Published in final edited form as:

J Proteomics. 2011 October 19; 74(11): 2395–2416. doi:10.1016/j.jprot.2011.07.014.

Oxidative stress induced carbonylation in human plasma

Ashraf G. Madian, Naomi Diaz-Maldonado, Qiang Gao, and Fred E. Regnier

Department of Chemistry, Purdue University, West Lafayette, Indiana, USA 47907

Abstract

The focus of this study was on the assessment of technology that might be of clinical utility in identification, quantification, characterization of carbonylation in human plasma proteins. Carbonylation is widely associated with oxidative stress diseases. Breast cancer patient samples were chosen as a stress positive case based on the fact that oxidative stress has been reported to be elevated in this disease. Measurements of 8-isoprostane in plasma confirmed that breast cancer patients in this study were indeed experiencing significant oxidative stress. Carbonyl groups in proteins from freshly drawn blood were derivatized with biotin hydrazide after which the samples were dialyzed and the biotinylated proteins subsequently selected, digested and labeled with iTRAQ™ heavy isotope coding reagent(s). Four hundred sixty proteins were identified and quantified, 95 of which changed 1.5 fold or more in concentration. Beyond confirming the utility of the analytical method, association of protein carbonylation was examined as well. Nearly one fourth of the selected proteins were of cytoplasmic, nuclear, or membrane origin. Analysis of the data by unbiased knowledge assembly methods indicated the most likely disease associated with the proteins was breast neoplasm. Pathway analysis showed the proteins which changed in carbonylation were strongly associated with Brca1, the breast cancer type-1 susceptibility protein. Pathway analysis indicated the major molecular functions of these proteins are defense, immunity and nucleic acid binding.

Keywords

oxidative stress; carbonylation; breast cancer; biotin hydrazide; iTRAQ™; knowledge assembly

1 Introduction

Pathological levels of oxidative stress (OS) have been implicated in a plethora of diseases ranging from *diabetes mellitus* [1] and neurodegenerative diseases (Alzheimer's disease [2], Parkinson's disease[3], and amyotrophic lateral sclerosis[4]) on to inflammatory diseases (atherosclerosis[5] and chronic lung disease[6]), cancer, and aging[7-9]. At the protein level, excessive OS leads to the oxidation of proteins in 35 or more ways, one of the more prominent being carbonylation [9, 10]. Carbonyl groups can be introduced into proteins either by i) direct oxidation of Pro, Arg, Lys, Thr, Glu, or Asp side chains or oxidative cleavage of the protein backbone, ii) introduction of 4-hydroxy-2-nonenal (HNE), 2-

© 2011 Elsevier B.V. All rights reserved.

Corresponding author: Fred E. Regnier, J. H. Law Distinguished Professor, Department of Chemistry, Purdue University, fregnier@purdue.edu, Phone: 765-494-3878, Office: WTHR 155/BIND 112.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

propenal or malondialdehyde from lipid peroxidation to a Cys, His or Lys residue, or iii) by formation of advanced glycation end-product adducts[11].

Given the prominent status of OS in so many diseases it would seem there should be numerous reports of increases in oxidized plasma proteins with disease progression. Surprisingly only alterations in isoforms of fibrinogen γ -chain precursor protein and of α -1-antitrypsin precursor were reported in studies related to heart failure [12] and Alzheimer's disease [13]. Perhaps the small number of proteins observed in these studies is due to lack of detection sensitivity. After derivatization of carbonylated proteins with 2,4-dinitrophenylhydrazine (DNP), samples were separated by two dimensional gel electrophoresis (2-DE) and the oxidized proteins detected by Western blotting using an anti-DNP antibody. The problem with this approach is that 2-DE suffers from low reproducibility, poor quantification, and limited dynamic range[14].

Preliminary studies with 32-36 year old human male subjects have shown that numerous oxidized proteins from a wide variety of cell types and organs are shed or released into plasma [10]. The objective of the work reported here was to explore the possibility that methods could be developed that measure the levels of oxidative stress induced post-translational modifications (OSi~PTMs) in blood proteins and identify the molecular function of these proteins or the biological processes with which they are associated. This was achieved at the analytical level by adapting highly selective proteomics methods that have been used with yeast [15], normal rat plasma [16], and normal human plasma [10] for the identification of carbonylated proteins [17-20].

Breast cancer was chosen as the source of oxidative stress for this study based on strong evidence that reactive oxygen species (ROS) play an important role in tumorigenesis [21, 22]. Overproduction of ROS and the associated OS have been reported to occur in several ways. One is by enhanced expression of enzymes such as thymidine phosphorylase and lactoperoxidase that elevate ROS production from within the tumor [23]. Another is by extracellular production of ROS through macrophage recruitment. Additionally, extracellular ROS entering tumor cells [23] are known to oxidatively damage mitochondria, nuclear DNA, ribosomal RNA, intercellular proteins, and lipids [24]. This damage further stimulates uncontrolled growth, ischemia, and glucose deprivation followed by a reduction in neovascularization and the production of even more OS [24]. Two recent studies have shown that the total concentration of carbonylated plasma proteins is strongly connected to breast cancer risk [25, 26].

Breast cancer patients and cancer free subjects in this study were compared to determine whether the elevated levels of OS occurring in the tumor impacted levels of oxidized plasma proteins. Qualitative and quantitative differences in putatively oxidized proteins found in the plasma of six breast cancer patients and matched controls were examined. This was achieved using a protocol recently described in OS studies of human plasma [10]. Biotin hydrazide (BH) was added to freshly prepared plasma samples derived from research subjects to derivatize carbonyl groups in oxidized proteins. The resulting Schiff bases were reduced with sodium cyanoborohydride and the samples were dialyzed to remove free BH. Biotinylated proteins were selected from plasma samples by avidin affinity chromatography and then trypsin digested. This proteolytic digest was then further fractionated by reversed phase chromatography and the peptides identified and quantified by tandem mass spectrometry.

2 Materials and Methods

2.1 Materials

Sodium cyanoborohydride, biotin hydrazide (BH), ultralinked immobilized monomeric avidin, D-biotin, protein A/G chromatography cartridges, UltraLink Biosupport™, Slide-A-Lyzer™ dialysis cassettes, biotinylated alkaline phosphatase, biotinylated horseradish peroxidase, biotinylated protein A and biotinylated protein G were purchased from Pierce (Rockford, IL, USA). Iodoacetamide, dithiothreitol (DTT), glycine, α -cyano-4-hydroxycinnamic acid (CHCA), proteomics grade N-p-tosyl-phenylalanine chloromethyl ketone (TPCK)-treated trypsin, ammonium bicarbonate, guanidine, and L-cysteine were obtained from Sigma Chemical Co. (St. Louis, MO, USA). Complete, Mini, EDTA-free, protease inhibitor cocktail tablets were purchased from Roche Diagnostics (Indianapolis, IN, USA). The ABI 4700 Proteomics Analyzer Calibration Mixture (4700 Cal Mix, bradykinin, angiotensin I, glu1-fibrinopeptide B, ACTH fragment 1-17, ACTH fragment 18-39, and ACTH fragment 7-38) and iTRAQ™ reagent multiplex kit were purchased from Applied Biosystems (ABI, Foster City, CA). Trifluoroacetic acid (TFA), and HPLC grade acetonitrile were purchased from Mallinckrodt Chemicals (Phillipsburg, NJ). OxiSelect™ 8-iso-Prostaglandin F2a Activity Assay Kit was purchased from Cell Biolabs Inc. (San Diego, CA, USA). BD Vacutainer™ venous blood collection tubes with EDTA were purchased from Fisher Scientific (Hanover Park, IL, USA). Sodium phosphate, sodium chloride and formic acid 88% were purchased from Mallinckrodt (St. Louis, MO, USA). Amicon Ultra-4™ centrifugal filter devices were purchased from Millipore (Billerica, MA, USA).

2.2 Plasma samples

Blood sample collection was carried out as part of the program on Clinical Proteomics Technology Assessment for Cancer (CPTAC) sponsored by NCI using protocols approved by Institutional Review Boards at UCSF and Purdue University. Samples were collected and derivatized at the UCSF. Breast cancer patient plasma samples were derived from newly diagnosed female subjects (five stage I and one stage II subjects) before any type of therapeutic intervention. The blood was also withdrawn before any breast biopsies performed to the donors. Six breast cancer patients and six normal matched controls donated blood (Table 1).

2.3 Biotinylation of plasma samples

Freshly drawn blood samples were collected in BD Vacutainer™ venous blood collection tubes coated with EDTA (Fisher Scientific, Hanover Park, IL, USA). A mixture of inhibitors for cysteine and serine proteases was added to inhibit intrinsic endoproteases as has been recently reported.[27]. Generally, each tablet of the protease inhibitor was dissolved in 1 ml of distilled water. The protease inhibitor solution was mixed with plasma in a 1:10 ratio (v/v) respectively. Because plasma samples were maintained at neutral pH during biotinylation and affinity selection there was no need to inhibit aspartate protease along with proteases that are only active at acidic pH. Carbonylated proteins form Schiff bases with lysine residues on other proteins during storage, even at -80°C . This means that carbonyl content will decline during storage and blood samples must be derivatized with biotin hydrazide (BH) before storage or transport. BH was added to plasma samples immediately after initial plasma preparation. Therefore, after centrifugation at $1500 \times g$ for 15min, the supernatant was removed, and then centrifuged a second time at $2000 \times g$ for 15 min. Immediately thereafter, 50mM of biotin hydrazide (BH) was added to the plasma to a final concentration of 5 mM and the reaction was allowed to proceed at room temperature for 2 hours. Sodium cyanoborohydride was then added to a final concentration of 15 mM and the solution incubated at 0°C for 1 hr. The samples were then dialyzed three times against at least 200 volumes of PBS buffer to remove unreacted BH.

2.4 Affinity chromatography

An Agilent series 1100 (Agilent Technologies) system was used for packing the immobilized monomeric avidin columns and the purification of biotinylated proteins. This ultralinked immobilized monomeric avidin (Pierce; Rockford, IL) was self packed in a PEEK column (4.6mm $\times$ 100mm). The column was then washed with PBS (0.15M pH 7.4) followed by 2 mM biotin to block any non-reversible biotin binding sites. This was followed by washing with a regeneration buffer (0.1M glycine, pH 2.5) and re-equilibration by PBS. The Bradford assay was used to estimate protein concentration in plasma. In the first phase of this study, plasma samples from breast cancer patient and their controls were affinity selected individually. Each time, a total of 5 mg of plasma proteins was applied to the avidin affinity column (about 2.5 ml of settled beads) with PBS mobile phase (0.15 M phosphate buffered saline, pH 7.4) at a flow rate of 0.5 mL/min for 320 column volumes. The bound proteins were then eluted with the regeneration buffer (0.1M glycine, pH 2.5). The affinity selected samples were then reduced, alkylated, digested, and labeled with the iTRAQ™ reagent as described below (Figure 1). In the 2nd phase of this study, in order to isolate the oxidized immunoglobulin from pooled breast cancer patients' plasma samples and their pooled controls, a protein A/G chromatography cartridge (1.3 $\times$ 3.8 cm) was used before the avidin purification. In that case the immunoglobulin fraction isolated using the protein A/G chromatography cartridge (Pierce; Rockford, IL) was purified using Amicon Ultra-4™ centrifugal devices (Millipore, Billerica, MA, USA) as described by the supplier's guide lines and applied on the avidin column (figure 2A). On the other hand, the flow-through fraction from the protein A/G was directed to the avidin purification (figure 2B). In each case, the sample was applied with PBS mobile phase (0.15 M phosphate buffered saline, pH 7.4) at a flow rate of 0.5 mL/min for 320 column volumes. The bound proteins were then eluted from either the protein A/G or avidin column with the regeneration buffer (0.1M glycine, pH 2.5). The purified proteins were then reduced, alkylated and analyzed using the LTQ-Orbitrap XL™ as described below.

2.5 Quantitative comparison of protein abundance with iTRAQ™ and MALDI/TOF/TOF analysis

Proteins that were purified using avidin affinity chromatography were resuspended in 0.5 M triethylammonium bicarbonate (TEAB), pH 8.5, and 0.1% (w/v) sodium dodecyl sulfate (SDS), reduced with 5 mM Tris(2-carboxyethyl)phosphine (TCEP) for 1 h at 60 °C, and alkylated with 10 mM s-methylmethanethiosulfonate (MMTS) at room temperature for 10 min. The samples were then digested and labeled with iTRAQ™ reagent. The supplier's guide lines (ABI) were followed for both trypsin digestion and labeling with iTRAQ™ reagent. The six breast cancer samples and their controls were labeled with the 117-dalton and 114-dalton iTRAQ™ labeling agents respectively. The labeled peptides were then desalted and fractionated using Agilent 1100 Series HPLC (Agilent Technologies). A Pepmap C18 trap column and a nano-column (Zorbax 300SB-C18, 3.5 μ m, 100 μ m i.d., 15 cm length, Agilent Technologies, Santa Clara, CA) were used. Two solvents were used for the reversed phase separation, solvent A composed of 0.1% TFA in deionized water and solvent B composed of 0.1% TFA in acetonitrile. The RPC separation was achieved using a 40 min linear gradient from 98% solvent A: 2% solvent B to 60% solvent A: 40% solvent B at a flow rate of 800 nL/min. A mixing tee was used to mix the peptides separated with (α -cyano-4-hydroxycinnamic acid, 4 mg/mL in 60% ACN/0.1% TFA) a matrix. A microfraction collector was used to spot the peptides matrix solution on the plate. The spotted peptides were analyzed using an ABI 4800 plus™ (4800 MALDI/TOF/TOF) Proteomics Analyzer equipped with a 200 Hz Nd:Yag laser in the positive ion mode. 4000 Explorer™ software controlled the automated acquisition of MS and MS/MS data. Protein identification based on the acquired MS/MS spectra was carried out with Protein Pilot software 2.0 using the Pro Group™ algorithm (ABI) for protein identification. Protein

Pilot™ software 2.0 with the Pro Group™ algorithm performed automated MS/MS data analysis for protein identification and quantification of iTRAQ™ reporter ions. Peaks thus generated were searched against *homo sapiens* species in the UniProt/Swiss-Prot database. Four missed cleavages were allowed. Variable modifications used were Cys alkylation with methyl methanethiosulfonate (MMTS) and all the biological modifications which include 126 modifications included in the ProteinPilot™ software. Only proteins with a confidence level of 95% or more were (unused score 1.3) were accepted. The false discovery rate (FDR) was 4.7% for the proteins identified.

2.6 Proteolysis

Tryptic digests of oxidized immunoglobulins and non-immunoglobulins were used to identify their oxidation sites. Samples were reconstituted in 6M guanidine HCl and 10 mM dithiothreitol and incubated for one hour at 70°C. Iodoacetamide was then added to the reaction to a final concentration of 10 mM and allowed to incubate for 30 minutes at 4°C. This was followed by a six fold dilution of the sample with 0.1 M ammonium bicarbonate (pH 8.0). Sequence grade trypsin (2%) was added and the reaction mixture incubated at 37°C for 18 hours. Proteolysis was stopped by addition of tosyl lysine chloro ketone (TLCK) (trypsin:TLCK ratio of 1:1 (w/w)). The tryptic peptides were then used for the characterization of the oxidation sites using a nanoHPLC-LTQ Orbitrap XL™ mass spectrometer.

2.7 LTQ Orbitrap-Based Identification and Characterization

The digested peptide mixtures resulting from trypsin proteolysis were separated on an Agilent 1100 HPLC system using a 75 $\mu\text{m} \times 120\text{ mm}$ C18 reversed phase chromatography (RPC) column packed with 5 μm C18 Magic beads. Peptide separations were achieved using a 60 min linear mobile phase gradient from 0.1% formic acid to 0.1% formic acid in acetonitrile at a flow rate of 0.3 $\mu\text{L}/\text{min}$. The HPLC system was coupled directly to the LTQ Orbitrap™ hybrid mass spectrometer (ThermoFisher, San Jose, CA). The LTQ-Orbitrap™ was equipped with a nanoelectrospray ion source (Proxeon Biosystems, Odense, Denmark). The MS was operated in the data-dependent mode, in which a survey full scan MS spectrum (from m/z 300 to 1600) was acquired in the Orbitrap™ with a resolution of 60,000 at m/z 400. This was then followed by MS/MS scans of the 3 most abundant ions with +2 to +3 charge states. Target ions already selected for MS/MS were dynamically excluded for 180s. The resulting fragment ions were recorded in the linear ion trap.

2.8 Mascot database searching

The MS data files (recorded with the Xcalibur™ software version 2.0.7, Thermo Fisher Scientific) obtained by LC/MS/MS analysis on the LTQ-Orbitrap XL™ instrument were converted to dta files using the in-house online LTQ_dta software supplied by Matrix Science Inc. Minimum scans per group parameter were set to 1. Charge of the precursor ions was determined automatically by the software. Files were then sent to an in-house MASCOT™ server (Version 2.2, Matrix Science). The human taxonomy in the Swiss-Prot/Uniprot database was searched. Mascot™ has the limitation of allowing only nine modifications per search. The database was searched three times. The “error tolerant” feature of Mascot was used in a first-pass search, finding oxidized methionine as the most likely modification on the proteins. The database was then searched for the oxidized methionine two more times in addition to carbonylation as the variable modifications. Carbamidomethyl cysteine was selected as a fixed modification in the first search with the variable modifications being biotinylated oxidized arginine, biotinylated oxidized lysine, biotinylated oxidized proline, biotinylated oxidized threonine and oxidized methionine. The second search included carbamidomethyl cysteine again as a fixed modification but with biotinylated 3-deoxyglucosone adduct, biotinylated HNE adduct, biotinylated glyoxal

adduct, biotinylated methyl glyoxal adduct and oxidized methionine as variable modifications. A decoy method was used to estimate the false discovery rate. The precursor mass tolerance was set to 5 ppm and the fragment mass tolerance was set to 0.6 Da. The maximum number of missed cleavages permitted was 4 and the instrument type selected was ESI-FTICR. The enzyme selected was trypsin. Monoisotopic mass values were used. Finally protein mass was left unrestricted. The decoy was 1.4% for the peptides identified. Strict criteria were used to eliminate false positive identification of any of the modifications identified. For all modifications except biotinylation, the modification had to have an expectation value less than 0.05 and was manually validated in order to be considered as a correct match. The fragmentation of biotin however produces noise (unassigned fragment ions) that negatively impacts the ion score [28-30]. Consequently, the biotinylated modifications were validated manually (Supporting Information 5). Oxidation and carbonylation sites detected in these searches were extracted and correlated with proteins quantification using the MALDI/TOF/TOF and listed in Table 2. A list of the masses of the oxidative modifications searched is included in Table 3

2.9 Gene ontology (GO) and pathway analysis

Only proteins with more than 1.5 fold changes were used in these analyses. To identify the role of these proteins in breast cancer, their Uniprot_Accession and SwissProt identifiers were uploaded into DAVID[31, 32] (The Database for Annotation, Visualization, and Integrated Discovery) for gene ontology analysis, and GeneGo™ (GeneGo Inc., Joseph, MI) for pathway analysis respectively. In both cases the background gene population was set to “Human”. With DAVID, proteins were analyzed using the functional annotation tool that provides GO by molecular function and GO by biological processes. Pie charts were made based on the output data from the gene ontologies. To study protein interactions and the relationship with diseases among our data set, we used the “build network” tool from GeneGo™. We selected the “network analysis” by “transcription factors” because a big portion of our list was involved in DNA binding (see GO by molecular function), immune system regulation and biological process regulation (see GO by Biological process). We also filtered the network by interactions. We selected all types of interactions with positive and negative effects and excluded non-specified interactions, pharmacological effects, toxic effects, and miRNA binding. To determine whether GeneGo™ could correlate our data set with specific diseases we used the GeneGo™ “disease biomarker networks” tool which contains disease biomarkers as seed nodes for the network. The networks were set to the default value of 0.05 of significant level which indicates a false positive value of no more than 5% for the list of significant networks. For cellular localization, the protein list was uploaded and analyzed using the Build Network tool from GeneGo™. The output was then filtered by cellular localization to determine where these proteins are normally found in the cell. For GeneGo™ GO distributions, we uploaded the protein list along with their calculated fold changes. All networks and distributions were sorted by statistical significance using a P-val of 0.05 as the cutoff.

2.10 Testing non-specific binding (NSB) to the support material

Proteins can non-specifically bind to the support material (UltraLink Biosupport™) on which avidin was immobilized. In order to test this, 300 mgs of UltraLink Biosupport™ was mixed with 1.0 M Tris, pH 8.0 buffer to block the reactive azalactone groups. This then went through three cycles of centrifugation at $1,200 \times g$ and resuspension. The beads were then resuspended in PBS pH 7.4 (5% Sodium azide) and the slurry self packed in a PEEK column (4.6mm $\times$ 100mm). Pooled breast cancer and control samples were then applied to columns and separated separately. Samples were applied with PBS mobile phase (0.15 M phosphate buffered saline, pH 7.4) at a flow rate of 0.5 mL/min for 320 column volumes. The bound proteins were then eluted with the regeneration buffer (0.1M glycine, pH 2.5).

Proteins eluted in this manner were then digested and identified using MALDI/TOF/TOF as described above.

2.11 8-Isoprostane measurement

Measurement of both free and esterified 8-iso-prostaglandin F_{2α} (8-iso-PGF_{2α}) were achieved by treating the six breast cancer patient plasma samples and their controls with 2N NaOH at 45 °C for 2 hours. The procedure for the enzyme immunoassay (EIA) was executed according to the instructions provided by Cell Biolabs Inc. (San Diego, CA, USA). Plasma samples (130 uL) were assayed twice by measuring absorbance at 450 nm. Concentrations of 8-isoprostanes at pg/ml levels were compared between control and cancer plasma samples using the Student t-test. A p-value of <0.05 on a 2-tailed test with 95% confidence intervals was considered statistically significant.

3 Results

3.1 Analytical strategy

The experimental approach taken in these studies was to exploit the fact that protein carbonylation is almost exclusively associated with oxidative stress and that carbonyl groups can be tagged with biotin hydrazide with high specificity. This allows carbonylated proteins in plasma thus labeled to be affinity selected and enriched by avidin affinity chromatography.

It is generally the practice to remove abundant proteins from blood samples in discovery projects; the thought being that abundant proteins interfere with the detection of low abundance proteins. While this is true in many cases, immunoaffinity techniques used to accomplish abundant protein removal also select large numbers of other proteins. At least 129 proteins were reported to have been captured during abundant protein removal in a recent study [33]. This phenomenon is the result of what is sometimes referred to as the abundant protein “sponge effect”. Native intermolecular complexes are an issue as well [34]. Affinity selection of any member of a complex captures all the members of the complex. A concern in these studies was that oxidized proteins would either complex with abundant proteins or covalently bind to them through Schiff base formation. Low abundance proteins could be removed by one of these phenomena.

For all the reasons noted above, abundant protein removal was not used in this study. Moreover, a strategy was built into the analytical protocol to remove weakly bound, non-affinity selected proteins from protein complexes captured by affinity columns. Mathematical modeling shows that washing a loaded affinity column with at least 15 column volumes of the loading buffer will elute many of the non-affinity selected components of protein complexes [35]. Proteins of low binding affinity dissociate from complexes and are eluted during this process. A 320 column volume wash was used in these studies. An indication that low affinity proteins had been eluted from columns was when absorbance returned to zero on a milli-absorbance scale (data not shown).

This study was modeled analytically after a previous study in which carbonylated proteins in normal human plasma [10] were derivatized with biotin hydrazide (BH) to form Schiff bases that were reduced with sodium cyanoborohydride and subsequently selected from blood by avidin affinity chromatography. Avidin immobilized on a hydrophilic copolymer of polyacrylamide and azalactone in a 4.6 × 100 mm column was used. Six mL samples of freshly drawn plasma were derivatized prior to centrifugation. These samples were aliquoted into fractions of 5 mg total protein and then stored at −80°C until the time of analysis. Freezing conveys the additional advantage of dissociating many protein complexes [36]. Samples were taken from six breast cancer patients at an average age of 45 years along with

an equal number of age matched control subjects. The initial phase of this work focused on a comparative analysis of oxidized proteins in the plasma proteome of breast cancer patients relative to controls. The affinity selected fractions of oxidized proteins from each donor were tryptic digested, the resulting peptides labeled with iTRAQ™ reagent, and then subjected to RPC-MS/MS analysis (Figure 1) for identification and quantification.

Initial analysis of the data (to be discussed later) indicated that a major portion of the biotinylated protein fraction was immunoglobulins. This observation was further confirmed through a second approach in which biotinylated plasma samples were subjected to tandem affinity chromatography. The first of the tandem dimensions employed a protein A/G column to select immunoglobulins from the blood proteome while an avidin affinity column was used in the second dimension to select oxidized forms of immunoglobulins that had been biotinylated. Protein A/G was immobilized on a cross-linked agarose column. Following trypsin digestion of the tandem affinity selected fraction the total hydrolysate was subjected to RPC-MS/MS analysis in an effort to identify these immunoglobulins and characterize their oxidation sites (Figure 2A). On the other hand, the large amount of oxidized immunoglobulins in the plasma of breast cancer patients suppressed the electrospray ionization of some of the non-immunoglobulins. Therefore, the immunoglobulin stripped flow through fraction from the protein A/G column was used for avidin affinity selection in the second dimension. The eluted fraction from the avidin column was then digested with trypsin and the peptides were subjected to RPC-MS/MS analysis to characterize the oxidation sites of non-immunoglobulins (Figure 2B).

Previous reports [9, 10, 19] from this laboratory indicated that using the protocol outlined above 1) multiple sites of oxidation are often found in the same protein, 2) oxidation can be non-stoichiometric across these sites, and 3) multiple forms of oxidation involving many amino acids can be found in a biotinylated protein. This means that a single protein can exist in many different oxidized forms. It is critical to note that the protocol outlined above neither captures nor recognizes all oxidized forms of a protein. Only those types of oxidation co-resident on a protein with a free carbonyl group that has been biotinylated will be selected and analyzed. Proteins associated with a biotinylated protein in a complex will be captured as well.

3.2 Comparative proteomics

In the initial phase of this study avidin affinity purification was used to isolate biotinylated proteins from the plasma of breast cancer patients and their controls. Proteins eluted from the avidin column were then digested and labeled with iTRAQ™ reagent followed by separation of the peptide fragments by reversed phase chromatography and analyzed with MALDI/TOF/TOF. This generated an average of 3540 spectra. A total of 460 proteins were identified and quantified in the six breast cancer patients and their controls. Among these proteins, 95 were found in cancer patients to differ 1.5 fold or more in concentration relative to controls. 11, 12, 23, 32, 39 and 13 proteins were found to differ 1.5 fold or more in concentration in breast cancer patients number 1, 2, 3, 4, 5 and 6 respectively. Of these proteins, 2 changed more than 1.5 folds in four donors relative to their controls, 6 changed more than 1.5 folds in 3 donors relative to their controls, and 13 changed more than 1.5 folds in two patients. (Table 2 and supporting information 1).

3.3 Knowledge assembly analysis

Diseases are best recognized through disease specific proteins, but that is not always possible. Use of biological functions and pathways unique to a disease provides another approach. Finding proteins in disease specific pathways that change in concert during disease progression is a strong indicator a subject has a disease even through the proteins

themselves are not unique to the disease. Assembling this knowledge from individual protein data is an important part of analytical method development.

Knowledge assembly is a term used here to describe the process of assembling new knowledge based on integration of many findings reported in the current study and the literature. An objective of this work was to examine whether organs or tissue undergoing elevated oxidative stress are shedding proteins into the circulatory system with oxidative stress induced modifications that provide a protein signature. Although a series of proteins were found to have undergone concentration changes exceeding 1.5 fold in cancer patient plasma relative to controls, they may not necessarily have come from the tumor. This issue was addressed by using pathway analysis software such as GeneGO™ and DAVID to determine whether an un-biased analysis of the scientific literature connects these proteins to cancer. GeneGO™ (St. Joseph, MI) is a comprehensive set of databases for pathway analysis that uses more than 50,000 signaling interactions based on 4.5 million reports collected from the literature over the last 5 years. On the other hand GeneGO™, the Database for Annotation, Visualization, and Integrated Discovery (DAVID) [37, 38] is a tool available at no cost from NIAID and provides 40 annotation categories such as protein-protein interactions, disease associations, bio-pathways, gene function, and tissue expression using the most relevant gene ontology (GO) associated with genes. The degree to which these two differing data analysis suites gave the same interpretation of the data was a strong factor in evaluating the validity of the analyses.

3.4 DAVID GO by molecular function

Gene ontology (GO) analysis of the molecular function (Figure 3) of the proteins identified showed that 23% function in defense and immunity (e.g. immunoglobulin heavy constant gamma). Nineteen percent are involved in nucleic acid binding (e.g. mitotin and splicing factor proline/glutamine-rich protein) while another 11% were related to DNA helicase (e.g. transthyretin and laminin gamma) activity. Other molecular functions include 10% being serine protease inhibitors (e.g. set binding factor 2), 8% are functioning as microtubule binding proteins (e.g. chromosome 20 open reading frame 23), 8% are related to RNA helicase activity (e.g. proteoglycan 4), 8% serve as transfer (carrier) proteins (e.g. apolipoprotein e), 7% act as actin binding proteins (e.g. titin), 3% are extracellular matrix proteins (e.g. coagulation factor) and 5% are cytoskeletal proteins (e.g. titin).

3.5 DAVID GO by biological processes

GO analysis of the biological processes involving these proteins shows (Figure 4) that 34% function in response to stimulus (e.g. complement factor h), 20% are involved in immune system processes (e.g. immunoglobulin kappa constant), 17% are part of protein transport (e.g. transferrin), 10% are used in positive regulation of biological process (e.g. katanin p80), 7% are part of acute inflammatory responses (e.g. complement component), 6% function in cell adhesion (e.g. zonadhesin), and 6% are involved in cytoskeleton organization and biogenesis (e.g. titin).

3.6 Network analysis

The Build Network tool from GeneGo™ was used to identify protein-protein interactions and biological pathways shared by the proteins with an oxidative stress induced post-translational modification (OSi-PTM) found in this work as shown in Figure 5. We built the network by allowing GeneGo™ to find the shortest path between our data set and transcription factors using the Analyze Network by Transcription Factors. This is because a large portion of the proteins identified were involved in DNA binding (see GO by Molecular function), immune system regulation, and biological process regulation (see GO by Biological process). From the list of networks obtained, the most statistically significant was

a network involved in “positive regulation of biological processes” with a P-val of 3.47×10^{-42} that correlated with results obtained by DAVID (Figure 5). Of critical importance is that this network includes proteins such as Brca1 (the breast cancer type-1 susceptibility protein), TGF β -types I and II (transforming growth factor receptor beta), and proteins from the MAPKK pathway, all of which are involved in cancer.

3.7 Cellular location

Figure 6 shows the cellular locations of proteins that changed in concentration more than 1.5 fold. Three were located in the nucleus, 19 in the cytoplasm, 2 in membranes and the rest were in the extracellular compartment. The detection of nuclear, cytoplasmic and the membrane proteins in blood is an important piece of information, indicating that these oxidized proteins were released by apoptosis (programmed cell death) or necrosis (death due to cell damage). They were not excreted. Excessive protein oxidation has been strongly associated with cell death in the literature [39-41].

3.8 Disease distribution

Searching disease connections with GeneGo (Figure 7) it was found that breast neoplasm was the most likely disease associated with the proteins identified as having changed more than 1.5 fold in concentration. Atherosclerosis was the 2nd most likely disease. Based on the fact that the “breast cancer patient” samples were known to have come from subjects diagnosed to have breast cancer and all of the blood donors were approximately 45 years of age, it is reasonable that GeneGo would associate these samples with breast neoplasia and atherosclerosis. The bars in Figure 7 represent breast cancer donors compared to controls.

3.9 GeneGo GO processes

The bars in Figure 8 represent each of the 6 BC patients compares their controls. The Figure indicates that the most commonly shared protein function between the 6 patients is immune system response and the proteins responsible for that appeared in all of the six breast cancer patients.

3.10 Analysis of protein oxidation

Amino acid side chain modifications identified in this work are outlined in Figure 9. Oxidized proteins were recognized in data analysis by the identification of a peptide bearing an OSi~PTM along with the site and structure of the modification. At least two un-oxidized peptides from a protein in addition to the presence of the OSi~PTM bearing peptide were required before an oxidation site and type of oxidation was considered to have been identified. No attempt was made to identify cross-linking sites. Mascot was used for the analysis of the mass spectra as described under METHODS.

3.11 Analysis of the oxidized immunoglobulins and their oxidation sites

Based on GeneGo and DAVID analysis the largest fraction of proteins associated with cancer was connected to immunity, both in terms of molecular function and biological processes involved. As a consequence, it was decided that identification of the participating putatively oxidized immunoglobulins might shed light on this process. A two dimensional fractionation scheme was developed for this study of biotin hydrazide derivatized samples. The total immunoglobulin fraction of plasma samples was selected in a first separation dimension by affinity chromatography with an immobilized protein A/G column. The affinity selected immunoglobulin fraction was then subjected to avidin affinity chromatography in which biotinylated immunoglobulins were selected in a second separation dimension. Putatively oxidized immunoglobulins thus selected were tryptic digested and identified as before.

3.12 Characterization of oxidation sites

This experiment was carried out by pooling the six breast cancer patient samples and control samples, respectively. Biotinylated immunoglobulins from these two samples were isolated and analyzed using the two dimensional affinity chromatography approach described above. Peptide identification was achieved using the LTQ-Orbitrap XL™ mass spectrometer. Twelve immunoglobulins were identified in the pooled breast cancer plasma sample while only 8 were identified in the pooled normal plasma sample (Table 4). Carbonylation sites were detected in seven of these immunoglobulins (Table 2). Carbonylation sites detected represented the 3 routes of carbonylation (direct carbonylation, formation of advanced glycation endproducts and formation of advanced peroxidation end products). Direct carbonylation was seen at T90, R97 and P99 from threonine, arginine, and proline oxidation in the immunoglobulin heavy chain region while the glyoxal adduct at K97 in IGHG1 was from reaction with an advanced lipid peroxidation end product and the deoxyglucosone adduct at K75 in immunoglobulin lambda light chain VLJ region was from advanced glycation end product formation.

Although there was a very substantial increase in immunoglobulin oxidation with breast cancer patients, there were few common structural features in the oxidative stress induced post-translational modifications of immunoglobulins between subjects in the cancer group. Each individual was different. We interpret this to mean that each individual is making a unique set of antibodies in response to their tumor and that oxidative modifications of the variable regions in these antibodies provide individually distinct, but not cancer specific signatures.

3.13 Characterization of the oxidation sites for non-immunoglobulin proteins

The large amount of putatively oxidized immunoglobulins in the plasma of breast cancer patients hindered detection of other proteins in many cases. This problem was circumvented through a modification of the two dimensional fractionation process described above. The total immunoglobulin fraction of biotinylated samples was removed in the first dimension with a protein A/G affinity column as before. The modification in the procedure was that the immunoglobulin stripped flow-through fraction from the protein A/G column was used for avidin affinity selection in the second dimension. The eluted fraction from the avidin purification was then digested and analyzed by RPC-MS using the LTQ-Orbitrap XL™ as describe above. This resulted in the detection of 67 non-immunoglobulin proteins while only 32 non-immunoglobulins were detected when immunoglobulins were not removed.

Forty proteins were shown to have at least one oxidation site (Table 2 and supporting information 2 and 3). A total of 21 carbonyls were detected. Fifteen of them were direct carbonylation products (oxidized proline, arginine, threonine and lysine). Six were formed by indirect carbonylation (glyoxal, methyl glyoxal, malondialdehyde and Amadori adducts). Proteoglycan-4 precursor carried four carbonylation sites, the largest number of carbonylation sites observed for any protein in this study. All of these four sites were the result of the direct ROS oxidation of arginine, proline and threonine residues. In addition to that, twenty nine methionine sulfoxides, one methionine sulfone, two hydroxyl leucines, one hydroxyl proline, one hydroxyphenylalanine, and one cysteine sulfenic acid were also detected.

Another sixty seven proteins with no detectable carbonyls were also seen (supporting information 2 and 3). Failure to detect carbonylation sites could have been for several reasons. One obvious reason is that there is not 100% sequence coverage of the proteins detected and oxidation sites were simply missed. A second could be poor ionization efficiency of the biotinylated peptides relative to other peptides in the protein. Another

reason could be that two other un-oxidized peptides from the proteins were not found as specified in our identification criteria. This strict evaluation criterion eliminated a large number of ambiguous carbonylation sites not confirmed by manual interpretation of the mass spectra. Still another possibility is that some were cross-linked or complexed with biotinylated proteins. We believe that most of these complexes have been already dissociated as we froze the samples during storage. Additionally, it is well known that the fragmentation of biotin produces noise in the spectrum which lowers the peptide scores [28-30]. Also, many of the proteins detected are of very low abundance as they are released into blood by apoptosis and necrosis. Some biotinylated peptides are likely to be present below the detection limit of the mass spectrometer. Non-specific binding to the avidin affinity column could have contributed as well.

3.14 Testing non-specific binding (NSB) to the support material

The possibility that some of the proteins detected were non-specifically bound to the affinity column support matrix was tested. Pooled biotinylated plasma samples from cancer patients along with the controls were applied to the column support without immobilized avidin. Proteins were eluted and identified using the same protocol as with the avidin affinity column. Only keratin contaminants and trypsin were identified (supporting information 4). This indicates that the UltraLink Biosupport™ matrix, used in this study, has a very low NSB, especially after washing the column with 320 column volumes of loading buffer. It is concluded that NSB was not responsible for any significant amount of protein capture in these studies.

3.15 8-Isoprostane measurement

Validation that the breast cancer patients in this study were in fact experiencing significant oxidative stress was determined by measuring 8-isoprostane in their plasma. 8-Isoprostane is formed specifically as a consequence of free radical induced lipid peroxidation [42] and is considered to be one of the most reliable indicators of *in vivo* oxidative stress[43]. It is seen in Figure 10 that the level of 8-isoprostane was significantly higher in breast cancer patient plasma than in the controls. This is in agreement with a study reporting a statistically significant increase in urinary isoprostanes of patients with invasive breast cancer relative to subjects with benign breast diseases and controls [44]. It was concluded in this study that elevation of reactive oxygen species in invasive breast cancer is a general phenomenon and that the elevation of 8-isoprostane observed resulted from free radical induced peroxidation of unsaturated fatty acids, one of the mechanisms by which protein carbonylation reported in these studies occurred. But elevated levels of 8-isoprostane can be impacted in other ways as well. Oxidative stress has been reported to be influenced by environmental exposure,[45] diet, exercise,[46] rest,[47] seasonal variability,[48] surgery,[49] smoking,[50] and drugs [51] [52]). Oxidative stress can vary between individuals as well. It is also important to note that the sample size used in these studies and the fact that 8-isoprostane was measured in blood instead of urine as in the previous report [46] caused some differences in the results, but the trend was the same.

4 Discussion

It has been reported in a previous study that oxidized proteins from a number of organs can be found in human plasma [10]. It's well known that oxidative stress leads to cell death by multiple mechanisms [53-55]. Upon cell death the soluble protein content inside the cells is released into plasma, some of which is oxidized at the time of cell lysis and reflects the redox environment inside the cell at the time of cell death. In effect, protein oxidation is providing a form of redox autopsy of cells that recently died. The interest in plasma is that it is so easily obtained. An aim of this study was to determine whether with a disease such as

cancer, known to be associated with increased oxidative stress and protein oxidation [25, 26, 44] along with enhanced apoptosis and various other forms of cell death, the concentration of carbonylated proteins would change in plasma. This was examined by i) quantifying carbonylation of specific proteins in plasma, ii) determine whether they were associated with tumorigenesis, and iii) evaluating whether when taken together they could serve as proteomic signature of early stage breast cancer.

The rationale for studying breast cancer was that recent studies have shown oxidative stress (OS) to be involved in the pathophysiology of breast cancer [22, 23, 56]. Moreover, the concentration of carbonylated proteins in plasma has been strongly correlated with breast cancer risk [25]. The data presented here corroborates these observations and shows that oxidation occurs at specific sites in a small group of proteins. Based on affinity chromatographic selection of oxidized proteins and proteins with which they are conjugated, a total of 460 proteins were identified. Among these proteins, 95 changed in concentration more than 1.5 fold. Although it was not established that all these proteins are directly associated with tumorigenesis, the concentration of putatively oxidized proteins in the plasma of cancer patients clearly differs from that in normal controls. Most increases in concentration but the level of some decreased, perhaps due to proteasomal degradation of oxidized proteins [57].

A further objective of this work was to determine whether these putatively oxidized proteins are of tumor origin. This was achieved by correlating putatively oxidized proteins with cancer specific pathways in which they are involved. Network analysis with GeneGo showed a clear involvement of the tumor suppressor breast cancer type 1 (BRCA1) gene pathway. Previous studies have shown that this pathway is responsible for maintaining the integrity of genes in response to DNA damage induced by several factors, including oxidative stress. To perform this function, BRCA1 interacts with various repair proteins [58]. Additionally, the loss or reduction of BRCA1 alters TGF- β growth, inhibiting activity during cellular response to oxidative stress [59]. Moreover, OS activates the migration of poorly invasive cancer cells through the activation of Erk signaling. Two of the pathways involved are the tumor necrosis factor (TNF) and transforming growth factor- β (TGF β) signaling pathways [24, 60], which were shown to interact with the proteins altered in our study. From this it is concluded that some portion of the putatively oxidized proteins in the plasma of breast cancer patients arose from tumor cells.

Further support comes from the gene ontology analyses in which the molecular function was found to be broadly in the areas of immunity, nucleic acid binding, DNA helicase, serine protease inhibitors, microtubule binding proteins, RNA helicase, carrier proteins, actin binding proteins, and extracellular matrix proteins. Going on to the function of these proteins through this unbiased *in silico* analysis it was found they were associated with response to immune system processes, protein transport, positive regulation of biological processes, acute inflammatory response, cell adhesion, cytoskeleton organization, and biogenesis. All of these processes are components of the Brca1 (Breast cancer type-1 susceptibility) pathway.

Beyond the experimental work reported here, there is a very substantial literature supporting these findings. ROS in the *in vivo* environment of adenocarcinoma are known to be continuously elevated by internal processes and extracellular macrophage [61]. Among the proteins putatively oxidized are; transcription factors that lose their ability to regulate transcription, the protease inhibitors that lead to increases in proteases that enhance cancer cell invasion and metastasis [61]. Overexpression of DNA helicases, several Rho family proteins, and microtubule end-binding protein 1 are additional cases of up-regulation that are induced by OS in breast cancer [62]. Microtubule end-binding protein 1 is one of the

microtubule binding proteins that assist in mediating cell migration, division, and morphogenesis. Actin is probably up-regulated as well, enhancing cancer cell motility as a result [63]. Elevation of DNA helicases can lead to enhanced DNA replication as needed in tumorigenesis [64]. RNA helicase are associated with splicing. It has been proposed that cancer is associated with the expression of specific alternatively spliced mRNA species but whether this plays a role in cancer is questionable [65].

Still another outcome of the high levels of OS in cancer cells is an elevation of DNA mutations, DNA strand breaks, DNA intra-strand adducts and DNA-protein crosslinking [66]. Along with lipid and protein damage this sets up a cycle that aggravates oxidative stress [66]. The outcome in many cases will be apoptosis [21] and even necrosis where cells lyse and release their contents into the extracellular fluids [66]. This probably accounts for how nuclear proteins were found in the plasma of cancer patients. They are certainly not excreted.

Finally there is the question of whether these proteins and their oxidation sites represent a breast cancer signature. The fact that immunoglobulins composed such a large fraction of the putatively oxidized proteins detected in this study isn't surprising. Tumor specific antigens arising from point mutations, aberrant post-translationally modified proteins, and proteins encoded by genes only expressed or overexpressed in tumors provide a means for the immune system to recognize tumors [67, 68]. The presence of large amounts of autoantibodies in cancer patient plasma was first reported in the 1970s [67]. But do these immunoglobulins provide a protein specific signature? Not according to these studies. There is a substantial increase in immunoglobulin expression but they seem not to be structurally related between individuals. On the other hand, the complete list of proteins that changed more than 1.5 fold was strongly correlated to breast neoplasia and breast cancer related signaling pathways. These proteins changed in concentration more than 1.5 fold in the plasma of six BC donors relative to their controls, despite some variability (Figure 7).

Findings from this study lead to the conclusion that extra- and intracellular oxidation of proteins increases in associated with increased oxidative stress in breast cancer patients and that the probable source of these proteins is the tumor itself. Moreover, as a result of cell lysis these proteins make their way into the circulatory system where then can be detected after selective biotinylation and enrichment by avidin affinity chromatography. Whether there is a clear biomarker signature in this set of proteins remains to be determined in larger, more diverse studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

REFERENCES

- [1]. Brownlee M. The pathobiology of diabetic complications: A unifying mechanism. *Diabetes*. 2005; 54:1615–25. [PubMed: 15919781]
- [2]. Butterfield DA, Perluigi M, Sultana R. Oxidative stress in Alzheimer's disease brain: New insights from redox proteomics. *Eur J Pharmacol*. 2006; 545:39–50. [PubMed: 16860790]
- [3]. Gotz ME, Freyberger A, Riederer P. Oxidative stress: a role in the pathogenesis of Parkinson's disease. *J Neural Transm Suppl*. 1990; 29:241–9. [PubMed: 2193108]
- [4]. Bergeron C. Oxidative stress: its role in the pathogenesis of amyotrophic lateral sclerosis. *J Neurol Sci*. 1995; 129(Suppl):81–4. [PubMed: 7595629]
- [5]. Anderson TJ. Oxidative stress, endothelial function and coronary atherosclerosis. *Cardiologia*. 1997; 42:701–14. [PubMed: 9270179]

- [6]. Saugstad OD. Chronic lung disease: the role of oxidative stress. *Biol Neonate*. 1998; 74(Suppl 1): 21–8. [PubMed: 9730588]
- [7]. Bokov A, Chaudhuri A, Richardson A. The role of oxidative damage and stress in aging. *Mech Ageing Dev*. 2004; 125:811–26. [PubMed: 15541775]
- [8]. Stadtman ER. Protein oxidation and aging. *Free Radical Res*. 2006; 40:1250–8. [PubMed: 17090414]
- [9]. Madian A, Regnier F. Proteomic Identification of Carbonylated Proteins and Their Oxidation Sites. *J Proteome Res*. 2010; 9:3766–80. [PubMed: 20521848]
- [10]. Madian AG, Regnier FE. Profiling Carbonylated Proteins in Human Plasma. *J Proteome Res*. 2010; 9:1330–43. [PubMed: 20121119]
- [11]. Madian AG, Regnier FE. Proteomic identification of carbonylated proteins and their oxidation sites. *J Proteome Res*. 2010; 9:3766–80. [PubMed: 20521848]
- [12]. Banfi C, Brioschi M, Barcella S, Veglia F, Biglioli P, Tremoli E, et al. Oxidized proteins in plasma of patients with heart failure: Role in endothelial damage. *Eur J Heart Failure*. 2008; 10:244–51.
- [13]. Choi J, Malakowsky CA, Talent JM, Conrad CC, Gracy RW. Identification of oxidized plasma proteins in Alzheimer's disease. *Biochem Biophys Res Commun*. 2002; 293:1566–70. [PubMed: 12054695]
- [14]. Greibrokk T, Pepaj M, Lundanes E, Andersen T, Novotna K. Separating proteins by pI-values - Can 2D LC replace 2D GE? LC-GC *Eur*. 2005; 18:355–6. 8–60.
- [15]. Mirzaei H, Regnier F. Affinity chromatographic selection of carbonylated proteins followed by identification of oxidation sites using tandem mass spectrometry. *Anal Chem*. 2005; 77:2386–92. [PubMed: 15828771]
- [16]. Mirzaei H, Baena B, Barbas C, Regnier F. Identification of oxidized proteins in rat plasma using avidin chromatography and tandem mass spectrometry. *Proteomics*. 2008; 8:1516–27. [PubMed: 18383005]
- [17]. Mirzaei H, Regnier F. Affinity chromatographic selection of carbonylated proteins followed by identification of oxidation sites using tandem mass spectrometry. *Anal Chem*. 2005; 77:2386–92. [PubMed: 15828771]
- [18]. Mirzaei H, Regnier F. Protein-RNA Cross-Linking in the Ribosomes of Yeast under Oxidative Stress. *J Proteome Res*. 2006; 5:3249–59. [PubMed: 17137326]
- [19]. Mirzaei H, Regnier F. Creation of Allotypic Active Sites during Oxidative Stress. *J Proteome Res*. 2006; 5:2159–68. [PubMed: 16944927]
- [20]. Mirzaei H, Regnier F. Identification of yeast oxidized proteins: chromatographic top-down approach for identification of carbonylated, fragmented and cross-linked proteins in yeast. *J Chromatogr A*. 2007; 1141:22–31. [PubMed: 17188699]
- [21]. Chua P-J, Yip GW-C, Bay B-H. Cell cycle arrest induced by hydrogen peroxide is associated with modulation of oxidative stress related genes in breast cancer cells. *Exp Biol Med* (Maywood, NJ, U S). 2009; 234:1086–94.
- [22]. Sener DE, Gonenc A, Akinci M, Torun M. Lipid peroxidation and total antioxidant status in patients with breast cancer. *Cell Biochem Funct*. 2007; 25:377–82. [PubMed: 16447143]
- [23]. Brown NS, Bicknell R. Hypoxia and oxidative stress in breast cancer. *Oxidative stress: its effects on the growth, metastatic potential and response to therapy of breast cancer*. *Breast Cancer Res*. 2001; 3:323–7. [PubMed: 11597322]
- [24]. Benz CC, Yau C. Aging, oxidative stress and cancer: paradigms in parallax. *Nat Rev Cancer*. 2008; 8:875–9. [PubMed: 18948997]
- [25]. Rossner P Jr, Terry MB, Gammon MD, Agrawal M, Zhang FF, Ferris JS, et al. Plasma protein carbonyl levels and breast cancer risk. *J Cell Mol Med*. 2007; 11:1138–48. [PubMed: 17979889]
- [26]. Zipprich J, Terry MB, Liao Y, Agrawal M, Gurvich I, Senie R, et al. Plasma Protein Carbonyls and Breast Cancer Risk in Sisters Discordant for Breast Cancer from the New York Site of the Breast Cancer Family Registry. *Cancer Res*. 2009; 69:2966–72. [PubMed: 19339271]
- [27]. Yi J, Kim C, Gelfand Craig A. Inhibition of intrinsic proteolytic activities moderates preanalytical variability and instability of human plasma. *J Proteome Res*. 2007; 6:1768–81. [PubMed: 17411080]

- [28]. Borisov Oleg V, Goshe Michael B, Conrads Thomas P, Rakov VS, Veenstra Timothy D, Smith Richard D. Low-energy collision-induced dissociation fragmentation analysis of cysteinyl-modified peptides. *Anal Chem.* 2002; 74:2284–92. [PubMed: 12038753]
- [29]. Han B, Stevens Jan F, Maier Claudia S. Design, synthesis, and application of a hydrazide-functionalized isotope-coded affinity tag for the quantification of oxylipid-protein conjugates. *Anal Chem.* 2007; 79:3342–54. [PubMed: 17385840]
- [30]. Mirzaei H, Regnier F. Protein:protein aggregation induced by protein oxidation. *J Chromatogr, B: Anal Technol Biomed Life Sci.* 2008; 873:8–14.
- [31]. Dennis G Jr, Sherman Brad T, Hosack Douglas A, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology.* 2003; 4:P3. [PubMed: 12734009]
- [32]. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols.* 2009; 4:44–57.
- [33]. Gong Y, Li X, Yang B, Ying W, Li D, Zhang Y, et al. Different Immunoaffinity Fractionation Strategies to Characterize the Human Plasma Proteome. *J Proteome Res.* 2006; 5:1379–87. [PubMed: 16739989]
- [34]. Issaq Haleem J, Xiao Z, Veenstra Timothy D. Serum and plasma proteomics. *Chem Rev.* 2007; 107:3601–20. [PubMed: 17636887]
- [35]. Cho W, Jung K, Regnier FE. Use of Glycan Targeting Antibodies To Identify Cancer-Associated Glycoproteins in Plasma of Breast Cancer Patients. *Anal Chem (Washington, DC, U S).* 2008; 80:5286–92.
- [36]. Gao Q, Madian AG, Liu X, Adamec J, Regnier F. Coupling protein complex analysis to peptide based proteomics. *Journal of Chromatography A.* 2010 accepted.
- [37]. Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biol.* 2003; 4
- [38]. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009; 4:44–57. [PubMed: 19131956]
- [39]. Janssen YM, Van Houten B, Borm PJ, Mossman BT. Cell and tissue responses to oxidative damage. *Laboratory investigation; a journal of technical methods and pathology.* 1993; 69:261–74.
- [40]. Orrenius S. Intracellular oxidative processes and cell killing. *European Journal of Pharmaceutical Sciences.* 1994; 2:65–6.
- [41]. Rumora L, Petrik J, Zanic-Grubisic T. Oxidative stress and cell death. *Biochimica Medica.* 2003; 13:75–82.
- [42]. Khosrowbeygi A, Zarghami N. Seminal plasma levels of free 8-isoprostane and its relationship with sperm quality parameters. *Indian Journal of Clinical Biochemistry.* 2008; 23:49–52.
- [43]. Vassalle C, Botto N, Andreassi Maria G, Berti S, Biagini A. Evidence for enhanced 8-isoprostane plasma levels, as index of oxidative stress in vivo, in patients with coronary artery disease. *Coronary artery disease.* 2003; 14:213–8. [PubMed: 12702924]
- [44]. Kedzierska M, Olas B, Wachowicz B, Jeziorski A, Piekarski J. The lipid peroxidation in breast cancer patients. *Gen Physiol Biophys.* 2010; 29:208–10. [PubMed: 20577033]
- [45]. Migliore L, Coppede F. Environmental-induced oxidative stress in neurodegenerative disorders and aging. *Mutat Res.* 2009; 674:73–84. [PubMed: 18952194]
- [46]. Bloomer RJ, Fisher-Wellman K. The role of exercise in minimizing postprandial oxidative stress in cigarette smokers. *Nicotine Tob Res.* 2009; 11:3–11. [PubMed: 19246436]
- [47]. Siu PM, Pistilli EE, Alway SE. Age-dependent increase in oxidative stress in gastrocnemius muscle with unloading. *J Appl Physiol.* 2008; 105:1695–705. [PubMed: 18801960]
- [48]. Rossner P Jr, Svecova V, Milcova A, Lnenickova Z, Solansky I, Sram RJ. Seasonal variability of oxidative stress markers in city bus drivers. Part I. Oxidative damage to DNA. *Mutat Res.* 2008; 642:14–20. [PubMed: 18436263]
- [49]. Kucukakin B, Gogenur I, Reiter RJ, Rosenberg J. Oxidative stress in relation to surgery: is there a role for the antioxidant melatonin? *J Surg Res.* 2009; 152:338–47. [PubMed: 18262562]

- [50]. Ozguner F, Koyu A, Cesur G. Active smoking causes oxidative stress and decreases blood melatonin levels. *Toxicol Ind Health*. 2005; 21:21–6. [PubMed: 15986573]
- [51]. Julka D, Sandhir R, Gill KD. Adriamycin-induced oxidative stress in rat central nervous system. *Biochem Mol Biol Int*. 1993; 29:807–20. [PubMed: 8508133]
- [52]. Wolf MB, Baynes JW. The anti-cancer drug, doxorubicin, causes oxidant stress-induced endothelial dysfunction. *Biochim Biophys Acta*. 2006; 1760:267–71. [PubMed: 16337743]
- [53]. Bresgen N, Jaksch H, Lacher H, Ohlenschlaeger I, Uchida K, Eckl PM. Iron-mediated oxidative stress plays an essential role in ferritin-induced cell death. *Free Radic Biol Med*. 2010; 48:1347–57. [PubMed: 20172024]
- [54]. Kimura N, Tsunoda S, Iuchi Y, Abe H, Totsukawa K, Fujii J. Intrinsic oxidative stress causes either 2-cell arrest or cell death depending on developmental stage of the embryos from SOD1-deficient mice. *Molecular Human Reproduction*. 2010; 16:441–51. [PubMed: 20223891]
- [55]. Li X, Feng J, Sun R. Oxidative stress induces reactivation of Kaposi's sarcoma-associated herpesvirus and death of primary effusion lymphoma cells. *Journal of Virology*. 2011; 85:715–24. [PubMed: 21068240]
- [56]. Nyaga SG, Jaruga P, Lohani A, Dizdaroglu M, Evans MK. Accumulation of oxidatively induced DNA damage in human breast cancer cell lines following treatment with hydrogen peroxide. *Cell Cycle*. 2007; 6:1472–8. [PubMed: 17568196]
- [57]. Jung T, Grune T. The proteasome and its role in the degradation of oxidized proteins. *IUBMB Life*. 2008; 60:743–52. [PubMed: 18636510]
- [58]. Huen MSY, Sy SMH, Chen J. BRCA1 and its toolbox for the maintenance of genome integrity. *Nat Rev Mol Cell Biol*. 11:138–48. [PubMed: 20029420]
- [59]. Li H, Sekine M, Seng S, Avraham S, Avraham HK. BRCA1 interacts with Smad3 and regulates Smad3-mediated TGF- β^2 signaling during oxidative stress responses. *PLoS One*. 2009; 4 No pp given.
- [60]. Yau C, Benz CC. Genes responsive to both oxidant stress and loss of estrogen receptor function identify a poor prognosis group of ER-positive primary breast cancers. *Breast Cancer Research*. 2008; 10 No pp given.
- [61]. Toyokuni S, Okamoto K, Yodoi J, Hiai H. Persistent Oxidative Stress in Cancer. *Febs Letters*. 1995; 358:1–3. [PubMed: 7821417]
- [62]. Dong X, Liu F, Sun L, Liu M, Li D, Su D, et al. Oncogenic function of microtubule end-binding protein 1 in breast cancer. *J Pathol*. 220:361–9. [PubMed: 19967727]
- [63]. Jiang P, Enomoto A, Takahashi M. Cell biology of the movement of breast cancer cells: Intracellular signalling and the actin cytoskeleton. *Cancer Lett (Shannon, Ire)*. 2009; 284:122–30.
- [64]. Sharma S, Doherty KM, Brosh RM Jr. DNA helicases as targets for anti-cancer drugs. *Curr Med Chem: Anti-Cancer Agents*. 2005; 5:183–99.
- [65]. Abdelhaleem M. Do human RNA helicases have a role in cancer? *Biochim Biophys Acta, Rev Cancer*. 2004; 1704:37–46.
- [66]. Ozben T. Oxidative stress and apoptosis: impact on cancer therapy. *J Pharm Sci*. 2007; 96:2181–96. [PubMed: 17593552]
- [67]. Caron M, Choquet-Kastylevsky G, Joubert-Caron R. Cancer immunomics using autoantibody signatures for biomarker discovery. *Mol Cell Proteomics*. 2007; 6:1115–22. [PubMed: 17376768]
- [68]. Tan HT, Low J, Lim SG, Chung MCM. Serum autoantibodies as biomarkers for early cancer detection. *Febs J*. 2009; 276:6880–904. [PubMed: 19860826]
- [69]. Amici A, Levine RL, Tsai L, Stadtman ER. Conversion of amino acid residues in proteins and amino acid homopolymers to carbonyl derivatives by metal-catalyzed oxidation reactions. *J Biol Chem*. 1989; 264:3341–6. [PubMed: 2563380]
- [70]. Requena JR, Chao CC, Levine RL, Stadtman ER. Glutamic and aminoadipic semialdehydes are the main carbonyl products of metal-catalyzed oxidation of proteins. *Proc Natl Acad Sci U S A*. 2001; 98:69–74. [PubMed: 11120890]

- [71]. Zhang Q, Ames JM, Smith RD, Baynes JW, Metz TO. A Perspective on the Maillard Reaction and the Analysis of Protein Glycation by Mass Spectrometry: Probing the Pathogenesis of Chronic Disease. *J Proteome Res.* 2009; 8:754–69. [PubMed: 19093874]
- [72]. Takeuchi M, Kikuchi S, Sasaki N, Suzuki T, Watai T, Iwaki M, et al. Involvement of advanced glycation end-products (AGEs) in Alzheimer's disease. *Current Alzheimer Research.* 2004; 1:39–46. [PubMed: 15975084]
- [73]. Fenaille F, Tabet J-C, Guy PA. Immunoaffinity purification and characterization of 4-hydroxy-2-nonenal- and malondialdehyde-modified peptides by electrospray ionization tandem mass spectrometry. *Anal Chem.* 2002; 74:6298–304. [PubMed: 12510752]

Figure 1.

A schematic of the strategy used to identify and quantify oxidized proteins in the plasma of breast cancer patients compared to controls. Biotinylated proteins were purified using avidin. Each set of purified samples from breast cancer patients and their controls were then labeled with iTRAQ™ and identified and quantified using an AB 4800 plus MALDI/TOF/TOF instrument.

Figure 2.
A schematic of the strategy used for the detection of the oxidation sites with immunoglobulins and non-immunoglobulins using an LTQ-Orbitrap XL™.

Figure 3.
A DAVID Gene Ontology (GO) analysis by molecular function of proteins that changed more than 1.5 fold in the plasma of BC patients compared to their controls.

Figure 4. DAVID gene ontology (GO) by biological process of proteins that changed more than 1.5 fold in the plasma of BC patients compared to their controls.

Figure 5.

Protein networks associated with proteins that changed more than 1.5 fold in the plasma of BC patients compared to their controls. The network was created by a direct interaction algorithm of GeneGo™ using the list of proteins from our dataset. Lines between nodes indicate the interaction between proteins with green being activation, red inhibition, and cyan canonical pathways. Shapes of the nodes represent the functional class of the proteins as shown at the bottom of the figure. Brown circles indicate the cancer related proteins that proteins from our data set interact with.

Figure 6. Cellular location of the proteins that changed more than 1.5 fold in the plasma of BC patients compared to their controls.

Figure 7.

The disease distribution of the proteins that changed more than 1.5 fold in the plasma of BC patients compared to their controls. The GeneGo histogram is ordered based on the most statistically significant disease. The Y-axis corresponds to each donor group based on the disease they represent. Variability of each group can be noted by the p-val each donor has on each group. The order of significance was sorted based on the donor with the lowest p-val for a particular disease. The color bars represent: BC donor 1 compared to her control (orange), BC donor 2 compared to her control (blue), BC donor 3 compared to her control (fuchsia), BC donor 4 compared to her control (green), BC donor 5 compared to her control (brown), and BC donor 6 compared to her control (purple).

Figure 8.

Gene ontology by processes for the proteins that changed more than 1.5 fold. The GeneGo histogram is ordered based on the most statistically significant processes. The Y-axis corresponds to each donor group based on the process they represent i.e. group 1 is immune system response. Variability of each group can be noted by the p-val each donor has on each group. The order of significance was sorted based on the donor with the lowest p-val for a particular process. The color bars represent: BC donor 1 compared to her control (orange), BC donor 2 compared to her control (blue), BC donor 3 compared to her control (fuchsia), BC donor 4 compared to her control (green), BC donor 5 compared to her control (brown), and BC donor 6 compared to her control (purple).

Figure 9.

Structures of the carbonylation products detected in this study. R refers to a polypeptide sequence. Glutamic semialdehyde is the oxidation product of proline and arginine. Aminoadipic semialdehyde is the oxidation product of lysine. 2-Amino-3-ketobutyric acid is the oxidation product of threonine. All other oxidation adducts are formed by the addition of glycation and advanced glycation end products (AGEs) or Advanced lipid peroxidation end products (ALEs) to the lysine residues

Figure 10. A scatter graph showing the plasma 8-isoprostane levels in breast cancer patients (n=6) compared to their controls (n=6).

Clinical information on the breast cancer subjects and their age and sex matched controls used in this study.

Table 1

Breast cancer subjects					Control subjects			
Donor number	Stage	Other diseases	Smoking (Current)	Smoking (Past)	Donor number	Other diseases	Smoking (Current)	Smoking (Past)
1	I	Hypertension	No	Used to in the past	1	hypertension	No	Used to in the past
2	II	None	No	Never	2	None	No	Never
3	I	None	No	Used to in the past	3	Reflux disease	No	Never
4	I	None	Yes	Yes	4	None	No	Never
5	I	None	No	Never	5	Hypothyroidism, diabetes	No	Never
6	I	None	No	Used to in the past	6	None	No	Never

Table 2

Proteins that changed more than 1.5 fold in the plasma of in any of the breast cancer patients compared to their controls. The ratios were calculated as fold change. Quantitation was achieved using an AB 4800 plus TM MALDI/TOF/TOF instrument while a nanoHPLC-LTQ Orbitrap XL TM mass spectrometer was used to characterize the oxidation sites.

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
1,4	2.02	5.3	O60318 MCM3A_HUMAN	80 kDa MCM3-associated protein	HUMAN	12	0.4	0.1	1.7E-02	ND	Nucleus
2,5	4.79	13.4	P01009 A1AT_HUMAN	Alpha-1-antitrypsin precursor	HUMAN	7	0.5	0.1	9.0E-01	K246: Biotinylated oxidized lysine (aminoadipic semialdehyde), R247: Biotinylated oxidized arginine (glutamic semialdehyde) and L377: Hydroxy-leucine	Extracellular
3,3	2.19	6.9	P01023 A2MG_HUMAN	Alpha-2-macroglobulin precursor	HUMAN	13	2.5	1.2	6.2E-02	F155: Hydroxyphenylalanine and M713, M1378: Methionine sulfoxide	Unspecified
4,5	2.48	19	P02652 APOA2_HUMAN	Apolipoprotein A-II precursor	HUMAN	3	1.9	0.3	4.2E-01	ND	Extracellular
5,3	4	12.6	P06727 APOA4_HUMAN	Apolipoprotein A-IV precursor _b	HUMAN	6	0.5	0.4	6E-01		
5,4	31.28	54.5	P06727 APOA4_HUMAN	Apolipoprotein A-IV precursor _b	HUMAN	20	0.5	0.2	2.5E-02	M322: Methionine sulfoxide	Extracellular
5,6	2	12.4	P06727 APOA4_HUMAN	Apolipoprotein A-IV precursor _b	HUMAN	5	0.5	0.1	1.0E-02		
6,4	3.54	5	P04114 APOB_HUMAN	Apolipoprotein B-100 precursor	HUMAN	22	1.8	0.2	5.0E-02	K293: Biotinylated methyl glyoxal adduct, K2425: Biotinylated glyoxal adduct, M564, M566, M1266, M812, M901, M1234, M2042, M2597, M3267, M3458: Methionine sulfoxide and M564: Methionine sulfone	Extracellular
7,4	4.29	30.1	P02654 APOC1_HUMAN	Apolipoprotein C-I precursor	HUMAN	3	0.4	0	4.0E-02	ND	Extracellular
8,2	2	21.2	P02656 APOC3_HUMAN	Apolipoprotein C-III precursor _b	HUMAN	2	0.2	0.1	2.0E-03	ND	Extracellular

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
8,5	2	19.2	P02656 APOC3_HUMAN	Apolipoprotein C-III precursor ^b	HUMAN	2	0.1	0.1	2.0E-04		
9,4	2	18.1	P55056 APOC4_HUMAN	Apolipoprotein C-IV precursor	HUMAN	2	2.2	0.4	4.9E-02	ND	Extracellular
10,1	22.21	57.7	P02649 APOE_HUMAN	Apolipoprotein E precursor ^b	HUMAN	17	2.4	0.5	8.0E-02	M153: Methionine sulfoxide	Extracellular
10,4	10.3	18.9	P02649 APOE_HUMAN	Apolipoprotein E precursor ^b	HUMAN	6	2	0.2	2.9E-01		
10,6	11.12	48.3	P02649 APOE_HUMAN	Apolipoprotein E precursor ^b	HUMAN	5	0.6	0.1	1.5E-01		
11,5	2.07	7	O94833 BPAAEA_HUMAN	Bullous pemphigoid antigen 1, isoforms	HUMAN	50	0.4	0.2	1.2E-01	ND	Cytoplasm
12,3	8.64	15.2	P04003 C4BP_HUMAN	C4b-binding protein alpha chain precursor	HUMAN	11	1.7	0.7	1E-02	M248: Methionine sulfoxide	Extracellular
13,5	6	13	O43866 CD5L_HUMAN	CD5 antigen-like precursor	HUMAN	4	2	0.5	1.0E-02	ND	Extracellular
14,4	2.84	4.2	P49454 CENPF_HUMAN	Centromere protein F	HUMAN	12	0.2	0.1	3.3E-02	ND	Cytoplasm
15,3	5.72	10	P00450 CERU_HUMAN	Ceruloplasmin precursor ^b	HUMAN	11	1.5	0.7	1.4E-03	ND	Extracellular
15,5	11.71	15.6	P00450 CERU_HUMAN	Ceruloplasmin precursor ^b	HUMAN	17	1.6	0.3	1.4E-01		
15,6	6.93	10.8	P00450 CERU_HUMAN	Ceruloplasmin precursor ^b	HUMAN	12	0.5	0.1	3.2E-01		
16,1	6	11.6	P05160 F13B_HUMAN	Coagulation factor XIII B chain precursor ^b	HUMAN	8	0.6	0.2	6.0E-02	ND	Extracellular
16,4	12.59	25	P05160 F13B_HUMAN	Coagulation factor XIII B chain precursor ^b	HUMAN	7	1.7	0.2	9.0E-02		
17,3	2.14	12.2	P02745 C1QA_HUMAN	Complement C1q subcomponent subunit A precursor	HUMAN	3	2	0.3	5.0E-02	M472: Methionine sulfoxide and M394: Methionine sulfoxide	Extracellular
18,6	4	14.7	P02747 C1QC_HUMAN	Complement C1q subcomponent subunit C	HUMAN	3	1.9	0.2	4.3E-02	ND	Extracellular

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
19,2	63.46	43.4	P01024 CO3_HUMAN	Complement C3 precursor	HUMAN	37	0.6	0.3	1.7E-05	K678: Biotinylated oxidized lysine (aminoaliphatic semialdehyde), P366: Hydroxyproline and M42, M201, M990.M1347.M1563: Methionine sulfoxide	Extracellular
20,5	17.23	18.2	P0C0L5 CO4B_HUMAN	Complement C4-B precursor	HUMAN	8	1.5	0.5	2.3E-01	ND	Extracellular
21,4	4.83	9.6	P08603 CFAH_HUMAN	Complement factor H precursor	HUMAN	11	2	0.1	5.9E-01	R1149: Biotinylated oxidized arginine (glutamic semialdehyde), T1151: Biotinylated oxidized threonine (2-amino-3-ketobutyric acid), P1160: Biotinylated oxidized proline (glutamic semialdehyde) and M162: Methionine sulfoxide	Extracellular
22,2	93.83	60.7	P02671 FIBA_HUMAN	Fibrinogen alpha chain precursor _b	HUMAN	123	0.4	0.3	2.7E-02	M537: Methionine sulfoxide	Extracellular
22,3	96.2	63.6	P02671 FIBA_HUMAN	Fibrinogen alpha chain precursor _b	HUMAN	114	4.5	2.3	5.0E-02		
22,4	51.54	41.7	P02671 FIBA_HUMAN	Fibrinogen alpha chain precursor _b	HUMAN	50	0.6	0.3	2.8E-02		
22,6	45.86	33.3	P02671 FIBA_HUMAN	Fibrinogen alpha chain precursor _b	HUMAN	47	2.1	0.6	1.1E-46		
23,2	77.51	66.2	P02675 FIBB_HUMAN	Fibrinogen beta chain precursor _b	HUMAN	115	2	0.3	8.7E-21	M468: Methionine sulfoxide	Extracellular
23,3	88.89	76.4	P02675 FIBB_HUMAN	Fibrinogen beta chain precursor _b	HUMAN	125	2	0.9	4.9E-10		
23,6	44.93	49.7	P02675 FIBB_HUMAN	Fibrinogen beta chain precursor _b	HUMAN	35	1.9	0.2	2.2E-13		
24,2	43.53	68.9	P02679 FIBG_HUMAN	Fibrinogen gamma chain precursor _b	HUMAN	38	0.6	0.2	3.1E-04	ND	Extracellular

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
24,3	40.56	52.5	P02679 FIBG_HUMAN	Fibrinogen gamma chain precursor ^b	HUMAN	37	2	0.6	2.7E-52		
24,4	22.18	46.1	P02679 FIBG_HUMAN	Fibrinogen gamma chain precursor ^b	HUMAN	15	0.6	0.2	1.4E-12		
24,6	14.99	30.9	P02679 FIBG_HUMAN	Fibrinogen gamma chain precursor ^b	HUMAN	7	1.7	0.4	4.5E-19		
25,4	183.52	50.8	P02751 FINC_HUMAN	Fibronectin precursor ^b	HUMAN	172	0.5	0.2	1.6E-04	P1759: Biotinylated oxidized proline (glutamic semialdehyde), K1586: Biotinylated amadori adduct and M926, M1548: Methionine sulfoxide	
25,5	143.08	45.2	P02751 FINC_HUMAN	Fibronectin precursor ^b	HUMAN	107	0.5	0.3	1.0E-04		Extracellular
25,6	79.06	33.9	P02751 FINC_HUMAN	Fibronectin precursor ^b	HUMAN	46	1.7	0.3	8.8E-04		
26,1	4	9.2	Q08380 LG3BP_HUMAN	Galectin-3-binding protein precursor	HUMAN	4	9	0	3.3E-65	ND	Extracellular
27,5	5.35	20.7	P00738 HPT_HUMAN	Haptoglobin precursor	HUMAN	7	1.6	0.4	6.0E-2	ND	Extracellular
28,4	10	45.6	P68871 HBB_HUMAN	Hemoglobin subunit beta ^b	HUMAN	6	1.6	0.8	1.3E-22	ND	Extracellular
28,6	4	23.8	P68871 HBB_HUMAN	Hemoglobin subunit beta ^b	HUMAN	3	1.9	0.6	8.0E-03		
29,5	6	27.2	P02042 HBD_HUMAN	Hemoglobin subunit delta	HUMAN	4	0.4	0.1	1.8E-01	ND	Extracellular
30,1	6.96	18.4	P02790 HEMO_HUMAN	Hemopexin precursor ^b	HUMAN	9	1.5	0.1	4.0E-02	ND	Extracellular
30,4	14.57	27.9	P02790 HEMO_HUMAN	Hemopexin precursor ^b	HUMAN	12	2.6	0.4	9.7E-03		
31,4	8.01	17.3	P04196 HRG_HUMAN	Histidine-rich glycoprotein precursor ^b	HUMAN	11	1.7	0.2	9.0E-02	ND	Extracellular
31,6	1.7	12	P04196 HRG_HUMAN	Histidine-rich glycoprotein precursor ^b	HUMAN	7	0.6	0.1	5.0E-01		

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
32,5	6.01	28.5	P13747 HLAE_HUMAN	HLA class I histocompatibility antigen, alpha chain E precursor	HUMAN	9	0.4	0.2	2.8E-02	ND	Membrane
33,5	2.12	3.9	Q96ED9 HOOK2_HUMAN	Hook homolog 2	HUMAN	4	2.6	0.4	3.3E-05	ND	Cytoplasm
34,1	2	2.9	Q14520 HABP2_HUMAN	Hyaluronan-binding protein 2 precursor	HUMAN	1	1.6	NA	NA	ND	Extracellular
35,4	12.48	18.4	P01876 IGHA1_HUMAN	Ig alpha-1 chain C _b region	HUMAN	14	2.6	0.7	6.2E-03	ND	Extracellular
35,6	8.21	18.1	P01876 IGHA1_HUMAN	Ig alpha-1 chain C _b region	HUMAN	8	1.8	0.3	7.0E-03		
36,2	5.1	17.4	P01877 IGHA2_HUMAN	Ig alpha-2 chain C region	HUMAN	8	1.6	0.5	3.5E-02	ND	Extracellular
37,3	20.36	54.8	P01857 IGHG1_HUMAN	Ig gamma-1 chain C _b region	HUMAN	22	0.6	0.5	5.0E-02	K97: Biotinylated glyoxal adduct M257: Methionine sulfoxide	Extracellular
37,4	25.19	58.5	P01857 IGHG1_HUMAN	Ig gamma-1 chain C _b region	HUMAN	21	1.6	1.2	1.4E-08		
37,5	39.84	75.2	P01857 IGHG1_HUMAN	Ig gamma-1 chain C _b region	HUMAN	27	2.1	0.6	7.6E-03		
38,3	6.65	51.2	P01859 IGHG2_HUMAN	Ig gamma-2 chain C region	HUMAN	22	1.9	0.3	4.1E-06		
39,2	2	64.8	P01860 IGHG3_HUMAN	Ig gamma-3 chain C region	HUMAN	6	2.4	0.5	3.4E-02		
40,1	3.2	31.8	P01861 IGHG4_HUMAN	Ig gamma-4 chain C region	HUMAN	5	0.3	0.4	2E-01		
41,2	4.03	20.5	P23083 HV103_HUMAN	Ig heavy chain V-I region V35 precursor	HUMAN	7	1.5	0.7	7E-01	T90: Biotinylated 2-amino 3-ketobutyric acid, R97 and P99: Biotinylated glutamic semialdehyde	Extracellular
42,4	2.16	37.1	P01781 HV320_HUMAN	Ig heavy chain V-III region GAL	HUMAN	1	1.9	NA	NA		Extracellular
43,5	2	20.6	P01772 HV311_HUMAN	Ig heavy chain V-III region KOL	HUMAN	3	1.7	1	3.0E-05		Extracellular
44,5	3.51	30.3	P01777 HV316_HUMAN	Ig heavy chain V-III region TEI	HUMAN	4	6.5	1.7	1.2E-03		Extracellular

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
45,5	2	26.2	P01762 HV301_HUMAN	Ig heavy chain V-III region TRO	HUMAN	4	2.1	0.5	1.5E-02		Extracellular
46,4	10.04	70.8	P01834 KAC_HUMAN	Ig kappa chain C region ^b	HUMAN	11	0.5	0.2	3.9E-02		Extracellular
46,5	23.45	84	P01834 KAC_HUMAN	Ig kappa chain C region ^b	HUMAN	25	3.7	1.7	6.8E-01		Extracellular
47,3	4.74	24.8	P04432 KV124_HUMAN	Ig kappa chain V-I region Daudi precursor	HUMAN	4	2.4	0.3	4.0E-02		Extracellular
48,5	2.59	28.7	P01610 KV118_HUMAN	Ig kappa chain V-I region WEA ^b	HUMAN	3	1.7	0.2	1.0E-04		Extracellular
48,6	2	31.5	P01610 KV118_HUMAN	Ig kappa chain V-I region WEA ^b	HUMAN	2	1.6	0.2	1.8E-01		Extracellular
49,5	2.36	42.7	P06309 KV205_HUMAN	Ig kappa chain V-II region GM607 precursor	HUMAN	5	0.3	0	3.7E-09	T220: Biotinylated 2-amino-3-ketobutyric acid	Extracellular
50,3	4.18	41.1	P04207 KV308_HUMAN	Ig kappa chain V-III region CLL precursor	HUMAN	6	2	0.5	1.6E-07		Extracellular
51,5	3.2	55.8	P18135 KV312_HUMAN	Ig kappa chain V-III region HAH precursor	HUMAN	9	3	1.4	3.1E-02		Extracellular
52,4	4.98	49.6	P18136 KV313_HUMAN	Ig kappa chain V-III region HIC precursor	HUMAN	6	2.2	0.4	1.3E-02		Extracellular
53,5	6.06	49.3	P06314 KV404_HUMAN	Ig kappa chain V-IV region B17 precursor	HUMAN	7	3.3	1.5	1.8E-03		Extracellular
54,3	2	33.1	P06313 KV403_HUMAN	Ig kappa chain V-IV region J1 precursor	HUMAN	6	1.7	0.1	2.1E-03		Extracellular
55,3	11.72	63.8	P01842 LAC_HUMAN	Ig lambda chain C region ^b	HUMAN	8	0.5	0.2	5.0E-02		Extracellular
55,5	14.03	67.6	P01842 LAC_HUMAN	Ig lambda chain C region ^b	HUMAN	10	3	0.7	1.2E-03	K75: Biotinylated deoxyglucosone adduct	Extracellular
56,4	3.71	34.2	P01701 LV103_HUMAN	Ig lambda chain V-I region NEW	HUMAN	3	1.7	0.2	4.1E-02		Extracellular
57,4	2	28	P01717 LV403_HUMAN	Ig lambda chain V-IV region H1 ^b	HUMAN	2	1.7	0.3	1.1E-03		Extracellular

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
57,6	4	28	P01717 LV403_HUMAN	Ig lambda chain V-IV region H1 ^b	HUMAN	3	1.6	0.1	6.5E-01		
58,2	31.42	44.9	P01871 MUC_HUMAN	Ig mu chain C region ^b	HUMAN	19	0.6	0.3	1.0E-02	M384: Methionine sulfoxide	Extracellular
58,5	48.88	62.6	P01871 MUC_HUMAN	Ig mu chain C region ^b	HUMAN	41	2.7	0.7	3.1E-01		
59,3	1.4	45	P04220 MUCB_HUMAN	Ig mu heavy chain disease protein	HUMAN	13	0.4	0.1	5.0E-02	ND	Extracellular
60,1	4	17.5	P01591 IGJ_HUMAN	Immunoglobulin J chain	HUMAN	3	1.6	0.3	5.0E-02	ND	Extracellular
61,5	4.01	3.7	P19827 ITIH1_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H1 precursor	HUMAN	3	0.1	0.1	4.9E-02	K902: Biotinylated methyl glyoxal adduct	Extracellular
62,5	1.31	6.6	Q9BVA0 KTNB1_HUMAN	Katanin p80 WD40-containing subunit B1	HUMAN	6	0.3	0.2	2.7E-02	ND	Cytoplasm
63,5	1.4	7.7	Q9C0H6 KLHL4_HUMAN	Kelch-like protein 4	HUMAN	10	0.3	0.2	2.6E-04	ND	Cytoplasm
64,4	1.7	3.9	Q8IXQ5 KLHL7_HUMAN	Kelch-like protein 7	HUMAN	2	0.5	0.2	7.9E-04	ND	Unspecifi
65,5	1.7	7.4	Q96L93 SNX23_HUMAN	Kinesin-like motor protein C20orf23	HUMAN	14	0.4	0.2	9.0E-02	ND	Cytoplasm
66,2	3.58	18.3	P01042 KNG1_HUMAN	Kininogen-1 precursor	HUMAN	8	1.9	1.1	5.0E-01	ND	Extracellular
67,4	1.57	4.6	Q13753 LAMC2_HUMAN	Laminin subunit gamma-2 precursor	HUMAN	6	0.1	0.2	5.0E-02	ND	Extracellular
68,5	2	3.2	Q8IUZ0 LRC49_HUMAN	Leucine-rich repeat-containing protein 49	HUMAN	2	0.4	0.2	5.0E-02	ND	Cytoplasm
69,1	2	2.3	O00187 MASP2_HUMAN	Mannan-binding lectin serine protease 2 precursor	HUMAN	1	9	NA	NA	ND	Extracellular
70,3	2.02	4.9	P27816 MAP4_HUMAN	Microtubule-associated protein 4	HUMAN	5	0.4	0.3	4.0E-02	ND	Cytoplasm
71,3	2	2.5	Q8NEV4 MYO3A_HUMAN	Myosin IIIA	HUMAN	7	0.6	0.1	1.4E-01	ND	Unspecified
72,5	2.04	6.9	Q9ULV0 MYO5B_HUMAN	Myosin-5B	HUMAN	20	0.6	0.1	5.0E-02	ND	Cytoplasm

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
73,3	2.07	7.5	Q86WG5 MTMRD_HUMAN	Myotubularin-related protein 13	HUMAN	20	4.7	1.5	1.0E-02	ND	Cytoplasm
74,5	2.1	5	Q8NF91 SYNE1_HUMAN	Nesprin-1	HUMAN	56	0.6	0.2	5.0E-02	ND	Cytoplasm
75,5	1.52	7.6	O43929 ORC4_HUMAN	Origin recognition complex subunit 4	HUMAN	4	0.6	0.1	2.0E-02	ND	Nucleus
76,5	1.53	5.3	P56715 RP1_HUMAN	Oxygen-regulated protein 1	HUMAN	18	0.6	0.2	1.6E-02	ND	Cytoplasm
77,1	4	23.2	P32119 PRDX2_HUMAN	Peroxiredoxin-2	HUMAN	3	0.1	0.1	4.8E-03	ND	Cytoplasm
78,5	2	4.9	O60486 PLXC1_HUMAN	Plexin-C1 precursor	HUMAN	12	0.2	0.1	7.8E-01	ND	Membrane
79,5	2.02	5.3	Q92954 PRG4_HUMAN	Proteoglycan-4 precursor	HUMAN	15	0.6	0.3	1.6E-02	K1048: Biotinylated oxidized lysine (aminoadipic semialdehyde), T1156: Biotinylated oxidized threonine (2-amino-3-ketobutyric acid), T1391: Biotinylated oxidized threonine (2-amino-3-ketobutyric acid) and R1392: Biotinylated oxidized arginine (glutamic semialdehyde)	Extracellular
80,4	2.35	13.3	P00734 THRB_HUMAN	Prothrombin precursor	HUMAN	11	0.5	0.1	1.2E-01	ND	Extracellular
81,5	2	5.2	Q9UHD2 TBK1_HUMAN	Serine/threonine-protein kinase TBK1	HUMAN	5	0.6	0.2	4.5E-02	ND	Cytoplasm
82,4	4	12	P02787 TRFE_HUMAN	Serotransferrin precursor	HUMAN	8	2.4	0.9	5.0E-02	M275: Methionine sulfoxide	Extracellular
83,4	49.75	50.7	P02768 ALBU_HUMAN	Serum albumin precursor	HUMAN	41	2.7	0.8	3.0E-02	C148: Sulfenic acid and M147, M353, M572: Methionine sulfoxide	Extracellular
84,5	10.05	17.7	P27169 PONI_HUMAN	Serum paraoxonase/arylesterase 1	HUMAN	8	1.6	0.3	2.2E-03	L14, L129: Hydroxy-leucine and M12: Methionine sulfoxide	Cytoplasm
85,4	1.7	2.5	Q9Y5W8 SNX13_HUMAN	Sorting nexin-13	HUMAN	2	0.3	0.1	1.5E-03	ND	Cytoplasm

Protein number, donor number	Unused ^a	% Cov	Accession #	Name	Species	Number of unique peptides	Breast cancer/control	SD	P-value	Oxidation sites detected	Cellular location
86,3	1.46	6.4	P02549 SPTA1_HUMAN	Spectrin alpha chain, erythrocyte	HUMAN	25	0.2	0.1	3.0E-02	ND	Cytoplasm
87,5	1.45	13.7	P23246 SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich	HUMAN	10	0.6	0.2	5.0E-02	ND	Cytoplasm
88,2	4.04	7.8	Q8WZ42 TTIN_HUMAN	Titin ^b	HUMAN	201	0.6	0.3	3.9E-03	ND	Cytoplasm
88,5	3.57	7.4	Q8WZ42 TTIN_HUMAN	Titin ^b	HUMAN	205	0.5	0.4	2.3E-12		
89,4	4	24.5	P02766 TTHY_HUMAN	Transferrin precursor	HUMAN	2	0.3	0.3	1.9E-01	ND	Extracellular
90,3	1.4	14.7	Q9Y333 LSM2_HUMAN	U6 snRNA-associated Sm-like protein LSM2	HUMAN	2	0.3	0.2	4.9E-05	ND	Unspecified
91,3	2.34	7.5	Q70EL4 UBP43_HUMAN	Ubiquitin carboxyl-terminal hydrolase 43	HUMAN	12	2.2	0.4	9.1E-03	ND	Unspecified
92,3	2.02	3.4	O75445 USH2A_HUMAN	Usherin precursor	HUMAN	27	0.4	0.2	1.7E-02	ND	Extracellular
93,3	1.4	5.3	Q5THI4 VP13D_HUMAN	Vacuolar protein sorting-associated protein 13D ^b	HUMAN	29	0.4	0.1	1.4E-03	ND	Unspecified
93,5	2.04	3.4	Q5THI4 VP13D_HUMAN	Vacuolar protein sorting-associated protein 13D ^b	HUMAN	22	0.4	0.3	4.1E-02		
94,4	33.26	49.4	P04004 VTNC_HUMAN	Vitronectin precursor	HUMAN	44	0.6	0.2	5.0E-02	M350, M359: Methionine sulfoxide	Extracellular
95,4	1.7	1.6	Q9Y493 ZAN_HUMAN	Zonadhesin precursor	HUMAN	4	0.1	0.1	2.3E-1	ND	Membrane

^aThe Unused Prot Score is a measure of all the peptide evidence for a protein that is not better explained by a higher-ranking protein. For proteins to be identified with >95%, the required Unused Prot Score is 1.3.

^bProteins that changed more than 1.5 fold in the plasma of in multiple breast cancer patients

Table 3

List of the oxidative modifications and the corresponding mass differences

Modification	Amino acid modified	Delta mass difference
Biotinylated oxidized arginine[15, 69]	R	199.066696
Biotinylated oxidized lysine[15, 70]	K	241.088501
Biotinylated oxidized proline[15, 69]	P	258.115051
Biotinylated oxidized threonine[15, 69]	T	240.104477
Biotinylated 3-deoxyglucosone adduct[9, 10, 71, 72]	K	388.17804
Biotinylated HNE adduct[9, 10, 73]	C,H, or K	398.235168
Biotinylated glyoxal adduct [9, 10, 71]	K	284.130692
Biotinylated methyl glyoxal [9, 10, 71, 72]	K	298.146347
Biotinylated malondialdehyde adduct [9, 10, 73]	K	281.1436
Biotinylated Amadori adduct [9, 10, 71, 72]	K	404.173

Table 4

List of the oxidized immunoglobulins detected in the normal and breast cancer plasma respectively (analyzed LTQOrbitrap XL™)

Normal plasma					Breast cancer plasma				
Protein accession number	Protein name	Protein score	Number of peptides used in the identification	percent coverage	Protein accession number	Protein name	Protein score	Number of peptides used in the identification	percent coverage
gi 229601	Ig G1 H Nie	503	7	18.8	gi 229601	Ig G1 H Nie	551	6	20.5
gi 11275302	anti TNF-alpha antibody light-chain Fab fragment	372	4	28.5	gi 11275302	anti TNF-alpha antibody light-chain Fab fragment	320	4	28.5
gi 442920	Chain B, Fab Fragment Of Humanized Antibody 4d5, Version 4	243	2	13.9	gi 442920	Chain B, Fab Fragment Of Humanized Antibody 4d5, Version 4	268	2	13.9
gi 54778900	immunoglobul in mu heavy chain	153	2	13.7	gi 54778900	immunoglobulin mu heavy chain	182	4	22.5
gi 46254112	immunoglobul in heavy chain	156	2	14.3	gi 37777926	immunoglobulin heavy chain variable region	176	2	20.1
gi 12054078	immunoglobul in heavy chain constant region gamma 4	246	2	11.9	gi 21669609	immunoglobulin lambda light chain VLJ region	137	3	8
gi 1185927	immunoglobul in kappa-chain VK-1	404	4	34.1	gi 10334541	immunoglobulin heavy chain	130	3	11.9
gi 60688640	IGHG1 protein	288	4	13.3	gi 54780728	immunoglobulin mu heavy chain	90	4	13.3
					gi 2765425	immunoglobulin lambda heavy chain	362	5	12.4
					gi 9857759	recombinant IgG4 heavy chain	40	1	4.1
					gi 127514	Ig mu chain C region	181	6	15.6
					gi 15779222	IGHG1 protein	379	12	34