion by collagenase, the N-domain shares another characteristic found in collagen. The fact that the N-domain did exhibit a thermal transition suggest that this portion of the polypeptide is not totally devoid of structure. This notion is also supported by the observation that prior denaturation of the N-domain structure by heating to  $40^\circ\text{C}$  resulted in a more rapid digestion of that portion of the polypeptide by collagenase.

While both the isolated N-domain and the intact polypeptide exhibited a low enthalpy transition, the  $T_m$  values of this transition were significantly different. Surprisingly, the transition temperature for the isolated N-domain ( $\sim 48^\circ\text{C}$ ) was higher than the corresponding value in the intact molecule ( $\sim 39^\circ\text{C}$ ). Although transition temperature changes have been reported when  $T_m$  values are compared



for intact proteins and isolated domains/fragments, the usual case is that the latter yield a lower transition temperature. For example, the NADP<sup>+</sup>-binding domain of C<sub>1</sub> - tetrahydrofolate synthase exhibited a transition at 49°C as an isolated fragment, whereas the corresponding value was 60°C in the intact enzyme (34). It would be of interest, therefore, to delineate the physical basis for the stabilization of the N-domain, when it is freed of the C-domain. One possibility is interactions between two molecules of N-domain, leading to transient or weakly-bound dimers. This hypothesis would be consistent with the results of Hsu *et al.* (31), implicating the N-domain in self-association of the human IgE-binding lectin.

The interactions responsible for the folding of the C-domain must be much more extensive, resulting in higher  $T_m$  and enthalpy values for its thermal denaturation. These values are comparable to the corresponding data reported for several globular proteins (35-37). It should be noted that, like these other systems, the present study analyzed the data in terms of equilibrium thermodynamics, even though the thermal denaturation of rCBP35 under the described conditions was irreversible — rescanning of once heated and cooled samples yielded no discernible transition from the baseline. There are, however, both theoretical and empirical bases for analyzing the calorimetric data using equilibrium thermodynamics, despite the apparent irreversibility. These justifications have been discussed by other investigators (35,36).

Finally, the binding of lactose by rCBP35 elevated the temperature of thermal denaturation (second transition) by  $\sim 10^\circ\text{C}$ . Similar results were also observed for the C-domain. This effect was specific inasmuch as the disaccharide

sucrose, which does not bind to CBP35, failed to yield the same effect. These results suggest that ligand binding by CBP35 is accompanied by a conformational change that significantly stabilizes the polypeptide against thermal denaturation.

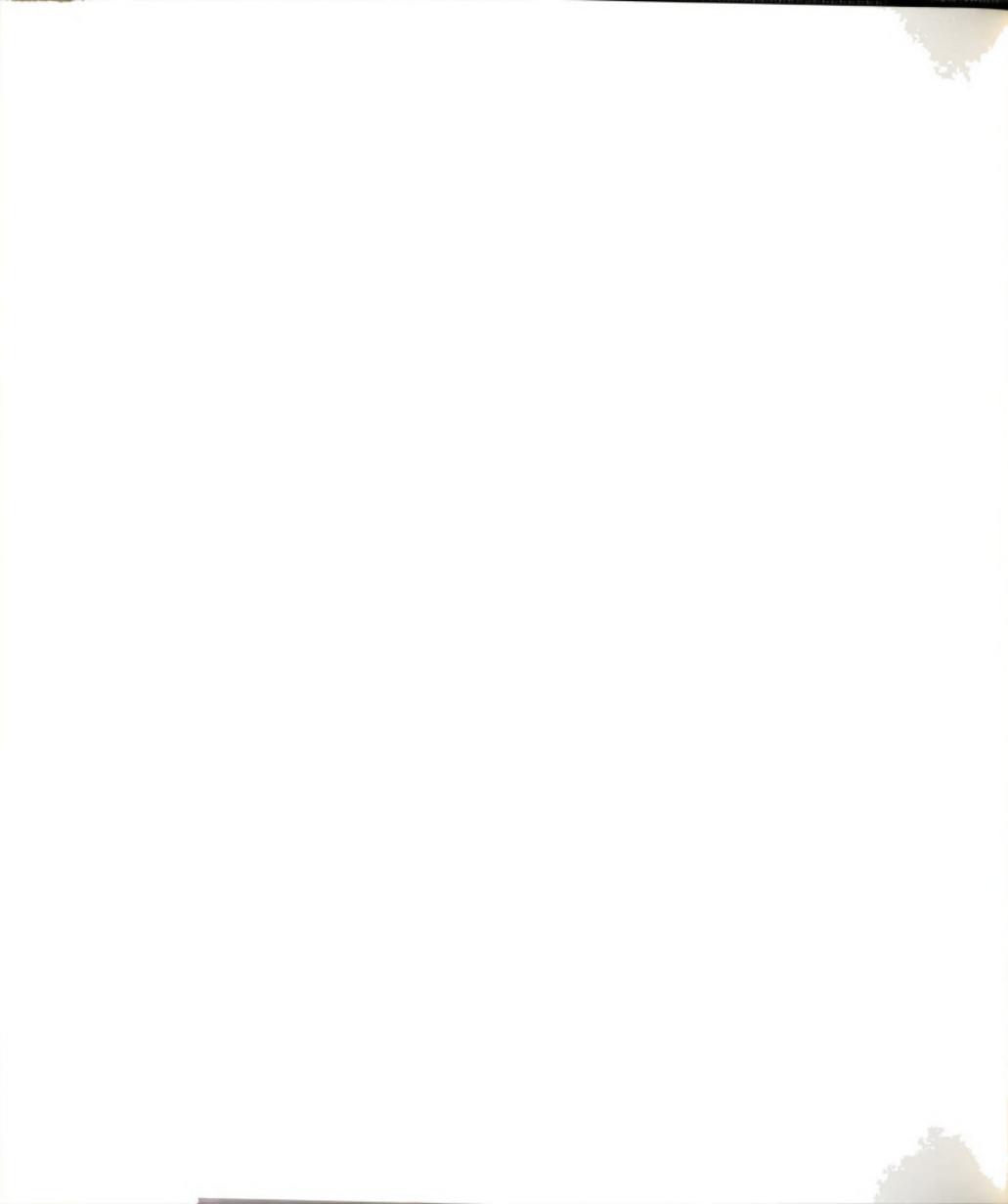
### ACKNOWLEDGMENTS

We thank Dr. Masayori Inouye for the gift of the pIN III *ompA1* and *A2* vectors and Dr. Kurt Drickamer for the gift of the JA221 strain of *E. coli*. We also thank Dr. John Wilson for his help with the differential scanning calorimetry analysis, Mrs. Patricia Voss for the preparation of the anti-Mac-2 monoclonal antibody, and Mrs. Linda Lang for the preparation of the manuscript.



## REFERENCES

1. Roff CF & Wang JL (1983) *J Biol Chem* **208**: 10657-10663.
2. Crittenden SL, Roff CF & Wang JL (1984) *Mol Cell Biol* **4**: 1252-1259.
3. Wang JL, Laing JG & Anderson RL (1991) *Glycobiology* **1**: 243-252.
4. Raz A, Pazerini G & Carmi P (1989) *Cancer Res* **49**: 3489-3493.
5. Raz A, Carmi P, Raz T, Hogan V, Mohamed A & Wolman SR (1991) *Cancer Res* **51**: 2173-2178.
6. Cherayil BJ, Weiner SJ & Pillai S (1989) *J Exp Med* **170**: 1959-1972.
7. Cherayil BJ, Chaitovitz S, Wang C & Pillai S (1990) *Proc Natl Acad Sci USA* **87**: 7324-7328.
8. Laing JG, Robertson MW, Gritzmacher CA, Wang JL & Liu F-T (1989) *J Biol Chem* **264**: 1907-1910.
9. Robertson MW, Albrandt K, Keller D & Liu F-T (1990) *Biochemistry* **29**: 8093-8100.
10. Leffler H, Masiarz FR & Barondes SH (1989) *Biochemistry* **28**: 9222-9229.
11. Oda Y, Leffler H, Sakakura Y, Kasai K & Barondes SH (1991) *Gene* **99**: 279-283.
12. Moutsatsos IK, Davis JM & Wang JL (1986) *J Cell Biol* **102**: 477-483.
13. Moutsatsos IK, Wade M, Schindler M & Wang JL (1987) *Proc Natl Acad Sci USA* **84**: 6452-6456.
14. Agrwal N, Wang JL & Voss PG (1989) *J Biol Chem* **264**: 17236-17242.
15. Jia S & Wang JL (1988) *J Biol Chem* **263**: 6009-6011.
16. Barondes SH (1988) *Trends Biochem Sci* **13**: 480-482.
17. Drickamer K (1988) *J Biol Chem* **263**: 9557-9560.
18. Ghrayeb J, Kimura H, Takahara M, Hsiung H, Masui Y & Inouye M (1984) *EMBO J* **3**: 2437-2442.



19. Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Vieira J & Messing J (1987) *Methods Enzymol* **153**: 3-11.
21. Kunkel TA, Roberts JD & Zakour RA (1987) *Methods Enzymol* **154**: 367-382.
22. Ho MK & Springer TA (1982) *J Immunol* **128**: 1221-1228.
23. Bradford MM (1976) *Anal Biochem* **72**: 248-254.
24. Laemmli UK (1970) *Nature* **227**: 680-685.
25. Wray W, Boulikas T, Wray V & Hancock R (1981) *Anal Biochem* **48**: 617-620.
26. Blum H, Beir H & Gross HJ (1987) *Electrophoresis* **8**: 93-99.
27. Bjerrum OJ & Schafer-Nielsen C (1986) *Analytical Electrophoresis* Dunn MJ, Ed, Verlag Chemie Weinheim.
28. Schein CH & Noteborn MHM (1988) *Biotechnology* **6**: 291-294.
29. Takagi H, Morinaga Y, Tsuchiya M, Ikemura H & Inouye M (1988) *Biotechnology* **6**: 948-950.
30. Frigeri LG, Robertson MW & Liu F-T (1990) *J Biol Chem* **265**: 20763-20769.
31. Hsu DK, Zuberi RI & Liu F-T (1992) *J Biol Chem* **267**: 14167-14174.
32. Miller EJ (1971) *Biochemistry* **10**: 1652-1659.
33. Berg RA & Prockop DJ (1973) *Biochem Biophys Res Comm* **52**: 115-120.
34. Villar E, Schuster B, Peterson D & Schirch V (1985) *J Biol Chem* **260**: 2245-2252.
35. Manly SP, Matthews KS, & Sturtevant JM (1985) *Biochemistry* **24**: 3842-3846.
36. Edge V, Allewell NM & Sturtevant JM (1985) *Biochemistry* **24**: 5899-5906.
37. White TK, Kim JY & Wilson JE (1990) *Arch Biochem Biophys* **276**: 510-517.

## CHAPTER IV

### CARBOHYDRATE BINDING PROTEIN 35:

#### **Preliminary Studies on the Transport of the Recombinant Polypeptide into the Nucleus\***

Neera Agrwal,<sup>†</sup> Richard S. Wozniak,<sup>#</sup> and John L. Wang<sup>†</sup>

<sup>†</sup>Department of Biochemistry

Michigan State University

East Lansing, MI 48824

and

<sup>#</sup>Laboratory of Cell Biology

The Rockefeller University

New York, NY 10021



**FOOTNOTES**

\*This work was supported by grants GM-38740 and GM-27203 from the National Institutes of Health and by a grant from the Howard Hughes Medical Institute to Dr. Günter Blobel.

<sup>1</sup>The abbreviations used are: CBP35, Carbohydrate Binding Protein 35; rCBP35, recombinant CBP35; Rh, rhodamine; NLS, nuclear localization signal of the SV40 large T antigen; FBS, fetal bovine serum; HSA, human serum albumin; DME, Dulbecco modified Eagle's medium; rCBPND, NH<sub>2</sub>-terminal domain of rCBP35; rCBPCD, COOH-terminal domain of rCBP35; PBS, phosphate-buffered saline; WGA, wheat germ agglutinin; RNP, ribonucleoprotein complex.



**SUMMARY**

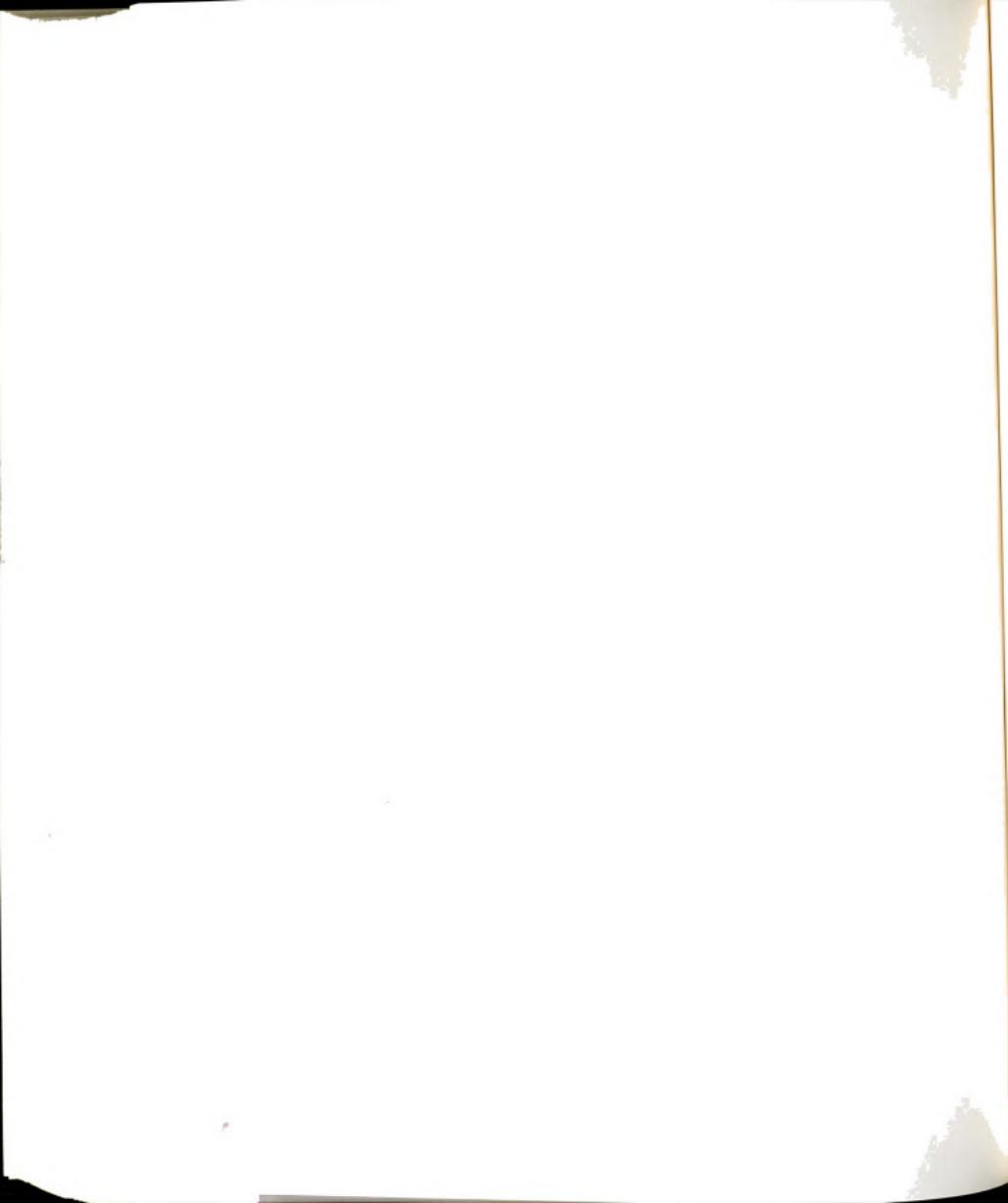
Recombinant Carbohydrate Binding Protein 35 (rCBP35) was conjugated directly with rhodamine and the resulting fluorescent derivative was microinjected into intact mouse 3T3 fibroblasts (*in vivo* assay) or was incubated with 3T3 cells permeabilized by digitonin (*in vitro* assay) to test for nuclear localization of the lectin. Microinjected rCBP35 accumulated in the nucleus of some recipient cells whereas fluorescently labeled human serum albumin failed to do so in any of the cells. In contrast to both of these cases, fluorescently-labeled human serum albumin conjugated with a peptide containing the nuclear localization signal of SV40 large T antigen was translocated into the nucleus in 100% of the microinjected cells. Thus, while rCBP35 introduced into 3T3 cells can be imported into the nucleus, its nuclear localization properties were different from those documented for the albumin substrate bearing the SV40 large T antigen nuclear localization signal. rCBP35 and human serum albumin containing nuclear localization signal were also different in terms of requirements and conditions of nuclear import in the *in vitro* assay. Little or no nuclear transport of rCBP35 could be demonstrated in this latter system. Thus, it appears that the S100 cytosolic fraction used in the *in vitro* assay could not supply the factor(s) necessary for the import of rCBP35 while the cytosol of microinjected cells could indeed fulfill the requirements of nuclear transport.



## INTRODUCTION

In previous studies (1,2), we had documented, using immunofluorescence and immunoblotting, the nuclear as well cytoplasmic localization of the galactose/lactose-specific lectin, designated Carbohydrate Binding Protein 35 (CBP35; Mr 35,000). In both the nucleus and the cytoplasm, CBP35 was found not as a free protein but in the form of a ribonucleoprotein complex (RNP) (3,4). The distribution of CBP35 between the two subcellular compartments was dependent on the proliferation state of the cell; the lectin was mostly cytoplasmic in quiescent cells whereas the proportion of nuclear CBP35 increased in proliferating cells. Moreover, when quiescent, serum-starved cultures of mouse 3T3 fibroblasts were stimulated by the addition of serum, there was an increase in the expression of CBP35, manifested by an accelerated rate of transcription of the gene, by elevated levels of its specific mRNA, and by higher levels of the polypeptide (2,5). Accompanying this increase in the expression of CBP35 was a dramatic translocation of the lectin from the cytoplasm into the nucleus, well before the onset of DNA synthesis in the stimulated cells. These results indicated that the subcellular localization of the lectin was strictly regulated and correlated with the proliferation state of the cell.

The availability of recombinant CBP35 (rCBP35), purified and in large amounts (6), allowed us to directly label the protein with the fluorescent rhodamine (Rh) group. Coupled with recent developments of nuclear import assays in both intact (7), as well as permeabilized cells (8), the availability of Rh-



rCBP35 in turn suggested the opportunity to study the translocation of the lectin between the intracellular compartments. On the basis of a comparison between the transport of CBP35 and human serum albumin (HSA) bearing the nuclear localization signal (NLS) of the SV40 large T antigen, it appears that the requirements and properties of nuclear import for the lectin are quite distinct from those synthetic substrates utilizing the NLS pathway. The results of our preliminary studies on this issue are documented in the present communication.



## MATERIALS AND METHODS

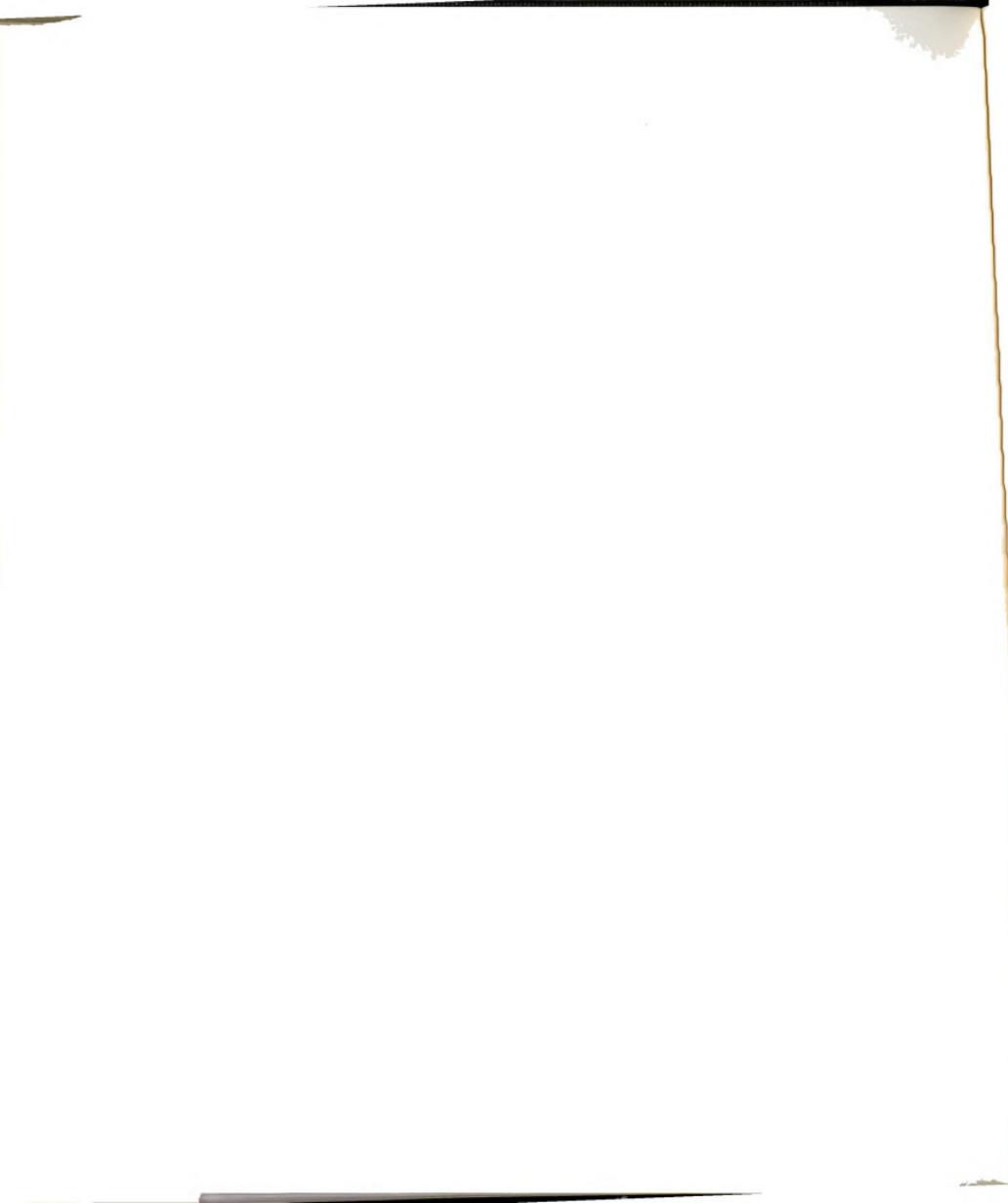
### Cell cultures

Balb/c 3T3 cells were used for microinjection experiments. The cells were cultured in a humidified incubator at 37°C in high glucose Dulbecco modified Eagle's medium (DME; Gibco) that was supplemented with 10% fetal bovine serum (FBS; Gibco), HEPES (final concentration of 20 mM), nonessential amino acids, and gentamycin (50 µg/ml). For experiments using asynchronous cultures, the 3T3 cells were trypsinized and replated onto glass coverslips at subconfluent densities 24-36 hours prior to microinjection. For synchronization of cells, the 3T3 cells were serum-starved for 48 hours by incubating in DME containing 0.2% FBS, and subsequently stimulated with DME containing 10% FBS.

The *in vitro* nuclear import assays used Swiss 3T3 cells, which were cultured in low glucose DME that was supplemented with 10% calf serum (Gibco) and penicillin-streptomycin (3). The cells were plated onto glass coverslips 16-24 hours prior to permeabilization and transport assay.

### Preparation of import substrates

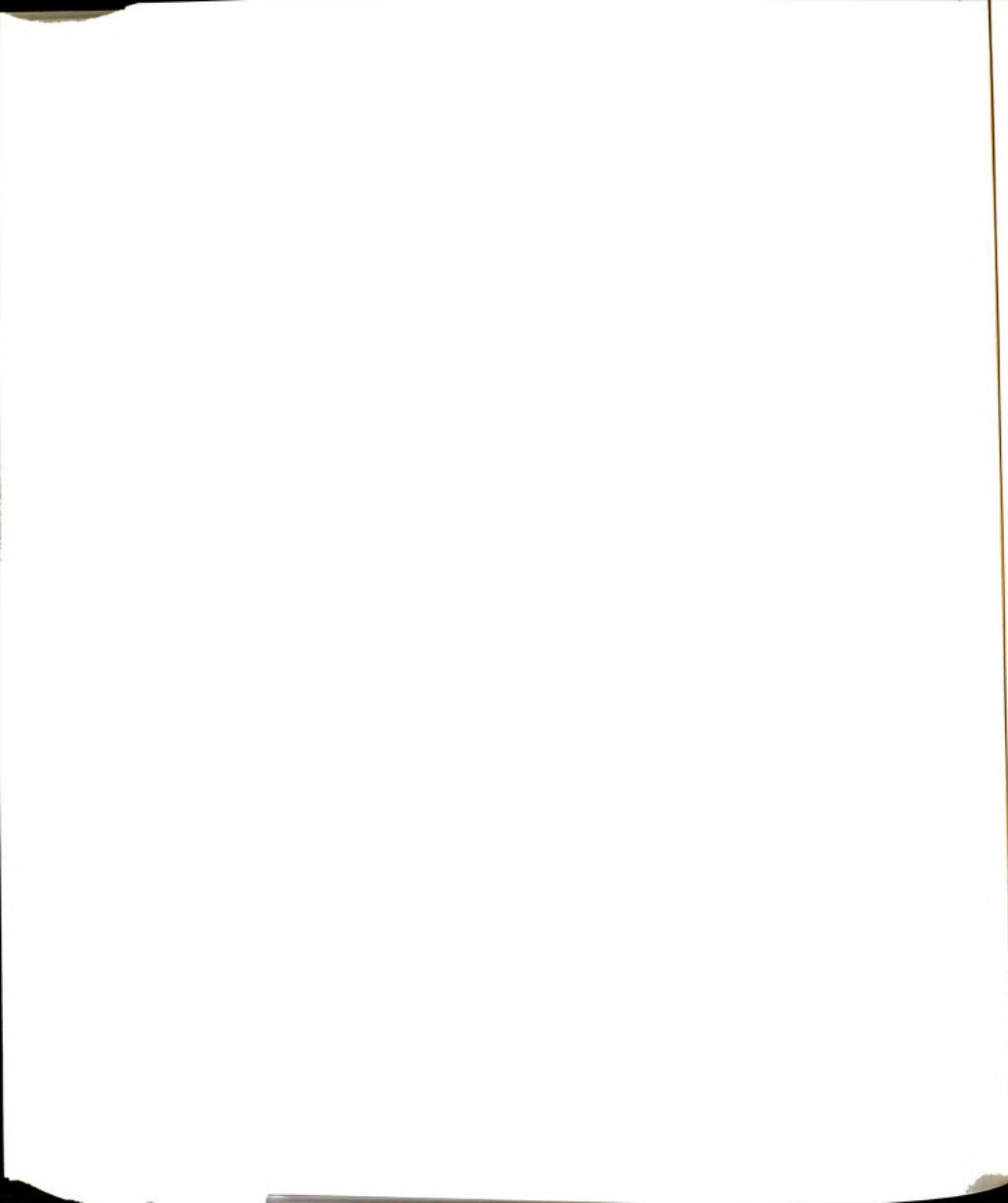
Essentially the same protocol was used to conjugate the fluorescent label Rh to the following proteins: HSA, rCBP35, the NH<sub>2</sub>-terminal domain of rCBP35 (rCBPND)(6), and the COOH-terminal domain of rCBP35 (rCBPCD) (6). The proteins (~ 1-2 mg/ml) were dialyzed overnight at 4°C against freshly prepared 0.1 M sodium bicarbonate. A 10 mg/ml solution of rhodamine succinimidyl ester



(Molecular Probes, # C1171) was prepared in dimethyl formamide and was added, with slow stirring, to the protein solution. The final molar ratio of dye to protein was  $\sim 2$ . The mixture was allowed to stir at room temperature for one hour, after which it was fractionated on a Sephadex G-25 column (18 x 0.8 cm) to separate unreacted dye from the protein. The conjugated protein was then dialyzed overnight at 4°C against transport buffer (20 mM HEPES-KOH, pH 7.3, 110 mM potassium acetate, 2 mM magnesium acetate, 1 mM EGTA, 2 mM dithiothreitol) and stored at -70°C. Trial preparations of Rh-rCBP35 by such a procedure showed that the saccharide-binding activity of the lectin was not affected by the derivatization. For short term use, an aliquot was thawed, microfuged to remove any particulate matter, and stored at 4°C.

Rhodamine-labeled HSA containing the NLS peptide (CYTPPKKKRKY) of SV40 large T antigen, hereafter designated Rh-HSA-NLS, was provided by Drs. M.S. Moore and G. Blobel of The Rockefeller University (9). This fluorescent import substrate contained 10 NLS peptides per HSA molecule.

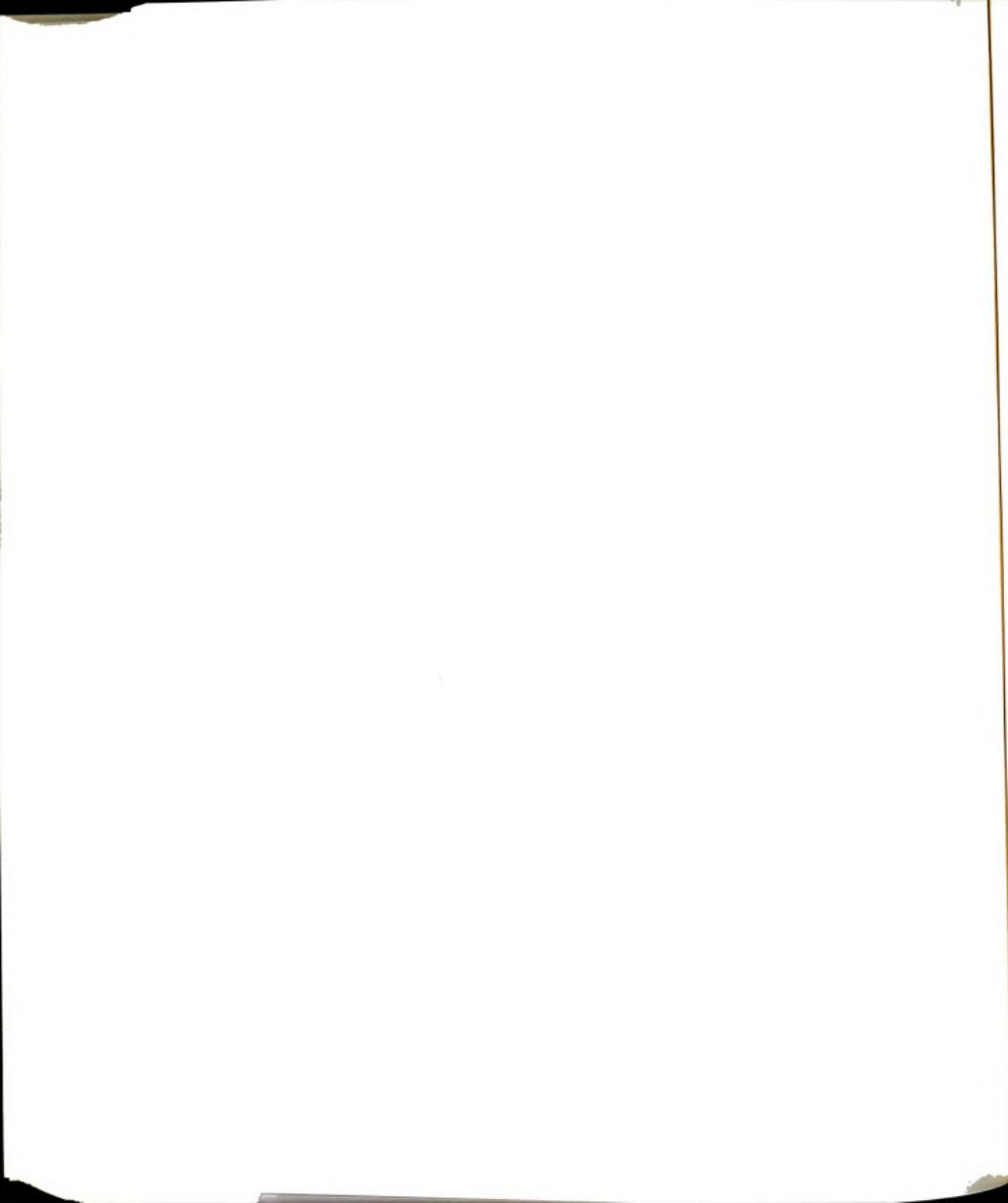
The following samples and concentrations were used for the microinjection assays: Rh-HSA-NLS at 2 mg/ml, Rh-HSA at 1.5 mg/ml, Rh-rCBP35 at 1.5 mg/ml, Rh-rCBPND at 0.3 mg/ml, and Rh-rCBPCD at 0.3 mg/ml. For the *in vitro* transport assays, Rh-HSA-NLS, Rh-HSA, and Rh-rCBP35 were used at 1:100 dilution of the stock solution used for microinjection assays.



### **Nuclear import assays**

The digitonin-permeabilized cell system as described by Moore and Blobel (9) was used for the *in vitro* transport assays. For permeabilization, the tissue culture plates containing the coverslips were placed on ice, the medium was aspirated and 1 ml of cold transport buffer containing leupeptin, aprotinin, and pepstatin (each at 1  $\mu\text{g/ml}$ ) and digitonin (35  $\mu\text{g/ml}$ ; Calbiochem) was added. The digitonin was diluted into the transport buffer from a 20 mg/ml stock solution in dimethyl sulfoxide immediately before use. The plates were incubated on ice for 5 minutes; then the plates were aspirated and cold transport buffer added. The coverslips containing the permeabilized cells were placed cell side down on top of a 20  $\mu\text{l}$  of transport assay mixture (see below) on parafilm. After a 30 minute incubation at room temperature, the reaction was terminated by the addition of cold transport buffer. The cells were then incubated with cold fixative, 3% paraformaldehyde in transport buffer, for 15 minutes on ice. The coverslips were then removed, blotted to remove excess moisture, and mounted in Slow-Fade (Molecular Probes). The edges of the coverslips were sealed with nail polish.

A typical transport assay mixture consisted of the following: 1  $\mu\text{l}$  of 20 mg/ml bovine serum albumin, 1  $\mu\text{l}$  of 20 mM ATP, 1  $\mu\text{l}$  of 100 mM phosphocreatine, 1  $\mu\text{l}$  of 400 U/ml creatine phosphokinase, 5  $\mu\text{l}$  of rabbit reticulocyte lysate (Gibco), 5  $\mu\text{l}$  of HeLa cell S100 cytosolic fraction (10 mg/ml protein), 1  $\mu\text{l}$  of the import substrate, and transport buffer to give a final volume of 20  $\mu\text{l}$ . For experiments involving inhibition by wheat germ agglutinin (WGA), coverslips were incubated in the transport reaction mixture containing WGA (final

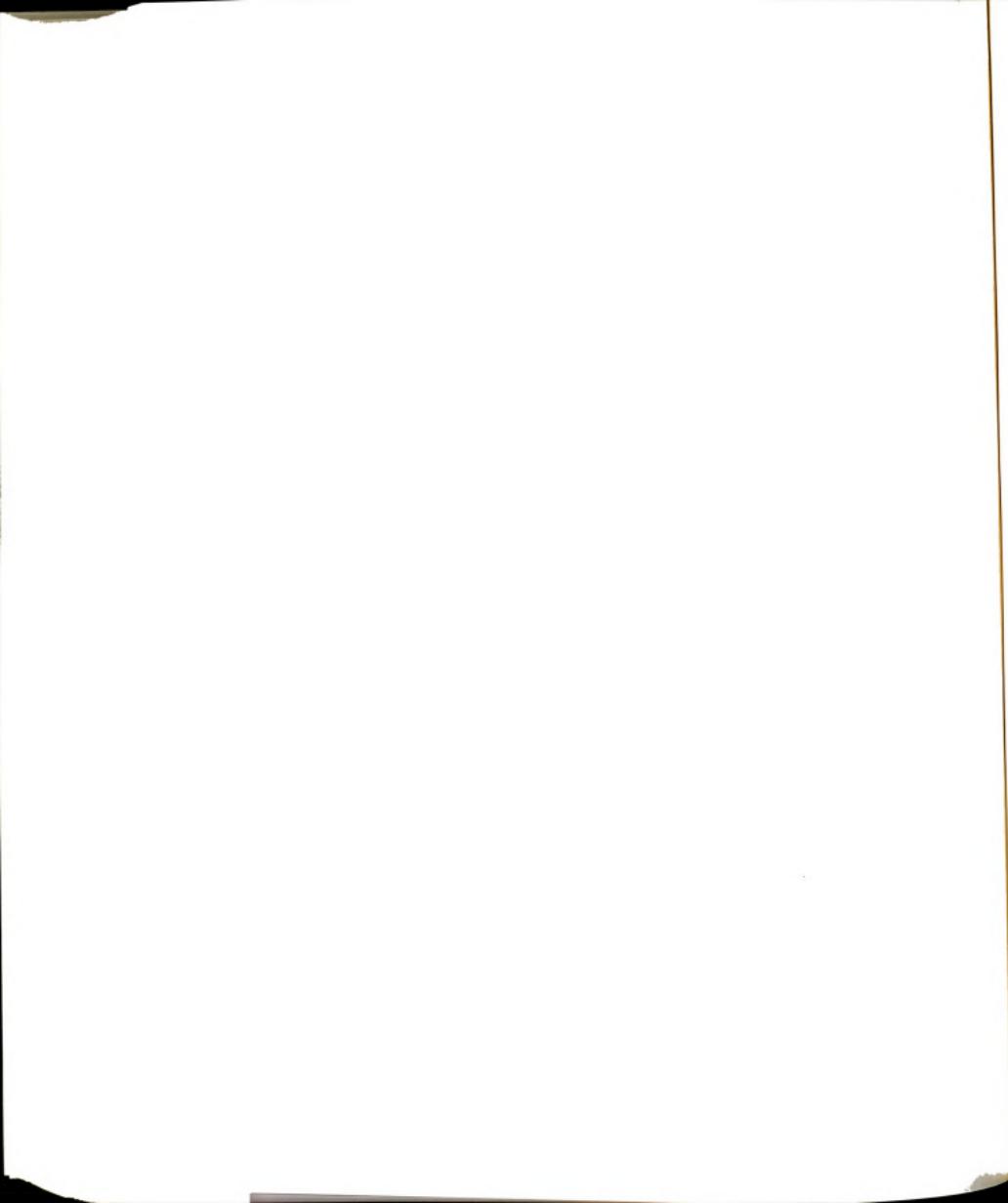


concentration of 1 mg/ml) for 15 minutes prior to the addition of the import substrate. The preparation of the cytosolic S100 fraction from HeLa cells has been described by Adam *et al.* (8).

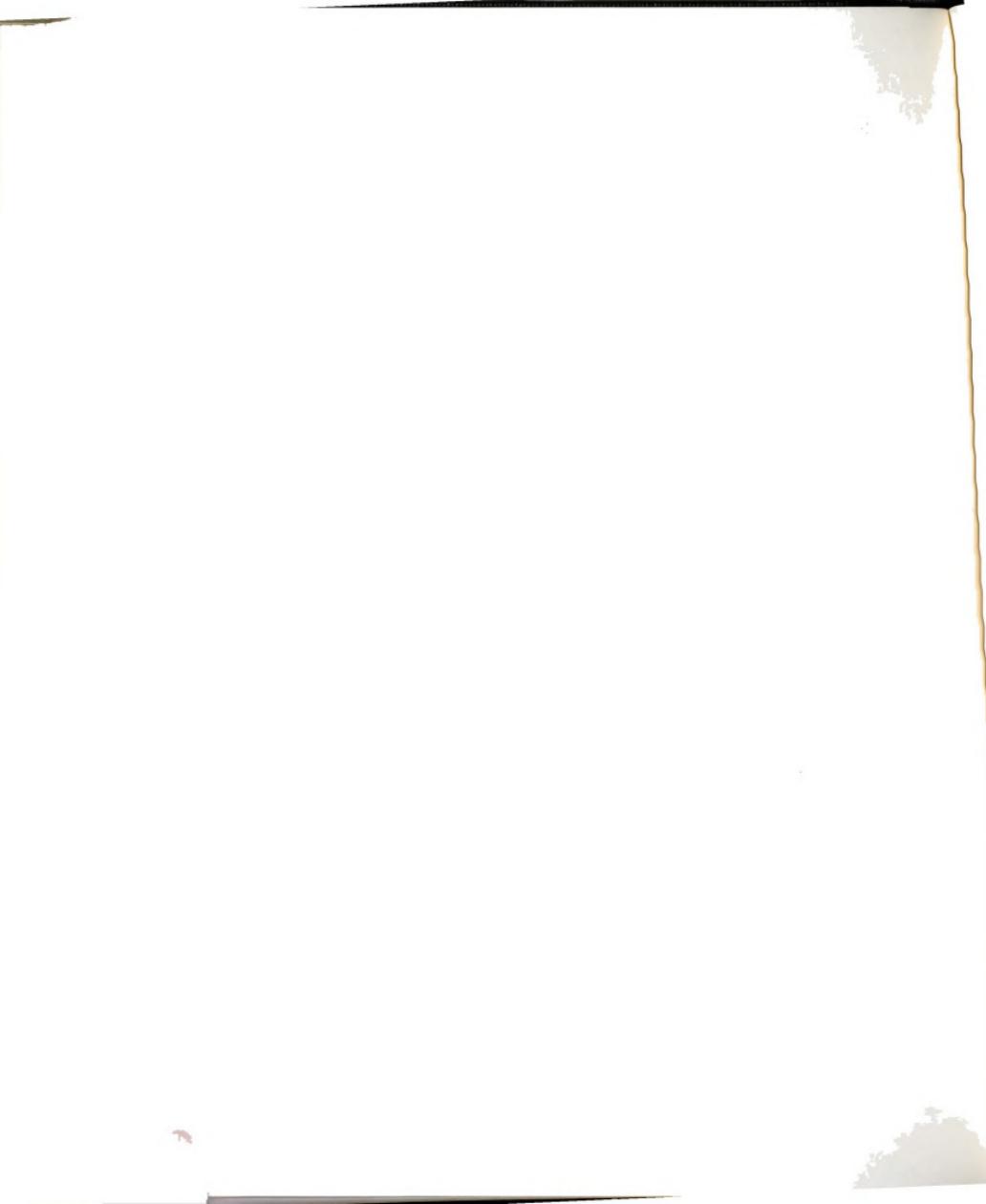
For assaying nuclear transport in intact cells (*in vivo* assay), glass coverslips containing the recipient cells were etched to facilitate subsequent identification of the field of injected cells. The import substrates were injected into the cytoplasm of the 3T3 cells at room temperature using an Eppendorf microinjector (model 5242) equipped with an Eppendorf micromanipulator. The cells were viewed during injection with a Zeiss Axiovert 10 inverted microscope equipped with a 32x objective. The needles (1 mm outer diameter, Sutton Scientific, Novato, CA) were pulled using a Narashige Scientific Instrument apparatus. Approximately 25-50 cells were injected on each coverslip using a continuous flow pressure of 40-60 hPa. After injection, the cells were incubated at 37°C for 30 minutes, washed with cold phosphate-buffered saline (PBS), and fixed for 15 minutes with cold 3% paraformaldehyde-PBS. The coverslips were then mounted on a drop of 10% PBS-90% glycerol containing 1 mg/ml phenylenediamine (Aldrich). The edges of the coverslips were sealed with nail polish. For WGA inhibition studies, the plant lectin was used at a final concentration of 1 mg/ml and injected concurrently with the import substrate.

### **Fluorescence microscopy**

The *in vitro* nuclear import assay coverslips were viewed on a Zeiss Axiophot microscope equipped with a 40x/0.75 Plan-Neofluar objective and



rhodamine fluorescence filters. The microinjected cells were examined on a Zeiss Axiophot microscope equipped with a 63x/1.4 immersion Plan-Apochromat lens and rhodamine fluorescence filters. The fluorescence photographs were taken on T-Max 3200 film (Eastman Kodak) developed at 1600 ASA.

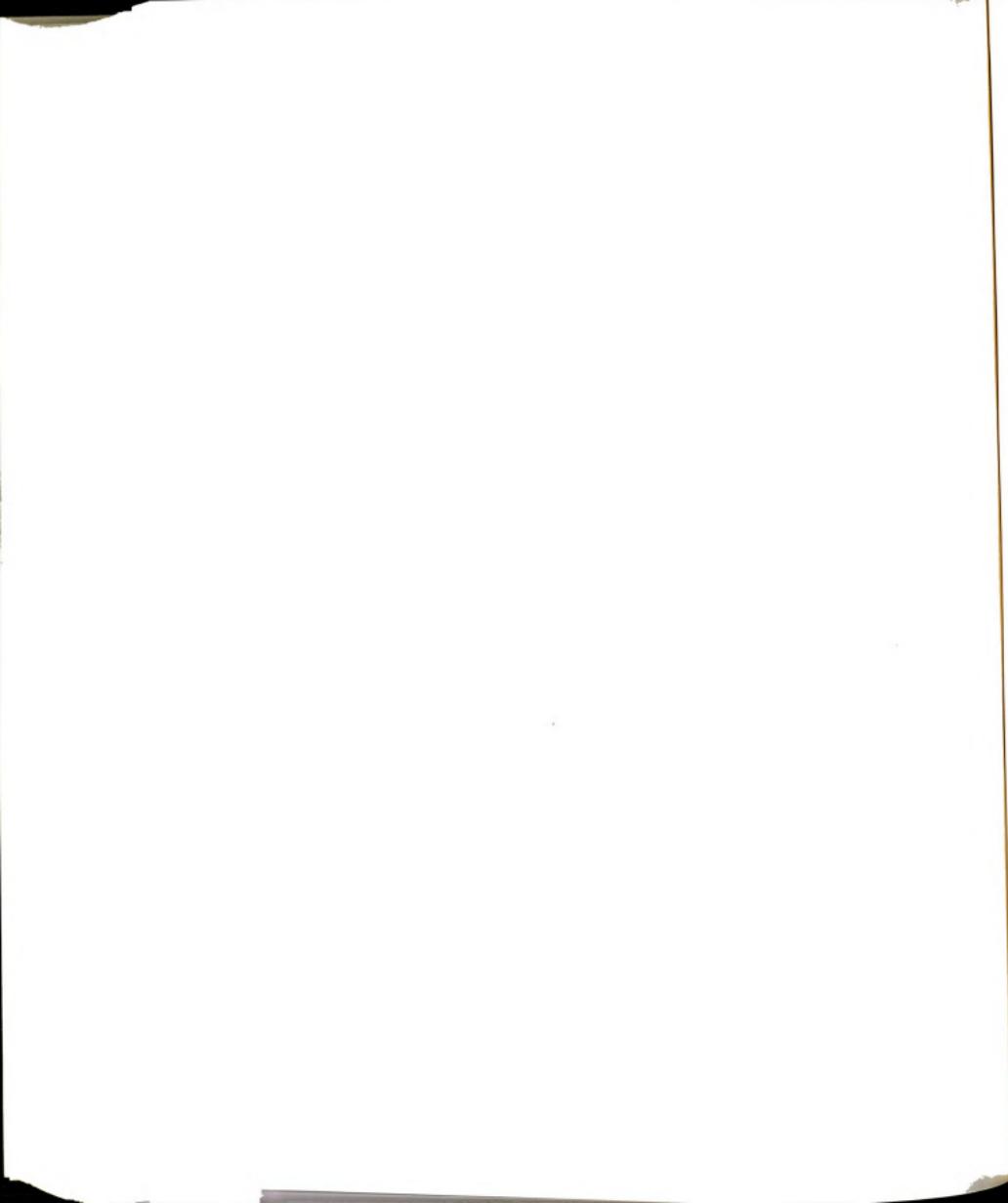


## RESULTS

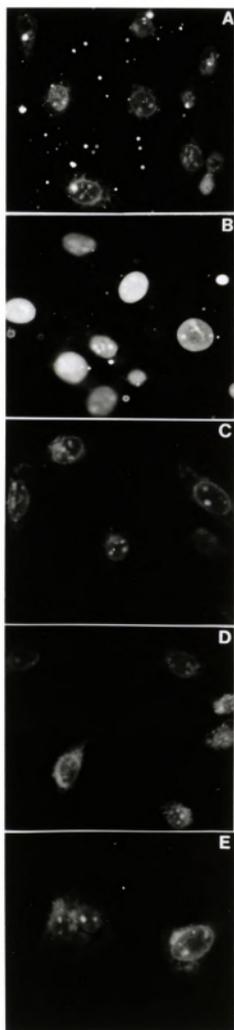
### Nuclear Transport of Rh-HSA-NLS Assayed with Permeabilized Cells

When mouse 3T3 fibroblasts were permeabilized with digitonin and then incubated with Rh-HSA-NLS in the presence or absence of the S100 fraction derived from HeLa cells, there was a clear S100-dependent transport of the fluorescent label into the nucleus. In the presence of the HeLa extract, intense staining of the nuclei was observed for all the cells (Fig. 1B). When S100 was omitted from the incubation, however, there was diffuse and faint staining of the cells; in some cells, there was distinct exclusion of the label from the nuclei (Fig. 1A). The nuclear localization was also dependent on NLS, since parallel experiments with Rh-HSA also resulted in diffuse distribution of fluorescence (Fig. 1D). As was observed with the analysis of Rh-HSA-NLS in the absence of S100 (Fig. 1A), many of the nuclei appeared to be "black holes," devoid of any fluorescence label (Fig. 1D). This fluorescence pattern was observed for Rh-HSA irrespective of whether S100 was included during the transport assay.

The plant lectin WGA binds to O-linked GlcNAc residues on nucleoporins and blocks protein import through the pore complexes (10-12). Consistent with these previously published results, the transport of Rh-HSA-NLS in our import assay was blocked by WGA (Fig. 1C). WGA, on the other hand, had no effect on the distribution of Rh-HSA (Fig. 1E compared to Fig. 1D). All of these results establish the assay system for our mouse 3T3 fibroblasts. More importantly, they demonstrate the functional activity of the human HeLa S100 cytosolic extract on



**Figure 1. *In vitro* nuclear import assay comparing the translocation of Rh-HSA-NLS with Rh-HSA.** Permeabilized 3T3 cells were incubated with transport reaction mixture containing the following substrates: A) Rh-HSA-NLS in the absence of S100 cytosolic extract; B) Rh-HSA-NLS in the presence of S100 cytosolic extract; C) Rh-HSA-NLS in the presence of S100 cytosolic extract and WGA (1 mg/ml); D) Rh-HSA in the presence of S100 cytosolic extract; and E) Rh-HSA in the presence of S100 cytosolic extract and WGA (1 mg/ml).



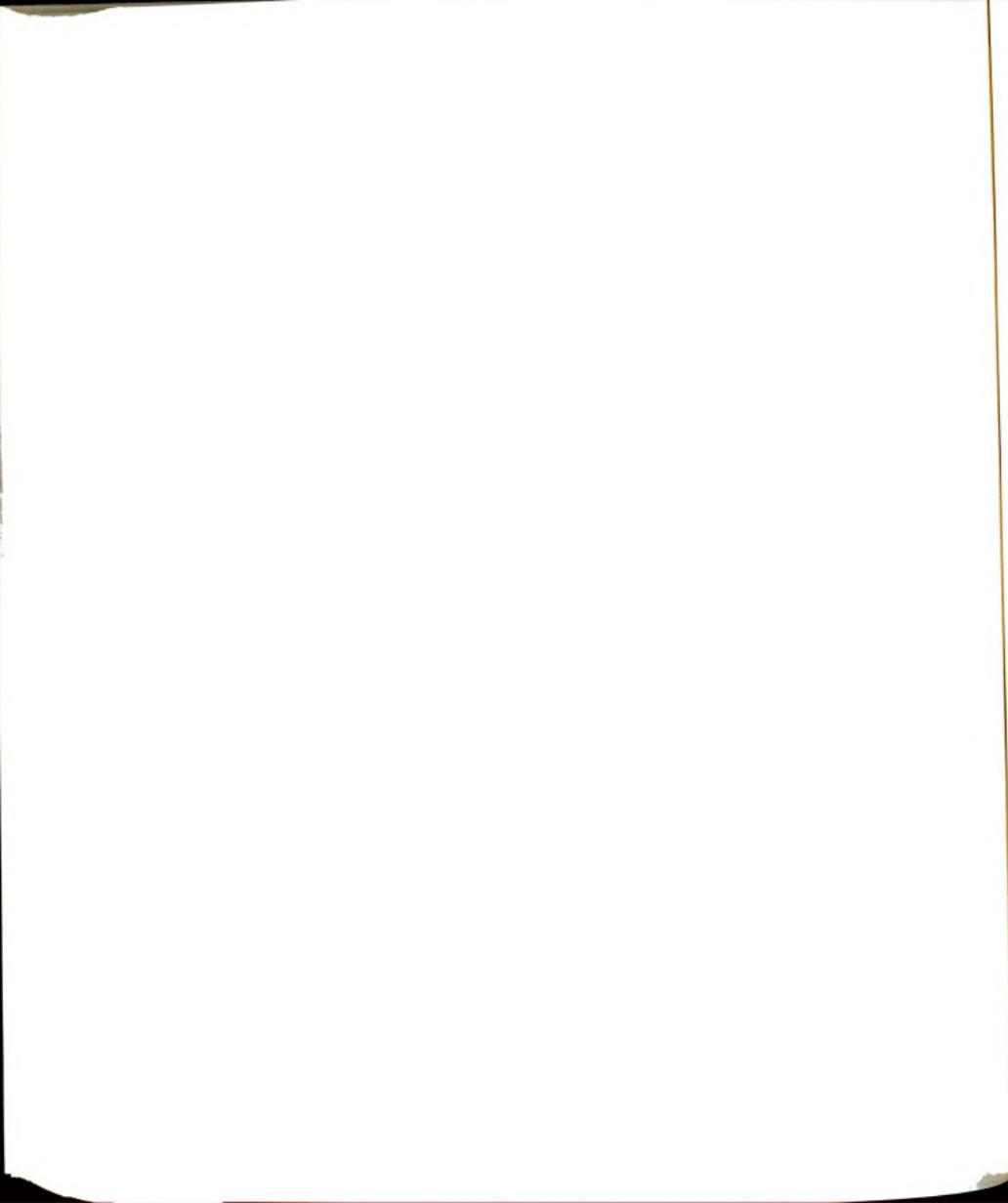


the heterologous mouse 3T3 recipient nuclei.

### **Behavior of Rh-rCBP35 in the Permeabilized Cell Assay System**

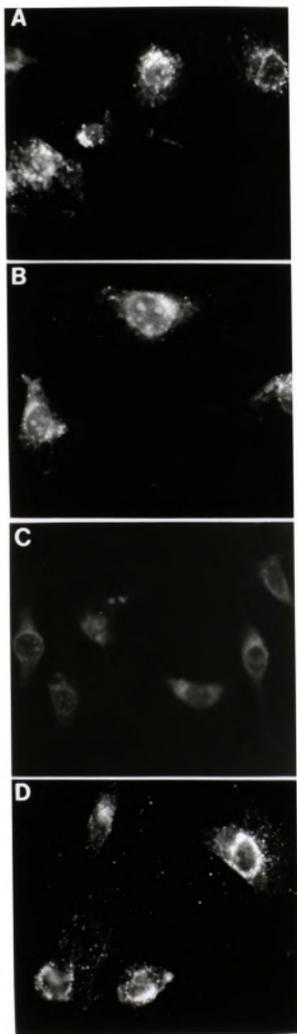
When the same *in vitro* transport assay was carried out in parallel using Rh-rCBP35, quite a distinct fluorescence pattern was observed. In the absence of S100, Rh-rCBP35 was found to label spots in certain portion of the cytoplasm; more strikingly, there was bright labeling of the nuclear periphery (Fig. 2A). The same pattern was obtained in the presence of S100 (Fig. 2B). There was no accumulation of the fluorescence label in the nuclei, as was observed for Rh-HSA-NLS incubated in the presence of the cytosolic extract (Fig. 1B).

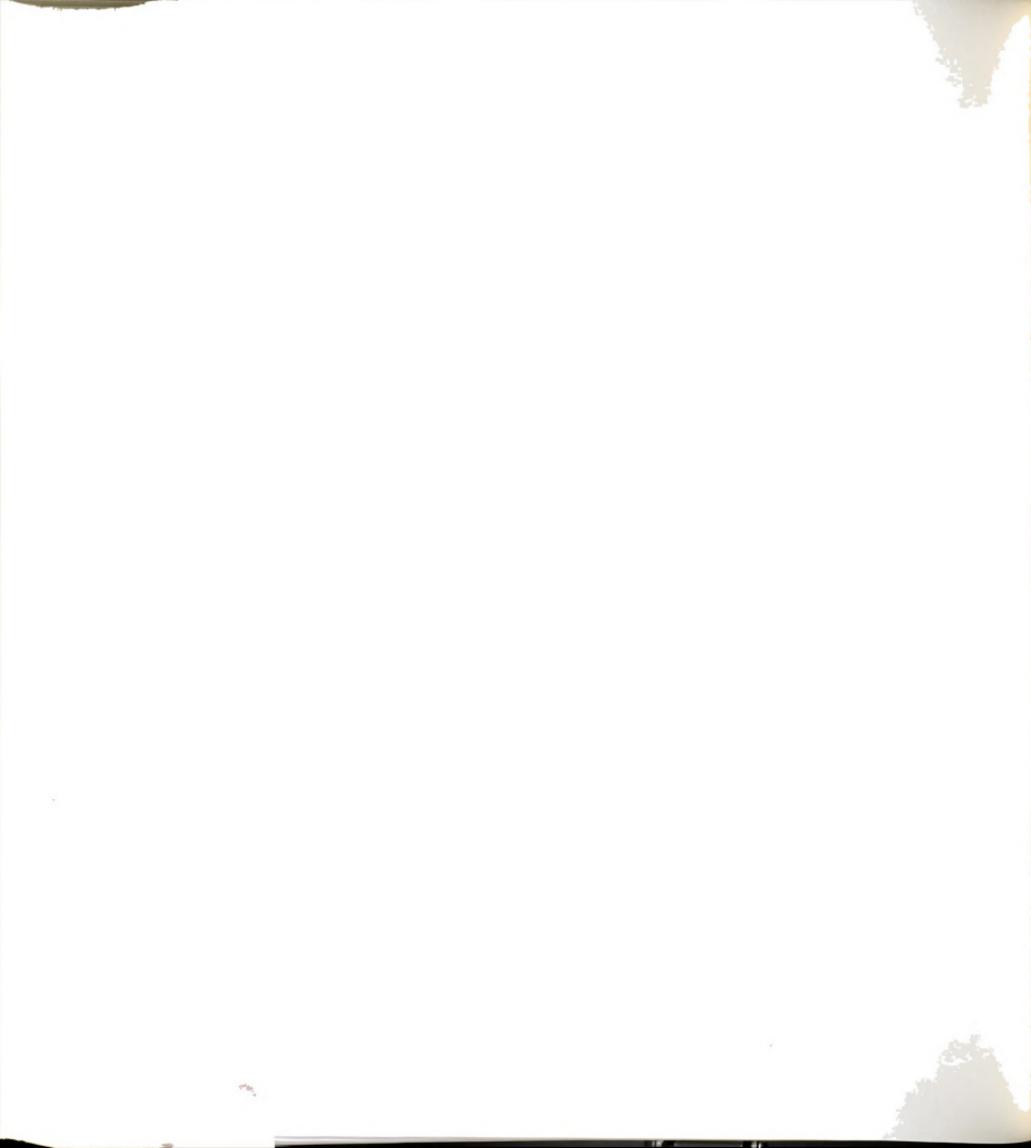
Because the Gal/Lac-specific CBP35 might be expected to bind to cell surface glycoconjugates (13), it was possible that nuclear transport might be blocked in this *in vitro* assay due to CBP35-glycoprotein interactions. To circumvent this potential problem, assays were carried out in the presence of 50 mM Lac. Under these conditions, however, there was a drastic reduction of the fluorescence due to Rh-rCBP35 (Fig. 2C). As least two factors (not mutually exclusive) could contribute to this decrease in fluorescence intensity: (a) Lac inhibition of Rh-rCBP35 binding to glycoproteins; and (b) Lac quenching of the fluorescence at 576 nm due to rhodamine. Attempts to overcome these difficulties, such as using higher concentrations of Rh-rCBP35 (Fig. 2D) or preincubating the permeabilized cells with unlabeled (non-fluorescent) rCBP35, did not alter the pattern of fluorescence, both in the presence and absence of S100 and in the presence and absence of Lac.





**Figure 2. *In vitro* nuclear import assay for Rh-rCBP35.** Permeabilized 3T3 cells were incubated with transport reaction mixture as follows: A) Rh-rCBP35 (15  $\mu\text{g/ml}$ ) in the absence of S100 cytosolic extract; B) Rh-rCBP35 (15  $\mu\text{g/ml}$ ) in the presence of S100 cytosolic extract; C) Rh-rCBP35 (15  $\mu\text{g/ml}$ ) in the presence of S100 cytosolic extract and lactose (50 mM); D) Rh-rCBP35 (120  $\mu\text{g/ml}$ ) in the presence of S100 cytosolic extract.





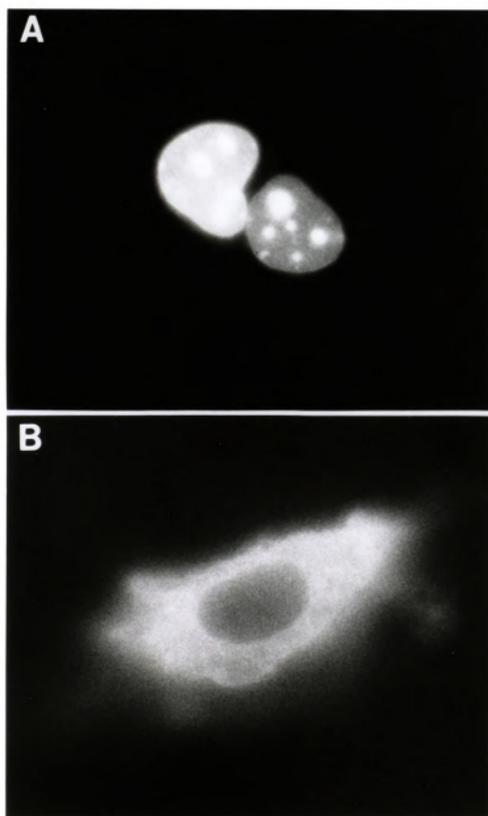
Thus, Rh-rCBP35 behaved quite differently when compared to Rh-HSA-NLS in the *in vitro* transport assay. No nuclear localization of the lectin could be observed, even under conditions when CBP35 endogenous to the cell could be found in the nuclei. One possibility is that the *in vitro* assay exposes CBP35 to glycoconjugates that the lectin would not normally encounter (intracellular lectin versus extracellular glycoconjugates). Another possibility is that the requirements for nuclear transport of CBP35 (e.g., phosphorylation) are different from those of HSA-NLS; HeLa S100 could provide the necessary factors for HSA-NLS transport but could not fulfill those required for CBP35. Therefore, microinjection of Rh-rCBP35 into intact cells was carried out to test if nuclear transport could be demonstrated for the labeled recombinant polypeptide.

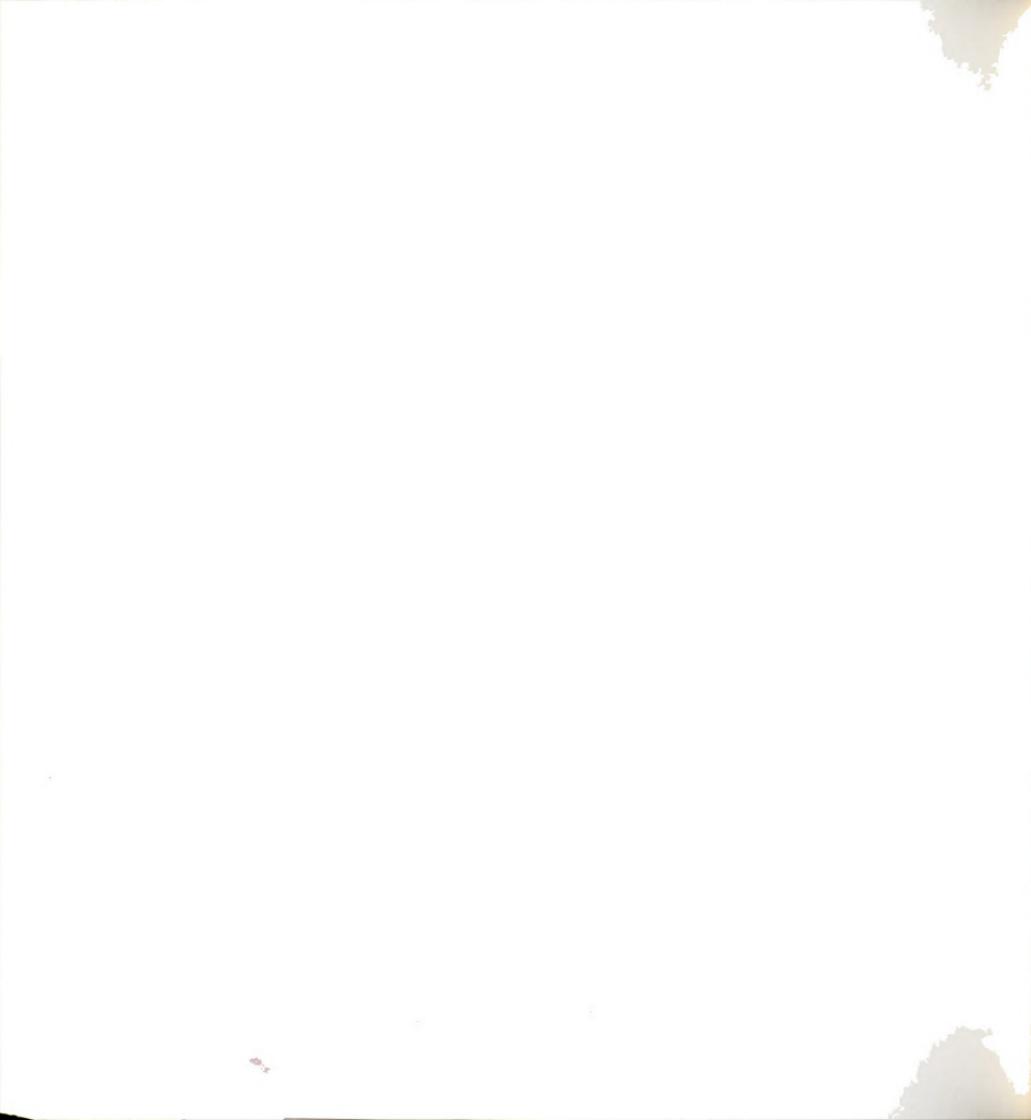
#### Nuclear transport in intact cells

Microinjection of Rh-HSA-NLS into mouse 3T3 fibroblasts resulted in the rapid accumulation of the fluorescent label in the nuclei of recipient cells. At the earliest time point tested (20 minutes post injection), there was intense staining of the nuclei (Fig. 3A). Some spots within the nucleus showed particularly bright staining; these possibly corresponded to nucleoli. Similar results were obtained at later times (40 and 60 minutes) following injection. This nuclear localization was inhibited when WGA was coinjected with the fluorescent label (Fig. 3B). In the absence of WGA, essentially 100% of the cells showed nuclear staining, as seen in Fig. 3A. In the presence of WGA, 80-90% of the microinjected cells showed cytoplasmic fluorescence, with little or no label in the nuclei as seen in Fig. 3B,



**Figure 3. *In vivo* nuclear import assay for Rh-HSA-NLS in the absence and presence of WGA.** Asynchronous 3T3 cells were microinjected with import substrate in: A) the absence of WGA; or B) in the presence of WGA (1 mg/ml).





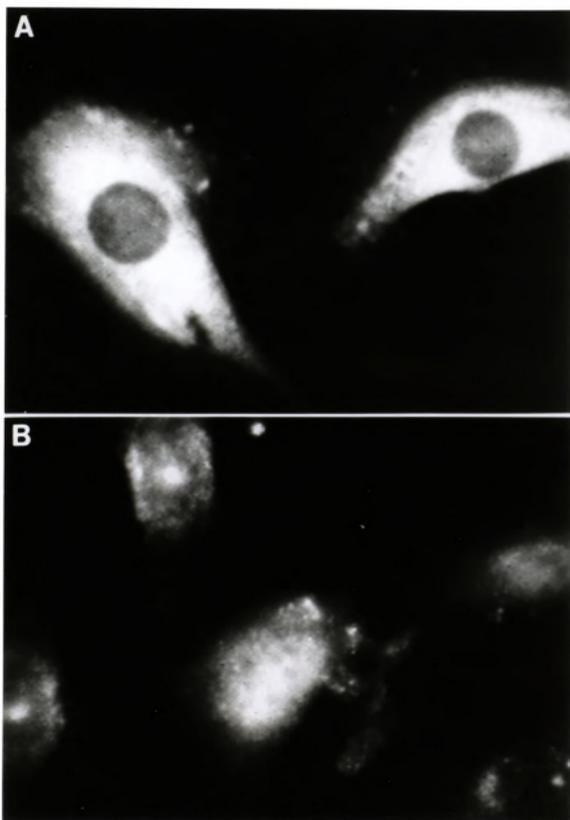
although some 10-20% of the cells still showed nuclear staining.

The requirement for NLS in the *in vivo* nuclear transport assay was established with Rh-HSA (Fig. 4A). In this case, greater than 95% of the recipient cells showed exclusive cytoplasmic staining and less than 5% showed any nuclear labeling. Thus, the results of our *in vivo* microinjection assay on Rh-HSA-NLS correspond to those previously published, showing strict requirement for NLS and sensitivity to inhibition by WGA.

When 3T3 cells were microinjected with Rh-rCBP35, there were cells showing clear nuclear localization of the fluorescent label (Fig. 4B). Indeed, Rh-rCBP35 behaved quite differently in the *in vivo* assay, compared to the *in vitro* assay (Fig. 2B). Either the intact cell could provide factors necessary for transport that were not fulfilled by the S100 fraction (e.g. by modifying the recombinant protein via phosphorylation), or microinjection obviated the difficulty of rCBP35 binding to cell surface glycoconjugates. Although some of the cells in the field shown in Fig. 4B revealed nuclear localization of rCBP35 (~40%), there were also cells that did not accumulate Rh-rCBP35 in their nuclei. A more detailed examination of the two labeling patterns was made on isolated single cells. In those cells showing nuclear localized Rh-rCBP35, the nucleus was stained except for regions corresponding to the nucleolus; there was also a diffuse distribution of fluorescence in the cytoplasm (Fig. 5A). On the other hand, in those cells in which the microinjected Rh-rCBP35 failed to localize in the nucleus, there was punctate staining of the cytoplasm and bright labeling of the nuclear periphery (Fig. 5B). In many respects, this latter staining pattern was reminiscent of the



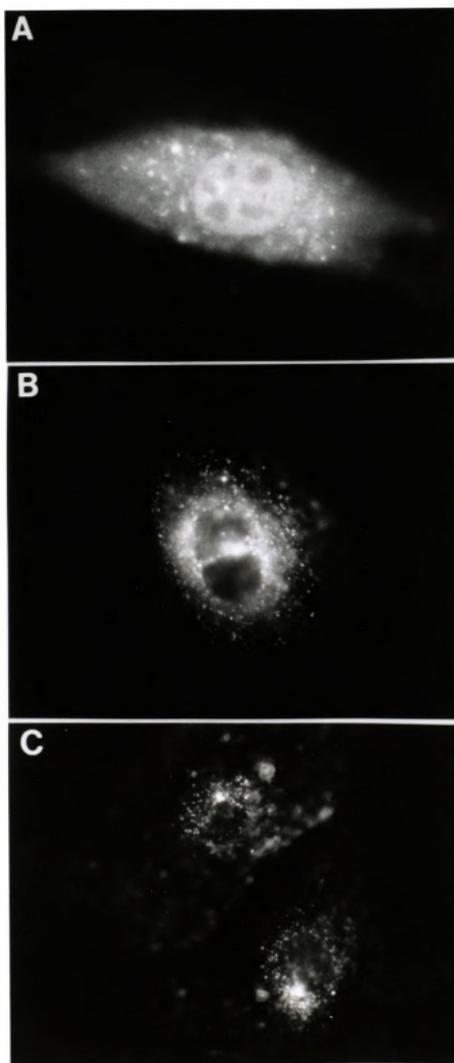
**Figure 4. *In vivo* nuclear import assay comparing the translocation of Rh-rCBP35 with Rh-HSA.** Asynchronous 3T3 cells were microinjected with: A) Rh-HSA; or B) Rh-rCBP35.



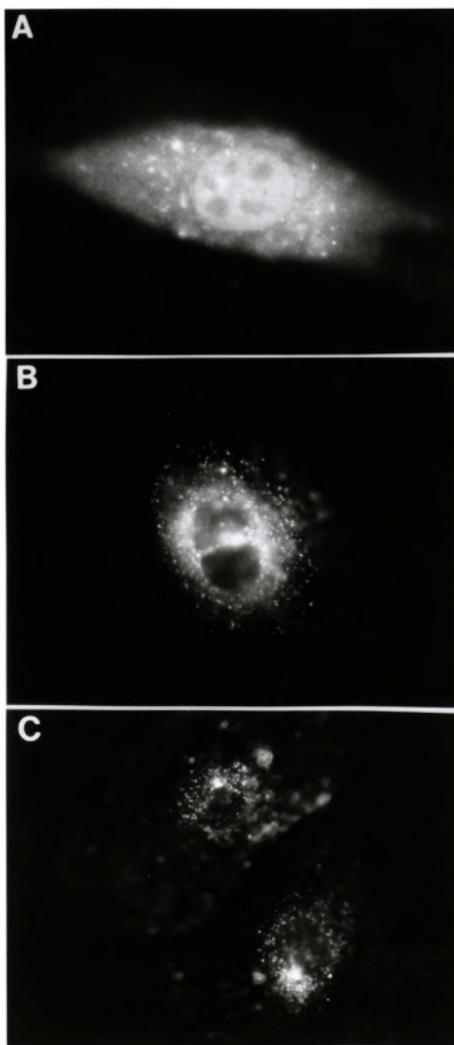




**Figure 5. *In vivo* nuclear import assay for Rh-rCBP35 showing the nuclear and non-nuclear localization of the protein.** Either asynchronous 3T3 cells (Panels A and B) or serum-starved quiescent 3T3 cells (Panel C) were microinjected with Rh-rCBP35. Panel A represents the nuclear localization seen in 20-30% of the cells. The remainder of the cells exhibit the type of localization seen in Panel B. The entire population of serum-starved cells (Panel C) exhibits no nuclear localization.



**Figure 5. *In vivo* nuclear import assay for Rh-rCBP35 showing the nuclear and non-nuclear localization of the protein.** Either asynchronous 3T3 cells (Panels A and B) or serum-starved quiescent 3T3 cells (Panel C) were microinjected with Rh-rCBP35. Panel A represents the nuclear localization seen in 20-30% of the cells. The remainder of the cells exhibit the type of localization seen in Panel B. The entire population of serum-starved cells (Panel C) exhibits no nuclear localization.





pattern observed when Rh-rCBP35 was used in the *in vitro* transport assay (Fig. 2A). We have also obtained the same results (Fig. 5A and 5B) when Rh-rCBP35 was microinjected into Buffalo rat liver cells.

Comparison of the behavior of Rh-rCBP35 and Rh-HSA-NLS in the *in vivo* nuclear transport assay revealed at least three key differences. First, when Rh-HSA-NLS localizes to the nucleus, there was hardly any fluorescence in the cytoplasm (Fig. 3A), whereas in cells with nuclear localized Rh-rCBP35, there was persistent fluorescent label in the cytoplasm (Fig. 5A). Second, while the nuclear transport of Rh-HSA-NLS could be inhibited by WGA (Fig. 3B), the nuclear localization of Rh-rCBP35 was not affected by coinjection of WGA. Finally, essentially all the cells microinjected with Rh-HSA-NLS showed nuclear labeling (in the absence of WGA) but not all of the cells receiving Rh-rCBP35 transported the lectin into their nuclei. In fact, the percent of cells showing nuclear localization versus cytoplasmic distribution (Fig. 5A versus Fig. 5B) depended on the time of observation after injection as well as on the condition of the culture under analysis.

### **Nuclear Transport of Rh-rCBP35 in quiescent and proliferating cells**

In previous studies (2), we had observed that serum-starved, quiescent cultures of 3T3 cells expressed low levels of CBP35 and what there is of the lectin was restricted to the cytoplasm of the cells. When serum-starved 3T3 cells were microinjected with Rh-rCBP35, practically all of the recipient cells showed punctate staining of the cytoplasm but little or no labeling of the cell nucleus (Fig.

5C). The fluorescence pattern was similar to those observed in asynchronous cultures of 3T3 cells in which the microinjected Rh-rCBP35 failed to localize to the nucleus (Fig. 5B). Thus, on the basis of correspondence of fluorescence patterns to Figure 5A (nuclear localized) versus Figure 5B (cytoplasmic localized), the percent of cells yielding nuclear transport of microinjected Rh-rCBP35 was essentially zero in the serum-starved cultures (Table I).

Microinjection of Rh-rCBP35 was also performed on 3T3 cultures that had been serum-starved and then stimulated for 5 hours. Under these conditions, approximately 20% of the cells showed nuclear localization of rCBP35, yielding fluorescence patterns resembling that shown in Figure 5A (Table I). The same experiment was carried out on 3T3 cells 16 hours post-stimulation, just at the onset of the S-phase of the cell cycle (2). Surprisingly, no recipient cells transported Rh-rCBP35 into their nuclei (Table I). These changes in nuclear localization properties of rCBP35 were not observed in a parallel analysis of microinjected Rh-HSA-NLS. In quiescent, as well as in stimulated 3T3 cells (5 hours and 16 hours post-stimulation of serum-starved cultures), essentially 100% of the microinjected cells accumulated Rh-HSA-NLS in their nuclei (Table I).

#### **Microinjection of NH<sub>2</sub>- and COOH-terminal domains of rCBP35**

The expression system for the production of large amounts of rCBP35 also led to the isolation of rCBPND and rCBPCD, corresponding to the NH<sub>2</sub>- and COOH-terminal domains, respectively, of the polypeptide. After fluorescent labeling with Rh, each individual domain was microinjected into 3T3 cells. There



TABLE I

Nuclear Localization in Serum-starved and Serum-stimulated 3T3 Cells

Protein Analyzed	Serum-starved	5 Hours After Serum Addition	16 Hours After Serum Addition
Endogenous CBP35 <sup>1</sup>	~ 2	~ 10	~ 40
Microinjected rCBP35 <sup>2</sup>	0	~ 20	0
Microinjected HSA-NLS <sup>2</sup>	100	100	100

<sup>1</sup> The cells were fixed with 3.7% formaldehyde, permeabilized with 0.5% Triton X-100, and stained with rabbit anti-CBP35 plus Rh-goat anti-rabbit immunoglobulin (2). Data represent percent of cells showing fluorescence in the nucleus.

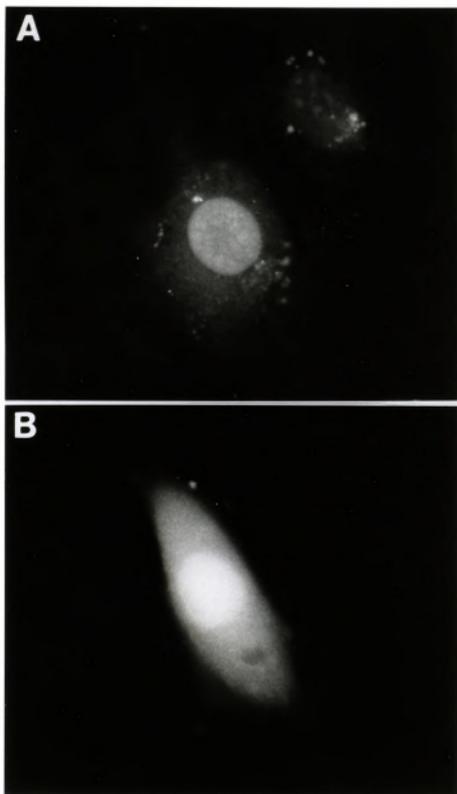
<sup>2</sup> The cells were microinjected with Rh-rCBP35 or Rh-HSA-NLS and analyzed directly for nuclear fluorescence. Data represent percent of cells showing fluorescence in the nucleus, on the basis of ~ 50 microinjected cells for each sample.



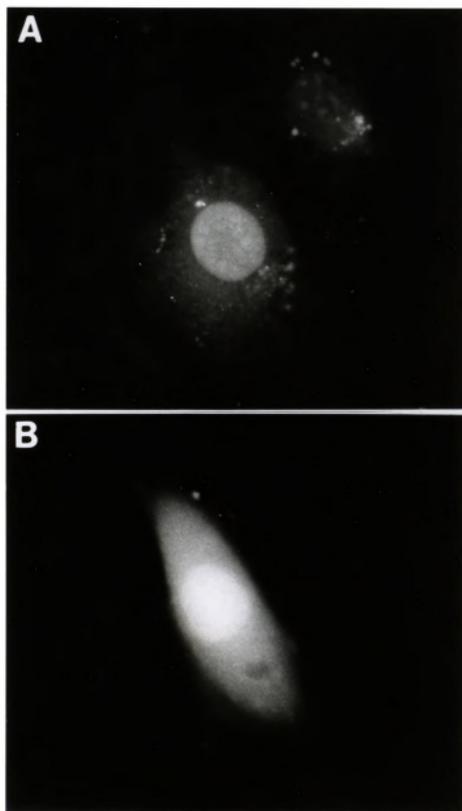
was clear-cut nuclear localization of Rh-rCBPND (Fig. 6A). As was seen for the full-length polypeptide, there was also population heterogeneity. Roughly half of the cells receiving Rh-rCBPND accumulated the fluorescent label in the nucleus while the other half of the population restricted the fluorescence to the cytoplasm. Similar results were also obtained for Rh-rCBPCD (Fig. 6B).



**Figure 6. *In vivo* nuclear import assay for the NH<sub>2</sub>- and COOH-terminal portions of rCBP35.** Asynchronous 3T3 cells were microinjected with: A) Rh-rCBPND; or B) Rh-rCBPCD.



**Figure 6. *In vivo* nuclear import assay for the NH<sub>2</sub>- and COOH-terminal portions of rCBP35.** Asynchronous 3T3 cells were microinjected with: A) Rh-rCBPND; or B) Rh-rCBPCD.





## DISCUSSION

The key findings of the present study include: (a) rCBP35 microinjected into intact cells could be translocated into the nucleus, while no such nuclear import could be observed with permeabilized cells; (b) under conditions where nuclear transport of rCBP35 was observed, the requirements and regulation of lectin transport were quite distinct from those observed for HSA-NLS; (c) serum-starved quiescent cultures of 3T3 fibroblasts failed to import rCBP35 into the nuclei, whereas serum-stimulated cells exhibit the capacity to translocate the lectin in a time-dependent fashion, with some transport observable during the early G<sub>1</sub> phase but no import at the onset of DNA synthesis; and (d) both the NH<sub>2</sub>- and COOH-terminal domains of rCBP35, when microinjected into 3T3 cells, were found to accumulate in the nucleus. We shall discuss each of the above items in turn.

Two main possibilities have been considered to explain the difference in the nuclear import properties of rCBP35 under *in vitro* versus *in vivo* assays. The first possibility is that the S100 cytosolic extract, while sufficient to facilitate the transport of HSA-NLS in the *in vitro* assay, could not provide a necessary factor for the transport of rCBP35. One candidate for such a specific requirement may be a protein kinase which normally phosphorylates CBP35 *in vivo*. The pI of the murine CBP35 polypeptide is 8.7, as determined by calculation from the deduced amino acid sequence and experimentally by isoelectric focusing of rCBP35 (14). When 3T3 cell extracts were subjected to two-dimensional gel electrophoresis and



immunoblotting, however, two spots are observed, corresponding to pI values of 8.7 and 8.2. The pI 8.2 form represents a posttranslational modification of the pI 8.7 polypeptide, by the addition of a single phosphate group (14). The phosphorylated form (pI 8.2) could be found in both the cytoplasm and the nucleus, whereas the pI 8.7 species was found exclusively in the nucleus. On the basis of these observations, it was hypothesized that phosphorylation of the rCBP35 polypeptide may be a requirement to traverse the nuclear pore complex during nucleo-cytoplasmic transport (15). Considered in these terms, the cytosol of the intact cell may be able to phosphorylate the rCBP35 while the S100 fraction of the *in vitro* assay could not do so.

An alternative possibility to account for the failure of rCBP35 to be translocated into the nucleus in digitonin-permeabilized cells may be due to the lectin's binding to glycoconjugates, which it would not encounter when microinjected. However, we could not overcome this difficulty either by inclusion of Lac to prevent saccharide-specific binding or by increasing the amount of rCBP35. Moreover, the staining pattern of rCBP35 in the *in vitro* assay, punctate cytoplasmic labeling with a distinct nuclear periphery, was similar to those observed in microinjected cells that failed to transport the lectin into the nuclei. Finally, this perinuclear staining pattern is similar to that observed as an initial step to nuclear localization after serum stimulation for a number of nuclear proteins, including c-Myc, DNA polymerase  $\alpha$ , and Proliferating Cell Nuclear Antigen (the auxiliary protein of DNA polymerase  $\delta$ ) (16). On the basis of these kinds of considerations, we presently favor the notion that the failure of rCBP35



to be transported in the *in vitro* assay may be due to a missing factor (kinase). This requirement may also account for the observation that only a fraction of the cells microinjected with CBP35 import the lectin into their nuclei and that this fraction varies with the proliferative state of the culture under analysis (see below).

From both the *in vitro* and *in vivo* assays, it seems evident that the requirements and regulation of nuclear import are quite different for rCBP35 and HSA-NLS. In microinjection experiments where nuclear translocation of both substrates was observed, the HSA-NLS transport was sensitive to inhibition by WGA, while that of rCBP35 was not affected by the plant lectin. Co-sedimentation and co-immunoprecipitation experiments (3; S.Y. Wang, unpublished observations) suggest that intracellular CBP35 is associated with a RNP, possibly including the U snRNPs. Thus, it is possible that the transport of rCBP35 would not be inhibited by the same concentration of WGA used to block entry of HSA-NLS, as had been reported for U1 and U5 snRNPs (17,18). We have not had the opportunity to test, as yet, the sensitivity of rCBP35 import to inhibition by trimethylguanosine cap (or by antibodies directed against trimethylguanosine cap), a structural marker that constitutes part of the nuclear localization signal for most of the U snRNPs (19,20).

The nuclear localization of HSA-NLS appears to be an all-or-none phenomenon. In any given cell, all of the Rh-HSA-NLS appeared to be accumulated in the nucleus, with little or no fluorescence in the cytoplasm. This was distinctly different from the case for CBP35, which showed both nuclear and

cytoplasmic staining. It should be noted that, in this respect, the behavior of microinjected Rh-rCBP35 mimics that of the lectin endogenous to 3T3 cells. Previous immunofluorescence staining of formaldehyde fixed and Triton X-100 permeabilized cells with anti-CBP35 had shown that the intracellular lectin could be found in both nuclear and cytoplasmic compartments of the same cell (2,3).

Irrespective of culture conditions, nearly all the cells microinjected with Rh-HSA-NLS translocated the label into the nucleus. In contrast, not all the cells receiving Rh-rCBP35 showed nuclear localized fluorescence. Again, this is more similar to the situation encountered with endogenous CBP35 (3), as well as other nuclear proteins such as c-Myc, DNA polymerase  $\alpha$ , and Proliferating Cell Nuclear Antigen (16). The fraction of cells exhibiting nuclear transport of microinjected rCBP35 appeared to depend on the proliferation state of the culture. The data on microinjected rCBP35 should be compared to the time course for the nuclear localization of endogenous CBP35 in serum-stimulated 3T3 fibroblasts (Table I). While the increase in fraction of cells showing nuclear localization of microinjected rCBP35 in early G<sub>1</sub> phase (5 hours post-stimulation) was consistent with an increase in nuclear CBP35 endogenous to the cell, the failure to transport microinjected rCBP35 at 16 hours post-stimulation presented a puzzle. One possibility is that the cell produces a limited amount of a factor required for nuclear import and the microinjected Rh-rCBP35 cannot effectively compete against endogenously synthesized CBP35 when the cell is producing the lectin maximally. Another possibility may be that by the onset of S-phase, the required transport factor is no longer available and that the high levels of endogenous



nuclear CBP35 reflects the accumulation of the lectin over the past 10-15 hours. Clearly, a more refined kinetic study assaying for both microinjected rCBP35 and the endogenous CBP35 would be required to sort out these preliminary findings.

One surprise of the present studies is that both the  $\text{NH}_2$ - and  $\text{COOH}$ -terminal domains of CBP35 could be found in the nucleus following microinjection. The molecular weights of the polypeptides corresponding to these two domains were, respectively, 15,000 and 16,000 as determined on SDS-PAGE (6). Gel filtration and sedimentation equilibrium studies indicate that they remain as monomers in nondenaturing buffers (unpublished observations). Although it has been suggested that nucleopores can accommodate the entry by passive diffusion of proteins with molecular weights of 20,000-40,000 (21,22), histones H1 (M, 21,000) and H2B (M, 13,800) are imported into the nucleus using mechanisms which are not dependent on simple diffusion (23,24). Moreover, if rCBPND and rCBPCD were freely diffusible through the nucleopores, it is not clear why they should stain the nuclei so intensely, relative to the cytoplasm. Finally, as was observed with intact rCBP35, there are cells microinjected with Rh-rCBPND and Rh-rCBPCD which clearly are not labeled in the nucleus. On the basis of these considerations, it does not appear likely that the individual domains entered the nucleus by diffusion. Thus, it would be of interest to determine if the same mechanisms that regulate the import of rCBP35 also apply to the  $\text{NH}_2$ - and  $\text{COOH}$ -terminal domains.



## ACKNOWLEDGMENT

We wish to thank Drs. G. Blobel and M.S. Moore for the generous gifts of Rh-HSA-NLS and for the use of the microinjection facilities and Dr. N. Raikhel for the use of the Zeiss Axiophot microscope.

## REFERENCES

1. Moutsatsos IK, Davis JM & Wang JL (1986) *J Cell Biol* **102**: 477-483.
2. Moutsatsos IK, Wade M, Schindler M & Wang JL (1987) *Proc Natl Acad Sci USA* **84**: 6452-6456.
3. Laing JG & Wang JL (1988) *Biochemistry* **27**: 5329-5334.
4. Wang JL, Werner EA, Laing JG & Patterson RJ (1992) *Biochem Soc Trans* **20**: 269-274.
5. Agrwal N, Wang JL & Voss PG (1989) *J Biol Chem* **264**: 17236-17242.
6. Agrwal N, Sun Q, Wang SY & Wang JL (1992) *J Biol Chem* (preceding chapter).
7. Greber UF & Gerace L (1992) *J Cell Biol* **116**: 15-30.
8. Adam SA, Marr RS & Gerace L (1990) *J Cell Biol* **111**: 807-816.
9. Moore MS & Blobel G (1992) *Cell* **69**: 939-950.
10. Featherstone C, Darby, MK & Gerace L (1988) *J Cell Biol* **107**: 1289-1297.
11. Finlay DR, Newmeyer DD, Price TM & Forbes DJ (1987) *J Cell Biol* **104**: 189-200.
12. Yoneda Y, Imamoto-Sonobe N, Yamaizumi M & Uchida T (1987) *Exp Cell Res* **173**: 586-595.



13. Wang JL, Laing JG & Anderson RL (1991) *Glycobiology* **1**: 243-252.
14. Cowles EA, Agrwal N, Anderson RL & Wang JL (1990) *J Biol Chem* **265**: 17706-17712.
15. Anderson RL & Wang JL (1992) *Trends Glycosci Glycotechnol* **4**: 43-52.
16. Vríz S, Lemaitre J-M, Leibovici M, Thierry N & Méchali M (1992) *Mol Cell Biol* **12**: 3548-3555.
17. Fischer U, Darzynkiewicz E, Tahara SM, Dathan NA, Lührmann R & Mattaj IW (1991) *J Cell Biol* **113**: 705-714.
18. Michaud N & Goldfarb D (1992) *J Cell Biol* **116**: 851-861.
19. Hamm J, Darzynkiewicz E, Tahara SM & Mattaj IW (1990) *Cell* **62**: 569-577.
20. Fischer U & Lührmann R (1990) *Science* **249**: 786-790.
21. Silver P (1991) *Cell* **64**: 489-497.
22. Garcia-Bustos J, Heitman J & Hall MN (1991) *Biochim Biophys Acta* **1071**: 83-101.
23. Moreland RB, Langevin GL, Singer RH, Garcea RL & Hereford LM (1987) *Mol Cell Biol* **7**: 4048-4057.
24. Breeuwer M & Goldfarb DS (1990) *Cell* **60**: 999-1008.



## CHAPTER V

### CONCLUDING STATEMENT

The expression and intracellular localization of CBP35 is linked to the proliferative state of the cell (1,2). In actively growing cells, CBP35 protein levels are elevated and the protein is in the nucleus. Therefore, the studies detailed in this dissertation were initiated to examine the control and kinetics of the expression of the CBP35 gene. Furthermore, it was our aim to delineate the mechanisms by which CBP35 can translocate to the nucleus.

The first question was answered by performing Northern blot analyses and nuclear run-off assays. The conclusions derived from these experiments establish that the protein is subject to transcriptional regulation in cells. Proliferating cells have a higher transcriptional rate for CBP35, as well as higher levels of CBP35 mRNA, as compared to quiescent cells. The mRNA for CBP35 is induced within 30 minutes after serum stimulation of quiescent cells. Moreover, the transcription of the CBP35 gene is independent of *de novo* protein synthesis. Together, these data provide a strong case for the classification of the CBP35 gene as a primary response gene.

In order to examine the nuclear translocation of CBP35, we required that sufficient quantities of CBP35 be available for derivatization with fluorescent



reagents. A procaryotic expression vector system was devised to produce large amounts of CBP35. The protein generated in this manner is designated as recombinant CBP35, or rCBP35.

When rCBP35 derivatized with rhodamine (Rh-rCBP35) was added to an *in vitro* nuclear import assay, no translocation was observed. This was in contrast to the results observed for the nuclear import substrate Rh-HSA-NLS. However, when rCBP35 was microinjected into the cytoplasm of live cells, nuclear import was evident. Thus, the conclusion that can be drawn from these data is that the nuclear import of Rh-rCBP35 cannot be achieved by a property intrinsic to the polypeptide. The import of Rh-rCBP35 may require that the protein undergo some type of modification, such as phosphorylation. On the other hand, Rh-rCBP35 may be co-transported with another nuclear molecule. A possible candidate for such a molecule would be the U snRNPs. It has been postulated that a possible ligand for the intracellular population of CBP35 is a ribonucleoprotein (3; S.Y. Wang, unpublished observations). The *in vivo* nuclear transport assay may, thus, not only allow us to examine the factors governing the nuclear entry of CBP35, but it may also be used to elucidate the ligand, and hence the function, for CBP35 in the cytoplasm and the nucleus.

#### REFERENCES

1. Moutsatsos IK, Davis JM & Wang JL (1986) *J Cell Biol* **102**: 477-483.
2. Moutsatsos IK, Wade M, Schindler M & Wang JL (1987) *Proc Natl Acad Sci USA* **84**: 6452-6456.
3. Laing JG & Wang JL (1988) *Biochemistry* **27**: 5329-5334.





MICHIGAN STATE UNIV. LIBRARIES



31293007924289