Levels of specific serum N-glycans identify breast cancer patients with higher circulating tumor cell counts

R. Saldova1, J. M. Reuben2, U. M. Abd Hamid1†, P. M. Rudd1* & M. Cristofanilli3‡

1NIBRT Dublin-Oxford Glycobiology Laboratory, The National Institute for Bioprocessing Research and Training, Conway Institute, University College Dublin, Belfield, Dublin, Ireland; 2Departments of Hematopathology, 3Breast Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, USA

Received 21 May 2010; revised 13 August 2010; accepted 16 August 2010

Background: Metastatic breast cancer (MBC) is currently an incurable condition that is primarily treated with palliative measures. Isolation of circulating tumor cells (CTCs) from the peripheral blood of these patients provides a predictive prognostic indicator, independent of the type of therapy, site of occurrence and biological characteristics of the primary disease. It has been well established that glycosylation processing pathways are disturbed in cancer, leading to alterations in the glycan content of glycoproteins.

Materials and methods: The bi-, tri- and tetraantennary glycans containing sialyl Lewis x (sLex) epitopes (A2F1G1, A3F1G1, A4F1G1 and A4F2G2) were quantified using normal phase high-performance liquid chromatography in combination with exoglycosidase array digestions in the glycan pools released from sera of 27 patients with advanced breast cancer (16 with CTCs <5/7.5 ml and 11 with CTCs ≥5/7.5 ml) and 13 healthy women.

Results: The levels of all these glycans were significantly higher in patients with CTCs ≥5/7.5 ml compared with patients with CTCs <5/7.5 ml.

Conclusions: As high levels of glycans containing sLex epitopes were associated with CTCs, their measurement may provide a new noninvasive approach for determining prognosis in women with MBC.

Key words: biomarker, breast cancer, circulating tumor cells, N-linked glycans, sLex

Introduction

Metastatic breast cancer (MBC) is currently an incurable condition that is primarily treated with palliative measures. Several clinical factors, including hormonal receptor status and site of metastasis, have been proposed to help in the prediction of long-term outcomes and efficacy of treatments, but these have never been validated. Recently, a cell sorting technique that measures circulating tumor cells (CTCs) in the peripheral blood of patients with MBC has been demonstrated to be a predictive prognostic indicator independent of type of therapy, site of occurrence and biological characteristics of the primary disease [1]. Furthermore, CTCs express biomarkers that are characteristic of stem cells, indicating a more heterogeneous disease in individual patients than has previously been reported [1–3]. Therefore, the description of the complex interplay between the heterogeneous cancer phenotypic features in each patient with breast cancer and the analysis of the immunobiology of the individual patients would facilitate a more accurate description of the metastatic process [4]. This could result in the development of more effective and personalized therapies. The challenging task of identifying individual heterogeneity within patients with MBC requires sophisticated and sensitive diagnostic tools that would provide novel information on cancer biology.

Due to their demonstrated involvement in intracellular communication between tumor cells, glycoproteins and glycolipids present at the tumor cell surface provide attractive molecular targets for diagnostic use, as well as for design of new therapeutic entities [5]. There is increasing evidence to show that a common feature of cancer is the production of tumor-associated carbohydrate antigens (TACAs) [5–7]. Aberrant glycosylation gives rise to TACAs that are modified glycan structures present on glycoproteins and glycolipids [5–7]. The TACAs that have been best characterized belong to one of two classes that are components of cell surface glycoproteins, including N-glycans and O-glycans and glycolipids of glycosphingolipids and glycolipids. The alterations in synthesis of A, B, H and Lewis blood group antigens represent other glycosylation changes that are characteristic of the disturbance of glycosylation pathways in carcinomas. Changes in protein glycosylation in serum have been detected in various cancer types, including breast cancer [8, 9]. Growing evidence suggests that this altered glycosylation contributes to cancer progression [10], especially to metastasis [8, 11–13].
The aim of this study was to determine whether the levels of specific glycans in serum are associated with CTCs in advanced breast cancer patients. A positive correlation between specific glycan levels and CTCs in MBC would provide better understanding of the molecular features of CTCs and possibly additional noninvasive approaches of establishing prognosis in these patients.

**materials and methods**

**serum samples**

Twenty-seven MBC patients and 13 healthy blood donors without known malignancy were included in the study (Table 1). The study subjects were involved in a large multicenter trial that evaluated the value of measurements of the levels of CTCs in predicting responses to therapy, progression-free survival and overall survival in patients with MBC. The median age for MBC patients was 48 years (range 28–82), and the median age for healthy donors was 50 years (range 35–60). The MBC blood samples were collected before provision of systemic therapy. Twenty patients (74%) had multiple sites of recurrences. In addition, the CTC levels of all samples were concurrently evaluated (CellSearch™, Veridex, LLC, Raritan, NJ). Ten patients (37% of patients) had 20 CTCs per 7.5 ml blood. This study was approved by the Institutional Review Board at the University of Texas MD Anderson Cancer Center (MDACC).

**isolation and enumeration of CTCs**

Operational details, precision, reproducibility and accuracy of CTC measurement using the The CellSearch™ System (Veridex, LLC) have been previously described [14]. In summary, blood samples were drawn into 10-ml CellSave™ collection tubes (Veridex, LLC), supplemented with cellular preservative, maintained at room temperature and subsequently processed within 72 h of collection using the CellSearch™ System. The system is composed of (i) a semiautomated sample preparation system and (ii) the CellSearch Epithelial Cell kit, which immunomagnetically enriches cells expressing the epithelial cell adhesion molecule (EPCAM) by incubating 7.5 ml of blood with ferrofluid coated with anti-EPCAM antibodies. The cell suspension is then placed in a magnetic field that enriches for epithelial cells that bind to the anti-EPCAM antibody-coated ferro particles. The isolated cells are then incubated with the nucleic acid dye 4',6-diamidino-2-phenylindole- and allylpyocyanin-labeled anti-CD45 monoclonal antibodies that target leukocytes and phycoerythrin-labeled antibodies to cytokeratins (CK)-8, -18 and -19 to identify epithelial cells. The resultant fluorescently labeled cells are then used to identify and enumerate CTCs using the CellSpotter® Analyzer (Immunicon Corporation, Huntingdon Valley, PA). The CellSpotter Analyzer is a semiautomated fluorescence microscopy system that allows for computer-generated reconstruction of cellular images [11]. Samples were maintained at room temperature and processed within 72 h after collection. All CTC evaluations were carried out in one central laboratory at MDACC.

The CellSpotter Analyzer generates images of potential candidate CTCs in a sample. CTCs were defined as cells that were oval or round in shape, containing a nucleus (identified by positive 4',6-diamidino-2-phenylindole staining) expressing cytoplasmic cytokeratin and lacking expression of the leukocyte antigen, CD45 [3, 4]. Results were expressed as number of cells per 7.5 ml of whole blood. Samples with less than five CTCs per 7.5 ml of whole blood and samples with five or more CTCs per 7.5 ml of whole blood were considered to be negative and positive for CTCs, respectively. This CTCs cut-off value for prognosis has been established in a prospective designed clinical trial including 177 patients [1]. Subsequent studies have confirmed this cut-off for prognosis in advanced disease. All CTC assessments were carried out in one central laboratory (MDACC, Houston, TX) by an experienced investigator.

**release and purification of N-glycans from serum in-gel block**

N-glycans were released from glycoproteins in the serum samples by digestion with N-glycosidase F (PNGase F; Roche Diagnostics GmbH, Mannheim, Germany) in-gel blocks as described previously [15]. Briefly, serum samples were reduced and alkylated and then set in sodium dodecyl sulfate gel blocks and the N-glycans were released with the addition of PNGase F.

**fluorescent labeling of N-glycans**

Released N-glycans were fluorescently labeled with 2-amino benzamide (2-AB) by reductive amination using a LudgerTag™ 2-AB labeling kit (Ludger Ltd, Abingdon, UK) [15].

**exoglycosidase digestion of 2-AB-labeled N-glycans**

Exoglycosidase digestions followed by normal phase (NP) high-performance liquid chromatography (HPLC) were used to confirm the structures of the N-glycans present in serum. The 2-AB-labeled glycans were digested in a volume of 10 μl for 18 h at 37°C in 50 mM sodium acetate buffer (pH 5.5) by sequentially adding the enzymes *Arthrobacter ureafaciens* sialidase (ABS; EC 3.2.1.18, 1 U/ml) and bovine testis β-galactosidase (BG; EC 3.2.1.23, 1 U/ml). Enzymes were supplied by Glyko (Novato, CA). Following digestion, enzymes were removed by filtration through protein-binding EZ filters (Millipore Corporation, Bedford, MA) [15].

**high-performance liquid chromatography**

Labeled glycans were separated on NP-HPLC, which was carried out using a TSK-Gel Amide-80 column (Anachem, Luton, UK) with a 20%–58% gradient of 50 mM ammonium formate (pH 4.4) versus acetonitrile. The system was calibrated using an external standard of hydrolyzed and 2-AB-labeled glucose oligomers that formed a dextran ladder, as previously described [15]. The NP-HPLC glycan chromatogram was used to quantitatively compare glycan peaks. Individual peaks were normalized to

### Table 1. Study subject demographics

<table>
<thead>
<tr>
<th></th>
<th>MBC patients (%</th>
<th>Controls (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>27 (67.5)</td>
<td>13 (32.5)</td>
</tr>
<tr>
<td>Median age in years</td>
<td>48 (28–82)</td>
<td>50 (35–60)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>17 (63)</td>
<td>9 (45)</td>
</tr>
<tr>
<td>Black</td>
<td>5 (18)</td>
<td>7 (35)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4 (15)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Asian</td>
<td>1 (4)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Unknown</td>
<td>0 (0)</td>
<td>–</td>
</tr>
<tr>
<td>ER/PR status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER/PR+</td>
<td>11 (41)</td>
<td>–</td>
</tr>
<tr>
<td>ER/PR−</td>
<td>16 (59)</td>
<td>–</td>
</tr>
<tr>
<td>Metastatic sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone/soft tissue</td>
<td>4 (15)</td>
<td>–</td>
</tr>
<tr>
<td>Visceral</td>
<td>3 (11)</td>
<td>–</td>
</tr>
<tr>
<td>Both</td>
<td>20 (74)</td>
<td>–</td>
</tr>
<tr>
<td>CTCs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥5/7.5 ml</td>
<td>11 (41)</td>
<td>–</td>
</tr>
<tr>
<td>≤5/7.5 ml</td>
<td>16 (59)</td>
<td>–</td>
</tr>
</tbody>
</table>

All subjects, N = 40.

MBC, metastatic breast cancer; ER, estrogen receptor; PR, progesterone receptor; CTC, circulating tumor cell.
the total peak area (whole glycome) and the percentage (%) area of the peaks was compared in all samples. The coefficient of variation in peak areas is between 2.5% and 25.6% for neutral and charged glycan peaks in the 3-h run, with the higher values arising from peaks that were not baseline separated [16]. The limit of detection (signal/noise ratio > 3:1) is 2 fmol [16]. The glucose unit (GU) values for neutral N-glycans are very reproducible with standard deviations of <0.03 between columns [17]. This allows direct comparison with database values collected from a range of instruments over a long period of time. For sialylated glycans, more variation is found (± 0.3 for disialylated biantennary) between columns and systems [15]. Our quantified peaks (A2F1G1, A3F1G1, A4F1G1 and A4F2G2) were well-separated neutral glycans.

**statistical analysis**

The one-sample Kolmonov–Smirnov test was carried out to determine normality of data distribution in each group and Levene’s test for equality of variances. The data were distributed normally and the variances were equal. Analysis of variance was applied to identify age-adjusted and t-test for unadjusted differentially expressed glycan peaks between different groups. P-value of <0.05 was considered as statistically significant. R version 2.10.1 was used. Boxplots were created using SPSS version 15.0.0 (SPSS Inc., Chicago, IL).

**results**

N-linked glycans from total serum glycoproteins of 27 patients with advanced breast cancer (16 with CTCs <5/7.5 ml and 11 with CTCs ≥5/7.5 ml) and 13 healthy women (Table 1) were analyzed by NP-HPLC in combination with sequential exoglycosidase array digestions and by making structural comparisons to a database of N-glycans (GlycoBase; http://glycobase.nibrt.ie:8080/database/show_nibrt.action). In this approach, glycans are labeled stoichiometrically in a structurally unbiased manner with 2-AB, allowing accurate quantitative measurements and relative comparison between samples [16]. This method enables the analysis of glycan isoforms based on sequence and linkage (e.g. core α1,6-fucosylation can be distinguished from α1,3-linked outer arm fucosylation). Glycan size and linkage result in a specific elution position that can be converted to GUs using a dextran hydrolyzate standard ladder. GUs are comparable between HPLC platforms and make interpretation easier as they can be used in conjunction with GlycoBase.

The whole-serum N-glycome has been fully characterized previously by Royle et al. [16]. In the current study, samples from MBC patients were compared with controls and alterations mostly in core-fucosylated biantennary, outer arm fucosylated and sialylated glycans were observed. These data were consistent with our previous study, in which glycans in MBC samples from an alternative cohort were analyzed and compared with controls [8]. Subsequently, breast cancer patients were separated into two groups, with CTCs ≥5/7.5 ml and with CTCs <5/7.5 ml, and the glycans’ levels in each group were investigated.

**levels of sialyl Lewis x are significantly increased in breast cancer patients with CTCs ≥5/7.5 ml compared with patients with CTCs <5/7.5 ml**

Significant differences in peaks containing highly sialylated outer arm fucosylated glycans [carrying sialyl Lewis x (sLe^x) epitope] were found. The separation of these biantennary, triantennary and tetaantennary glycans carrying the sLe^x epitope from other glycans was achieved by digesting the released glycan pool with an array of exoglycosidases enabling segregation and quantification of these potential markers. Following digestion with sialidase and β-galactosidase, the sialylated α1,3-fucosylated biantennary, triantennary and tetaantennary sialylated structures formed the α1,3-fucosylated monogalactosylated biantennary structure (A2F1G1), α1,3-fucosylated monogalactosylated triantennary structure (A3F1G1), α1,3-fucosylated monogalactosylated tetaantennary structure (A4F1G1) and α1,3-fucosylated digalactosylated tetaantennary structure (A4F2G2). These digestion products elute at GU 7.1, 7.5, 8.1 and 9.1, respectively, as clear baseline-separated peaks, allowing accurate quantification by % area under each peak (Figure 1). Total sLe^x is the sum of the four peaks that contain sLe^x epitope (A2F1G1 + A3F1G1 + A4F1G1 + A4F2G2).

The levels of all of these individual glycans and total sLe^x, except A2F1G1 alone, were increased significantly in breast cancer patients with CTCs ≥5/7.5 ml when compared with patients with CTCs <5/7.5 ml (Table 2, Figure 2). Increase in the levels of A4F1G1 was more significant than increase in the levels of A3F1G1. Glycans containing one sLe^x epitope and more antennae were more significantly increased in breast cancer patients with CTCs ≥5/7.5 ml (A4F1G1 > A3F1G1 > A2F1G1).

**discussion**

To date, even the use of modern high-resolution imaging technologies does not enable us to detect tumor cell metastasis at the single cell level, and we are typically accustomed to initiate medical treatment only after clinical demonstration of the disease. This approach has generally been unsuccessful and has stimulated research to detect and characterize CTCs, as these cells are viewed as early accessible cellular determinants of subsequent overt metastasis. In addition, such investigation brings the potential to curatively approach this stage of metastasis by early intervention.

The CellSearch™ system is currently the only assay approved by the USA Food and Drug Administration for detection of CTCs for prognostic and predictive assessment of patients with advanced disease [14, 18]. Several trials on breast cancer patients have documented the prognostic implications of CTCs [1, 19–23]. A key trial by Cristofanilli et al. [1] demonstrated the prognostic and predictive value of CTCs in 177 patients with MBC before initiation of a new systemic treatment. A cut-off of five or more CTCs per 7.5 ml separated patients with higher CTC counts and both shorter progression-free survival (2.7 versus 7.0 months, P = 0.0001) and overall survival (10.9 versus 21.9 months, P < 0.0001). These data have subsequently been confirmed in additional retrospective studies. Moreover, the number of detectable CTCs and the molecular characterization of such cells provide important prognostic and biological information in patients with advanced prostate and colon cancer [24, 25].

Clearly, the detection and enumeration of CTCs in the peripheral blood also represent a less invasive alternative to
tissue biopsies and may enable earlier detection/diagnosis, monitoring of disease progression and ultimately treatment selection. In order to implement such modality for the latter purpose, we need to be able to better define the biology of CTCs. In that respect, the CellSearch does not detect CTCs in 30%–35% of MBC patients mostly due to acquisition of a mesenchymal phenotype (loss of expression of epithelial markers), a critical phenomenon in cancer progression. Altered cell surface glycosylation is a prominent feature of malignant tumor cells and define their invasive and/or metastatic properties in general [26]. SLex is a tetrasaccharide (NeuAcα2→3Galβ1→4[Fucα1→3]GlcNAcβ1→R) that is involved in selectin-mediated adhesion of cancer cells to vascular endothelium and this determinant is thought to be closely associated with hematogenous metastases of cancer [27].

Alpaugh et al. [28] have described a cooperative role between E-cadherin and sLex in the passive dissemination of tumor emboli and in the genesis of the lymphovascular embolus of inflammatory breast cancer. Recent studies in canine mammary tumors showed an inverse relationship between the expression of the two molecular markers suggesting that acquisition of epithelial mammary tumor and aggressive phenotype is associated with increased sLex expression [29].

Abd Hamid et al. [8] previously reported that glycans containing sLex epitope were significantly increased in advanced breast cancer patients with metastases (stage IV) when compared with healthy controls and positively correlated with disease progression and metastases. Glycans containing the sLex epitopes were also previously found to be elevated in patients with aggressive early breast cancer [30]. Additional
reports using mass spectrometry (MS) found specific glycans in serum associated with breast cancer [9, 10]. Kirmiz et al. [9] separated four breast cancer patients (two with advanced disease) and four controls using principal component analysis of serum glycan levels. Kyselova et al. [10] also used MS to compare glycan levels in sera from breast cancer and controls. This study contained samples from 27 healthy controls and 82 breast cancer patients, with the majority having metastatic disease. Eight N-glycans were described as being characteristic of breast cancer and increases in sialylation and fucosylation indicated breast cancer progression.

In this study, we demonstrated that levels of N-glycans containing the sLex epitope (A2F1G1, A3F1G1, A4F1G1 and A4F2G2) and total sLe\(^\alpha\), except A2F1G1 alone, were significantly higher in patients with CTCs ≥5/7.5 ml when compared with those with CTCs <5/7.5 ml (Table 2, Figure 2). A2F1G1 is a biantennary glycan, A3F1G1 a triantennary glycan and A4F1G1 a tetraantennary glycan, all containing one sLe\(^\alpha\) epitope. A4F2G2 is a tetraantennary glycan containing two sLe\(^\alpha\) epitopes. All our results were normalized for age, as it is known to influence glycosylation [31].

These specific changes in glycosylation in patients with CTCs ≥5/7.5 ml may reflect changes in fucosylation and sialylation pathways. In order to generate sLe\(^\alpha\) structures, the precursor core structure has to firstly be sialylated by α2,3-sialyltransferase forming sialyl-N-acetyl-D-lactosamine and then fucosylated by α(1,3/1,4)-fucosyltransferases [32, 33]. Increased levels of sLe\(^\alpha\) have been correlated to decreased expression of α1,2-fucosyltransferase, which competes with α2,3-sialyltransferase for the same substrate [32] and increased expression of α(1,3/1,4)-fucosyltransferases in human pancreatic cancer cells [32]. Therefore, increased levels of sLe\(^\alpha\) suggest a change in the regulation of fucosyltransferases in the liver hepatocytes.

Interestingly, the more branched glycans with one sLe\(^\alpha\) epitope are more significantly increased in MBC patients with CTCs ≥5/7.5 ml (A4F1G1 > A3F1G1 > A2F1G1). This fact could be due to increased branching observed in cancer [34].

Since levels of serum sLe\(^\alpha\) were significantly elevated in patients with CTCs ≥5/7.5 ml in comparison to those with CTCs <5/7.5 ml, their measurement in serum may provide a new noninvasive approach to determining prognosis in women with MBC. The increased levels might be due to an acute phase response, resulting in increased expression of abundant glycoproteins such as α1-acid glycoprotein, α1-antichymotrypsin or haptoglobin. These acute phase proteins contain high levels of the sLex epitope that increase with disease progression and metastases in breast cancer [8]. Increases in the sLex epitope on acute phase proteins have been found in several cancers and chronic inflammatory conditions and also after stimulation with certain inflammatory cytokines [34].

We found no significant correlations of our N-glycans containing the sLe\(^\alpha\) epitope with clinical data available (Table 1), but this could also be due to low numbers in some of the patients groups (e.g. most of the patients had multiple sites of metastasis and only three visceral and four bone/soft tissue). Correlations of CTCs detection and prognosis with molecular subtypes and site of metastasis was published in retrospective

Figure 2. Boxplots illustrating % A2F1G1, % A3F1G1, % A4F1G1, % A4F2G2 and % Total sialyl Lewis x (sLex\(^x\)) in whole serum from 13 healthy controls and 27 advanced breast cancer patients [16 with circulating tumor cells (CTCs) <5/7.5 ml and 11 with CTCs ≥5/7.5 ml]. Boxes represent the 25th and 75th percentile with the median indicated. The bars indicate the 10th and 90th percentile.
study Giordano et al. [35]. CTCs strongly predict survival in different immunohistochemical subtypes of MBC, except for the HR−/HER-2 (hormone receptor/human epidermal growth factor receptor-2) and CTCs are a strong prognostic factor in patients without or with metastatic visceral involvement [35].

Furthermore, we believe that future molecular characterization of CTCs should focus not only on the definition of molecular pathways associated with oncogene activation but additionally on evaluating membrane protein glycosylation as new model for improving detection and enrichment methods.

In conclusion, future studies should focus on validating these results and evaluating in more detail the glycosylation status of patients with CTCs ≥5/7.5 ml, in order to generate more appropriate detection methods and therapeutic interventions.

funding
State of Texas Rare and Aggressive Breast Cancer Research Program.

acknowledgements
We would like to thank Mr Yue Fan for his help with statistical analysis and Dr Eugene Dempsey for careful reading of this manuscript.

disclosure
The authors declare no conflict of interest.

references
