
Received 18 December 1991

**Nuclear and cytoplasmic localization of a lectin–ribonucleoprotein complex**

John L. Wang,* Elizabeth A. Werner,** James G. Laing* and Ronald J. Patterson†

*Department of Biochemistry and †Department of Microbiology, Michigan State University, East Lansing, MI 48824, U.S.A.

**Introduction**

Carbohydrate-binding protein 35 (CBP35; Mr ~35,000) was initially purified from extracts of mouse 3T3 fibroblasts on the basis of its binding to galactose-containing glycoconjugates [1]. The affinity of the lectin for the disaccharide lactose is 60–100 times greater than that for galactose. Moreover, the binding affinity is further enhanced ~10-fold by the incorporation of an acetamido group at position 2 of Glc (i.e. to yield N-acetyllactosamine). An even greater enhancement of binding activity results from the addition of GalNAcα1 at position 3 of the Gal moiety of lactose derivatives (e.g., as in blood group A tetrasaccharide; see [2, 3]).

Immunoblotting analysis, using a polyclonal rabbit antiserum raised against 3T3 cell-derived CBP35, showed that the lectin and its homologues in other species are found in a variety of tissue types and various cell lines [2, 3]. The amino acid sequence (264 residues), deduced from the nucleotide sequence of a cDNA clone [4], showed that the CBP35 polypeptide consisted of two domains (Fig. 1a): an N-terminal half that contains repeats of the sequence Pro-Gly-Ala-Tyr-Pro-Gly followed by three other amino acids (Pro-, Gly-rich domain) and a C-terminal half that is homologous to other galactose-binding proteins classified as S-type lectins by Drickamer [5]. Lectins are grouped into S-type family on the basis of conserved amino acid sequence, 15 amino acid residues are invariant due sequence, 15 amino acid residues are invariant within a 39-residue sequence, 15 amino acid residues are invariant in the CRDs of all S-type lectins (Fig. 1a). The S-type lectins are divided in turn into two subgroups: L-14 and L-30. In general, the polypeptide of the L-14 group consists of a single domain, the CRD (Mr, 10,000–16,000). The L-30 group poly-
Fig. 1
Domain content and structural features in the L-14 and L-30 subgroups of S-type lectins

(a) The 15 invariant amino acid residues that occur in a 39-residue sequence in the CRDs of all of the L-14 and L-30 S-type lectins are shown. Also shown is the consensus sequence (nine residues) that is repeated in the proline-, glycine-rich domain of L-30. The letter \( n \), designating the number of repeats, ranges from five in the human Mac-2 sequence to ten in the rat eBP sequence. A single variable residue is denoted by a hyphen (-). Sequences of two or more variables residues are denoted by the symbol \((-\))

(b) Nomenclature of proteins belonging to the L-30 subgroup of S-type lectins. Abbreviations used: L-34, tumour cell lectin; HL-29 and RL-29, lectins from human and rat lungs, respectively.

<table>
<thead>
<tr>
<th>L-14</th>
<th>L-30</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-\text{PGAYPG-})_n--H-NPRF-V-N-WG-E-R-FPF-G--</td>
<td>(-\text{H-NPRF-V-N-WG-E-R-FPF-G-}--)</td>
</tr>
</tbody>
</table>

peptide contains two domains \((M, 29,000-35,000)\); the CRD is coupled to another domain (Pro-, Gly-rich domain), presumably responsible for mediating some as yet unidentified effector function (Fig. 1a).

Thus, the relative molecular mass and the sequences within the CRD classify CHP35 in the L-30 subgroup of the S-type lectin family [2, 3].

Amino acid sequence information on several other proteins, originally isolated from various sources and designated different names, showed that these must also be L-30 S-type lectins (Fig. 1b). Indeed, L-34 [6], HL-29 [7] and RL-29 [8] were isolated on the basis of their galactose-specific binding activities from mouse, human and rat cells, respectively. On the basis of sequence information, laminin-binding protein (LBP) [9], macrophage cell-surface antigen (Mac-2) [10] and IgE-binding protein (eBP) [11] were also tested for carbohydrate-binding activity, with the results showing galactose-specific binding. In addition to these proteins, a galactose-specific lectin \((M, \sim 30,000)\) has also been purified from baby hamster kidney (BHK21 C13) cells [12]. It is likely that this protein will be a member of the L-30 group of S-type lectins; at present, however, no sequence information has been reported. The available sequence information also indicates that all of the L-30 proteins listed in Fig. 1 are either identical (within a given species) or are homologous (between species). Thus, there appears to be a single member of the L-30 subgroup of the S-type animal lectins.

The variety of perspectives from which this single gene product has been studied is quite striking, particularly in terms of localization to subcellular compartments. The predominant proportion of CHP35 was found to be intracellular by immunofluorescent staining of fixed and permeabilized cells [13]. There was prominent labelling of the nucleus and variable staining of the cytoplasm in proliferating cultures of 3T3 cells. By immunofluorescence, L-34 has been found at the cell surface and in intracellular pools of various murine tumour cells (melanoma B16-F1 and K-1735, fibrosarcoma UV-2237-IP3 and carcinoma HeLa-S3 cell lines). On the other hand, the Mac-2 antigen and LBP were both identified in mouse macrophages on the basis of their cell-surface localization. The former was originally described by a monoclonal antibody that stained the cell surface of thioglycollate-elicited peritoneal macrophages to a much higher extent than resident macrophages [14]. The latter was described as a macrophage protein \((M, \sim 35,000)\) isolated by laminin–Sepharose affinity chromatography, that did not belong to the integrin family of laminin receptors. LBP bound to laminin–Sepharose could be eluted by the addition of galactose or lactose [9]. Using the same approach, eBP was originally identified in rat basophilic leukaemia cells.
on the basis of its isolation on a rat IgE affinity column [15]. Subcellular fractionation studies on rat basophilic leukaemia cells showed that the majority of eBP is intracellular, including the nucleus, and that a small proportion is on the cell surface [16]. The cDNAs identified for L-34, Mac-2 and eBP have revealed no obvious N-terminal signal sequence for sequestration into the lumen of the endoplasmic reticulum. Therefore, the lectin found in the extracellular compartment probably did not follow the classical endomembrane pathway of secretion.

The nucleus and cytosol of the intracellular compartment can communicate through nuclear pores and are thought to be topologically continuous (through an aqueous channel), although functionally distinct. However, the extracellular medium and the cell surface represent the extracellular compartment, which is topologically continuous with the lumen of intracellular, membrane-enclosed vesicles, including the endoplasmic reticulum and Golgi apparatus, but is topologically distinct from the intracellular compartment. In this respect, the L-30 S-type lectin represents an example of a growing list of proteins with dual intracellular and extracellular localization [2, 3]. In this paper, we discuss the intracellular localization of CBP35 in mouse 3T3 fibroblasts, emphasizing the association of the lectin with ribonucleoprotein (RNP) complexes in both the cytoplasm and the nucleus.

**Association of CBP35 with nuclear and cytoplasmic ribonucleoprotein complexes**

Although rabbit anti-CBP35 detected a small amount of the lectin on the surface of the 3T3 cells (by immunofluorescent staining, by agglutination of live cells, and by immunosolubilization of 125I-labelled lectin after surface iodination), the predominant portion of the lectin was found to be intracellular by immunofluorescent staining of fixed and permeabilized cells [13]. There was prominent labelling of the nucleus and variable staining of the cytoplasm. Identical results have been obtained using the monoclonal antibody generated against the Mac-2 antigen [14] to stain 3T3 fibroblasts.

Several lines of evidence suggest that CBP35 is associated with RNP s in the nucleoplasm. First, treatment of permeabilized 3T3 cells with RNAase A removed the immunofluorescence due to CBP35 in the nucleus, whereas similar treatment with DNAase I failed to yield the same effect [17]. An example of such an experiment is shown in Fig. 2. Without prior fixation, Triton X-100 extracted most of the cytoplasmic CBP35. The CBP35 remained in the nucleus, however, and yielded a punctate intranuclear staining pattern (Fig. 2, 2a). Digestion of the Triton X-100 permeabilized cell with DNAase I, followed by ammonium sulphate extraction, removed most of the DNA, as indicated by the loss of Hoechst dye staining (Fig. 2, 2b). The chromatin-depleted nuclei retained, however, the immunofluorescence due to CBP35 (Fig. 2, 2b). Finally, RNAase A digestion of the Triton X-100-extracted and chromatin-depleted nuclear residue resulted in the loss of nuclear CBP35 (Fig. 2, 2c).

Secondly, when nucleoplasm was fractionated by density gradient centrifugation, immunoblotting analysis localized CBP35 in fractions with densities corresponding to those reported for heterogeneous nuclear ribonucleoprotein complexes (hnRNP) (1.30–1.35 g/ml on cesium sulphate gradients and 40S on sucrose gradients). Thirdly, fractionation of nucleoplasm on Gal-Sepharose or anti-CBP35 polyacrylamide beads resulted in binding of the lectin as well as RNA. When the anti-CBP35 immunoprecipitate was subjected to immunoblotting, CBP35 and the B/B' polypeptides of the Sm antigens of small nuclear ribonucleoprotein particles (snRNP s) were detected by specific antibody reagents. Conversely, an autoimmune serum reactive with the Sm antigens co-immunoprecipitated both the Sm antigens (B/B' polypeptides) and CBP35. All of these results suggest that CBP35 in the nucleoplasm is associated with RNP complexes; in particular, the association of CBP35 with hnRNP and snRNP indicates that the lectin may be a component of the spliceosome and suggests specific tests of its physiological activity (see below).

The CBP35 in the cytoplasm is also complexed with RNP. When the cytoplasm of 3T3 cells was fractionated on sucrose gradients, immunoblot analysis showed that CBP35 was found in fractions corresponding to particles of greater than 40S. In contrast, parallel immunoblotting of the fractions for the cytosolic marker lactate dehydrogenase showed that the soluble enzyme concentrated at the top of the gradient. Recombinant CBP35, purified from an expression system in *Escherichia coli*, also sedimented at the top of the gradient. Finally, when the cytoplasm of 3T3 cells was treated first with micrococcal nuclease and then subjected to sedimentation on sucrose gradients, most of the RNA (as detected by absorbance at 260 nm) disappeared from the bottom of the gradient. Under these conditions, parallel immunoblotting of the fractions for CBP35 showed that the lectin also shifted to the top of the gradient. Thus, it appears that CBP35 in the cyto-
Effect of extraction and enzyme treatments on the localization of CBP35 in the nucleus of 3T3 cells

(a) Triton X-100. The cells were permeabilized in 20 mM-Tris (pH 7.2), 5 mM-KCl, 1 mM-MgCl₂, 1 mM-phenylmethanesulphonyl fluoride (TKM buffer) containing 1 mM-vanadyl adenosine, 250 mM-ammonium sulphate, and 0.5% Triton X-100 for 30 min at 4°C. (b) DNAase I. Cells permeabilized in (a) were incubated in TKM buffer containing DNAase I (100 μg/ml) and 1 mM-vanadyl adenosine for 30 min at 21°C. The cells were then extracted with 250 mM-ammonium sulphate. (c) RNAase A. Chromatin-depleted nuclear residue prepared in (b) was incubated with TKM buffer containing RNAase A (25 μg/ml) for 30 min at 21°C. After the treatment(s) in (a), (b) or (c), the cells were fixed in 3.7% (v/v) formaldehyde for 15 min before staining and light microscopy. Row 1: phase contrast microscopy. Row 2: immunofluorescence staining with rabbit anti-CBP35 (1:10 dilution of antiserum) and rhodamine-conjugated goat anti-rabbit immunoglobulin (1:30 dilution) observed with a 580 nm barrier filter. Row 3: fluorescence staining with the DNA-specific dye Hoechst 33258 (10 μg/ml) observed by using a 430 nm barrier filter. The bar indicates 10 μm.

In quiescent cultures of 3T3 fibroblasts, CBP35 is found primarily in the cytoplasm, whereas in proliferating cultures, it increases in amount and is located predominantly in the nucleus. The addition of serum growth factors to quiescent 3T3 cells increases the expression of CBP35; this increase occurs within 30 min of mitogenic stimulation, well before the onset of the first S-phase of the cell cycle [18, 19]. These conclusions are derived from: (a) analysis at the single-cell level by immunofluorescence; (b) analysis at the protein level by Western blotting; (c) analysis at the mRNA level by Northern blotting; and (d) analysis of the transcription of the CBP35 gene in nuclear run-off experiments. At the mRNA accumulation and gene transcription levels, the increases in the expression of CBP35 occur even in the presence of cycloheximide. In this respect, the expression of CBP35 resembles that of c-fos, whose transcription is also activated in response to serum growth factors, independent of the synthesis of other proteins [20].

At the protein level, the expression of CBP35 was further analysed by subcellular fractionation, two-dimensional PAGE to separate the isoelectric variants, and immunoblotting [21]. 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 recombinant CBP35. However, when cell extracts are subjected to two-dimensional PAGE and immunoblotting, two spots are observed, corresponding to pI values of 8.7 and 8.2. The pI 8.2 form represents a post-translational modification of the pI 8.7 polypeptide by the addition of a single plasmid is also in a RNP complex. The density of this complex, as determined on cesium sulphate gradients, was ~ 1.40 g/ml.

Nuclear and cytoplasmic distribution of CBP35
In quiescent cultures of 3T3 fibroblasts, CBP35 is found primarily in the cytoplasm, whereas in proliferating cultures, it increases in amount and is located predominantly in the nucleus. The addition of serum growth factors to quiescent 3T3 cells increases the expression of CBP35; this increase occurs within 30 min of mitogenic stimulation, well before the onset of the first S-phase of the cell cycle [18, 19]. These conclusions are derived from: (a) analysis at the single-cell level by immunofluorescence; (b) analysis at the protein level by Western blotting; (c) analysis at the mRNA level by Northern blotting; and (d) analysis of the transcription of the CBP35 gene in nuclear run-off experiments. At the mRNA accumulation and gene transcription levels, the increases in the expression of CBP35 occur even in the presence of cycloheximide. In this respect, the expression of CBP35 resembles that of c-fos, whose transcription is also activated in response to serum growth factors, independent of the synthesis of other proteins [20].

At the protein level, the expression of CBP35 was further analysed by subcellular fractionation, two-dimensional PAGE to separate the isoelectric variants, and immunoblotting [21]. The pI of the murine CBP35 polypeptide is 8.7, as determined both by calculation from the deduced amino acid sequence and experimentally by isoelectric focusing of recombinant CBP35. However, when cell extracts are subjected to two-dimensional PAGE and immunoblotting, two spots are observed, corresponding to pI values of 8.7 and 8.2. The pI 8.2 form represents a post-translational modification of the pI 8.7 polypeptide by the addition of a single
phosphate group, probably O-linked. The phosphorylated (pI 8.2) form of CBP35 is found both in the cytosol and nucleus, whereas the unmodified (pI 8.7) species is found exclusively in the nucleus. Quiescent populations have primarily the phosphorylated form, but it is at a low level and is located predominantly in the cytoplasm. Serum-stimulated cells have an increased level of the phosphorylated form, both in the cytosol and the nucleus, but the amount of the unmodified (pI 8.7) form increases dramatically and is exclusively nuclear. The significance and mechanism of the differential expression and localization of the two isolectric variants of CBP35 remain as some of the many intriguing, but unanswered, questions regarding this lectin.

**Perturbations of the nuclear CBP35–RNP complex**

Analysis of CBP35 in the nucleoplasm suggested its association with hnRNP and snRNP complexes. Inasmuch as these complexes participate in the formation of the spliceosome and play a role in the processing, packaging and transport of precursor RNA in the nucleus to functional mRNA in the cytoplasm, a cell-free assay for splicing offered the opportunity to test for effects of saccharides and antibodies that bind to CBP35. The pre-mRNA substrate used was an SP6 transcript generated from a plasmid containing the human β-globin gene. Nuclear extract, prepared from HeLa cells [22], was incubated with the pre-mRNA at 30°C for a 60 min reaction. The RNA components of the splicing mixture were extracted and resolved on a 10% acrylamide/8.3 M urea gel system [23]. This gel system resolved the pre-mRNA substrate (exon 1 and exon 2 separated by an intervening sequence), the completed mRNA product (exon 1 spliced onto exon 2), the lariat and debranched intervening sequence, as well as both of the intermediates (free exon 1 and lariat–exon 2).

Using the assay, it was found that saccharide ligands of CBP35 (lactose, thiodigalactoside and an A-tetrasaccharide human serum albumin conjugate) inhibited the formation of the spliced mRNA product (Table 1). In contrast, control reagents that did not bind to CBP35 (sucrose, glucose, BSA) failed to yield the same effect. Moreover a rat monoclonal antibody reactive with CBP35 (anti-Mac-2) also showed inhibition, whereas an isotype-matched control rat monoclonal antibody (anti-transferrin receptor) did not. These results provoke the intriguing possibility that CBP35 may play a role in the processing of mRNA precursors in the nucleus.

**Table 1. Inhibition of splicing in a cell-free assay**

<table>
<thead>
<tr>
<th>Reagent</th>
<th>IC₅₀ (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-tetrasaccharide (serum albumin conjugate)</td>
<td>&lt;0.17</td>
</tr>
<tr>
<td>Lactose</td>
<td>55</td>
</tr>
<tr>
<td>Thiodigalactoside</td>
<td>90</td>
</tr>
<tr>
<td>N-Acetyl-D-glucosamine</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Glucose</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Inositol</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Mannose</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Ribose</td>
<td>&gt;150</td>
</tr>
<tr>
<td>Sucrose</td>
<td>&gt;150</td>
</tr>
</tbody>
</table>

The rigorous tests of this hypothesis, depletion of CBP35 from the nuclear extract with concomitant loss of splicing activity and reconstitution of the activity upon addition of recombinant CBP35, remain to be performed.

It should be emphasized that even with such a demonstration of CBP35 involvement in RNA splicing, the perturbation effects of lactose, thiodigalactoside, and A-tetrasaccharide do not necessarily imply a role for carbohydrate recognition in the process. For one thing, we have, as yet, to identify a carbohydrate ligand with which the nuclear lectin will complex. For another, previous experiments have failed to show that saccharides such as lactose can release CBP35 from the RNP complex in the nucleus [17]. Therefore, it remains as a challenge to demonstrate that the CRD of CBP35 is utilized in any physiological activity of the protein within the cell.

This work was supported by grants GM-38740 and GM-27203 from the National Institutes of Health. We thank Mrs Linda Lang for her help in the preparation of the manuscript.

Volume 20

Blood group-related oligosaccharides are ligands in cell-adhesion events*  
Ten Feizi

Glycoconjugates Section, M.R.C. Clinical Research Centre, Watford Road, Harrow HA1 3UJ, Middlesex, U.K.

Introduction

Understanding the biological roles of the diverse oligosaccharides of glycoproteins and glycolipids has been a major challenge in cell biology. Observations with monoclonal antibodies, that there are temporal and positional patterns in the display of specific oligosaccharide antigens at the surface of cells and in extracellular matrices during stages of embryonic development and cellular differentiation, have raised the possibility [1-4] that such carbohydrate differentiation antigens may be recognition structures for proteins (endogenous lectins) which determine the way cells migrate or respond in various microenvironments. Prominent among such oligosaccharides are a large family which include the major blood group antigens. Moreover, the transient expression of blood group antigens in certain organs of the developing human foetus had been documented in earlier studies using polyclonal antibodies [5].

The structures of the blood group antigens were elucidated by pioneering work in the laboratory of Kabat in New York [6] and that of Morgan and Watkins in London [7, 8], predominantly using glycoproteins from ovarian cyst fluids as an abundant source of these antigens. Shown in Fig. 1 is a composite structure proposed in 1968 ([9, 10] by the guest of honour at the Host Colloquium of this 641st Meeting of the Biochemical Society, Elvin A. Kabat). Depicted here are the interrelationships of (a) the blood group A, B and H antigenic determinants, based both on Type-1, Galβ1-3GlcNAc- and Type-2, Galβ1-4GlcNAc- backbones, (b) the Lea and Leb antigens based on Type-1 backbones, and (c) the products of another (then new) gene that gives rise to mono- and difucosyl isoformers of the Lea

A composite structure proposed by Kabat depicting the interrelationship of various blood group determinants

Adapted from [51].

Abbreviations used: SSEA-1, mouse stage-specific embryonic antigen; CHO, Chinese hamster ovary; CHO-ELAM-1, CHO cells transfected to express full-length E-selectin.

* Dedicated to Elvin A. Kabat.