Structural Study of GCDFP-15/gp17 in Disease versus Physiological Conditions Using a Proteomic Approach†

Emilia Caputo,*,‡,§ Alessandra Camarca,§,|| Ramy Moharram,‡ Peter Tornatore,⊥ Bradley Thatcher,# John Guardiola,§ and Brian M. Martin‡

Unit on Molecular Structures, LNT, NIMH, NIH, 10 Center Drive, Building 10 3N309, Bethesda, Maryland 20892-1262, Institute of Genetics and Biophysics-I.G.B., A. Buzzati-Traverso-, CNR, Via G. Marconi 10, I-80125 Naples, Italy, Ciphergen Biosystems, Fremont, California, and Protein Solutions, 347 Maple Avenue, Stratford, ON, Canada N5A 7P6

Received January 9, 2003; Revised Manuscript Received March 24, 2003

ABSTRACT: Gross cystic disease fluid protein (GCDFP-15), also known as prolactin-inducible protein (PIP), is a specific breast tumor marker. GCDFP-15/PIP is also identified as gp17 and/or seminal actin-binding protein (SABP) from seminal vesicles and as extraparotid glycoprotein (EP-GP) from salivary glands. It is an aspartyl proteinase able to specifically cleave fibronectin (FN), suggesting a possible involvement in mammary tumor progression and fertilization. Other functions were attributed to this protein(s) on the basis of its ability to interact with an array of molecules such as CD4, actin, and fibrinogen. We investigated the structure of the protein(s) under disease versus physiological conditions by RP-HPLC chromatography, ProteinChip technology, and QStar MS/MS mass spectrometry. The proteins behaved differently when examined by RP-HPLC chromatography and surface-enhanced laser desorption ionization time-of-flight (SELDI-TOF) mass spectrometry, suggesting different conformations and/or tissue-specific posttranslational modifications of the proteins, although their primary structure was identical by MS/MS analysis. Both showed a single N-glycosylation site. A different N-linked glycosylation pattern was observed in pathological GCDFP-15/PIP as compared with physiological gp17/SABP protein by coupling enzymatic digestion and ProteinChip technology. Furthermore, taking advantage of ProteinChip technology, we analyzed the interaction of both proteins with CD4 and FN. We observed that the physiological form was mainly involved in the binding to CD4. Moreover, we defined the specific FN binding-domain of this protein. These data suggested that, depending on its conformational state, the protein could differently bind to its various binding molecules and change its function(s) in the microenviroments where it is expressed.

The gross cystic disease fluid protein-15 (GCDFP-15),[1] also known as prolactin-inducible protein (PIP), is a breast tumor marker (*1*−*3*). It is expressed specifically in human mammary gland diseases, such as various mastopathies, and in breast carcinomas, characterized by apocrine metaplastic changes in breast epithelium (*4*, *5*). GCDFP-15/PIP is not found in normal breast epithelium, but it is uniformly expressed in all normal apocrine glands (*4*). It also has been independently identified as gp17 and/or seminal actin-binding protein (SABP), a 17 kDa glycoprotein expressed in the glandular epithelium of seminal vesicles, and as extraparotid glycoprotein (EP-GP) in submandibular and sublingual glands (*6*−*8*).

Several functions have been attributed to this protein (*3*, *9*−*11*). We previously demonstrated that this factor is an aspartyl proteinase with specific fibronectin-degrading ability, suggesting its potential role in processes such as mammary tumor progression and fertilization (*3*). In addition, GCDFP-15/gp17 is found on the post-acrosomal region of ejaculated spermatozoa and remains bound to the sperm cell surface after capacitation, thus further suggesting a possible role in fertilization (*9*). It is also reported to be a potent inhibitor of T lymphocyte apoptosis induced by sequential cross-linking of CD4 and T cell receptor (*11*). Moreover, the functional relevance of its interaction with actin, fibrinogen, keratin, tropomyosin, and hydroxyapatite (a major component of tooth enamel) as well as with some strains of oral bacteria and *Staphylococcus hominis* from the skin and ear canal is presently not clear (*8*, *12*, *13*).

Previous studies from our laboratory demonstrated immunological similarities and differences between GCDFP-15/PIP and gp17/SABP/EP-GP from different sources (*14*),

† This work was partly supported by a grant from the Italian Association for Cancer Research (AIRC) to J.G.

* To whom correspondence should be addressed. Fax: 301 480 0198. E-mail: caputo@iigbna.iigb.na.cnr.it.

‡ NIH.

§ Institute of Genetics and Biophysics-I.G.B.

|| A.C. and E.C. contributed equally to this work.

⊥ Ciphergen Biosystems.

Protein Solutions.

[1] Abbreviations: GCDFP-15, gross cystic disease fluid protein; PIP, prolactin-inducible protein; SABP, seminal actin binding protein; EP-GP, extra parotid glycoprotein; BCF, breast cyst fluid; HSP, human seminal plasma; SELDI-TOF MS, surface-enhanced laser desorption/ionization−time-of-flight mass spectrometry; SPA, sinapinic acid; CHCA, α-cyano-4-hydroxy-cinnamic acid; IAA, iodoacetic acid; ACTH, adrenocorticotropic hormone; PNGase F, Peptide N-glycosidase F; DTT, dithiothreitol; Gal, galactose; GlcNAc, N-acetyl glucosamine; MALDI-TOF, matrix-assisted laser desorption/ionization−time-of-flight.

10.1021/bi034038a CCC: $25.00
Published on Web 04/29/2003

although the purified proteins from BCF and HSP exhibited an identical electrophoretic mobility, by SDS−PAGE. Each showed two bands, under reducing conditions, in positions corresponding to an apparent molecular mass of about 17 and 13 kDa, representing the glycosylated and unglycosylated species, respectively (*6*, *7*).

In this report, to elucidate the differences between GCDFP-15 and gp17, expressed in pathological versus physiological tissues respectively, we analyzed the protein from different sources by coupling RP-HPLC chromatography, ProteinChip technology (*15*), and MS/MS (*16*).

We found that the protein from pathological versus physiological sources exhibited a disparate behavior by RP-HPLC chromatography and SELDI-TOF MS analysis, suggesting that the different structural properties may be a consequence of varying conformations and/or tissue-specific posttranslational modifications. Furthermore, the direct mass analysis by SELDI-TOF MS of glycopeptides derived from GCDFP-15/PIP and gp17/SABP revealed a different N-linked glycosylation pattern. A preliminary structural analysis of the N-linked glycans in the two glycoproteins from pathological and normal sources was also carried out by mass analysis of their tryptic glycopeptides before and after treatment with various exoglycosidases and of permethylated N-glycans. The majority of complex oligosaccharides in gp17/SABP presented monosialylated diantennary structures with different degrees of fucosylation, whereas triantennary structures were also detected in pathological GCDFP-15/PIP.

In addition, the interaction with molecules such as CD4 and fibronectin was examined by using ProteinChip technology, showing a differential binding ability of protein from the two sources. In particular we identified the fibronectin binding domain of this protein in two regions which did not overlap with the CD4-binding domain. The first region contained the amino acid residues 42−57 and the second contained the C-terminal residues 109−118.

MATERIALS AND METHODS

Materials. Trypsin and fibronectin were purchased from Roche Molecular Biochemicals (Mannheim, Germany). Recombinant soluble CD4 molecules were obtained from Genentech (South San Francisco, CA). H4 and PS10 ProteinChips and the calibration standard molecules for the SELDI-TOF mass spectrometer were purchased from Ciphergen Biosystems (Fremont, CA). Sinapinic acid (SPA), α-cyano-4-hydroxy-cinnamic acid (CHCA), and iodoacetic acid (IAA) were obtained from Sigma (St. Louis, MO). Peptide N-glycosidase F (PNGase F) was obtained from NEN Research Products (Boston, MA). Dithiothreitol (DTT) was from GIBCO BRL (Grand Island, NY). Sialidase from *Arthrobacter ureafaciens* was purchased from Oxford GlycoSciences (Bedford, MA). All other chemicals used were the highest purity available.

Purification of GCDFP-15/PIP and gp17/SABP from BCF and HSP, Respectively. The protein GCDFP-15/PIP was isolated from breast cyst fluid obtained by needle aspiration from women with gross cystic disease. The breast cyst fluid proteins were ultrafiltered and dialyzed against 20 mM monoethanolamine buffer, pH 9.4 (buffer Q) by Pellicon XL filters (Millipore, Bedford, MA), with a 10 000 Da cutoff. The proteins were loaded onto a Q-Fast Flow column (Pharmacia, Uppsala, Sweden), and the material was eluted with a linear gradient of 0−0.5 M NaCl in buffer Q at a flow rate of 1 mL/min. The material eluted between 0.25 and 0.3 M NaCl was pooled, dialyzed against 20 mM monoethanolamine buffer, pH 9.4, and loaded onto a mono-Q HR 5/5 (Pharmacia). GCDFP-15/PIP, obtained from this second chromatographic step, was then purified by gel filtration on Superdex 75 (HL 16/60) column (Pharmacia) by using a FPLC apparatus from Pharmacia, as previously reported (*17*). The gp17/SABP was purified from seminal fluids obtained from healthy donors, using the same GCDFP-15/PIP purification procedure.

Reverse-Phase HPLC Chromatography. Both proteins, after gel filtration, were loaded on RP-HPLC C18 column (150 × 2.1 mm, 300 Å pore size, Vydac, Hesperia, CA). A Beckman system Gold (Fullerton, CA) composed of a model 168 Diode Array Detector and a Model 126 pump module were used. A linear gradient from 5% to 65% B solvent over 60 min, with 5 min 5% B initial condition (where solvent A was 0.12% aqueous trifluoroacetic acid and solvent B was 0.1% trifluoroacetic acid in acetonitrile) was performed at flow rate 0.2 mL/min to analyze GCDFP-15/PIP and gp17/SABP from the two sources. The peptides deriving from tryptic digestion were fractionated by a linear gradient from 5% to 70% B solvent over 70 min preceded by a 5 min hold at 5% B, where A and B were the same solvents described above. All peaks were manually collected, lyophilized in a Speed-Vac centrifuge (Savant, Holbrook, NY), numbered according to the elution time and stored at −20 C until their subsequent analysis.

Surface-Enhanced Laser Desorption Ionization Time-of-Flight (SELDI-TOF) Mass Spectrometry. The instrument was externally calibrated using the $[M + H]^+$ ion peaks of the bovine insulin at *m*/*z* 5733.6 and human angiotensin at *m*/*z* 1296.5 for peptide analysis and bovine cytochrome C at *m*/*z* 12230.9 and bovine superoxide dismutase at *m*/*z* 15591.4 for protein analysis. All mass spectra were recorded in the positive-ion mode using a Ciphergen PBS II ProteinChip Array reader, a linear laser desorption/ionization-time-of-flight mass spectrometer with time-lag focusing (*15*). Prior to SELDI-TOF MS analysis, either 1 μL of CHCA or SPA solution (in 0.1% trifluoroacetic acid in 50% aqueous acetonitrile) was added to each feature of ProteinChip surface for peptide or protein analysis, respectively. Raw data were analyzed using the computer software provided by the manufacturer and are reported as average masses.

Digestion of Proteins in Solution. The proteins, after RP-HPLC chromatography on the C18 column, were subjected to digestion with trypsin. Initially, they were denatured in 6 M GuCl and subjected to a reduction and alkylation reaction by treatment with 45 mM dithiothreitol (DTT) for 20 min at 55 C and then 100 mM iodoacetic acid (IAA) for 15 min at room temperature in the dark. Following removal of GuCl by dialysis, using dialysis membranes with 3500 Da cutoff (Pierce, Rockford, IL), the reduced and alkylated proteins were cleaved with trypsin by using an Enzyme/Substrate ratio of 1:100. The tryptic digestion was carried out in 50 mM ammonium bicarbonate, pH 7.8, at 37 C overnight. In solution PNGase F digestion of glycopeptides was performed in 50 μL of 100 mM sodium phosphate, pH 8.5 for 18 h at 37 C, by using 0.4 units of enzyme.

SELDI Quadrupole Time-of-Flight Tandem Mass Spectrometry. Tandem mass spectrometric peptide sequencing was accomplished using an Applied Biosystems/Sciex QStar triple quadrupole time-of-flight instrument (Toronto, Canada) equipped with a Ciphergen SELDI ProteinChip Interface. The instrument was calibrated externally using an acquired MS/MS spectrum of ACTH $_{(18-39)}$ peptide at *m*/*z* 2465.2, where four fragment ions and the parent were used as calibration points. All mass spectra were acquired in positive-ion mode with a source collisional cooling gas pressure of approximately 100 mTorr. MS/MS data were acquired using collision energies following the approximate rule of 50 eV/kDa parent peptide mass. Matrix conditions were identical to SELDI-TOF analysis described previously. Raw data were analyzed using the instrument's Analyst software.

Amino Acid Sequencing of GCDFP-15/gp17 by Edman Degradation. N-terminal sequence analysis of the proteins and peptides was performed on an ABI Procise protein sequencer (Foster City, CA), according to the manufacter's pulse-phase protocol.

N-Glycans Analysis. Total N-linked glycans released by PNGase F were separated from peptides by reverse phase chromatography on Sep-Pak cartridges (Waters, Milford, MA) equilibrated with 5% acetic acid. The oligosaccharides were eluted in the aqueous phase and freeze-dried. Glycans were permethylated according to the procedure described by Dell (*18*). In short, 0.5−1 mL of dimethyl sulfoxide/NaOH solution was added to the oligosaccharide sample followed by about 0.5 mL of methyl iodide. The mixture was shaken for 10 min at room temperature and then quenched with 1 mL of water. The products were extracted into chloroform then washed several times with water. The chloroform layer was dried under a stream of nitrogen, and the residue was dissolved in methanol for mass analysis, by using a Voyager DE MALDI-TOF mass spectrometer (ABI), equipped with delayed extraction.

Chemical Reaction on ProteinChip Surfaces. The on-chip exoglycosidase digestion was performed after the glycopeptides had been spotted on the H4 ProteinChip surface and analyzed. In short, for PNGase digestion, 4 μL of 100 mM sodium phosphate, pH 8.5, and 0.1 unit of enzyme were used for the reaction on each feature of ProteinChip array; for the hydrolysis reaction of the glycopeptides with sialidase from *A. ureafaciens*, 4 μL of 50 mM ammonium acetate, pH 5.5 and 0.1 unit of sialidase were used. All hydrolysis reactions were carried out for 1 h in a humid chamber. After incubation, the reaction buffers on each feature of the surface was removed by washing with 5% aqueous acetonitrile solution and 1 μL of CHCA solution was added for the subsequent analysis by SELDI-TOF-MS.

Interaction Analysis of GCDFP-15/PIP and gp17/SABP with CD4 and Fibronectin on the ProteinChip Surface. CD4 and FN were covalently attached to PS10 preactivated ProteinChip surface arrays. These capture surfaces were prepared following the manufacturer's instructions. In short, CD4 and FN (in PBS) were incubated on the PS10 preactivated surface for 18 h in a humid chamber. A solution of PBS was used to prepare the blank capture surfaces without CD4 and FN covalently attached, to check for the binding specificity of this assay. After the coupling reactions of CD4, FN, and their solvent, the activated groups on each feature of the array were blocked by treatment with 1 M ethanolamine in PBS for 30 min in a humid chamber. After washing with PBS buffer, GCDFP-15/PIP and gp17/SABP, obtained from the gel-filtration purification step, were added on each feature of CD4 or FN or blank specific capture surfaces and incubated for 4 h in a humid chamber. After incubation, the surfaces were washed twice with PBS containing 0.05% Tween 20 and twice with PBS. The interaction of the tryptic peptides with CD4 and FN was also analyzed using the same procedure. SPA and CHCA solutions were used as matrix.

RESULTS

RP-HPLC and SELDI-TOF Mass Spectrometry Analysis of GCDFP-15/PIP and gp17/SABP from BCF and HSP. GCDFP-15/PIP and gp17/SABP were previously isolated from breast cystic fluid (BCF) and human seminal plasma (HSP) and appeared as dimeric and/or tetrameric forms (*13*, *19*−*20*). Although both exhibited identical electrophoretic mobility by SDS−PAGE, we have previously demonstrated differential immunological reactivity between the proteins from different sources, as well as their differential ability to bind to molecules such as CD4 and actin by Biacore analysis (*14*).

To obtain further structural information, we analyzed these proteins by RP-HPLC chromatography and by SELDI-TOF mass spectrometry. To this end, we purified GCDFP-15/PIP and gp17/SABP from BCF and HSP by using a combination of ion-exchange and gel-filtration chromatography as previously reported (*17*) and as described in Materials and Methods. As shown in Figure 1, we observed that gp17/SABP isolated from HSP eluted as a sharp peak from C18 RP-HPLC column, while GCDFP-15/PIP from BCF appeared as two broad peaks. These molecular species were analyzed by SELDI-TOF mass spectrometry on the H4 ProteinChip surface (Figure 2). The analysis showed the gp17/SABP (the sharp peak, obtained from RP-HPLC) as a molecular species with *m*/*z* ratio of 15902 and two minor species with *m*/*z* values of 14596.8 and 16406. The initial peak (49.5 min) of GCDFP-15/PIP from BCF presented as a broad peak centered at *m*/*z* 15940 and minor molecular species with *m*/*z* 14032.5. Interestingly, the GCDFP-15/PIP exhibited a decreased ability to ionize, although the same amount of protein was examined by mass spectrometry. Furthermore, the GCDFP-15/PIP molecular species that eluted from C18 as a broad shoulder peak at 54.8 min (Figure 1) showed the same molecular species found in the initial peak upon mass spectrometric analysis, suggesting a complex heterogeneity of this pathological species compared to gp17/SABP deriving from HSP.

Both gp17/SABP and GCDFP-15/PIP eluted from Superdex 75 gel filtration column at retention times which supported their identification as dimeric species. However, when examined on SELDI-TOF, only GCDFP-15/PIP showed a minor molecular species at *m*/*z* value of 30 880.3, while gp17/SABP and the majority of GCDFP-15/PIP gave *m*/*z* around 15 000 (data not shown).

Amino Acid Sequence Analysis of GCDFP-15/PIP from BCF. Although identical cDNA clones have been isolated from a cDNA library derived from steroid- and prolactin-treated T47D human breast cancer cell line (*2*), from a human salivary gland cDNA library (*21*), and from human seminal vesicles (*22*), conflicting data at protein level are reported.

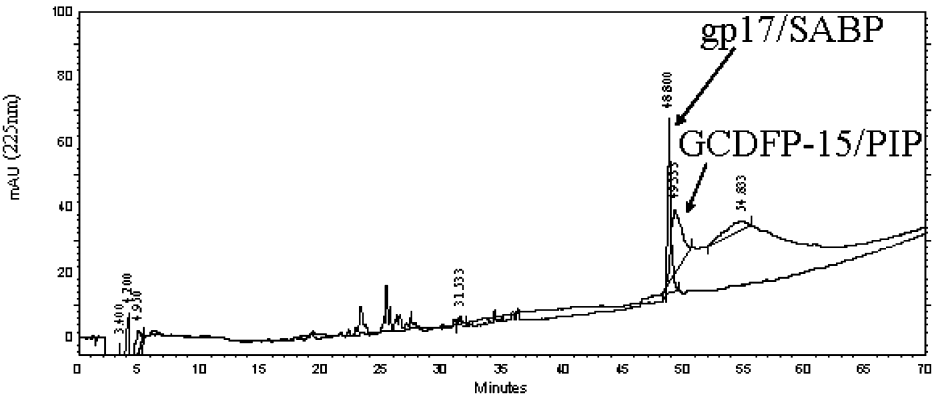

FIGURE 1: HPLC profiles of GCDFP-15/PIP and gp17/SABP proteins. 50 μg of protein from different sources was loaded on a C18 reverse-phase HPLC column. After injection, isocratic conditions were applied for 5 min at the initial 5% solvent B, followed by a gradient to 65% solvent B over 60 min.

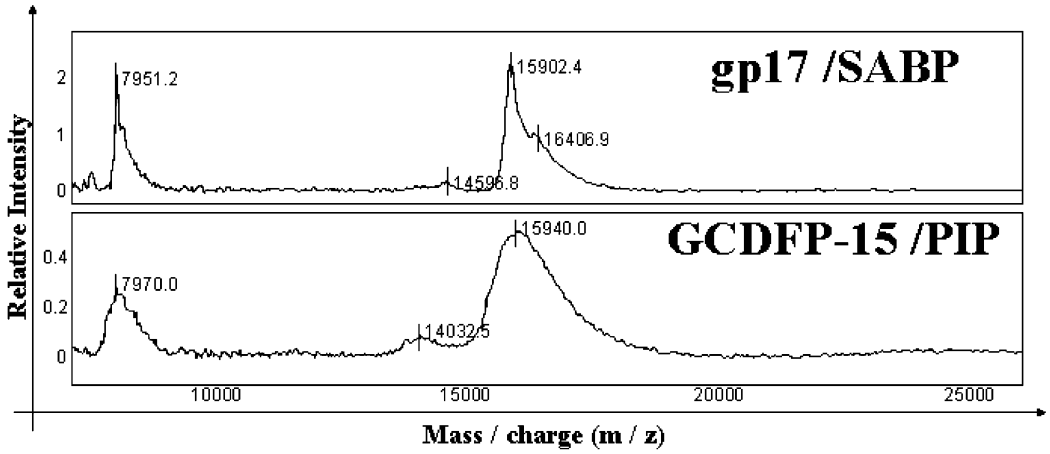

FIGURE 2: SELDI-TOF MS spectra of GCDFP-15/PIP and gp17/SABP proteins. Each protein from the major peak obtained by elution on C18 RP-HPLC column was spotted on H4 ProteinChip and allowed to air-dry. After washing with 5% acetonitrile, each feature of array was allowed to air-dry. 1 μL of saturated sinapinic acid was added and air-dried and the sample examined by SELDI-TOF mass spectrometry. The mass/charge (*m*/*z*) values of each detected species were reported.

Analysis of the C-terminus of gp17/SABP from HSP gave a Glu residue (*12*, *20*), in contrast to a Val residue for GCDFP-15/PIP from BCF, obtained by Vandewalle et al. (*23*).

To investigate the primary structure of the GCDFP-15/PIP protein expressed in pathological tissues, 50 μg of protein was subjected to tryptic digestion, and the resulting peptides were separated by RP-HPLC chromatography (Figure 3A) as described in Materials and Methods. The fractions, eluted from RP-HPLC C18 column, were examined by SELDI-TOF MS (Figure 3 B). We observed peptides with molecular masses consistent with the expected tryptic peptides, except for the *m*/*z* 1255 molecular species from fraction 14, which could be attributed to the partially cleaved C-terminal 109−118 peptide. Taking advantage of the ProteinChip Interface (*15*), we directly analyzed on the ProteinChip surface the sequence of peptides by MS/MS, using a Qstar mass spectrometer.

As shown in Figure 3C, the obtained collision-induced dissociation (CID) spectra of GCDFP-15/PIP derived tryptic peptide with a *m*/*z* value of 1254.7 revealed the sequence FYTIEILKVE, identical to the C-terminal 109−118 peptide from human seminal plasma (*12*, *20*). This finding demonstrated that both proteins had identical C-terminal amino acid sequences. The sequences from the peptides with molecular mass of *m*/*z* 1554.3 and 1283.8 were also determined by MS/MS and corresponded to the 91−105 and 79−90 peptides, respectively (data not shown). The *m*/*z* 1817.3 peptide sequenced by Edman degradation analysis was identified as the 42−57 peptide.

SELDI-TOF Mass Spectrometry Analysis of GCDFP-15/PIP and gp17/SABP-Derived Glycopeptide$_{(69-78)}$. SELDI-TOF MS analysis of tryptic digestion products from both proteins, previously fractionated, showed that fraction 10 (Figure 3A) contained molecular species with mass/charge in a range between 2600 and 4200 (Figure 4A,B). These species could be partial tryptic digestion fragments and/or glycosylated forms of the protein.

To investigate this point, we subjected these molecular species to digestion with PNGase F to remove any associated N-glycans. To this end, we added PNGase F directly on the chip surface, where the molecules were previously spotted and the reaction was carried out in a humid chamber for 1 h, as described in Materials and Methods. As shown in Figure 4C, after the deglycosylation reaction of both fractions, we detected a major molecular species with *m*/*z* 1413.8, which is consistent with the mass expected for the deglycosylated tryptic peptide$_{(69-78)}$.

Characterization of N-Glycans from GCDFP-15/PIP and gp17/SABP. The differences observed in the spectra obtained

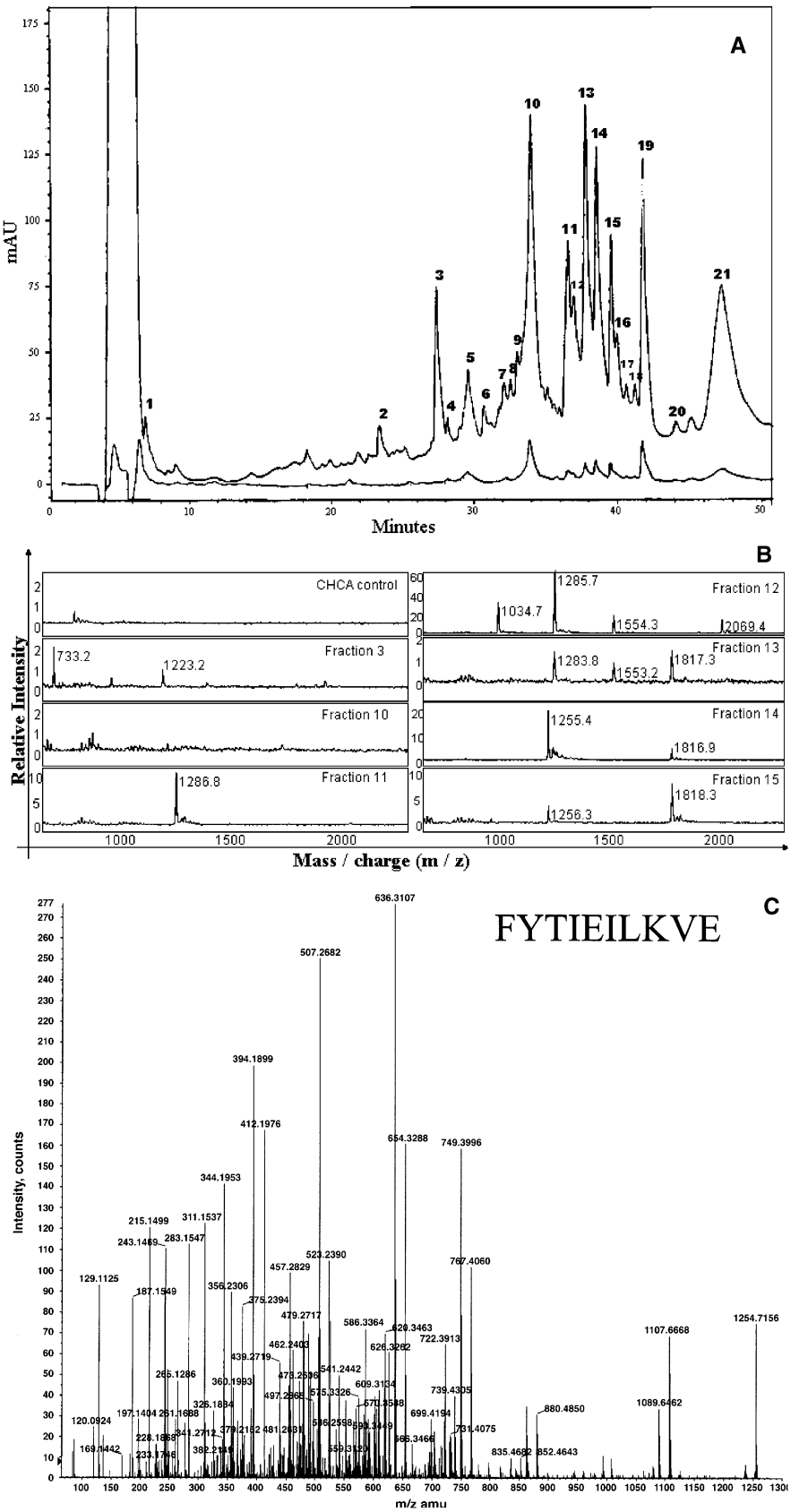

FIGURE 3: RP-HPLC separation, SELDI-TOF MS, and CID spectra of GCDFP-15/PIP derived tryptic peptides. Separation of peptides generated by cleavage with trypsin of reduced and alkylated GCDFP-15/PIP from BCF. The tryptic peptides were separated on a C18 column using a linear acetonitrile gradient, as described in Materials and Methods. The upper trace and the lower trace indicated the absorbance at 220 and 280 nm, respectively (A). The fractions, obtained by RP-HPLC chromatography, as described in the text, were spotted on H4 ProteinChip and allowed to air-dry. After washing with 5% acetonitrile, each feature of ProteinChip array was allowed to air-dry. 1 μL of CHCA saturated solution was added, and after it was air-dried, SELDI-TOF mass spectrometry was performed for each fraction. The mass/charge (*m*/*z*) values of each detected species were reported (B). Collision-induced dissociation (CID) spectra from 1254 (*m*/*z*) molecular species with the determined sequence presented (C).

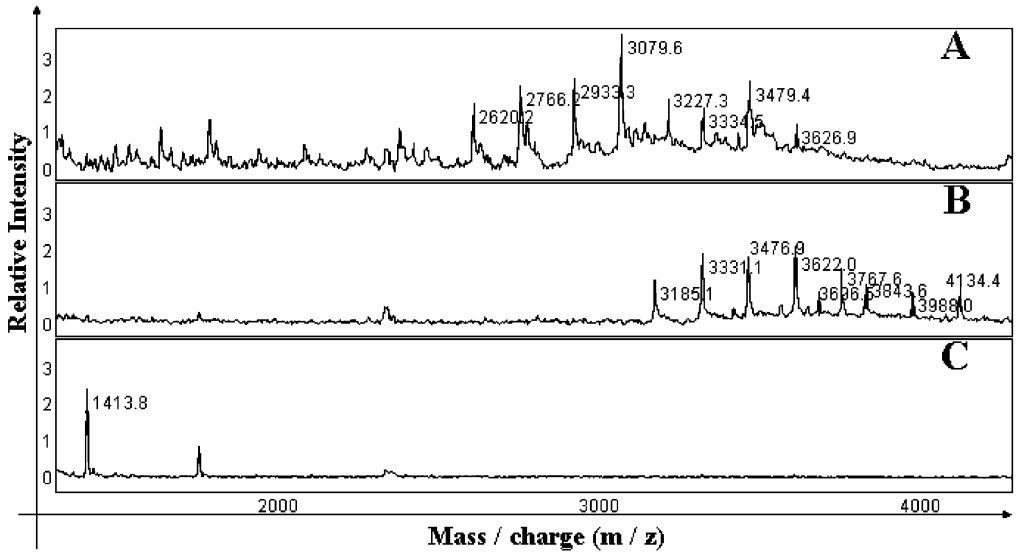

FIGURE 4: PNGase F hydrolysis on ProteinChip surface. SELDI-TOF MS spectra of fraction 10 of gp17/SABP and GCDFP-15/PIP derived tryptic peptides before (A and B) and after (C) treatment with PNGase F glycosidase on H4 ProteinChip array. The mass/charge (*m*/*z*) values of each detected species were reported.

Table 1: Oligosaccharide Structures Observed in the MALDI Spectra of the Permethylated Glycans Released from Gp17/SABP and GCDFP-15/PIP Glycopeptide$_{(69-78)}$

	MNa^+	structure
gp17/SABP	2430.8	mono-sialylated diantennary
	2605.2	fucosylated mono-sialylated diantennary
	2779.6	difucosylated mono-sialylated diantennary
	2953.7	trifucosylated mono-sialylated diantennary
GCDFP-15/PIP	2429.8	mono-sialylated diantennary
	2604.1	fucosylated mono-sialylated diantennary
	2791.4	disialylated diantennary
	2965.1	fucosylated disialylated diantennary
	3227.8	difucosylated mono-sialylated triantennary

from GCDFP-15/PIP and gp17/SABP-derived glycopeptide$_{(69-78)}$ (Figure 4A,B) prompted us to further investigate their posttranslational modification.

To this end, the structure of the overall glycosidic moiety of GCDFP-15/PIP and gp17/SABP from both sources was examined by mass analysis of the permethylated oligosaccharide mixture (*18*). The intact N-linked oligosaccharides from both glycopeptides were released by PNGase F treatment in solution, as described in Materials and Methods. The glycan mixture was separated from the peptides by a Sep-Pak reverse phase chromatographic step, permethylated and analyzed by matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry. The different glycan structures were identified by their masses on the basis of the known biosynthetic pathway of N-linked glycan structures that are reported in Table 1. N-linked glycans from both proteins constituted a heterogeneous mixture of complex type structures. In particular, the abundant species detected from GCDFP-15/PIP were consistent with mono- and disialylated, fucosylated, and nonfucosylated diantennary structures. Minor components corresponding to fucosylated and nonfucosylated, monosialylated triantennary structures were also detected. However, the most abundant species from gp17/SABP consisted essentially of fucosylated diantennary structures.

Exoglycosidase Digestion of Glycopeptide$_{(69-78)}$ from GCDFP-15/PIP and gp17/SABP on ProteinChip Surface. To confirm the hypothetical structures, we examined the glycopeptides deriving from the proteins from both sources before and after treatment with exoglycosidases by SELDI-TOF MS. In particular, both glycopeptides were subjected to hydrolysis with sialidase from *Arthrobacter ureafaciens*. Initially, GCDFP-15/PIP derived glycopeptide$_{(69-78)}$ was analyzed by SELDI-TOF MS on a H4 ProteinChip surface (Figure 5, panel A). The hydrolysis reaction was carried out for 1 h in a humid chamber in the appropriate reaction buffer, as described in Materials and Methods. After incubation, 1 μL of CHCA was added, and the sample was analyzed. After hydrolysis of GCDFP-15/PIP derived glycopeptide$_{(69-78)}$ with sialidase, we observed the disappearance of major molecular species at *m*/*z* value of 4131.3, 3753.4, 3627.1, 3477.9, 3424.3, and 3327.5 and the appearance of molecular species at *m*/*z* of 3841.9, 3183.9, and 3139.3 in the spectra, consistent with the removal of sialic acid residue(s) from the glycoforms examined (Figure 5, panel B). This further supported the data from our mass analysis of permethylated glycans (Table 1).

The glycopeptide$_{(69-78)}$ from gp17/SABP was also examined using the same procedure. A single sialic acid residue on these glycoforms was detected (data not shown) and consistent with the mass analysis of permethylated sugars (reported in Table 1).

Analysis of CD4 Interaction with GCDFP-15/PIP and gp17/SABP by ProteinChip Technology. We have previously demonstrated a differential binding ability to CD4 for GCDFP-15/PIP and gp17/SABP from different sources by using Biacore analysis (*14*).

To investigate the molecular species of both proteins involved in the binding to CD4, we analyzed this interaction by ProteinChip technology. To this end, we covalently coupled CD4 molecules on PS10 chip surface, as described in Materials and Methods. GCDFP-15/PIP and gp17/SABP were incubated on this specific CD4-chip surface for 4 h in a humid chamber. The surfaces were washed, and 1 μL SA was added to each spot and the bound material was analyzed by SELDI-TOF MS (*15*). As shown in Figure 6 A, all the molecular species from both proteins were detected after

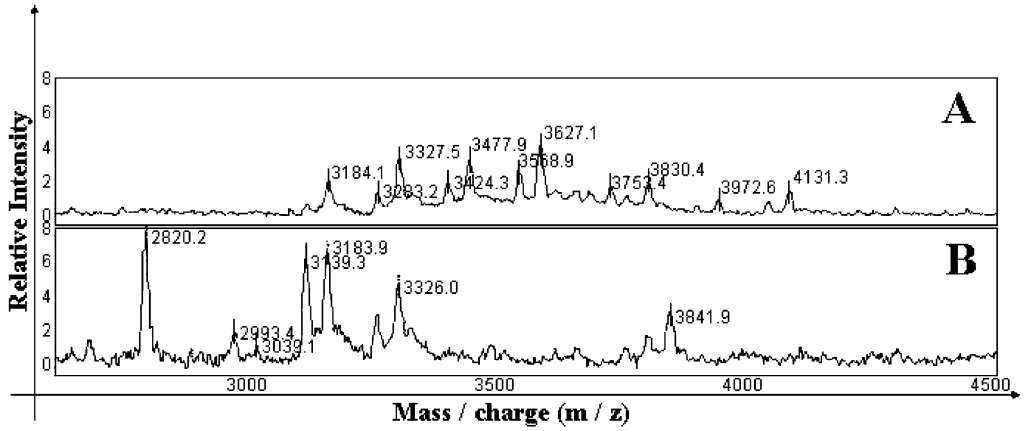

FIGURE 5: SELDI-TOF MS spectra of GCDFP-15/PIP derived glycopeptide$_{(69-78)}$ before and after treatment with sialidase from *A. ureafaciens* on ProteinChip surface. 1 μL of sample was spotted on H4 ProteinChip and allowed to air-dry. After washing with 5% acetonitrile, each feature of array was allowed to air-dry. 1 μL of saturated CHCA was added, and after it was air-dried, the sample was examined by SELDI-TOF mass spectrometry (panel A). After the analysis, the reaction buffer and enzyme were added to the sample, and the hydrolysis was carried out as described in the text. After hydrolysis, the buffer reaction and enzyme were removed, and 1 μL of saturated CHCA was added, and after it was air-dried, the sample was analyzed by SELDI-TOF MS (panel B). The mass/charge (*m*/*z*) values of each detected species were reported.

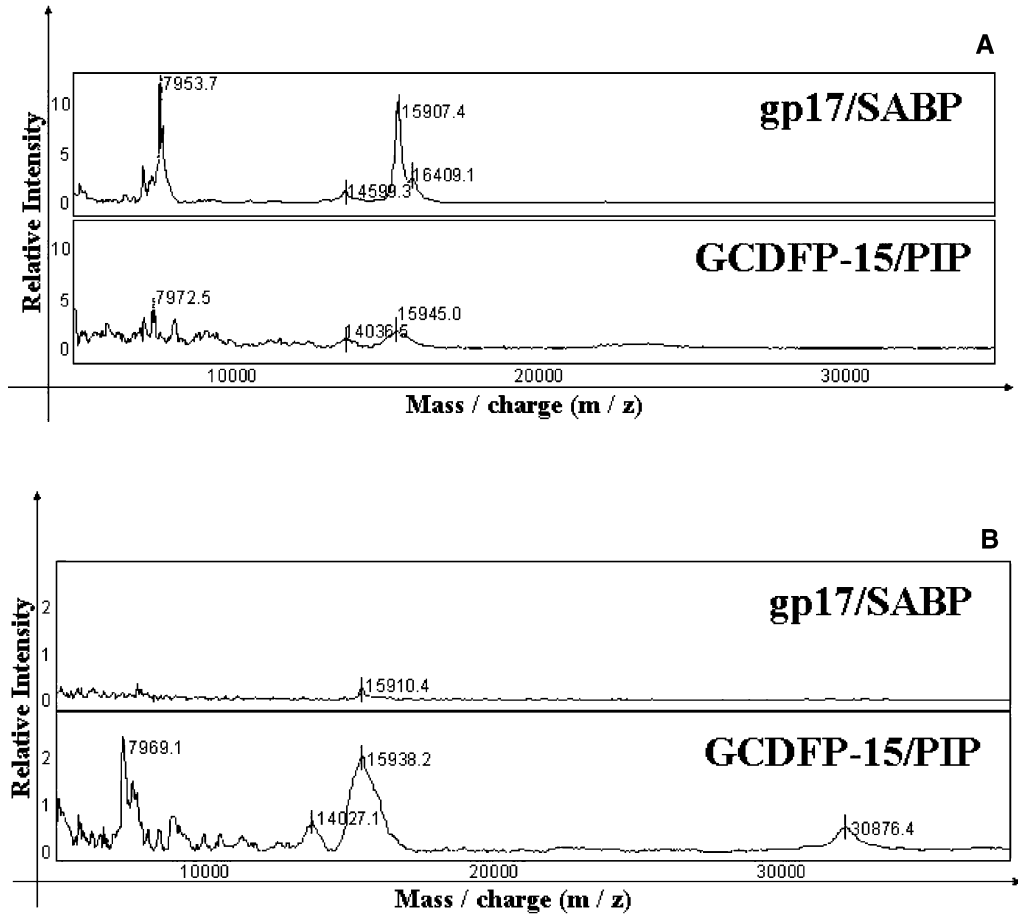

FIGURE 6: SELDI-TOF MS spectra of GCDFP-15/PIP and gp17/SABP on CD4 capture ProteinChip Surface. 100 ng of GCDFP-15/PIP and gp17/SABP were incubated on CD4 ProteinChip capture-surface, as described in the text. After appropriate washing, the bound material was detected by SELDI-TOF mass spectrometry (A). The unbound material was spotted on H4 ProteinChip and allowed to air-dry. After appropriate treatment SELDI-TOF mass spectrometry was performed (B). The mass/charge (*m*/*z*) values of each detected species were reported.

binding to CD4-chip surface. The binding specificity was tested by incubation of the same amount of protein on a blank PS10 ProteinChip surface without CD4, prepared as described in Materials and Methods. No bound material was detected on this suface by SELDI-TOF MS analysis.

The unbound material from the CD4 ProteinChip surface was spotted on the H4 ProteinChip surface and revealed that GCDFP-15/PIP was partially able to bind CD4, while gp17/SABP appeared to be more completely bound to CD4 (Figure 6 B). Interestingly, the unbound material from GCDFP-15/

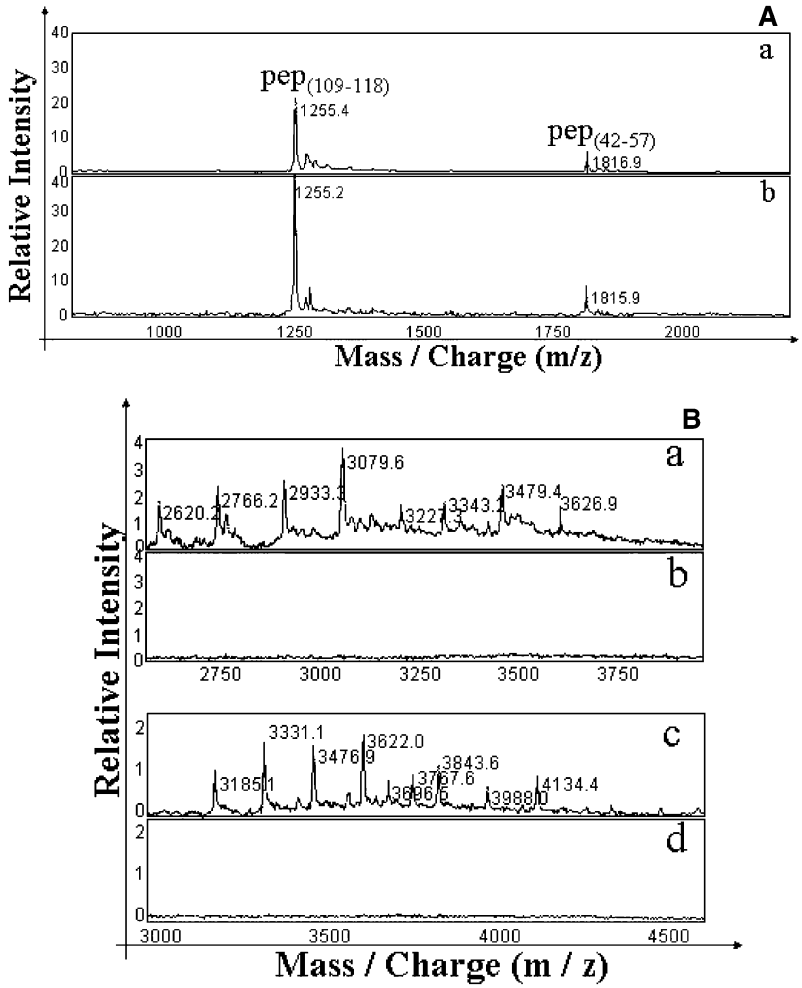

FIGURE 7: SELDI-TOF MS spectra of GCDFP-15/PIP- and gp17/SABP-derived tryptic peptides on FN capture ProteinChip Surface. The $pep_{(109-118)}$ and $pep_{(42-57)}$ peptide mixture (panel a) were incubated on FN (panel b) capture ProteinChip surface as described in the text. After appropriate washing, the bound material was examined by SELDI-TOF MS analysis (A). The gp17/SABP- (panel a) and GCDFP-15/PIP- (panel d) derived $glycopeptide_{(69-78)}$ were tested for their ability to bind to FN (panel c and e), respectively (B).

PIP showed a similar molecular species pattern to that observed after Superdex 75. In contrast, both proteins appeared to effectively bind to the FN molecule (data not shown).

GCDFP-15/PIP and gp17/SABP Mapping Region of Binding to FN on ProteinChip Surface. To get more information about the region on GCDFP-15/PIP and gp17/SABP involved in the binding to FN molecules, the tryptic peptides from both proteins were assayed for their ability to bind to these molecules by SELDI-TOF MS.

FN molecule was covalently coupled on the PS10 ProteinChip surface (*15*). After fractionation of both sets of tryptic peptides by RP-HPLC chromatography, they were dried, resuspended in PBS buffer, and then incubated on specific FN ProteinChip arrays for 4 h in a humid chamber. After incubation, each feature of array was washed as described in Materials and Methods and allowed to air-dry. A 1 μL sample of CHCA was added to each spot, and the material bound was detected by SELDI-TOF MS analysis. We observed that the 109−118 (1254.4 *m*/*z*) and 42−57 (1816 *m*/*z*) peptides (Figure 7A, panel a) exhibited the ability to interact with FN (Figure 7A, panel b). In addition, our assay showed that both the gp17/SABP and GCDFP-15/PIP derived $glycopeptides_{(69-78)}$ (Figure 7B, panels a and c, respectively) did not bind to FN (Figure 7B, panels b and d, respectively).

DISCUSSION

GCDFP-15/PIP/gp17/SABP/EP-GP is an aspartyl proteinase. GCDFP-15/PIP is found in the cystic fluid secretions from patients with gross cystic disease and is expressed in primary and metastatic breast tumors with apocrine characteristics (*4*, *5*). Gp17/SABP/EP-GP is secreted from epithelium cells of all apocrine glands such as lacrimal, submandibular, as well as seminal vesicles (*3*, *6*−*8*, *13*). Many different properties have been attributed to these proteins based on their peculiar ability to bind to different molecules such as CD4, actin, fibrinogen, and fibronectin, although their function is not clearly understood (*3*, *8*−*13*).

We have previously demonstrated similarities and differences among the proteins isolated from different sources such as human seminal plasma and breast cyst fluid of patients having gross cystic disease (*14*). These differences could help to explain the function of this protein in different situations. In this study, we addressed the analysis of its structural properties by using a combination of analytical tools, such as RP-HPLC chromatography, ProteinChip technology (*15*), and QStar MS/MS mass spectrometry (*16*).

Chromatographic ProteinChip surfaces provided the best platform for protein identification and characterization by direct enzymatic digestions for several reasons (*15*). They retained the substrate, trapped reaction products, facilitated sample clean up, and allowed the sample to be placed in the PBS II or QSTAR (via the interface). Little or no modification of standard solution methods was required when using ProteinChip surfaces. Reactions with changes in mass could be followed by SELDI-TOF, providing a complete experimental platform; reaction, detection and identification.

We demonstrated that both proteins, isolated from different biological fluids using the same purification procedure, shared the same primary structure. In particular, we analyzed the tryptic peptides derived from GCDFP-15/PIP from BCF by MS/MS analysis on QStar mass spectrometer and Edman degradation analysis. We found C-terminal sequence identity between the pathological and physiological form of the protein.

In addition, we found differences between the proteins when examined by RP-HPLC chromatography and ProteinChip technology. We observed a different elution profile from a reverse phase column and different spectra from SELDI-TOF MS analysis between the GCDFP-15/PIP and gp17/SABP proteins. These differences could be attributed to different conformations of the proteins and/or their molecular heterogeneity (*14*). We identified different glycosylation patterns between proteins by direct SELDI-TOF MS analysis of glycopeptides deriving from both. This may explain the different behavior observed between two species.

Interestingly, the analysis of interaction with CD4 of both showed a differential ability to bind to this molecule. GCDFP-15/PIP showed poor binding to CD4 compared with the gp17/SABP physiological form further supporting the previously reported Biacore data (*14*). This would appear unrelated to the different glycoforms of the two proteins, since we observed that their glycopeptides were not directly involved in the binding to CD4.

Previously, the CD4-binding domain of gp17/SABP was identified as the peptides, encompassing the 1−35 amino acids N-terminal region and 78−105 region by Spot method combined with mutagenesis (*24*). In this study, we identified the region of the two proteins involved in the interaction with fibronectin. We demonstrated that the C-terminal tryptic peptide 109−118 and the peptide 42−57 were able to bind to FN. This specific ability to interact with its binding partners may be relevant toward further investigation of its function. The above-mentioned peptides able to bind the fibronectin and not CD4 could be used as natural ligands for example in biological assay to test their ability to inhibit the GCDFP-15/gp17 proteolytic activity.

In addition, the analysis of permethylated N-glycans together with the hydrolysis by specific exoglycosidases provided the first structural analysis of N-glycosylation of this protein. Our preliminary data showed monosialylated diantennary structures with different degrees of fucosylation in gp17/SABP, as a major complex oligosaccharide component, whereas triantennary structures were also observed for its pathological counterpart. In particular, the spectra of GCDFP-15/PIP derived glycopeptide after hydrolysis with sialidase still showed the molecular species at *m*/*z* 3325.0, consistent with mono-sialylated diantennary structure after 1 h of incubation. This finding suggests the possibility that a 1 h incubation is insufficient for complete removal of sialic acid residues. Another explanation may be that the removal of one sialic acid residue from the molecular species of *m*/*z* 3622 and/or the *m*/*z* 3331 molecular species could carry an α, 2−6 linked sialic acid residue, for which the sialidase kinetic reaction is reduced. In this case, this finding may suggest a disequilibrium between the α 2−3(Gal β 1−3GalNAc) or α 2−6(Gal β 1−4GlcNAc) sialyltransferases in pathological conditions as a consequence of their differential expression and/or altered enzymatic activity. A correlation between sialyltransferase expression and/or activity with disease progression has been reported (*25*, *26*, *27*). In addition, the high electronegative and hydrophilic sialic acid residues on the protein in pathological conditions may protect the glycoprotein from proteolytic degradation and influence the uptake and processing of the sialylated glycoprotein antigen from the cells of immune system.

The specific N-glycosylation of GCDFP-15/PIP may be useful to provide an accurate criterion for breast tumor diagnosis, which is reported for alpha-fetoprotein in patients with neoplastic disease of liver (*28*, *29*).

The different N-glycosylation may be among the important factors involved in the stabilization of dimeric structure of the protein in pathological as compared to physiological conditions. This may enhance the proteolytic activity of GCDFP-15/gp17 versus fibronectin (*3*) and the metastatic potential of tumor cells (*30*). In addition, these data raised the possibility that GCDFP-15/PIP could represent a cancer-specific antigen against which an immune response can be mounted (*31*). The different N-glycosylation and/or conformation of protein may allow presentation to the immune system of previously cryptic determinants, as demonstrated for MUC1 and altered oncoproteins (*32*, *33*). Antibodies against these epitopes coupled with cytotoxic agents could be considered as anticancer vaccines (*34*).

ACKNOWLEDGMENT

We are grateful to Drs. G. Tartaglione and M. Scibelli for providing seminal plasma and Drs. R. Pasquinelli and A. Anzisi for providing breast cyst fluids. The authors wish to thank Drs. C. D'Ambrosio, F. Dal Piaz, A. Salzano, A. Amoresano, and P. Pucci for the analysis of permethylated oligosaccharides and their critical suggestions. Thanks are due to Dr. S. Markey and E. Unsworth for their support and critical discussion.

REFERENCES

1. Haagensen, D. E., Jr., Dilley, W. G., Mazoujian, G., and Wells, S. A (1990) *Ann. N.Y. Acad. Sci. 586*, 161−173.
2. Murphy, L. C., Tsuyuki, D., Myal, Y., and Shiu, R. P. (1987) *J. Biol. Chem. 262*, 15236−15241.
3. Caputo, E., Manco, G., Mandrich, L., and Guardiola, J. (2000) *J. Biol. Chem. 275*, 7935−7941.
4. Mazouujian, G. G. S., Pinkus, S. D., and Haagensen, D. E., Jr. (1983) *Am. J. Pathol. 110*, 105−111.
5. Viacava, P., Naccarato, A. G., and Bevilacqua G. (1997) *Virchow Arch. 431*, 205−209.
6. Autiero, M., Abrescia, P., and Guardiola, J. (1991) *Exp. Cell. Res. 197*, 268−271.
7. Akiyama, K., and Kimura, H. (1990) *Biochim. Biophys. Acta 1040*, 206−210.
8. Schenkels, L. C, Schaller, J., Walgreen-Weterings, E., Schadee-Eestermans, I. L., Veerman, E. C. I., and Nieuw Amerongen, A. V. (1994) *Biol. Chem. Hoppe Seyler 375*, 609−615.

9. Bergamo, P., Balestrieri, M., Cammarota, G., Guardiola, J., and Abrescia, P. (1997) *Hum. Immunol. 58*, 30−41.
10. Autiero, M., Gaubin, M., Mani, J. C., Castejon, C., Martin, M., el Marhomy, S., Guardiola, J., and Piatier-Tonneau, D. (1997) *Eur. J. Biochem. 245*, 208−213.
11. Gaubin, M., Autiero, M., Basmaciogullari, S., Metivier, D., Mis hal, Z., Culerrier, R., Oudin, A., Guardiola, J., and Piatier-Tonneau, D. (1999) *J. Immunol. 162*, 2631−2638.
12. Schaller, J., Akiyama, K., Kimura, H., Hess, D., Affolter, M., and Rickli, E. E. (1991) *Eur. J. Biochem. 196*, 743−751.
13. Rathman, W. M., Van Zeyl, M. J., Van den Keybus, P. A., Bank, R. A., Veerman, E. C., and Nieuw Amerongen, A. V. (1989) *J. Biol. Buccale 17*, 199−209.
14. Caputo, E., Autiero, M., Mani, J. C., Basmociugullari, S., Piater-Tonneau, D., and Guardiola, J. (1998) *Int. J. Cancer 78*, 76−85.
15. Merchant, M., and Weinberger, S. R. (2000) *Electrophoresis 21*, 1164−1167.
16. Baldwin, M. A., Medzihradszky, K. F., Lock, C. M., Fisher, B., Settineri, T. A., and Burlingame, A. L. (2001) *Anal. Chem. 73*, 1707−1720.
17. Caputo, E., Carratore, V., Ciullo, M., Tiberio, C., Mani, J. C., Piater-Tonneau, D., and Guardiola, J. (1999) *Eur. J. Biochem. 265*, 664−670.
18. Dell, A. (1990) *Methods Enzymol. 193*, 647−660.
19. Haagensen, D. E., Jr., Mazoujian, G., Dilley, W. G., Pedersen, C. E., Kister, S. J., and Wells, S. A. (1979) *J. Natl. Cancer Inst. 62*, 239−245.
20. Autiero, M., Cammarota, G., Friedlein, A., Zulauf, M., Chiappetta, G., Dragone, V., and Guardiola, J. (1995) *Eur. J. Immunol. 25*, 1461−1464.
21. Osawa, M., Seto, Y., Yukawa, N., Saito, T., and Takeichi, S. (1996) *Arch. Androl. 36*, 29−30.
22. Autiero, M., Bouchier, C., Basmaciogullari, S., Zaborski, P., El Marhomy, S., Martin, M., Guardiola, J., and Piatier-Tonneau, D. (1997) *Immunogenetics 46*, 345−348.
23. Vandewalle, B., Hornez, L., Vennin, P., Peyrat, J. P., and Lefebvre, J. (1986) *Biochimie 68*, 649−656.
24. Basmaciogullari, S., Autiero, M., Culerrier, R., Mani, J. C., Gaubin, M., Mishal, Z., Guardiola, J., Granier, C., and Piatier-Tonneau, D. (2000) *Biochemistry 39*, 5332−5340.
25. Sotiropoulou, G., Kono, M., Anisowicz, A., Stenman, G., Tsuji, S., and Sager, R. (2002) *Mol. Med. 8*, 42−55.
26. Recchi, M. A., Harduin-Lepers, A., Boilly-Marer, Y., Verbert, A., and Delannoy, P. (1998) *Glycoconjugate J. 15*, 19−27.
27. Wang, P. H., Lo, W. L., Hsu, C. C., Lin, T. W., Lee, W. L., Wu, C. Y., Yuan, C. C., and Tasi, Y. C. (2002) *Eur. J. Gynaecol. Oncol. 23*, 221−226.
28. Aoyagi, Y., Suzuki, Y., Igarashi, K., Saitoh, A., Oguro, M., Yokota, T., Mori, S., Nomoto, S., Isemura, M., and Asakura H. (1991) *Cancer 17*, 2390−2394.
29. Aoyagi, Y., Saitoh, A., Suzuki, Y., Igarashi, K., Oguro, M., Yokota, T., Mori, S., Suda, T., Isemura, M., and Asakura, H. (1993) *Hepatology 17*, 50−52.
30. Laidler, P., and Litynska, A. (1997) *Acta Biochim. Pol. 44*, 343−357.
31. Stockert, E., Jager, E., Chen, Y. T., Scanlan, M. J., Gout, I., Karbach, J., Arand, M., Knuth, A., and Old, L. J. (1998) *J. Exp. Med. 187*, 1349−1354.
32. Galli-Stampino, L., Meinjohanns, E., Frische, K., Meldal, M., Jensen, T., Werdelin, O., and Mouritsen, S. (1997) *Cancer Res. 57*, 3214−3222.
33. Boon, T., Coulie, P. G., and Van Den Eynde, B. (1997) *Immunol. Today 18*, 267−271.
34. Kudryashov, K., Kim, H. M., Ragupathy, G., Danishefsky, S. J., Livingston, P. O., and Lloyd, K. O. (1998) *Immunotherapy 45*, 281−286.

BI034038A