ELSEVIER

FEBS Letters 337 (1994) 114–118

FEBS 13509

FEBS LETTERS

The human HIP gene, overexpressed in primary liver cancer encodes for a C-type carbohydrate binding protein with lactose binding activity

Laurence Christa[a,*], Murielle Felin[b], Olivier Morali[a], Marie-Thérèse Simon[a], Chantal Lasserre[a], Christian Brechot[a], Annie-Pierre Sève[b]

[a]*INSERM U370 and Liver Unit, CHU Necker, 156 rue de Vaugirard, 75742 Paris Cedex 15, France*
[b]*INSERM U180, 45 rue des Saints-Pères, 75270 Paris, Cedex 06, France*

Received 29 November 1993

Abstract

HIP was originally identified as a gene expressed in primary liver cancers, and in normal tissues such as pancreas and small intestine. Based on gene data base homologies, the HIP protein should consist of a signal peptide linked to a single carbohydrate recognition domain. To test this hypothesis HIP and the putative carbohydrate recognition domain encoded by the last 138 C-terminal amino acids, were expressed as glutathione-*S*-transferase proteins (GST-HIP and GST-HIP-142, respectively). Both recombinant proteins were purified by a single affinity purification step from bacterial lysates and their ability to bind saccharides coupled to trisacryl GF 2000M were tested. Our results show that HIP and HIP-142 proteins bind to lactose, moreover the binding requires divalent cations. Thus the HIP protein is a lactose-binding lectin with the characteristics of a C-type carbohydrate recognition domain of 138 amino acids in the C-terminal region.

Key words: Primary liver cancer; Lectin; Lactose binding protein; Fusion protein; HIP

1. Introduction

Differential screening of a human hepatocellular carcinoma complementary DNA library using subtracted probes allowed us to identify a gene, named HIP, which was abundantly expressed (at the mRNA level) in 7 of 29 primary liver cancers (PLC) and not in non-tumorous liver tissues [1]. HIP mRNA expression is tissue specific since it is present in the normal pancreas and small intestine while it is not detectable in normal adult and fetal liver, colon, brain, kidney and lung. The derived amino acid sequence of HIP has 49% homology with the human reg protein. Reg is thought to be a pancreatic islet regenerating factor and is identical to the pancreatic stone (PSP) or pancreatic thread (PTP) proteins. Human HIP also has 68.5% homology with the bovine pancreatic thread protein (BPTP) which is in fact the bovine homologue of the HIP gene [1]. Finally, Orelle et al. [2] recently reported the overexpression of the human pancreatic associated protein (PAP-H) during acute pancreatitis, and PAP-H is in fact the same as HIP according to published cDNA sequences.

Detailed analysis of the predicted structure of the HIP protein strongly suggested that HIP (or PAP-H), BPTP and PSP/PTP/reg should define a new family sharing the carbohydrate recognition domain (CRD) of the calcium dependent C-type lectins with their structural characteristics.

Recently, with the increase in data available on primary structures of lectins, a new classification for lectins has become possible based on evolutionary characteristics. Drickamer was the first to organize animal lectins into two categories, so-called C- and S-lectins [3]. C-type lectins constitute a family which has the common property of only being able to ligate sugars in the presence of calcium. Sequence alignements of more than 50 CRDs revealed the presence of a common sequence motif of approximatively 130 amino acids long, consisting of a pattern of 14 invariant residues (including 4 invariable half-cystines which form two intrachain disulfide bridges) and 18 conserved residues. C-type CRDs have specificities for various different saccharides and are associated with a variety of effector domains in a large number of proteins.

The HIP CRD presents 10/14 invariant (including the 4 cysteines) and 13/18 highly conserved residues (Fig. 1). In contrast to the others C-type lectins, the predicted structures of HIP and reg related proteins are that they consist of a single CRD linked to a signal peptide and should represent the group VII in the Drickamer's classification of C-type lectins [4].

* Corresponding author. Fax: (33) (1) 4061-5581.

Abbreviations: PLC, primary liver cancer; GST, glutathione-*S*-transferase; Lac, β-D-lactopyranosyl; CRD, carbohydrate recognition domain.

0014-5793/94/$7.00 © 1994 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(93)E1426-M

In view of the potential importance of the HIP protein in liver cell differentiation and/or proliferation, the aim of the present study was to demonstrate that the HIP protein is really a carbohydrate binding protein, as suggested by nucleotide sequence analysis. The synthesis of the human HIP protein in *E. coli*, its purification and characterization as a C-type lectin are reported.

2. Materials and methods

2.1. Expression of proteins in Escherichia coli

Expression of the recombinant HIP protein was performed using the pGex vector (Pharmacia) in DH5α™ *E. coli* (Gibco BRL/Life Technologies, Gaithersburg, MD, USA). The complete HIP cDNA and the cDNA encoding the last 138 amino acids of the C-terminal domain, which potentially corresponds to the carbohydrate-recognition region, were expressed as glutathione-*S*-transferase proteins (GST-HIP and GST-HIP-142, respectively).

DNA fragments containing a *Sma*I site 5′ to the start codon and an *Eco*RI site downstream of the stop codon were generated by PCR using cDNA as template. The primer sequence were:

5′CGCCCCGGGATGCTGCCTCCCATGGCCCTG3′

5′CGCGAATTCCTAGTCAGTGAACTTGCAGACATA3′

PGex-3x and amplified HIP cDNA were digested with *Sma*I and *Eco*RI, and ligated with T4 DNA ligase (Pharmacia). For expression of HIP-142, HIP cDNA was digested with *Nsp*BII and *Eco*RI to clone a 412nt cDNA fragment into the PGex-2T vector. DNA sequencing confirmed the correct orientation, the restored open reading frame of the fusion construct and the coding sequences.

Transformed DH5α™ *E. coli* were cultured in one litre of LB medium at 37°C until the OD_{600} reached 0.7. Isopropyl-1-thio-β-D-galactoside (IPTG) (Boehringer-Mannheim) was then added to a final concentration of 0.1 mM, and the induced cells were grown for an additional 3 h. The cells were then centrifuged (7000 × *g*), washed once with PBS (pH 7.4) and frozen at −80°C.

Subsequent protein purification steps were typically done after resuspension of the bacterial pellet in 20 ml of PBS containing EDTA (2 mM), DTT (5 mM) and Triton X-100 (1%). Cells were then lysed by mild sonication and centrifuged (10,000 × *g*, 30 min). Fifty μg of fusion protein was then purified from the supernatant by affinity chromatography (using glutathione Sepharose 4B and elution with glutathione (12.5 mM) in Tris buffer (50 mM, pH 8.00)).

It is still possible to recover HIP protein (50 μg/l), from the pellet, by guanidine solubilization: the pellet from 250 ml was solubilized in 1 ml of 6 M guanidine and centrifuged 13,000 × *g* for 30 min at +4°C. The supernatant was diluted drop by drop in 50 ml of cold Tris-HCl, 50 mM pH 8.00, 25 mM $CaCl_2$ buffer. Renaturation was performed overnight at +4°C and denatured proteins were removed by centrifugation (5000 × *g*). Supernatant was directly used for affinity chromatography.

2.2. Production of HIP antibodies

A peptide situated in the central loop of the protein (YVWIGLHDPTQGTEPNGEG named 316) was chemically synthesized (Laboratoire de Microchimie, Institut Gustave Roussy, Villejuif, France), coupled to bis-diazotized benzidine and used as an immunogen. A New-Zealand rabbit was injected subcutaneously with 150 μg of the immunogen (in 1 ml mixed with 1 ml of CFA) five times at 10–15 days intervals. Six days after the last injection, blood was collected by cardiac puncture.

2.3. Affinity chromatography

Affinity chromatography was performed according to the procedures described by Sève et al. [5], using the following sugars immobilized on Trisacryl GF 2000M: α-D-glucopyranosyl, α-L-fucose, α-D-galactopyranosyl, *N*-acetyl-β-D-glucosamine, and β-D-lactopyranosyl (Lac). Total bacterial extract containing GST-HIP protein (the extraction was performed with PBS and Triton X-100 (1%) without EDTA and DTT) was incubated overnight at 4°C with 1 ml sugar-Trisacryl GF 2000M under batch conditions, in the presence of 1 mM $CaCl_2$ and 1 mM $MgCl_2$. The mixture was then packed into 2 ml disposable syringe and the flow-through material was immediately collected. The column was successively washed with 60 ml complete PBS, and then with 5 ml complete PBS containing 0.2 M of the sugar homologous to that immobilized on Trisacryl GF 2000M. Heterologous elution was performed for Lac-column by 5 ml of 0.2 M α-L-fucose followed by 5 ml α-D-glucopyranosyl and finally by 5 ml of Lac.

Two ml of purified GST-HIP, GST-HIP-142 or GST proteins (50 μg), in 50 mM Tris buffer, pH 8.00, containing 12.5 mM glutathione, were incubated overnight at 4°C with 1 ml of Lac-Trisacryl under batch conditions, in the presence or in the absence of 1 mM $CaCl_2$ and 1 mM $MgCl_2$. In the presence of divalent cations, the elution was performed by 0.2 M Lac containing the divalent cations and then by 1 M NaCl in order to test that no material remained bound to the column. If the binding to Lac-Trisacryl was done without divalent cations, the elution was performed first with 0.2 M Lac and then with 1 M NaCl.

2.4. Electrophoresis

Proteins were resolved by SDS-PAGE on 12.5% acrylamide mini-slab running gels under denaturing conditions according to Laemmli's procedure. The polypeptide bands were visualized by silver staining. Marker proteins was from Biorad.

2.5. Immunodetection by Western blot

Following electrophoresis and transfer to Immobilon-P paper as previously described [5], the filters were incubated with the antiserum against peptide 316, diluted to 1/5000. Protein was visualized using alkaline phosphatase-conjugated mouse anti-rabbit monoclonal antibodies as previously described [5]. Preimmune serum from the same rabbit was used as the negative control

3. Results

3.1. Expression and purification of HIP and HIP-142 proteins as GST fusion proteins

A cDNA fragment (coding for the total coding region) was used to produce the entire HIP protein including 3 putative domains: a peptide signal of 22 amino acids, a propeptide of 15 amino acids and a 138 amino acids sequence of the COOH-terminal (Fig. 1). The latter segment was chosen for expression in a separate vector to demonstrate that it corresponds to the functional CRD of HIP protein (= HIP-142). This fragment includes 10/14 invariant amino acids (including the 4 half-cystines that form two intrachain disulfide bridges in all C-type CRDs) and 13/18 highly conserved residues (Fig. 1). After purification on glutathione-Sepharose 4B and electrophoresis on SDS-PAGE, HIP and HIP-142 expressed as GST fusion proteins, were visualized by silver staining (Fig. 2A) and immunoblotting analysis (Fig. 2B), as single polypeptides of 46 and 42 kDa, respectively.

3.2 Binding of sugars to GST-HIP and GST-HIP-142

Total soluble bacterial extract containing HIP protein was incubated overnight at 4°C with 1 ml of α-D-glucopyranosyl-, α-L-fucose-, α-D-galactopyranosyl-, *N*-acetyl-β-D-glucosamine-, and β-D-lactopyranosyl-Trisacryl GF 2000M in the presence of 1 mM $CaCl_2$ and 1 mM $MgCl_2$. After electrophoresis, silver staining (Fig. 3-1) and Western blotting (Fig. 3-a) showed that GST-HIP protein binds only to Lac-Trisacryl and was eluted by

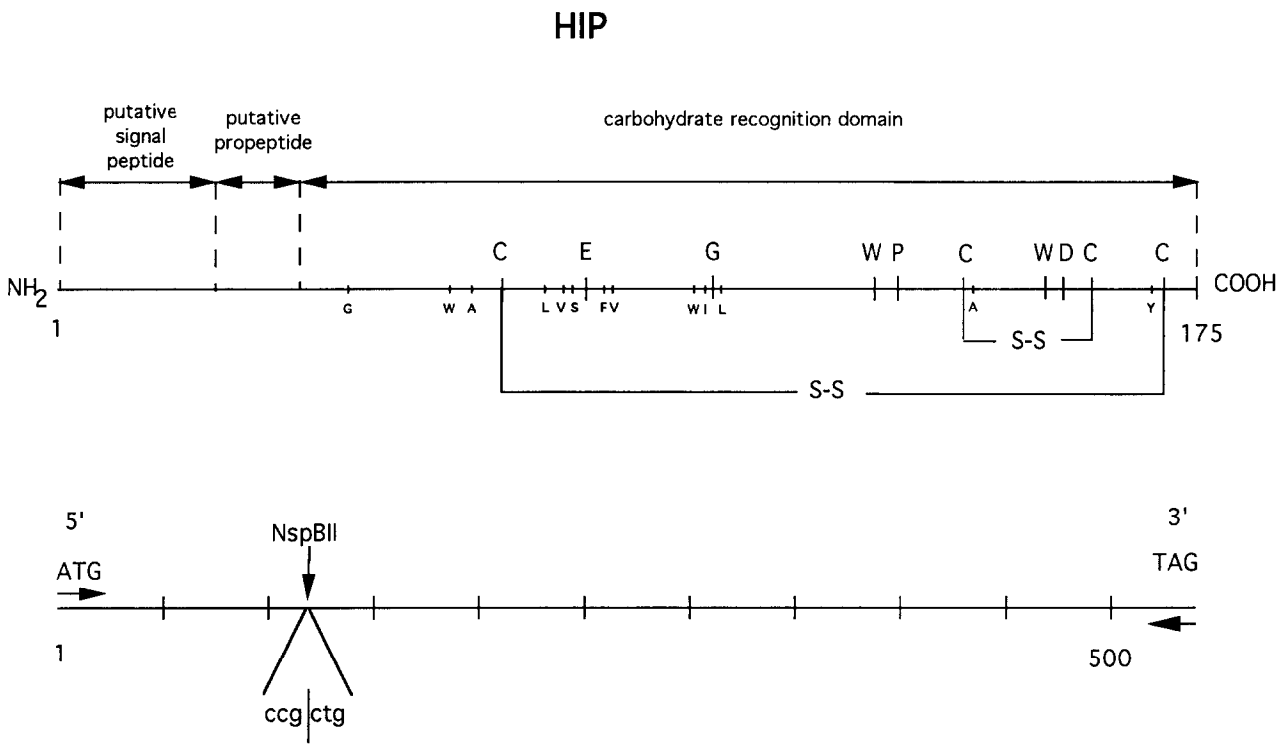

Fig. 1. Schematic map of HIP protein (above) and the corresponding cDNA. Large and small capitals indicated invariant and highly conserved residues respectively in the C-type lectin domain [4]. Disulfide bonds are indicated as –S–S–. Putative signal peptide and propeptide are deduced from gene data base homologies. Horizontal small arrows indicate the sites of the primers used to amplify the cDNA

Lac 0.2 M. After electrophoresis, silver staining showed that the lactose elution was complete: the addition of 1 M NaCl did not elute further proteins (Fig. 3-2). To verify the HIP specificity for Lac, heterologous elution of Lac-columns was led first by α-L-fucose, followed by α-D-glucopyranosyl and finally by Lac. No material was eluted with α-L-fucose, and α-D-glucopyranosyl, and GST-HIP was specifically eluted by Lac (Fig. 4).

In an attempt to determine which part of the protein is the CRD, total soluble bacterial extract containing GST-HIP-142 protein was tested for its ability to bind lactose in the presence of divalent cations. After electrophoresis, silver staining showed that GST-HIP-142 binds to Lac-Trisacryl (Fig. 3-3) and the elution is complete with 0.2 M Lac (Fig. 3-4). Thus, HIP-142 and HIP have the same Lac binding activity.

Purified GST was incubated with Lac-Trisacryl to test the non specific binding. Silver staining of SDS-PAGE showed that GST was completely recovered in the flow-through fraction of the Lac-column and no material was eluted by Lac elution (Fig. 5).

C-Type lectins constitute a family which have the common property of ligating sugars in the presence of divalent cations only. Thus, the binding of purified HIP to Lac columns in the presence or in the absence of divalent cations was tested: 100 μg of purified HIP protein was divided into two fractions to be loaded on two Lac columns with or without 1 mM $CaCl_2$ and 1 mM $MgCl_2$. Silver staining (Fig. 6, A-1, A-3) and immunoblot analysis (Fig. 6, B-1, B-3) showed that HIP protein bind to Lac-columns only in the presence of divalent cations. The elution is complete with buffer containing 0.2 M Lac plus the divalent cations (Fig. 6, A-2, B-2).

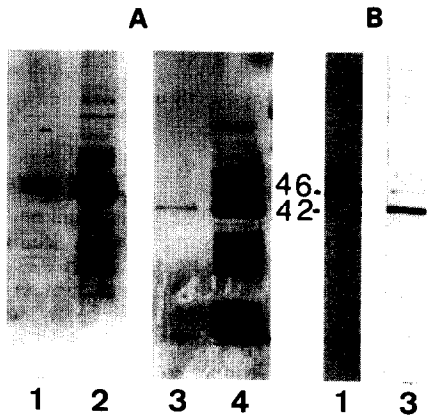

Fig. 2. Glutathione-Sepharose purification of GST-HIP and GST-HIP-142. Proteins were resolved on 12.5% acrylamide gels and either stained with silver (A) or analysed by immunoblotting (B). Tracks: soluble bacterial extracts producing (2) GST-HIP and (4) GST-HIP-142; (1) purified GST-HIP and (3) purified GST-HIP-142.

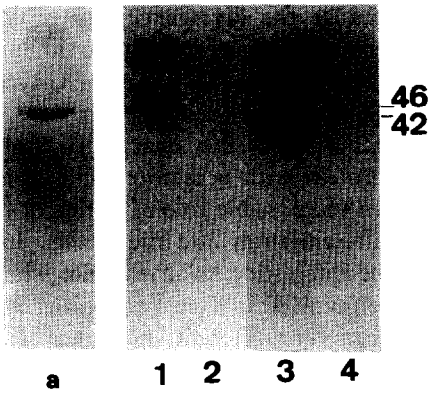

Fig. 3. Lac affinity chromatography of bacterial extracts. Proteins are specifically eluted by 0.2 M Lac: (lane a) immunoblotting analysis of GST-HIP; (lanes 1 and 3) silver staining of electrophoresed GST-HIP and GST-HIP-142, respectively. Washing with NaCl (1 M), after Lac elution: (lanes 2 and 4) for GST-HIP and GST-HIP-142, respectively.

4. Discussion

Our results demonstrate that human HIP cDNA, recently isolated in primary liver cancer [1], does encode a C-type carbohydrate binding protein. This result is in agreement with the deductions from the amino acid sequence and further points to the potential role of lectins in liver carcinogenesis.

Data bank homologies have shown that HIP protein should consist of a N-terminal signal peptide of 22 amino acids, plus a propeptide of 15 amino acids followed by a single CRD of 138 amino acids which is similar to the C-type CRD described by Drickamer [4]. Thus, two expression vectors were constructed to direct the synthesis of either the entire HIP protein, or the putative CRD of the HIP protein (HIP-142). Results from affinity chromatography indeed indicate that HIP is a divalent cation dependent lactose-binding protein, with the characteristics of a C-type CRD corresponding to the 138 C-terminal amino acid (HIP-142). Thus HIP protein can be classified as belonging to the C-type lectin family.

HIP protein had a greater affinity for Lac than for galactose which is the same general specificity of other S-Lac lectins, suggesting that all of them bind to one side of the lactose molecule involving parts of the galactose and glucose residues. Accordingly, CBP35 can distinguish the galactose from the Lac with an affinity of Lac for CBP35 60–100 times higher than that of galactose [6].

In addition, our results showed that HIP protein can be purified from bacterial extracts with lactose-Trisacryl affinity chromatography.

In vertebrates, HIP is the smallest protein (19 kDa) reported among the C-type lectins and comprises a single CRD linked to a signal peptide. The other C-type carbohydrate-recognition proteins may have several sugar combining sites on different [7] or on the same polypeptide chain [8]. Their CRDs are found in association with a variety of other protein domains which confer the specific functions on the lectins, e.g. those implied in endocytosis of glycoproteins [9], cell adhesion and metastasis [9], reduction of lung surface tension [10], bacterial defences [7]. In contrast, HIP protein does not have a second functional domain. That is apart from the signal peptide which would be involved in secretion of the protein, and is not required for binding to lactose. In this respect, Orelle et al. [2] showed that PAP-H (or HIP) is secreted in serum of patients with acute pancreatitis and suggested that PAP-H protein might be an acute phase protein. However, as it comprises a single CRD linked

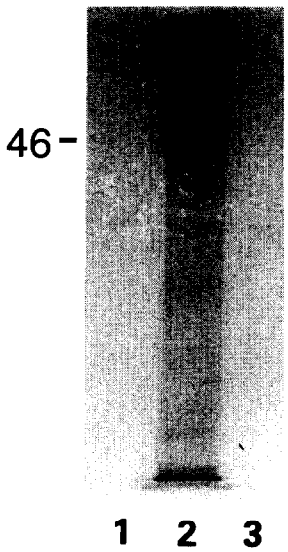

Fig. 4. Lac affinity chromatography of bacterial extracts followed by heterologous and homologous sugars elution and analysis by silver staining SDS-PAGE. Elution was successsively performed by (lanel) 0.2 M α-L-fucose, (lane 3) α-D-glucopyranosyl and (lane 2) 0.2 M Lac.

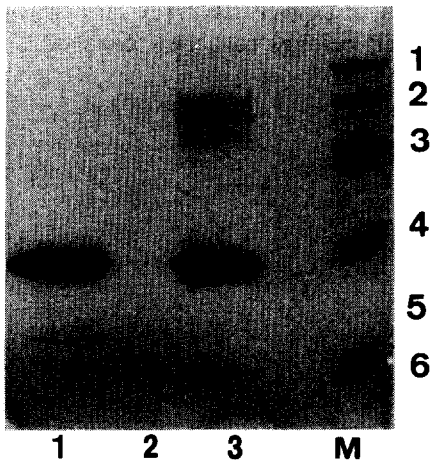

Fig. 5. Lac affinity chromatography of purified glutathione-Sepharose GST expressed in *E. coli* from pGEX-2T. Silver staining after electrophoresis of: (lane 1) control GST before Lac-column; (lane 2) Lac elution, (lane 3) GST excluded from the Lac-column by washing with PBS. M = marker proteins: 1 = M_r 97,400; 2 = 66,200; 3 = 45,000; 4 = 31,000; 5 = 21,500; 6 = 14,400.

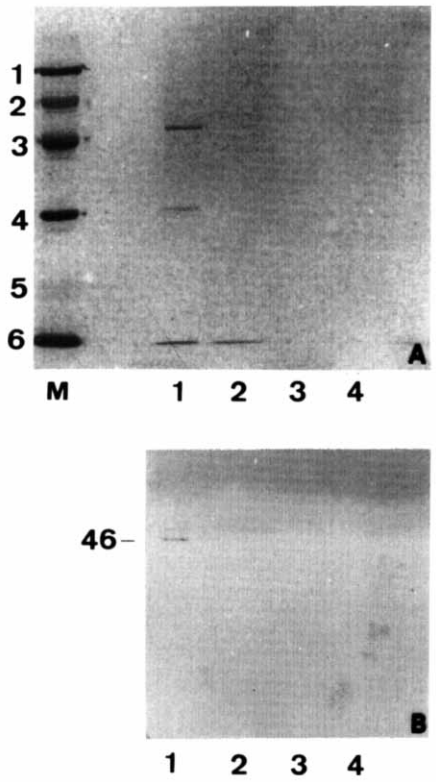

Fig. 6. Lac affinity chromatography of purified glutathione-Sepharose GST-HIP. (A) Silver staining of acrylamide gels. (B) Immunoblotting analysis. Binding to Lac-Trisacryl GF 2000M and elution by 0.2 M Lac in the presence (lane 1) or in the absence (lane 3) of divalent cations. Washing with NaCl (1 M) after Lac elution (lane 2 and 4 in the presence or in the absence of divalent cations, respectively).

to a signal peptide, HIP is closely related to invertebrate lectins which are important in the development of these organisms, e.g. sea-urchin lectin, acorn barnacle lectin, tunicate lectin, or flesh-fly lectin [1]. Therefore, it is plausible that HIP belongs to a new class of C-lectins in vertebrates because the divergence between HIP and PSP/PTP/reg is in an order of magnitude similar between to that found in the other subclasses of C-types lectins described by Drickamer.

The purification and the characterization of the human HIP protein with a lactose binding activity is important given the possible role of lectins in cell to cell interaction, differentiation and metastasis. In this respect, it will be interesting to investigate whether HIP protein may be involved in liver, pancreatic, and intestinal cell proliferation or differentiation since HIP mRNA is specifically expressed in three tissues derived from common embryonic origin, namely pancreas, small intestine and liver.

Acknowledgements: We thank Frederic Troalen and Dominique Bellet for the synthetic peptide, Marie-Agnes Doyennette-Moyne for preparation and use of sugars-affinity beads. This work was supported by INSERM, ARC, LNC, Fondation de France.

References

[1] Lasserre, C., Christa, L., Simon, M.T., Vernier, P. and Bréchot, C. (1992) Cancer Res. 52, 5089–5095.
[2] Orelle, B., Keim, V., Masciotra, L., Dagorn, J.C. and Iovanna, J.L. (1992) J. Clin. Invest. 90, 2284–2291.
[3] Drickamer, K. (1988) J. Biol. Chem. 263, 9557–9560.
[4] Drickamer, K. (1993) Current Opinions in Structural Biology 3, 393–400.
[5] Sève, A.P., Felin, M., Doyennette-Moyne, M.A., Sahraoui, T., Aubery, M. and Hubert, J. (1993) Glycobiology 3, 23–30.
[6] Wang, J.L., Laing, J.G. and Anderson, R.L. (1991) Glycobiology 1, 243–252.
[7] Weis, I.E., Drickamer, K. and Hendrickson, W.A. (1992) Nature 360, 127–134.
[8] Taylor, M.E. and Drickamer, K. (1993) J. Biol. Chem. 268, 399–404.
[9] Springer, T.A. and Lasky, L.A. (1991) Nature 349, 196–197.
[10] Childs, R.A., Wright, J.R., Ross, G.F., Yuen, C.T., Lawson, A.M., Chai, W., Drickamer, K. and Feizi, T. (1992) J. Biol. Chem. 267, 9972–9979.