Nucleolin Undergoes Partial N- and O-Glycosylations in the Extranuclear Cell Compartment†

Mathieu Carpentier, Willy Morelle, Bernadette Coddeville, Alexandre Pons, Maryse Masson, Joël Mazurier, and Dominique Legrand*

Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche N8576 du Centre National de la Recherche Scientifique, Institut Fédératif de Recherche 118, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France

Received October 8, 2004; Revised Manuscript Received February 17, 2005

ABSTRACT: Nucleolin is an ubiquitous, nonhistone nucleolar phosphoprotein involved in fundamental aspects of transcription regulation, cell proliferation, and growth. Nucleolin was primarily found in the nucleus, but it was also proposed as a possible shuttle between the nucleus, cytoplasm, and cell membrane. We report here that part of the extranuclear nucleolin undergoes complex N- and O-glycosylations. A band with higher molecular mass (113 kDa) than the 105-kDa classical major nucleolin band was detected on SDS−PAGE gel that cross-reacted with specific anti-nucleolin antibodies and was identified as a nucleolin isoform by mass spectrometry. The presence of *N*-glycans was first suggested by sensibility of the 113-kDa nucleolin isoform to tunicamycin treatment. Determination of monosaccharide composition by heptafluorobutyrate derivation followed by gas-chromatography mass spectrometry indicated the presence of *N*- and *O*-glycans. The structures of *N*- and *O*-glycans were first investigated using specificity of binding to lectins. This approach allowed a partial characterization of *N*-glycan structures and revealed *O*-glycan structures that could otherwise go unnoticed. Further study of *N*-glycans by mass spectrometry using direct exoglycosidase treatment on MALDI−TOF target allowed the complete definition of their structures. Finally, the use of peptide mass fingerprinting with sinapinic acid allowed identification of N317 and N492 as the two N-glycosylation sites. N317 and N492 belong to RNA-binding domains 1 and 3 of nucleolin, respectively, that suggests a role of glycosylation in regulating the function of the protein.

Nucleolin (C23)[1] is an ubiquitous, nonhistone nucleolar phosphoprotein of exponentially growing eukaryotic cells involved in fundamental aspects of transcription regulation, cell proliferation, and growth (*1*−*3*). The protein consists of three domains through which it controls the organization of nucleolar chromatin, packaging of pre-RNA, rDNA transcription, and ribosome assembly. The negatively charged N-t domain controls rDNA transcription by inducing nucleolar chromatin decondensation through ionic interaction of the acidic amino acids with histone H1 and binding to nontranscribed spacer regions in DNA that separate the rDNA repeats (*4*). The central domain contains four ∼80 amino acid RNA-binding domains (RBDs), which specifically interact with nascent 47S rRNA, thus leading to the recruitment of other factors, such as U3 snoRNP, required for the cleavage of RNA into mature 18S, 28S, and 5.8S rRNA species. The C-terminal domain containing methylated RGG repeats controls unstacking of bases and the unfolding of the RNA secondary structure but also mediates interactions with ribosomal proteins (*5*). Such interactions suggest that nucleolin may play a role in the import of ribosomal proteins to the nucleus and the assembly and the export of ribosomal subunits to the cytosol. In addition, recent evidence of nucleolin at the surface of cells suggests that the molecule is a shuttling receptor between the cell surface and nucleus (*6*−*8*). The binding and nuclear targeting of several ligands such as the anti-HIV cytokine midkine (*9*, *6*) and lactoferrin (*10*) have been recently described. Hence, nucleolin was proposed as a mediator for the extracellular regulation of nuclear events (*3*). Except for the presence of a nuclear localization sequence (NLS) upstream from the first RBD (*11*), the structural features that govern the putative shuttling activity of nucleolin in the cell are not well-defined. Post-

† This investigation was supported in part by the Centre National de la Recherche Scientifique (UMR N8576; Glycobiologie Structurale et Fonctionnelle; Director, Dr. J. C. Michalski, Université des Sciences et Technologies de Lille, France) and the Institut Fédératif de Recherche 118 (IFR 118; Modifications Post-traductionnelles, Glycobiologie et Protéomique; Director, Dr. J. Mazurier). The mass spectrometry facility used in this study was funded by the European Community (FEDER), the région Nord-Pas de Calais (France), the CNRS, and the Université des Sciences et Technologies de Lille.

* To whom correspondence should be addressed: Unité de Glycobiologie Structurale et Fonctionnelle, UMR N8576 du CNRS, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq Cedex, France. Telephone: 33 3 20 33 72 38. Fax: 33 3 20 43 65 55. E-mail: dominique.legrand@univ-lille1.fr.

[1] Abbreviations: C-23, nucleolin; ConA, concanavalin A; DHB, dihydroxybenzoic acid; DSA, *Datura stramonium* agglutinin; GC/MS, gas chromatography/mass spectrometry; GNA, *Galanthus nivalis* agglutinin; HMG, high-mobility group protein; MAA, *Maackia amurensis* agglutinin; MS, mass spectrometry; Neuase, neuraminidase; NLS, nuclear localization sequence; PBS, phosphate-buffered saline; PNA, peanut agglutinin; PNGase F, peptide *N*-glycosidase F from *Flavobacterium meningosepticum*; RBD, RNA-binding domain; RPMI-FCS, RPMI 1640 containing 10% (v/v) heat-inactivated fetal calf serum; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; SNA, *Sambucus nigra* agglutinin; TBS, Tris-buffered saline; WGA, wheat germ agglutinin.

10.1021/bi047831s CCC: $30.25
Published on Web 03/24/2005

translational modifications of nucleolin such as phosphorylation with CKII, cdc2, and PKC-ξ kinases, arginine methylation, and autodegradation probably play critical roles in the function and traffic of the protein (*12*−*15*), but no glycosylation was clearly evidenced.

Many functions of particular glycosylations have been reported. These functions include modulation of enzyme and hormone activity, regulation of intracellular traffic, control of protein folding, ligand recognition, and cell−cell interactions (*16*−*18*). In the particular case of O-N-acetylglucosaminylation which is, together with phosphorylation, a common dynamic regulatory event of key cytoplasmic/nuclear proteins, a specific glycosyl transferase was evidenced in the cytosol and nucleus (*19*). Complex N- and/or O-glycosylations of nuclear and cytoplasmic proteins were also described for a set of proteins (*20*−*24*). For example, N-glycosylation was detected on the α subunit of the sodium pump (Na^+,K^+-ATPase) from dog kidney and on the so-called high-mobility group (HMG) proteins, although HMG proteins have no consensus sites for N-glycosylation (*21*, *22*). Even though these studies are widely cited as evidence for cytoplasmic and nuclear *N*-glycans, they also suffer from a lack of definitive structural data. More interestingly, complex N-glycosylation of the NFIC transcription factor during early mammary gland involution was recently reported (*24*).

In the present paper, we report that part of the extranuclear nucleolin undergoes complex O- and N-glycosylations. In the extranuclear fraction, a band with higher molecular mass (113 kDa) than the 105-kDa major nucleolin band was detected on sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE). The 113-kDa band cross-reacted with specific anti-nucleolin antibodies and was identified as a nucleolin isoform by mass spectrometry. Use of classical approachs such as endoglycosidase treatment or lectin screening and of up-to-date techniques such as gas chromatography/mass spectrometry (GC/MS) or mass spectrometry (MS) analysis allowed us to identify two *O*- and two *N*-glycan structures. Residues N317 and N492 were identified as the two glycosylation sites.

EXPERIMENTAL PROCEDURES

Extraction and Purification of Nucleolin from Jurkat Cells. Jurkat cells were grown at 37 °C in a humidified atmosphere of 95% air and 5% CO_2 in RPMI 1640 containing 10% (v/v) heat-inactivated fetal calf serum (RPMI-FCS). A total of $1.5-1.8 \times 10^9$ Jurkat cells grown in the exponential phase (600 000 cells/mL) were centrifuged at 900*g* for 15 min. Pellets were pooled and washed twice with 25 mL of phosphate-buffered saline (PBS). Pellet was resuspended in 25 mL of buffer E [20 mM Tris/HCl at pH 7.6, 150 mM NaCl, 7.5 mM $MgCl_2$, 5 mM β-mercaptoethanol, 0.5% Triton X-100, and 1 mM Pefabloc (4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride), Roche, Basel, Switzerland] and incubated on ice for 15 min. Nuclei were pelleted by centrifugation at 1000*g* for 15 min. The collected supernatant refers to the extranuclear fraction. Nuclei were washed once with buffer E and resuspended in buffer I (20 mM Tris/HCl at pH 7.6, 50 mM KCl, 400 mM NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol, 1% Triton X-100, 20% glycerol, and 1 mM Pefabloc) on ice for 15 min. The mixture was centrifuged at 1000*g* for 15 min. The supernatant was collected and refers to the nuclear fraction. Remaining insoluble material was removed from both nuclear and extranuclear fractions by centrifugation at 12000*g* for 20 min. Antibodies directed against nucleoporin p62 (nucleus marker) and actin (cytoplasm marker) were used on Western-blotted aliquots to control the identity of the nuclear and extranuclear fractions, respectively, and the absence of cross-contamination (not shown). Both fractions were either used immediately or stored at −80 °C for less than 3 months.

A rapid two-step chromatography procedure was used to purify nucleolin from nuclear and extranuclear fractions (*10*). All steps were performed at 4 °C using ice-cold columns and buffers in the presence of 1 mM Pefabloc and complete protease inhibitor cocktail (Roche). The nuclear or extranuclear extract of Jurkat cells (25 mL) was diluted 10-fold with 20 mM sodium phosphate at pH 7.0 and passed through a 23 mL DEAE-Sepharose Fast Flow column (Amersham Pharmacia Biotech, Uppsala, Sweden). After the column was washed with 600 mL of 20 mM sodium phosphate at pH 7.0, elution of the adsorbed proteins was performed with 45 mL of the same buffer containing 1 M NaCl. The eluant was diluted 10-fold with 50 mM Tris/HCl at pH 7.9, 5 mM $MgCl_2$, 0.1 mM EDTA, and 1 mM β-mercaptoethanol (buffer A) and loaded onto a 4 mL Heparin-Sepharose 6 Fast flow column (Amersham Pharmacia Biotech) equilibrated with the same buffer. The gel was washed with 80 mL of buffer A containing 0.2 M ammonium sulfate, and proteins were eluted in 200 μL fractions with 10 mL of PBS containing 1 M NaCl. The presence of nucleolin in each fraction was checked by SDS−PAGE on 4−15% polyacrylamide Phastgels (Amersham Pharmacia Biotech) stained with Coomassie blue and immunorevealed with rabbit anti-nucleolin polyclonal antibodies. Nucleolin-containing fractions were pooled and dialyzed against PBS containing 1 mM Pefabloc at 4 °C for 2 h before storage at −80 °C. Storage did not exceed 1 month to avoid degradation, presumably autoproteolysis (*8*, *15*, *25*).

Identification of Nucleolin Isoforms. Purified nuclear and extranuclear nucleolins were loaded on a 7.5% polyacrylamide gel in the presence of SDS. After electrophoresis, proteins were either stained with colloidal blue Coomassie G250 (*26*) or immunoblotted with rabbit anti-nucleolin polyclonal antibodies directed against the C-terminal region of nucleolin (residues 345−706) (*10*) or with monoclonal anti-C-23 antibody (Santa-Cruz Biotechnology Inc., Santa Cruz, CA).

Immunoidentification was further confirmed by MS. Coomassie-stained protein bands were cut from the SDS−PAGE gel and treated according to ref *27*. The pieces of gel were washed with 400 μL of a 100 mM ammonium bicarbonate/acetonitrile 1:1 (v/v) solution. The wash solution was discarded, and the pieces were dried using a Speed Vac concentrator (Eppendorf AG, Hamburg, Germany). Enzymatic cleavage was initiated by reswelling the gel in 50 mM ammonium bicarbonate containing trypsin (Promega, Madison, WI) (20 μg/1.5 mL). After absorption of the protease solution, aliquots of pure water were added sequencially. The digestion was carried out for 12−16 h at 30 °C. The resulting peptides were recovered through extraction with solutions containing acetonitrile/formic acid at 45:10 and 95:5%. The final extract was dried and resuspended in 1 μL of high-

purity water. The peptide-containing solution was mixed with 1 μL of matrix solution [10 mg of 2,5-dihydroxybenzoic acid (DHB)/mL of 0.1% (v/v) trifluoroacetic acid and 30% (v/v) acetonitrile]. After air-drying, MS measurements were realized on a Voyager DE-STR MALDI−TOF instrument (Applied Biosystems, Foster City, CA) in reflectron mode and using an accelerating voltage of 20 kV. Mass analysis was calibrated using the average mass of the three autolysis trypsin fragments at *m*/*z* 842.51, 1045.56, and 2211.10. Proteins were identified according to their peptide mass fingerprint after database searching using Protein Prospector (http://prospector.ucsf.edu).

Characterization of Cell-Surface Nucleolin. Cell-surface biotinylation was performed by the method described in ref *28* but with some modifications. Jurkat cells grown in the exponential phase (600 000 cells/mL) were centrifuged at 900*g* for 15 min and washed 3 times with DPBS. Two pellets containing 35×10^6 cells were obtained, one of which was used as a negative control (no biotinylation). Each pellet was resuspended in 1.3 mL of DPBS and incubated or not with 0.5 mg/mL biotin-sulfo-NHS (Sigma−Aldrich, St Louis, MO) for 1 h at 4 °C. Biotinylation was stopped by replacing the labeling solution with RPMI-FCS. After incubation for 5 min at 4 °C, cells were washed twice with DPBS. Cells were solubilized in 1.3 mL of DPBS at pH 7.4, 1% Triton X-100, 2 mM EDTA, 0.5 mM β-mercaptoethanol, and 1 mM Pefabloc for 20 min on ice with shaking. Cellular fragments were eliminated by centrifugation at 15000*g* for 15 min. The 1.3 mL supernatants were incubated with 70 μL avidin-Agarose (Sigma) for 4 h at 4 °C with gentle shaking. Avidin-Agarose was washed once with 1 mL of DPBS at pH 7.4, 1% Triton X-100, 5 mM EDTA, and 0.2% SDS, 3 times with DPBS at pH 7.4, 1% Triton X-100, and 5 mM EDTA, and once with water. The cell-surface proteins bound to agarose were eluted by heating at 95 °C for 10 min in 150 μL of reducing Laemmli buffer. The 150 μL eluates were loaded on a 6.5% polyacrylamide gel in the presence of SDS and immunoblotted with anti-C-23 monoclonal antibody (Santa-Cruz Biotechnology Inc.).

Characterization of Nucleolin from Subcellular Fractions. Subcellular fractionation of Jurkat cells was performed using a protocol adapted from ref *29*. A total of 35×10^6 cells were washed 3 times with DPBS and resuspended in 1.5 mL of 50 mM Tris/HCl at pH 7.6, 0.35 M saccharose, 25 mM KCl, 10 mM $MgCl_2$, and 1 mM Pefabloc. Cells were subjected to Dounce homogenization on an ice bath. Nuclear fraction was pelleted by centrifugation at 1000*g* for 20 min. Heavy membrane pellets and mitochondria were removed by three centrifugation steps at 10000*g* for 15 min. The resulting supernatant was centrifuged at 105000*g* for 2 h to separate the microsomal and cytosolic fractions. The total microsomal fraction and only a tenth of the nuclei and the cytosolic fractions were loaded on a 7.5% polyacrylamide gel in the presence of SDS and immunoblotted with anti-C-23 monoclonal antibody (Santa-Cruz Biotechnology Inc.).

Analysis of Nucleolin in Tunicamycin-Treated Jurkat Cells. Jurkat cells were diluted at 150 000 cells/mL in fresh culture medium containing or not containing 10 μg/mL tunicamycin and cultured at 37 °C as described above. Cell culture volumes corresponding to 500 000 cells were collected either immediately (D0) or after 24 h (D1) and centrifuged at 1000*g* for 10 min. Viability of cells in pellets was checked by Trypan Blue staining and was >95% for all aliquots. Cell counting was performed to ensure identical numbers of cells per aliquot (500 000). The cell pellets were washed twice with DPBS, resuspended in 50 μL of buffer E, and processed as described above to collect the extranuclear fractions. Each fraction was added to 100 μL of sample buffer containing 2% SDS and 5% β-mercaptoethanol and submitted to SDS−PAGE. Detection of nucleolin was performed by immunoblotting using monoclonal anti-C-23 antibody.

Interaction of Nucleolin with Gel-Immobilized Lectins. Nuclear and extranuclear fractions of Jurkat cells were incubated at 4 °C for 1−2 h on ice in the presence of 1 mM Pefabloc with Wheat germ agglutinin coupled to Agarose (WGA-Agarose) in 20 mM Tris-HCl at pH 7.4, 0.5 M NaCl, or Concanavalin A coupled to Sepharose (ConA-Sepharose) in 20 mM Tris-HCl at pH 7.4, 0.5 M NaCl, 1 mM $MnCl_2$, and 1 mM $CaCl_2$. As a control, binding of nucleolin to WGA-Agarose and ConA-Sepharose was inhibited by 500 mM GlcNAc and 500 mM α-methylglucopyranoside, respectively. After incubation, gels were washed 4 times with the corresponding lectin buffers. Glycoproteins retained on gels were eluted by boiling in 100 μL of sample buffer containing 2% SDS and 5% β-mercaptoethanol. Proteins were separated by 7.5% SDS−PAGE and blotted on nitrocellulose. Nucleolin was then immunostained with anti-nucleolin polyclonal antibodies (*10*).

Lectin-Screening Studies. Purified extranuclear nucleolin (100 ng) was submitted to SDS−PAGE, blotted to nitrocellulose membrane, and analyzed with the DIG glycan differentiation kit according to the instructions of the manufacturer (Roche). The membrane was incubated in blocking solution, washed twice with Tris-buffered saline (TBS) (0.05 M Tris/HCl at pH 7.5 and 0.15 M NaCl), and incubated with the appropriated digoxigenin-labeled lectin: *Galanthus nivalis* agglutinin (GNA), *Sambucus nigra* agglutinin (SNA), *Maackia amurensis* agglutinin (MAA), or *Datura stramonium* agglutinin (DSA). The blot was washed 3 times with Tris-buffered saline (TBS). Recognized glycoproteins were revealed by incubation with anti-digoxigenin antibodies conjugated with alkaline phosphatase. The presence of O-glycosylation was explored using peanut agglutinin (PNA) after treatment of purified nucleolin with 20 milliunits of Neuraminidase (Roche) in 50 mM sodium acetate at pH 5 containing 1 mM Pefabloc during 2 h at 37 °C.

GC/MS Analysis of Nucleolin Monosaccharides. Purified nucleolin isoforms were separated by electrophoresis and electrotransferred onto a poly(vinylidene difluoride) (PVDF) membrane. After Ponceau S staining, the 105- and 113-kDa bands were excised. The bands were washed and dried under a stream of nitrogen. The procedure described by ref *30* was then used. Briefly, samples were submitted to acid-catalyzed methanolysis (20 h at 80 °C in 500 μL of anhydrous methanol containing 0.5 M gaseous HCl). After the samples were dried under a stream of nitrogen, they were supplemented with 200 μL of acetonitrile and 25 μL of heptafluorobutyric anhydride and heated for 30 min at 150 °C in a sand bath. Under these conditions, all monosaccharides are recovered as *O*-methyl-glycosides, except the particular GlcNAc residue forming the N-glycosidic bond (*31*).

After the reagents were evaporated, samples were dissolved in 200 μL of dried acetonitrile and 1 μL was injected

in the Ross injector (260 °C) of a Carlo Erba GC 8000 gas chromatograph equipped with a 25 m × 0.32 mm CP-Sil5 CB low bleed/Ms capillary column, 0.25 μm film phase (Chrompack, les Ulis, France). The temperature program started at 90 °C for 3 min, followed by an increase (5 °C/min) until 260 °C. The column was coupled to a Finnigan Automass II mass spectrophotometer (Finnigan, San Jose, CA). Analyzes were performed in the electron impact mode (ionization energy,70 eV; source temperature, 150 °C). Quantitation of various constituents was performed using the total ion count of the MS detector and the Xcalibur software (Finnigan).

Glycopeptides Characterization by MS. The 105- and 113-kDa nucleolin isoforms were separated by SDS−PAGE. The bands were excised and treated by trypsin as described above. The final extract was dried and resuspended in 1 μL of high-purity water. The peptide-containing solution was mixed with 1 μL of matrix solution suitable for glycopeptides obtained by saturating a water−acetonitrile 50:50 (v/v) and 3% trifluoroacetic acid (v/v) solution with sinapinic acid (*32*). A total of 1 μL of each sample was spotted onto the target, air-dried, and analyzed on a Voyager DE-STR MALDI−TOF instrument in the linear positive-ion mode by delayed extraction using an accelerating voltage of 25 kV.

Exoglycosidase Digestion of Nucleolin Glycans. A solution of highly concentrated nucleolin was needed for exoglycosidase digestion of glycans and further analysis by MS. Because the presence of sulfate ammonium in purified nucleolin fractions was found inhibitory for acid or solvent precipitation of proteins (personal data), nucleolin isoforms were first immunoprecipitated. Protein A-Sepharose (150 mg) (Sigma−Aldrich) was incubated with anti-nucleolin polyclonal antibodies (200 μL) in 40 mM Hepes at pH 7.5 and 0.1% ovalbumin during 90 min at 4 °C. Protein A-Sepharose coupled to antibodies was washed 3 times with DPBS and then incubated in batch with extranuclear nucleolin (5−10 mg) purified above, in 50 mM Tris/HCl at pH 7.4, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA, and 1 mM Pefabloc during 90 min at 4 °C. The gel was washed twice with the same buffer. Nucleolin was eluted with 100 mM Gly/HCl at pH 2.8. The fraction containing nucleolin was then acido-precipitated on ice with trichloroacetic acid to a 10% final concentration. After centrifugation (10000*g*) at 4 °C for 15 min, the pellet was washed 3-fold with ethanol and vacuum-dried.

Enzymatic release and sequencing of *N*-glycans were directly performed on the plate according to ref *33*. Extranuclear nucleolin amounts close to 1 μg were deposited on the plate and reconstituted in 1 μL of reaction buffer (10 mM sodium phosphate at pH 6.5). A total of 30 milliunits of PNGase F was added to each spot. β-Galactosidase (0.3 milliunit) was added to the second spot. β-Galactosidase and *N*-acetyl-β-D-glucosaminidase (0.3 milliunit) were added to the third spot. The plate was then placed in a humidified atmosphere at 40 °C for 4 h. High-purity water was continuously added to the spots to prevent drying. A control was performed with nuclear nucleolin immunoprecipitated in the same conditions.

Samples were allowed to dry first, before 1 μL of 2.5-dihydroxybenzoic acid matrix solution [10 mg/mL disolved in CH_3OH/H_2O (50:50, v/v)] was added. Mass spectra were acquired on a Voyager DE-STR MALDI−TOF instrument (AME Bioscience Ltd., Toroed, Norway) operating in the positive-ion reflectron mode.

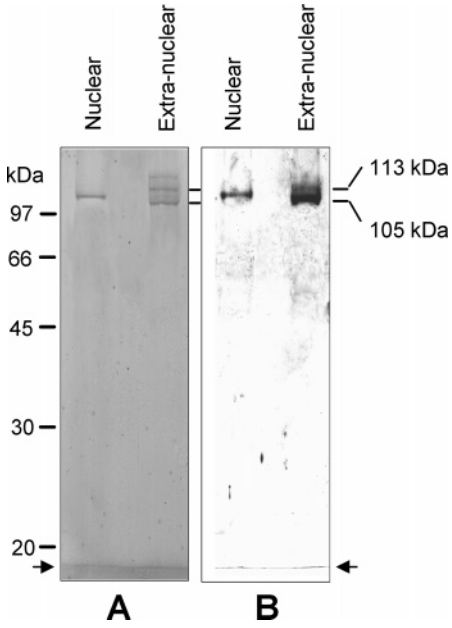

FIGURE 1: Presence of two nucleolin isoforms in Jurkat cells. Nuclear and extranuclear fractions were prepared from Jurkat cells and submitted to chromatography on DEAE-Sepharose and heparin-Sepharose as described in the Experimental Procedures. Purified proteins (2−5 μg) were separated by SDS−PAGE and either stained with colloidal Coomassie blue (A) or blotted onto nitrocellulose membrane and revealed by immunoluminescence with anti-nucleolin polyclonal antibodies and peroxidase-conjugated anti-rabbit IgG (B). Arrows indicate the front dye. The 113- and 105-kDa labels show the positions of the two putative nucleolin isoforms.

RESULTS

Evidence for Two Extranuclear Nucleolin Isoforms. Nucleolin was purified from Jurkat cells as previously reported (*10*). An absence of cross-contamination between the nuclear and extranuclear fractions was checked using antibodies against actin, specific for cytoplasm and p62 and specific for the nucleus (data not shown). After SDS−PAGE, three 105-, 113-, and 125-kDa protein bands were revealed by Coomassie blue staining in the extranuclear fraction, while only one 105-kDa band was detected in the nuclear fraction (Figure 1A). The 105-kDa band corresponds to that formerly described for human nucleolin (*34*), and thus, it was strongly recognized by polyclonal antibodies, specific for nucleolin (*10*) (Figure 1B). Interestingly, the 113-kDa band but not the 125-kDa band was also recognized by either polyclonal (Figure 1B) or monoclonal anti-nucleolin antibodies (not shown), although to a lower extent than the 105-kDa band. Because the 125-kDa band appears to be a contaminant in the nucleolin preparation, it has not been further characterized. Identification of the 105- and 113-kDa bands as nucleolin isoforms was confirmed using tryptic peptide mass fingerprinting (Table 1). A total of 22 and 17 peptides were detected from the 105- and 113-kDa bands, respectively, that correspond to a total of 25 peptides overlapping nucleolin sequence 56−624. Identification of both the 113- and 105-kDa protein bands as nucleolin was ensured by the high MOWSE score (protein Prospector) with a low mass tolerance (±43.7 and ±60.4 ppm) and 10 and 13% coverage of the total sequence, respectively (Table 2). The relatively low

Table 1: Mass Comparison of the Tryptic Peptides from the 105- and 113-kDa Proteins that Fit the Theoritical Mass of Nucleolin Peptides[a]

measured mass of the isoform (Da)				
105 kDa	113 kDa	theoritical mass (Da)	position	sequence
501.23	501.26	501.27	275−279	EAPGK
	530.31	530.29	111−116	GATPGK
541.32	541.42	541.37	224−228	VVPVK
545.35	545.33	545.36	394−398	TLLAK
577.37	577.15	577.31	343−347	IGMTR
601.35	601.39	601.30	562−567	GSPNAR
603.35	603.39	603.31	547−551	EIEGR
608.30		608.29	605−610	ETGSSK
616.59		616.37	97−102	TVTPAK
	619.31	619.34	56−62	AAATSAK
629.39	629.31	629.36	89−95	AAATPAK
634.37		634.36	111−116	NLPYK
806.36		806.44	371−377	VFGNEIK
812.47		812.46	555−561	LELQGPR
832.40	832.44	832.44	404−410	VTQDELK
875.39	875.41	875.42	450−457	QGTEIDGR
884.42	884.40	884.45	514−521	VPQNQNGK
	937.43	937.50	325−333	TGISDVFAK
995.46	995.47	995.44	478−486	N*STWSGESK
1000.53		1000.54	334−342	NDLAVVDVR
1057.63	1057.65	1057.60	633−645	VTLDWAKPK
1160.52		1160.58	458−467	SISLYYTGEK
1561.61	1561.69	1561.68	611−624	GFGFVDFNSEEDAK
1648.64	1648.73	1648.73	349−362	FGYVDFESAEDLEK
2311.28		2312.16	298−318	VEGTEPTTAFNLFVGNLNFN*K

[a] Asterisks indicate the position of potential N-glycosylation sites.

Table 2: Identification of the 105- and 113-kDa Proteins Following Database Searching with Protein Prospector (http://prospector.ucsf.edu)

protein (kDa)	mass tolerance (ppm)	MOWSE score	masses matched	coverage (%)	identification (top of the list)
105	±60.4	4.007 E+5	22	13	human nucleolin
113	±43.7	1.25 E+4	17	10	human nucleolin

number of nucleolin peptides that were detected could be related to the abundance of post-translational modification occurring on nucleolin (*2*). We indeed chose to limit the research to unmodified peptide to reduce the risk of mismatch. Interestingly, peptides 298−318 and 478−486 (2312.16 and 995.44 Da, respectively), which contain two of the three potential N-glycosylation sites of nucleolin (N317 and N478) (*4*) were detected from the 105-kDa band, while only peptide 478−486 (995.44 Da) could be detected from the 113-kDa band. The lack of detection of the 2311-Da peptide in the 113-kDa isoform could be explained either by the lower amount of the 113-kDa nucleolin isoform, as compared to the 105-kDa protein, or by post-translational modification of this peptide, presumably by N-glycosylation at position 317. The third potential glycosylation site 492 should be contained in a 2501.26-Da peptide, which was not detected in either the 105- or 113-kDa isoform.

Characterization of Cell-Surface Nucleolin. Surface biotinylation of Jurkat cells and extraction of labeled surface proteins were performed as described in the Experimental Procedures. Western blot analysis of surface proteins with anti-nucleolin antibodies showed both 105- and 113-kDa protein bands, thus indicating the presence of the glycosylated nucleolin isoform (Figure 2). The 113/105-kDa protein band ratio was very similar to that observed in the extranuclear fraction. However 76-, 53-, and <25-kDa protein bands were revealed that indicate partial proteolytic degradation of nucleolin. The degradation could be attributable to the number and length of preparation steps. The specificity of cell-surface nucleolin capture was confirmed by the absence of the signal from the nonbiotinylated cell control.

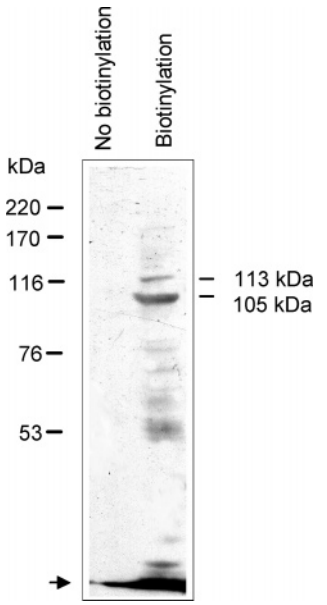

FIGURE 2: Characterization of cell-surface nucleolin. Jurkat cells (35×10^6) were surface-biotinylated and solubilized as described in the Experimental Procedures. The cell lysates were centrifuged, and the supernatants were incubated with avidin-Agarose. After extensive washes of agarose beads, the trapped biotinylated proteins were eluted with reducing Laemmli buffer, entirely loaded on a 6.5% polyacrylamide gel in the presence of SDS and immunoblotted with monoclonal anti-C-23 antibody (Santa-Cruz Biotechnology Inc.) (lane "Biotinylation"). A control was made with nonbiotinylated cells submitted to the same protocol (lane "No biotinylation"). The arrow indicates the dye front.

Nucleolin Characterization in Subcellular Fractions. As shown in Figure 3, the 105-kDa nucleolin band was only

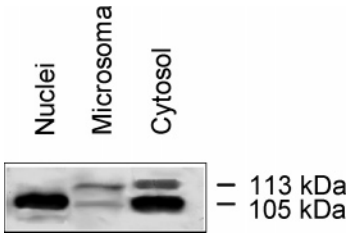

FIGURE 3: Characterization of nucleolin in subcellular fractions. A total of 35×10^6 Jurkat cells were washed 3 times with DPBS and resuspended in 1.5 mL of 50 mM Tris/HCl at pH 7.6, 0.35 M saccharose, 25 mM KCl, 10 mM $MgCl_2$, and 1 mM Pefabloc. Cells were subjected to Dounce homogenization on ice. Nuclear fraction was pelleted by centrifugation at 1000*g* for 20 min. Heavy membrane pellets and mitochondria were removed by three centrifugations at 10000*g* for 15 min. The resulting supernatant was centrifuged at 105000*g* for 2 h to separate the microsomal and cytosolic fractions. The total microsomal fraction and only a tenth of the nuclei and of the cytosolic fraction were loaded on a 7.5% polyacrylamide gel in the presence of SDS and immunoblotted with anti-C-23 monoclonal antibody.

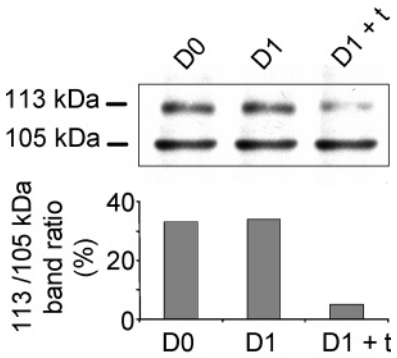

FIGURE 4: Sensibility of the 113-kDa nucleolin isoform to tunicamycin. The figure shows the results of a typical experiment in which the extranuclear fractions of 500 000 viable Jurkat cells aliquots were separated by SDS−PAGE, Western-blotted, and stained by immunoluminescence using anti-C-23 monoclonal mouse antiboby (Santa Cruz) and peroxidase-conjugated anti-rabbit IgG. The lanes show nucleolin isoforms from cells incubated in the presence of 10 μg/mL tunicamycin during 24 h (D1 + t) or without tunicamycin during 0 h (D0) and 24 h (D1). The histogram shows the 113/105-kDa band ratio, as estimated by densitometry using the BioRad GS-710 imaging densitometer and Quantity one 4.1.0 software (BioRad, Hercules, CA).

observed in the nuclear fraction. In the microsomal fraction containing secretory vesicles, both 105- and 113-kDa nucleolin bands were detected, which is in agreement with our results showing that both nucleolin isoforms are expressed at the surface of cells. More surprisingly, the 113-kDa isoform can also be observed in the cytosolic fraction.

Tunicamycin Prevents Formation of the 113-kDa Nucleolin Isoform. Tunicamycin is an antibiotic that blocks the first step of N-linked glycoprotein synthesis. This is due to the blockage of the transfer of the 14-residue core oligosaccharide (two GlcNAc, nine Man, and three Glc) from a dolichol phosphate donor molecule to Asn residues on the proteins (*35*). Hence, exposing Jurkat cells to tunicamycin has been used to investigate its effects on the synthesis of the two 113- and 105-kDa nucleolin isoforms. As described in the Experimental Procedures, cells were treated with 10 μg/mL tunicamycin during 24 h and the presence of the two nucleolin isoforms was investigated by immunostaining at times 0 and 24 h. Figure 4 shows that, as compared to the control cells, cells incubated with tunicamycin exhibited a markedly decreased 113/105-kDa protein band ratio. Although indirect effects of tunicamycin could contribute to the 113-kDa band decrease, these results suggest that the 113-kDa isoform undergoes N-glycosylation steps common to all N-linked glycoproteins.

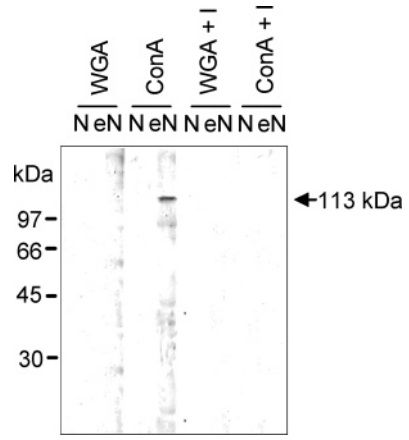

FIGURE 5: Specific binding of the 113-kDa nucleolin isoform to ConA-Sepharose but not to WGA-Agarose. Nuclear (N) and extranuclear (eN) fractions of Jurkat cells were incubated with WGA-Agarose (WGA) or ConA-Sepharose (ConA) as described in the Experimental Procedures. Nucleolin eluted from the immobilized lectins was separated by SDS−PAGE, Western-blotted, and stained by immunoluminescence using anti-C23 monoclonal mouse antiboby (Santa Cruz) and peroxidase-conjugated anti-rabbit IgG. Nonspecific binding was assessed in the presence of inhibitors (+ I): 500 mM GlcNAc and 500 mM α-methylglucopyranoside for WGA and ConA, respectively.

Binding of the 113-kDa Nucleolin Isoform to Concanavalin A-Sepharose. Nuclear and extranuclear fractions of Jurkat cells and purified extranuclear nucleolin were incubated in the presence of Wheat germ agglutinin (*Triticum vulgaris*)-Agarose (WGA-Agarose) or Concanavalin A-Sepharose (ConA-Sepharose). WGA-Sepharose is specific for $GlcNAc_2$ or NeuAc residues (*36*), and binding to WGA can be inhibited in the presence of a large excess of GlcNAc. Affinity to WGA is commonly used to detect the presence of O-GlcNac residues on proteins (*37*). ConA-Agarose specifically binds to Man and Glc residues of polysaccharides and glycoproteins, and that binding can be inhibited in the presence of a large excess of α-methylglucopyranoside (*38*). After incubation with the nucleolin-containing fractions, lectin matrixes were boiled in sample buffer with SDS and β-mercaptoethanol. Eluted proteins were separated by SDS−PAGE, Western-blotted, and immunostained using anti-nucleolin antibodies. As shown in Figure 5, no nucleolin isoform from nuclear and extranuclear fractions bound to WGA, thus suggesting that nucleolin is not O-N-acetylglucosaminylated. This result was confirmed by the use of RL2 antibodies (Alexis Biochemicals) specific for the O-GlcNac motif (data not shown). Sumoylation with SUMO-1, another potential common post-translational modification that could account for larger nucleolin isoforms, was not detected using specific antibodies (anti-SUMO-1 FL-101, Santa Cruz, CA; data not shown). In addition, Figure 5 shows that the 113-kDa nucleolin isoform from the extranuclar fraction was retained on ConA-Agarose, while no band was detected from the nuclear fraction. From these results, it can be assumed that the 113-kDa nucleolin isoform from the extranuclear fraction is glycosylated by Man or Glc

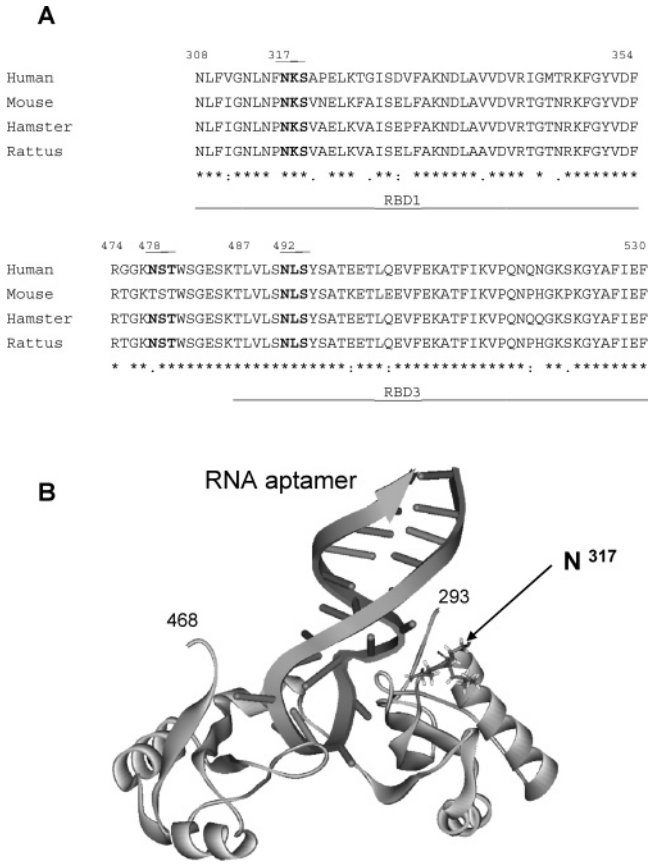

FIGURE 6: Location of the three potential N-glycosylation sites on nucleolin. (A) Alignment of sequences of RBDs 1 and 3 of human, mouse, hamster, and rattus nucleolin. Numbering refers to the sequence of human nucleolin (*4*). The locations of RBD1 and 3, as well as the three potential N-glycosylation sites (N317, N478, and N492), are indicated with horizontal brackets. "*", ":", and "." indicate identical, homologous, and nonhomologous residues, respectively. (B) Structure of chinese hamster nucleolin RBD1 (residues 293−468) in complex with a RNA 22-base aptamer GGCCGAAAUCCCGAAGUAGGCC (*56*). The arrow indicates the position of potentially glycosylated N317 on the edge of the RNA-binding cleft.

residues, while nuclear nucleolin is not glycosylated. Binding of the 113-kDa isoform to ConA was abolished by 500 mM α-methylglucopyranoside, confirming the specificity of recognition.

Potential N- and O-Glycosylation Sites of Nucleolin. Using NGlyc-Prediction (http://www.expasy.ch) and as previously reported (*4*), three potential N-glycosylation sites were proposed: N317, N478, and N492. As shown in Figure 6A, sequence alignment of human, mouse, hamster, and rattus nucleolins reveals that, unlike N317 and N492, N478 is not conserved in all mammalian species. Hence, the ability of N478 to be glycosylated and its functional role thereof are unlikely. This is suggested by our peptide mass fingerprinting of the 113-kDa nucleolin isoform (Table 1) that shows the presence of peptide ^{478}NSTWSGESK^{486} and rules out any glycosylation of N478. Conversely, residues N317 and N492 are conserved in mammalian nucleolin sequences (Figure 6A), and they are strong candidates for N-glycosylation. As illustrated in Figure 6A, N317 and N492 are located in domains RBD1 and RBD3, respectively. The peptide ^{298}VEGTEPTTAFNLFVGNLNFNK^{318} comprising N317 was not detected in the 113-kDa isoform when it is in the 105-kDa isoform, suggesting that N317 could be N-glycosylated. A complete 3D structure of nucleolin is not elucidated, but the structure of RBD1, that includes N317, was recently defined by NMR (PDB 1FJ7). As shown in Figure 6B, residue N317 is close to the RNA aptamer-binding site. It may thus be hypothesized that N-linkage of a complex glycan to N317 generates a steric hindrance to RNA recognition. A similar position of residue N492 inside RBD3 can be observed (data not shown).

Potential O-glycosylation sites were predicted using NetOGlyc (http://www.expasy.ch). Five potential O-glycosylation sites were proposed in basic N-terminal patterns (TPXKK) of nucleolin: T84, T92, T105, T106, and T113 (data not shown).

Lectin-Screening Studies. Specific binding of lectins to glycans was used to study the structure of nucleolin oligosaccharides. Purified extranuclear and nuclear nucleolins and control glycoproteins were submitted to SDS−PAGE and then blotted onto nitrocellulose membrane. The lectins used in the screening kit are conjugated with the steroid hapten digoxigenin, which enables immunological detection of the bound lectins. GNA recognizes terminal Man (α1−3), (α1−6), or (α1−2) linked to Man, and it is thus suitable to identify high mannose or hybrid *N*-glycan chains (*39*). As shown in Figure 7, the carbohydrate structure of control glycoprotein carboxypeptidase Y (63 kDa) was recognized (*40*), while nucleolin was not detected. MAA recognizes sialic acid linked (α2−3) to Gal in *N*-glycans and (α2−3)-linked sialic acid in *O*-glycans. The control glycoprotein fetuin (68 kDa), which contains sialic acid (α2−3) Gal motifs, was recognized (*41*, *42*), while nucleolin was not detected (Figure 7). SNA recognizes sialic acid terminally linked (α2−6) to Gal in *N*-glycans or GalNAc in *O*-glycans. As expected, the carbohydrate structure of control transferrin (80 kDa) was recognized (Figure 7) because it presents sialic acid (α2−6) Gal motifs (*43*). The control glycoprotein fetuin (68, 65, and 61 kDa) was also recognized because fetuin presents (α2−6)-linked sialic acid in addition to (α2−3)-linked sialic acid. Although a diffuse and weak signal was sometimes observed for the 113-kDa nucleolin isoform, most of our experiments showed no significant signal (Figure 7). The result was not conclusive and pointed out limits to use lectins for determining glycan structures. The possibility that the 113-kDa nucleolin isoform possesses sialic acid led us to investigate whether sialylation occurs on *N*- or *O*-glycans. To verify possible O-glycosylation, extranuclear nucleolin was treated with neuraminidase before electrophoresis and PNA detection. PNA recognizes the core disaccharide Gal(β1−3)GalNAc and is thus suitable for identifying O-glycosidically linked carbohydrate chains (*44*). As shown in Figure 7, both the control glycoprotein asialofetuin (61 kDa) and the 113-kDa extranuclear nucleolin isoform treated with neuraminidase reacted positively with PNA. This result confirms the presence of the Gal(β1−3)-GalNAc motif O-linked in 113-kDa nucleolin. On the contrary, extranuclear nucleolin either untreated or incubated in lectin-free buffer (pH 5) was not detected by PNA, thus demonstrating that sialic acid was specifically removed by the neuraminidase allowing recognition of the Gal(β1−3)-GalNAc motif. These results clearly indicate that the extranuclear 113-kDa nucleolin isoform is O-glycosylated by Gal(β1−3)GalNAc with GalNAc substituted by sialic acid (α2−6). DSA recognizes the Gal(β1−4)GlcNAc motif in

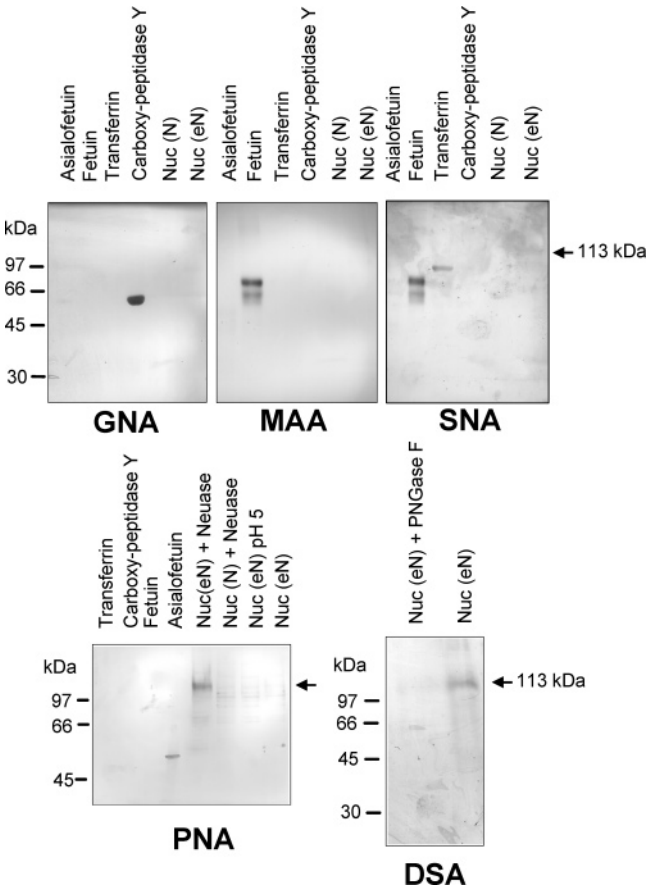

FIGURE 7: Partial characterization of carbohydrates of the 113-kDa nucleolin isoform with lectins. Lectin screening with GNA, MAA, SNA, PNA, or DSA (DIG glycan differentiation kit from Roche) was performed on nucleolin purified from nuclear (N) and extranuclear (eN) fractions of Jurkat cells, submitted to SDS−PAGE and electroblotted onto nitrocellulose membrane. Asialofetuin, fetuin, transferrin, and carboxy-peptidase Y were used according to the instructions of the manufacturer as glycoprotein controls. GNA recognizes N-glycosidically linked high mannose or hybrid structures present on carboxypeptidase Y. MAA binds to sialic acid terminally linked (α2−3) to Gal present on fetuin. SNA recognizes sialic acid terminally linked (α2−6) to Gal present on transferrin and fetuin. PNA binds to the Gal(*β*1−3)GlcNac motif that is present on asialofetuin and forms the core unit of *O*-glycans. As described in the Experimental Procedures, nucleolin was pretreated with neuraminidase (Neuase) to expose the (*β*1−3)GlcNac core. DSA binds to the Gal(*β*1−4)GlcNac motif. Specificity of binding between DSA and the 113-kDa nucleolin isoform was checked by pretreatment of extranuclear nucleolin with PNGase F.

complex and hybrid *N*-glycans and GlcNAc in *O*-glycans (*45*). Extranuclear 113-kDa nucleolin isoform was detected by PNA, suggesting the presence of GlcNAc in *O*-glycans and/or Gal(*β*1−4)GlcNAc motif in *N*-glycans (Figure 7). Because treatment of extranuclear nucleolin with PNGase F prevented detection by DSA (Figure 7), it may be assumed that the Gal(*β*1−4)GlcNAc motif is present on *N*-glycans.

Determination of the Monosaccharide Composition of the 113-kDa Nucleolin Isoform. GC/MS analysis of the HFB derivatives from the 113-kDa nucleolin isoform allowed the qualitative and quantitative determination of its monosaccharide composition (Table 3). According to this method (*31*), all monosaccharides are recovered as *O*-methyl-glycosides except the particular GlcNAc residue forming the N-glycosidic bond that is converted to glucosamine (GlcN). The GlcN peak (retention time of 12.39 min) was used as a reference, and its corrected area (6 253 775) was considered as corresponding to one monosaccharide residue. As shown in Table 3, our results support the presence of both N- and O-linked glycans on the 113-kDa nucleolin isoform. Only 3 Man residues were detected that suggests the presence of a complex *N*-glycan structure rather than high mannose or hybrid structures. We detected a 0.84 GlcNAc residue that presumably corresponds to 1 GlcNAc residue linked to the N-linked GlcNAc but also to the Gal(*β*1−4)GlcNAc motif previously identified by DSA. Interestingly, the number of Gal residues (3.33 that can be rounded to 3) indicates that the number of GlcNAc is probably underestimated. Two additional GlcNAc residues are indeed necessary for 2 Gal(*β*1−4)GlcNAc motifs involving 2 Gal residues. Hence, the third Gal residue could be involved in the *O*-glycan structure whose presence is revealed by the detection of 0.75 (1) GalNAc (Table 3). Similar amounts of GalNAc and GlcN residues indicate that the 113-kDa isoform possesses equal numbers of *N*- and *O*-glycans. Finally, a 1.15 NeuAc residue corresponding to 1 sialic acid was found. The presence of sialic acid on *N*- and *O*-glycans of the 113-kDa isoform was indeed suspected by lectin screening with SNA and PNA after neuraminidase treatment (Figure 7). Glc, a contaminating monosaccharide, was also detected through GC/MS analysis from the HFB derivatives of nucleolin isoforms. This only monosaccharide was detected in the 105-kDa isoform used as a control.

Table 3: Monosaccharide Composition of the 113-kDa Nucleolin Isoform as Determined by Heptafluorobutyrate Derivation and GC/MS

monosaccharide	number of residues
GlcN	1.00
Gal	3.33
Man	3.00
GalNac	0.75
GlcNac	0.84
NeuAc	1.15
total	10

Identification of N317 and N492 as the N-Glycosylated Sites of Nucleolin. MALDI−TOF analysis was used to localize the N-glycosylation sites from the 113-kDa nucleolin and to determine the structure of *N*-glycans. Matrix solution used to visualize glycopeptides and large peptides was obtained by saturating a water−acetonitrile trifluoroacetic acid solution with sinapinic acid (*32*). As shown in Figure 8A, mass spectra analysis of tryptic peptides from the 105-kDa nucleolin isoform revealed masses 2311.0 and 2499.5 Da corresponding to the nonglycosylated peptides: ^{298}VEGTEPTTAFNLFVGNLNFNK^{318} (2312.16 Da) and ^{487}TLVLSNLSYATEETLQEVFEK^{508} (2501.27 Da). On the contrary and as shown in Figure 8B, neither the 2311.0- nor 2499.5-Da mass were detected on the mass spectra of tryptic peptides from the 113-kDa nucleolin isoform. Two new mass peaks appeared at 3649.0 and 3842.9 Da instead, which could correspond to 1339.0- and 1343.4-Da mass increments of mass peaks 2311.0 and 2499.5 Da, respectively. Interestingly, 1339.0 Da is the characteristic mass of the *N*-glycan structure: $GlcNAc_2Man_3GlcNAc_2$ (Figure 8). Use of the linear positive-ion mode for mass determination of glycopeptides impaired precision of this determination and can justify the difference of 4.4 Da between the 1339.0-Da theoretical mass and the 1343.4-Da observed mass. From

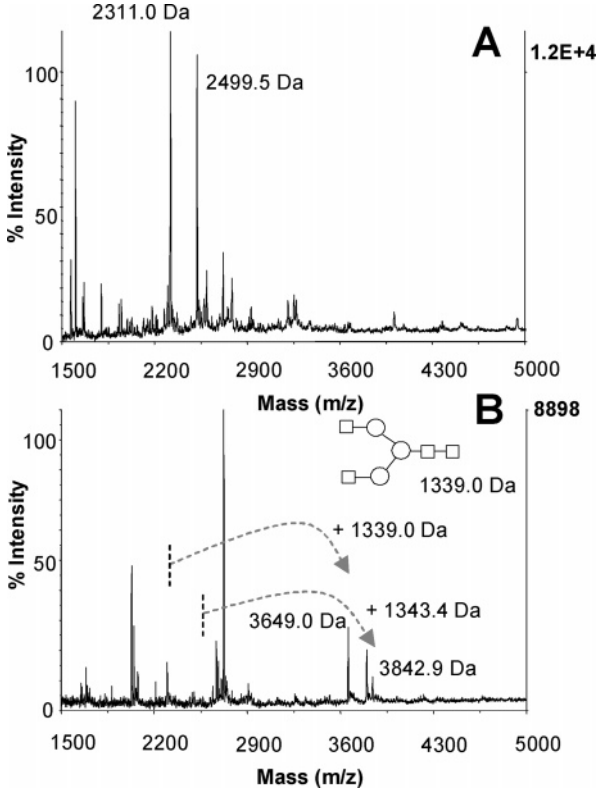

FIGURE 8: Identification of potential N-glycosylated peptides from the 105-kDa isoform (A) and N-glycosylated peptides from the 113-kDa isoform (B). As described in the Experimental Procedures, nucleolin isoforms were separated by SDS−PAGE. The bands were excised and treated by trypsin. Tryptic fragments were mixed with a sinapinic acid matrix. Each sample was analyzed on a Voyager DE-STR MALDI−TOF instrument in the linear, positive-ion mode using an accelerating voltage of 20 kV. The arrows indicate the mass shifts observed between the nonglycosylated (2311.0 and 2499.5 Da) and the glycosylated (3649.0 and 3842.9 Da) peptides. Putative *N*-glycan structure corresponding to the 1339.0-Da mass increment is indicated [GlcNAc (□) and Man (○)].

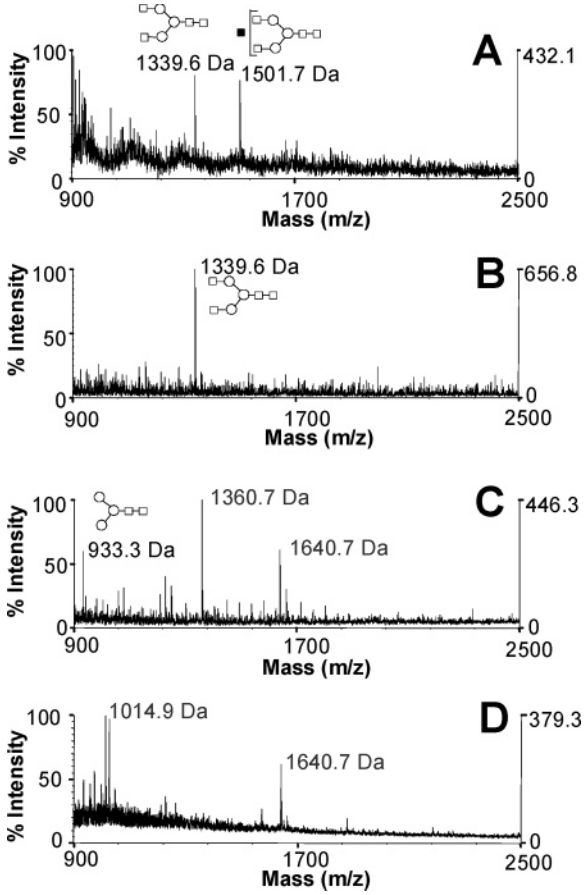

FIGURE 9: MALDI−TOF mass spectra of the N-linked oligosaccharide released from the 113-kDa nucleolin isoform. MS profiles result from digestion of 5 μg of purified extranuclear nucleolin with PNGase F (A), PNGase F and β-galactosidase (B), and PNGase F, β-galactosidase, and β-*N*-acetyl glucosaminidase (C). A control MS profile was obtained after treatment with PNGase F of nuclear nucleolin immunopurified in similar conditions than extranuclear nucleolin (D). The structures of N-linked oligosaccharides deduced from the corresponding mass peaks are indicated [GlcNAc (○), Man (□), and Gal (■)].

these results, it can be assumed that both residues N317 and N492 are substituted in the $GlcNAc_2Man_3GlcNAc_2$ structure. However, this structure represents a truncated form of that previously deduced from lectin screening and GC/MS monosaccharide composition studies.

Mass Spectra of N-Linked Oligosaccharides. PNGase F is able to release intact *N*-glycans from glycoproteins (*46*). Sequential use of PNGase F with β-galactosidase and *N*-acetyl-β-D-glucosaminidase in combination with MALDI−TOF analysis allowed identification of the *N*-glycan chain. In Figure 9A, the *m*/*z* values of 1339.6 and 1501.7 Da indicated that two *N*-glycan structures were released by PNGase F. The 1339.6-Da mass could correspond to the $GlcNAc_2Man_3GlcNAc_2$ structure previously identified in glycopeptides and the 1501.7-Da mass to the $GalGlcNAc_2$-$Man_3GlcNAc_2$ structure deduced from the lectin-screening studies. Simultaneous sample treatment with PNGase F and β-galactosidase confirmed these results (Figure 9B). The only 1339.6-Da mass peak was indeed detected, thus confirming the (β1−4) linkage of the 1501.7-Da peak with 1 Gal residue. Finally, simultaneous sample treatment with PNGase F, β-galactosidase, and *N*-acetyl-β-D-glucosaminidase yielded a 933.3-Da mass peak corresponding to the above-mentioned structure but deprived of 2 GlcNAc residues (β1−2) linked to the three mannosidic structure (Figure 9C). The 1360.7- and 1640.7-Da mass peaks correspond probably to contamination by peptides during enzymatic digestion. When these results are taken together, they revealed the presence of $GalGlcNAc_2Man_3GlcNAc_2$ and $GlcNAc_2Man_3GlcNAc_2$ in the 113-kDa nucleolin isoform. A negative control was performed with nuclear nucleolin immunopurified in the same conditions than extranuclear nucleolin. Spectra before (data not shown) and after treatment with PNGase F (Figure 9D) were similar. This demonstrates that *N*-oligosaccharides were actually released from extranuclear nucleolin and not from contaminating antibodies. The mass peaks in Figure 9D most probably correspond to proteolytic autodegradation of nucleolin.

DISCUSSION

Owing to its intracellular traffic, as well as its numerous functions and molecular partners (for a review, see ref *3*), nucleolin is a very intriguing molecule. In particular, its possible role as a shuttle between the cell surface and nucleus, which we recently pointed out as a requisite for nuclear targeting of lactoferrin (*10*), indicates that nucleolin takes atypical pathways in the cell. Nonclassical secretion pathways were indeed evidenced (*8*). Many post-translational

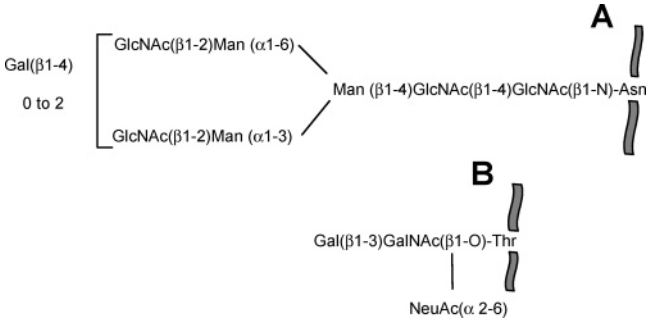

FIGURE 10: Structures of the N-linked (A) and O-linked (B) oligosaccharides identified in the 113-kDa extranuclear nucleolin isoform.

modifications but not glycosylation that might regulate both functions and traffic of nucleolin were clearly evidenced (*3*).

In the present paper, we demonstrate that part of the extranuclear nucleolin, identified as a 113-kDa isoform, is N- and O-glycosylated. Two N-glycosylated sites were defined at positions N317 and N492, which are localized on the edge of the RNA-binding cleft of domains RBD1 and RBD3 (Figure 6). Using MS analysis of glycopeptides, monosaccharide composition of released glycans, and lectin screening, the structure of the N-linked glycans was defined as complex asialo-biantennary glycans shown in Figure 10A. The estimated number of (β1−4)-linked Gal differed from one experimental approach to the other. Gal residues were indeed detected using either lectin screening or MS analysis of released *N*-glycans, i.e., techniques requiring fewer preparation and analysis steps than MS glycopeptide analysis. Hence, this discrepancy could most probably be due to degradation by exoglycosidases during glycoprotein preparation. Such degradation could be compared to complex glycan catabolism (*47*) and could be favored by the absence of sialic acids on nucleolin glycans. Additionally, monosaccharide composition and lectin-screening results brought evidence for O-glycosylation of the 113-kDa nucleolin isoform. From these results, it may be assumed that nucleolin has two *O*-glycans sialyled in (α2−6) to GalNAc O-linked to a Thr residue. Gal is (β1−3)-linked to GalNAc to form the 2,6-sialyl-T antigen (Figure 10B). Five potential O-glycosylation sites were predicted: T84, T92, T105, T106, and T113. Unfortunately, we were not able to localize the glycosylated Thr residues by MS because of the minute amounts of glycosylated nucleolin.

Preliminary results indicate that the glycosylated nucleolin isoform is not present in the nucleus. It seems located at the surface of cells and, more surprisingly, in the cytosol. The exact subcellular distribution of nucleolin is now under investigation by using recombinant nucleolin mutated at the N-glycosylation sites, subcellular fractionation, and confocal microscopy.

Glycosylation such as O-N-acetylglucosaminylation is, together with phosphorylation, a common post-translational modification of cytoplasmic and nuclear proteins that regulate their intracellular traffic (*48*). In particular, Lefebvre et al. (2003) (*49*) recently established a direct relationship between O-GlcNAc glycosylation, phosphorylation, and nuclear import of human microtubule-associated Tau proteins implicating nuclear lectins. However, despite the fact that nucleolin is tightly regulated by phosphorylation with CKII, cdc2, and PKC-ζ kinases (*12*, *13*), no O-N-acetylglucosaminylation could be detected (this study). Complex glycosylation of nucleolin was totally unexpected, although preliminary studies suggested glycosylation of surface nucleolin (*50*, *51*).

In some rare instances, complex glycosylation of cytosolic or nuclear proteins was evidenced. The first example was given by the HMG proteins, which are important structural components of chromatin (*22*). Highly purified preparations of HMG 14 and HMG 17 were indeed found to contain GlcNAc, Man, Gal, Glc, Fuc, and possibly Xyl (*22*). Nevertheless, even though these HMG studies are widely cited as evidence for nuclear *N*-glycans, they are highly controversial. First, any consensus site for N-glycosylation is present in HMG proteins. Second, Medina and Haltiwanger (*52*) have not been successful in verifying the presence of complex glycans of any type on the HMG proteins. Other examples were given with some N-glycosylated acidic nonhistone proteins, the sodium pump (Na^+,K^+-ATPase) from dog kidney, and human transcription factor NFIC (*21*, *24*, *53*). Actually, N-glycosylated acidic nonhistone proteins were induced to translocate from the cytoplasm to the nucleus in lymphocytes (*53*). Elsewhere, N-linked glycans with terminal GlcNAc residues were found in the cytosolic domain of the α subunit of the sodium pump (Na^+,K^+-ATPase) (*21*). Finally, N-glycosylation was recently described for the transcription factor NFIC during early mammary-gland involution (*24*). These studies may be questionable because, contrary to the present study, they suffered from the lack of site mapping of the putative glycans and the absence of accurate structural data. Furthermore, even if the effect of glycosylation on the traffic of some molecules was demonstrated (*24*, *53*), the possible impact of glycosylation on their function was not documented. Of course, these results involve the existence of not yet evidenced nonconventional biosynthesis pathways.

In the present paper, we show that, unlike glycosylated NFIC, which was exclusively located in the nucleus of cells (*24*), the glycosylated nucleolin isoform could only be detected outside the nucleus. This occurs despite the presence of the NLS in the nucleolin isoform and suggests that glycosylation influences the fate of the isoform. Interestingly, the observation that the glycosylated 113-kDa nucleolin isoform is not targeted to the nucleus ties well with the hypothesis that glycosylation may interfere with the function of nucleolin. Figure 6 indeed shows that the two N-glycosylated sites are localized on the edge of the RNA-binding cleft of domains RBD1 and RBD3. Hence, it may be hypothesized that, owing to the steric hindrance of the glycan moieties, the N-glycosylated nucleolin is no longer able to bind RNA. Furthermore, although we were not able to experimentally identify the two O-glycosylation sites, bioinformatic tools support the presence of potential sites in the TPXKK N-terminal region of nucleolin (*3*, *12*). Because the TPXKK region shares common phosphorylation sites, it may be assumed that O-glycosylation may also modify the function of the molecule.

As for other cytosolic and nuclear proteins whose complex glycosylation was reported (*21*, *22*, *24*, *51*), the biosynthesis pathways of the glycosylated nucleolin isoform are not known. It was previously hypothesized that cytosolic and nuclear glycoproteins containing oligomannose-type *N*-glycans could come from the endoplasmic reticulum by retrotranslocation through membrane protein complexes, al-

lowing the export of glycoproteins from the reticulum to cytosol (*54*, *55*). However, there is no clear evidence that mature N-glycoproteins with terminal Gal added in the Golgi apparatus could leave the secretory pathway to reach the cytosol. This is definitively not the case for nucleolin lacking a signal sequence that would allow it to enter the endoplasmic reticulum. Nevertheless, the presence of nucleolin on the surface of cells supports the hypothesis that nucleolin may follow nonclassical pathways in the cell. Hovanessian et al. (*8*) indeed showed that part of the cytoplasmic nucleolin is located in small vesicles that translocate nucleolin to the surface, although conventional inhibitors of the intracellular glycoprotein transport have no effect. Furthermore, surface nucleolin was shown to be internalized, alone or together with bound ligands (*3*, *6*−*10*). It may hypothesized that, during these processes, part of nucleolin takes unknown retrotranslocation pathways permitting N- and O-glycosylations.

In conclusion, our results give biochemical evidence for partial complex glycosylation of nucleolin. Although one may wonder about the synthesis pathway of the glycosylated form and the physiological relevance of such a modification, it is likely that glycosylation modifies the traffic and function(s) of nucleolin. Further investigations using recombinant mutated nucleolins are undertaken to elucidate the importance of glycosylation in the activity of the molecule, notably in its shuttle function between the cell surface and nucleus. A possible role of nucleolin glycosylation in the regulation of cell proliferation and growth will also be investigated.

ACKNOWLEDGMENT

We are grateful to the RIO proteomic platform of Lille for its technical support.

REFERENCES

1. Lapeyre, B., Bourbon, H., and Amalric, F. (1987) Nucleolin, the major nucleolar protein of growing eukaryotic cells: An unusual protein structure revealed by the nucleotide sequence, *Proc. Natl. Acad. Sci. U.S.A. 84*, 1472−1476.
2. Tuteja, R., and Tuteja, N. (1998) Nucleolin: A multifunctional major nucleolar phosphoprotein, *Crit. Rev. Biochem. Mol. Biol. 33*, 407−436.
3. Srivastava, M., and Pollard, H. B. (1999) Molecular dissection of nucleolin's role in growth and cell proliferation: New insights, *FASEB J. 13*, 1911−1922.
4. Srivastava, M., Fleming, P. J., Pollard, H. B., and Burns, A. L. (1989) Cloning and sequencing of the human nucleolin cDNA, *FEBS Lett. 250*, 99−105.
5. Bouvet, P., Diaz, J. J., Kindbeiter, K., Madjar, J. J., and Amalric, F. (1998) Nucleolin interacts with several ribosomal proteins through its RGG domain, *J. Biol. Chem. 273*, 19025−19029.
6. Saïd, E. A., Krust, B., Nisole, S., Svab, J., Briand, J. P., and Hovanessian, A. G. (2002) The anti-HIV cytokine midkine binds the cell surface-expressed nucleolin as a low affinity receptor, *J. Biol. Chem. 277*, 37492−37502.
7. Semenkovich, C. F., Ostlund, R. E., Jr., Olson, M. O., and Yang, J. W. (1990) A protein partially expressed on the surface of HepG2 cells that binds lipoproteins specifically is nucleolin, *Biochemistry 29*, 9708−9713.
8. Hovanessian, A. G., Puvion-Dutilleul, F., Nisole, S., Svab, J., Perret, E., Deng, J. S., and Krust, B. (2000) The cell-surface-expressed nucleolin is associated with the actin cytoskeleton, *Exp. Cell Res. 261*, 312−328.
9. Take, M., Tsutsui, J., Obama, H., Ozawa, M., Nakayama, T., Maruyama, I., Arima, T., and Muramatsu, T. (1994) Identification of nucleolin as a binding protein for midkine (MK) and heparin-binding growth associated molecule (HB-GAM), *J. Biochem. 116*, 1063−1068.
10. Legrand, D., Vigie, K., Saïd, E. A., Elass, E., Masson, M., Slomianny, M. C., Carpentier, M., Briand, J. P., Mazurier, J., and Hovanessian, A. G. (2004) Surface nucleolin participates in both the binding and endocytosis of lactoferrin in target cells, *Eur. J. Biochem. 271*, 303−317.
11. Schmidt-Zachmann, M. S., and Nigg, E. A. (1993) Protein localization to the nucleolus: A search for targeting domains in nucleolin, *J. Cell. Sci. 105*, 799−806.
12. Peter, M., Nakagawa, J., Doree, M., Labbe, J. C., and Nigg, E. A. (1990) Identification of major nucleolar proteins as candidate mitotic substrates of cdc2 kinase, *Cell 60*, 791−801.
13. Zhou, G., Seibenhener, M. L., and Wooten, M. W. (1997) Nucleolin is a protein kinase C-ζ substrate. Connection between cell surface signaling and nucleus in PC12 cells, *J. Biol. Chem. 272*, 31130−31137.
14. Najbauer, J., Johnson, B. A., Young, A. L., and Aswad, D. W. (1993) Peptides with sequences similar to glycine, arginine-rich motifs in proteins interacting with RNA are efficiently recognized by methyltransferase(s) modifying arginine in numerous proteins, *J. Biol. Chem. 268*, 10501−10509.
15. Fang, S. H., and Yeh, N. H. (1993) The self-cleaving activity of nucleolin determines its molecular dynamics in relation to cell proliferation, *Exp. Cell. Res. 208*, 48−53.
16. Kukuruzinska, M. A., and Lennon, K. (1998) Protein N-glycosylation: Molecular genetics and functional significance, *Crit. Rev. Oral Biol. Med. 9*, 415−448.
17. Sairam, M. R. (1989) Role of carbohydrates in glycoprotein hormone signal transduction, *FASEB J. 3*, 915−926.
18. Wells, L., and Hart, G. W. (2003) O-GlcNAc turns twenty: Functional implications for post-translational modification of nuclear and cytosolic proteins with a sugar, *FEBS Lett. 546*, 154−158.
19. Kreppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats, *J. Biol. Chem. 72*, 9308−9315.
20. Galland, S., Degiuli, A., Frot-Coutaz, J., and Got, R. (1998) Transfer of *N*-acetylglucosamine to endogenous glycoproteins in the nucleus and in non-nuclear membranes of rat hepatocytes: Electrophoretic analysis of the endogenous acceptors, *Biochem. Int. 7*, 59−67.
21. Pedemonte, C. H., Sachs, G., and Kaplan, J. H. (1990) An intrinsic membrane glycoprotein with cytosolically oriented N-linked sugars, *Proc. Natl. Acad. Sci. U.S.A. 87*, 9789−9793.
22. Reeves, R., and Chang, D. (1983) Investigations of the possible functions for glycosylation in the high mobility group proteins. Evidence for a role in nuclear matrix association, *J. Biol. Chem. 258*, 679−687.
23. Roquemore, E. P., Chou, T. Y., and Hart, G. W. (1994) Detection of O-linked *N*-acetylglucosamine (O-GlcNAc) on cytoplasmic and nuclear proteins, *Methods Enzymol. 230*, 443−460.
24. Kane, R., Murtagh, J., Finlay, D., Marti, A., Jaggi, R., Blatchford, D., Wilde, C., and Martin, F. (2002) Transcription factor NFIC undergoes N-glycosylation during early mammary gland involution, *J. Biol. Chem. 277*, 25893−25903.
25. Callebaut, C., Blanco, J., Benkirane, N., Krust, B., Jacotot, E., Guichard, G., Seddiki, N., Svab, J., Dam, E., Muller, S., Briand, J. P., and Hovanessian, A. G. (1998) Identification of V3 loop-binding proteins as potential receptors implicated in the binding of HIV particles to CD4(+) cells, *J. Biol. Chem. 273*, 21988−21997.
26. Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250, *Electrophoresis 6*, 255−262.
27. Vercoutter-Edouart, A. S., Lemoine, J., Le Bourhis, X., Louis, H., Boilly, B., Nurcombe, V., Revillion, F., Peyrat, J. P., and Hondermarck, H. (2001) Proteomic analysis reveals that 14−3−3σ is down-regulated in human breast cancer cells, *Cancer Res. 1*, 76−80.
28. Altin, J. G., and Pagler, E. B. (1995) A one-step procedure for biotinylation and chemical cross-linking of lymphocyte surface and intracellular membrane-associated molecules, *Anal. Biochem. 224*, 382−389.
29. Palade, G. E., and Siekevitz, P. (1956) Liver microsomes: An integrated morphological and biochemical study, *J. Biophys. Biochem. Cytol. 2*, 171−200.

30. Pons, A., Richet, C., Robbe, C., Herrmann, A., Timmerman, P., Huet, G., Leroy, Y., Carlstedt, I., Capon, C., and Zanetta, J. P. (2003) Sequential GC/MS analysis of sialic acids, monosaccharides, and amino acids of glycoproteins on a single sample as heptafluorobutyrate derivatives, *Biochemistry 42*, 8342−8353.
31. Maes, E., Strecker, G., Timmerman, P., Leroy, Y., and Zanetta, J. P. (1999) Quantitative cleavage of the N-glycosidic bond under the normal conditions of methanolysis used for the analysis of glycoprotein monosaccharides, *Anal. Biochem. 267*, 300−308.
32. Geng, M., Zhang, X., Bina, M., and Regnier, F. (2001) Proteomics of glycoproteins based on affinity selection of glycopeptides from tryptic digests, *J. Chromatogr., B: Biomed. Sci. Appl. 752*, 293−306.
33. Mechref, Y., and Novotny, M. V. (1998) Mass spectrometric mapping and sequencing of N-linked oligosaccharides derived from submicrogram amounts of glycoproteins, *Anal. Chem. 70*, 455−463.
34. Chen, C. M., Chiang, S. Y., and Yeh, N. H. (1991) Increased stability of nucleolin in proliferating cells by inhibition of its self-cleaving activity, *J. Biol. Chem. 266*, 7754−7758.
35. Heifetz, A., Keenan, R. W., and Elbein, A. D. (1979) Mechanism of action of tunicamycin on the UDP-GlcNAc:dolichyl-phosphate Glc-NAc-1-phosphate transferase, *Biochemistry 18*, 2186−2192.
36. Lotan, R., Beattie, G., Hubbell, W., and Nicolson, G. L. (1977) Activities of lectins and their immobilized derivatives in detergent solutions. Implications on the use of lectin affinity chromatography for the purification of membrane glycoproteins, *Biochemistry 16*, 1787−1794.
37. Hanover, J. A., Cohen, C. K., Willingham, M. C., and Park, M. K. (1987) O-linked *N*-acetylglucosamine is attached to proteins of the nuclear pore. Evidence for cytoplasmic and nucleoplasmic glycoproteins, *J. Biol. Chem. 262*, 9887−9894.
38. Yahara, I., and Edelman, G. M. (1972) Restriction of the mobility of lymphocyte immunoglobulin receptors by concanavalin A, *Proc. Natl. Acad. Sci. U.S.A. 69*, 608−612.
39. Shibuya, N., Goldstein, I. J., van Damme, E. J., and Peumans, W. J. (1988) Binding properties of a mannose-specific lectin from the snowdrop (*Galanthus nivalis*) bulb, *J. Biol. Chem. 263*, 728−734.
40. Hashimoto, C., Cohen, R. E., Zhang, W. J., and Ballou, C. E. (1981) Carbohydrate chains on yeast carboxypeptidase Y are phosphorylated, *Proc. Natl. Acad. Sci. U.S.A. 78*, 2244−2248.
41. Nilsson, B., Norden, N. E., and Svensson, S. (1979) Structural studies on the carbohydrate portion of fetuin, *J. Biol. Chem. 254*, 4545−4553.
42. Edge, A. S., and Spiro, R. G. (1987) Presence of an O-glycosidically linked hexasaccharide in fetuin, *J. Biol. Chem. 262*, 16135−16141.
43. Finne, J., and Krusius, T. (1979) Structural features of the carbohydrate units of plasma glycoproteins, *Eur. J. Biochem. 102*, 583−588.
44. Goldstein, I. J., and Hayes, C. E. (1978) The lectins: Carbohydrate-binding proteins of plants and animals, *Adv. Carbohydr. Chem. Biochem. 35*, 127−140.
45. Crowley, J. F., and Goldstein, I. J. (1982) *Datura stramonium* lectin, *Methods Enzymol. 83*, 368−373.
46. Takahashi, T., and Nishibe, H. (1981) Almond glycopeptidase acting on aspartylglycosylamine linkages. Multiplicity and substrate specificity, *Biochim. Biophys. Acta 657*, 457−467.
47. Michalski, J. C. (1996) Normal and pathological of glycoprotein, *New Compr. Biochem. 30*, 55−97.
48. Iyer, S. P., and Hart, G. W. (2003) Dynamic nuclear and cytoplasmic glycosylation: Enzymes of O-GlcNAc cycling, *Biochemistry 42*, 2493−2499.
49. Lefebvre, T., Ferreira, S., Dupont-Wallois, L., Bussiere, T., Dupire, M. J., Delacourte, A., Michalski, J. C., and Caillet-Boudin, M. L. (2003) Evidence of a balance between phosphorylation and O-GlcNAc glycosylation of Tau proteins−A role in nuclear localization, *Biochim. Biophys. Acta 1619*, 167−176.
50. Salazar, R., Brandt, R., Kellerman, J., and Krantz, S. (2000) Purification and characterization of a 200 kDa fructosyllysine-specific binding protein from cell membranes of U397 cells, *Glycoconjugate J. 17*, 713−716.
51. Krantz, S., Salazar, R., Kellerman, J., and Lottspeich, F. (1995) Purification and partial amino acid sequencing of a fructosyllysine-specific binding protein from cell membranes of the monocyte-like cell line U397, *Biochim. Biophys. Acta 1266*, 109−112.
52. Medina, L., and Haltiwanger, R. S. (1998) Calf thymus high mobility group proteins are non-enzymatically glycated but not significantly glycosylated, *Glycobiology 8*, 191−198.
53. Polet, H., and Molnar, J. (1988) Demonstration that some of the nonhistone proteins, inducible to translocate into the nucleus, are glycosylated, *J. Cell. Physiol. 135*, 47−54.
54. Yoshida, Y., Chiba, T., Tokunaga, F., Kawasaki, H., Iwai, K., Suzuki, T., Ito, Y., Matsuoka, K., Yoshida, M., Tanaka, K., and Tai, T. (2002) E3 ubiquitin ligase that recognizes sugar chains, *Nature 418*, 438−442.
55. Tsai, B., Ye, Y., and Rapoport, T. A. (2002) Retro-translocation of proteins from the endoplasmic reticulum into the cytosol, *Nat. Rev. Mol. Cell. Biol. 3*, 246−255.
56. Allain, F. H. T., Gilbert, D. E., Bouvet, P., and Feigon, J. (2000) Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target, *J. Mol. Biol. 303*, 227−236.

BI047831S