Site-Specific GlcNAcylation of Human Erythrocyte Proteins

Potential Biomarker(s) for Diabetes

Zihao Wang,[1] Kyoungsook Park,[1] Frank Comer,[1] Linda C. Hsieh-Wilson,[2] Christopher D. Saudek,[3] and Gerald W. Hart[1]

OBJECTIVE—*O*-linked *N*-acetylglucosamine (O-GlcNAc) is upregulated in diabetic tissues and plays a role in insulin resistance and glucose toxicity. Here, we investigated the extent of GlcNAcylation on human erythrocyte proteins and compared site-specific GlcNAcylation on erythrocyte proteins from diabetic and normal individuals.

RESEARCH DESIGN AND METHODS—GlcNAcylated erythrocyte proteins or GlcNAcylated peptides were tagged and selectively enriched by a chemoenzymatic approach and identified by mass spectrometry. The enrichment approach was combined with solid-phase chemical derivatization and isotopic labeling to detect O-GlcNAc modification sites and to compare site-specific O-GlcNAc occupancy levels between normal and diabetic erythrocyte proteins.

RESULTS—The enzymes that catalyze the cycling (addition and removal) of O-GlcNAc were detected in human erythrocytes. Twenty-five GlcNAcylated erythrocyte proteins were identified. Protein expression levels were compared between diabetic and normal erythrocytes. Thirty-five O-GlcNAc sites were reproducibly identified, and their site-specific O-GlcNAc occupancy ratios were calculated.

CONCLUSIONS—GlcNAcylation is differentially regulated at individual sites on erythrocyte proteins in response to glycemic status. These data suggest not only that site-specific O-GlcNAc levels reflect the glycemic status of an individual but also that O-GlcNAc site occupancy on erythrocyte proteins may be eventually useful as a diagnostic tool for the early detection of diabetes. ***Diabetes*** **58:309–317, 2009**

From the [1]Department of Biological Chemistry, School of Medicine, The Johns Hopkins University, Baltimore, Maryland; the [2]Howard Hughes Medical Institute and Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California; and the [3]Department of Medicine, Division of Endocrinology and Metabolism, Johns Hopkins University School of Medicine, Baltimore, Maryland.

Corresponding author: Gerald W. Hart, gwhart@jhmi.edu.

Received 22 July 2008 and accepted 16 October 2008.

Published ahead of print at http://diabetes.diabetesjournals.org on 4 November 2008. DOI: 10.2337/db08-0994.

F.C. is currently affiliated with the Cell and Molecular Biology Group, Wellstat Therapeutics, Gaithersburg, Maryland.

The dynamic, enzyme-catalyzed modification of nucleocytoplasmic proteins by *O*-linked *N*-acetylglucosamine (O-GlcNAc) has extensive cross talk with phosphorylation (1) and serves as a nutrient sensor to regulate signaling, transcription, proteasomal activity, and stress responses (2–4). GlcNAcylation is highly sensitive to nutrients and to cellular stress (5–9). Therefore, we hypothesize that the extent of GlcNAcylation can be used to evaluate the glucoregulatory status of people with both subtle and overt glucose dysregulation, perhaps to identify normal, pre-diabetic individuals and overtly diabetic individuals (10,11). GlcNAcylation is nearly as ubiquitous as phosphorylation in all multicellular eukaryotes and, in many cases, competes with phosphorylation for the same or adjacent hydroxyl groups on serine or threonine residues (1,5). The donor substrate for GlcNAcylation, uridine diphosphate (UDP)-GlcNAc, occurs within cells at up to millimolar concentrations—levels approaching that for ATP. In fact, between 2 and 5% of all of the glucose used by cells is consumed by the hexosamine biosynthetic pathway (HBP) with UDP-GlcNAc as the major end product (7). Studies from many laboratories have shown that the HBP, and O-GlcNAc in particular, plays a key role in insulin resistance and in glucose toxicity (4–7). Increased GlcNAcylation in adipocytes blocks insulin signaling (12), preventing both glucose uptake and the activation of glycogen synthase (13,14). Targeted overexpression of O-GlcNAc transferase (OGT), the enzyme that catalyzes the addition of O-GlcNAc, in muscle and adipose tissue causes insulin resistance and hyperleptinemia in mice (15). The extent of GlcNAcylation on nucleocytoplasmic proteins is highly sensitive to the concentrations of glucose and other nutrients surrounding cells and to nearly all types of cellular stress. The catalytic activity of OGT is highly sensitive to the intracellular level of UDP-GlcNAc over a broad range of concentrations (nanomolar to >100 mmol/l) (16). Cycling of O-GlcNAc on many nucleocytoplasmic proteins occurs rapidly at a time scale similar to phosphorylation and is tightly regulated. Cycling of O-GlcNAc on the same protein may occur at widely different rates for different attachment sites. Based on these findings, we hypothesize that changes in the O-GlcNAc levels on some erythrocyte proteins may be used diagnostically to monitor the history of cellular exposure to changes in nutrients, especially glucose, and to oxidative stress. Because O-GlcNAc on some proteins turns over rapidly and on others cycles more slowly, it is possible that both the severity and duration

of glucose dysregulation in individuals can be estimated by monitoring the levels of O-GlcNAc simultaneously at specific sites on several key proteins in erythrocytes.

Here, we report the exploratory phase of a project that aims at developing an O-GlcNAc–based, clinically useful diagnostic tool for early detection of diabetes. We show that a number of human erythrocyte proteins are modified by O-GlcNAc. By using chemoenzymatic tagging approaches combined with solid-phase chemical derivatization, we enriched, identified, and quantified O-GlcNAc occupancy ratios on an array of O-GlcNAc sites on erythrocyte proteins from both diabetic and normal individuals. The data generated in this study not only unambiguously show that differentially regulated GlcNAcylation exists in diabetic erythrocytes but also lay the basis for future studies, including validation of the O-GlcNAc dynamics using targeted mass spectrometry and the development of site-specific O-GlcNAc antibodies to be used as diagnostic tools.

RESEARCH DESIGN AND METHODS

Blood collection and processing. Blood samples were obtained from normal and diabetic volunteers at the Johns Hopkins Diabetes Center with written consent. The research was approved by the institutional review board, consistent with the Helsinki Declaration. Subjects gave written informed consent. The identity of subjects was masked to those doing assays and analyzing data, but all authors had access to the primary data. Blood samples were drawn and collected into a vial containing EDTA. O-GlcNAcase inhibitor PUGNAc was added into the vial directly before blood collection to yield a final concentration of ~10 μmol/l. Blood cells were fractionated to isolate erythrocytes using Histopaque-1077 (Sigma-Aldrich) according to the manufacturer's instruction. Erythrocytes were lysed by sonication and centrifuged. Supernatant was recovered, and hemoglobin was partially depleted by HemogloBind resin (Biotech Support Group) following the manufacturer's instructions.

Immunoblotting and immunoprecipitation. Fifty micrograms of hemoglobin-depleted (partially) erythrocyte proteins were resolved by SDS-PAGE, transferred to nitrocellulose membrane, and blotted by O-GlcNAc antibody (CTD 110.6) (1:5,000) (17). Signals were visualized by enhanced chemiluminescence (Amersham, Piscataway, NJ). For immunoprecipitation, 1 mg lysates was incubated overnight with protein A/G beads (Santa Cruz Biotechnology) and antibodies against band 3, catalase, peroxiredoxin 2, or HSP90 α (Abcam, Cambridge, MA). After 5× washing with the lysis buffer, bound proteins were eluted by boiling for 5 min in 2× Laemmili sample buffer.

Chemoenzymatic tagging and enrichment of O-GlcNAcylated proteins. A previously described protocol was modified and followed to isolate O-GlcNAc–modified proteins from erythrocytic lysates (18). Briefly, labeling of terminal O-GlcNAc by mutant galactose transferase (GalT1) (19) was performed overnight at 4°C in the presence of 5 mmol/l $MnCl_2$, 0.5 mmol/l UDP-Gal-ketone, and 2,000 units/ml PNGase F (New England Biolabs, Ipswich, MA). The reaction mixture was then dialyzed into denaturing buffer (5 mol/l urea, 50 mmol/l NH_4HCO_3, and 100 mmol/l NaCl, pH 7.8). The pH was adjusted to 4.8 by 0.3 mol/l NaOAc. After removing the insoluble by centrifugation, 3 mmol/l aminoxy biotin (Dojindo, Gaithersburg, MD) was added to the supernatant and incubated for 24 h at room temperature. The reaction was quenched by adjusting the pH to 7.9. The reaction buffer was again dialyzed into denaturing buffer, followed by 50 mmol/l NH_4HCO_3 and 10 mmol/l NaCl, pH 7.8. After preclearing with Sepharose 6B beads, the mixture was incubated with agarose-conjugated streptavidin (Pierce, Rockford, IL) for 2 h. The beads were extensively washed by low-salt buffer (0.1 mol/l Na_2HPO_4, 0.15 mol/l NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, and 0.1% SDS, pH7.5) and high-salt buffer (0.1 mol/l Na_2HPO_4, 0.5 mol/l NaCl, and 0.2% Triton X-100, pH 7.5). Bound proteins were eluted by boiling the beads in 50 mmol/l Tris-HCl, 2.5% SDS, 100 mmol/l dithiothreitol (DTT), 10% glycerol, and 5 mmol/l biotin. O-GlcNAc proteins were resolved in SDS-PAGE and in-gel digested by trypsin as previously described (20). Peptides were extracted for mass spectrometric analysis.

Chemoenzymatic tagging and enrichment of O-GlcNAc peptides. Normal and diabetic erythrocytic lysates (1 mg each) were in-solution digested overnight at 37°C by 40 μg trypsin. Trypsin was removed by filtering the solution through a 5-kDa cutoff membrane (Millipore, Billerica, MA). Fifty units of calf intestine phosphatase (New England Biolabs) were added and incubated for 4 h in the presence of 1 mmol/l $MgCl_2$. UDP-GalNAz (Invitrogen, Carlsbad, CA) was added (~2× in excess) and incubated overnight with mutant GalT1 and 2,000 units/ml PNGase F in 50 mmol/l NH_4HCO_3. After reaction, excess UDP-GalNAz was removed by passing the mixture through a C18 spin column (Nestgroup, Southborough, MA). Peptides were eluted in 80% acetonitrile and lyophilized. Cycloaddition reaction was performed in a volume of 20 μl containing biotin–polyethylene glycol (PEG)–alkyne (~3× in excess, dissolved in DMSO; Invitrogen), 2 mmol/l Tris (2-carboxyethyl) phosphine hydrochloride, 2 mmol/l Tris [(1-benzyl-1H-1,2,3-triazol-4-yl) methyl] amine, and 2 mmol/l $CuSO_4$. The reaction mixture was incubated for 12 h at room temperature with gentle shaking. The mixture was diluted into cation exchange loading buffer. Cation exchange was performed on a strong cation exchange (SCX) spin column (Nestgroup) according to the manufacturer's instruction. Peptides were eluted in one fraction by high-salt buffer (5 mmol/l KH_2PO_4, 10% acetonitrile, and 300 mmol/l KCl, pH 3.0). The elutant was allowed to bind to agarose-conjugated streptavidin for 2 h at room temperature, followed by extensive washing.

Chemical derivatization and fractionation of enriched peptides. Eight times the bead volume of BEMAD buffer (1.5% triethylamine and 20 mmol/l DTT, pH adjusted to 12.0–12.5 by NaOH) was added to the washed avidin beads and allowed to incubate at 52–54°C for 4 h with shaking. The reaction was quenched by neutralizing the pH by 2% trifluoroacetic acid. The supernatant was desalted by C18 spin column as described above. The derivatized peptides were fractionated by SCX using a polysulfoethyl A column (0.32 × 100 mm, 5 μm, 300 Å; Column Technology, Fremont, CA) coupled to an Agilent 1100 series high-performance liquid chromatography (HPLC) (Agilent Technology, Santa Clara, CA). Fractionation was performed with a 40-min linear gradient of 0–350 mmol/l KCl (10 mmol/l KH_2PO_4 and 25% acetonitrile, pH 2.8) at a flow rate of 5 μl/min. 15 fractions (10 μl each) were collected.

iTRAQ labeling and fractionation of peptides. Peptides from the flow-through and three washes of the avidin columns were pooled, desalted, and dried down by speed vacuum. The peptides were resuspended and differentially labeled by iTRAQ reagents (Applied Biosystems, Foster City, CA) according to the manufacturer's instruction. After labeling, the peptides were combined and fractionated similarly by SCX as described above except that a different polysulfoethyl A column (2.1 × 100 mm, 5 μm, 300 Å; PolyLC, Columbia, MD) was used instead.

Mass spectrometry. Enriched O-GlcNAcylated proteins were identified by analysis on an LCQ ion trap mass spectrometer coupled to Magic 2002 HPLC (Michrom BioResources) and nanospray interface (Proxeon). The instrument was set in a information-dependent acquisition mode with three MS/MS (tandem mass spectrometry) followed by one full survey scan. Derivatized O-GlcNAc peptides were analyzed either on a Qstar Pulsar mass spectrometer (Applied Biosystems-MDS Sciex, Foster City, CA) or an LTQ-Orbitrap XL, both coupled with an Eksigent nano–liquid chromatography system (Dublin, CA). Peptides were desalted on a precolumn (75 μm inner diameter, 3 cm length, packed with irregular size particles 5–15 μm, 120 Å), and separated on an RF analytical column packed with 10 cm of C18 beads (5 μm, 120 Å; YMC ODS-AQ; Wather, Milford, MA). The main HPLC gradient was 5–40% solvent B (A, 0.1% formic acid; B, 90% acetonitrile and 0.1% formic acid) in 60 min at a flow rate of 300 nl/min. For Qstar, each survey scan was acquired from *m/z* 350–1,200 followed by MS/MS of up to three most intense precursors. For LTQ-Orbitrap, each survey scan (Fourier transform-MS, 60,000 resolution) of *m/z* 400–2,000 was followed by collision-assisted dissociation (CAD) MS/MS (ion trap-MS) of up to five most intense precursor ions. Dynamic exclusion was enabled with a repeat count of 2 and exclusion duration of 60 s.

Mass spectrometric data analysis. For protein identification, peak lists of LCQ raw files were extracted and submitted to the Mascot search engine (version 2.2.0) with the following parameters: SwissProt as database, human as species, trypsin as enzyme with up to one missed cut, carbamidomethyl (C) as fixed modification, and oxidation (M) as variable modification. Mass tolerance was set at 1.2 amu (atomic mass units) for precursors and 0.8 amu for fragment ions. Raw data from derivatized O-GlcNAc peptides were similarly searched against SwissProt database using Mascot except that DTT (ST), DTT-H6(ST), deamination, and oxidation (M) were used as variable modification, and no fixed modification was selected. Precursor and fragment ion mass tolerances were 0.3 and 0.15 amu for Qstar and 0.1 and 0.8 amu for LTQ-Orbitrap, respectively. Quantitation was performed manually by averaging peak areas over the time of elution of given ion pairs. Mass spectrometry spectra originating from iTRAQ-labeled samples were extracted and searched against SwissProt database using ProteinPilot software (version 2.0; Applied Biosystems) with Paragon algorithm. Peptide identifications were further processed by the Pro Group algorithm (Applied Biosystems), which determines the minimal set of proteins that can be reported. Protein abundance ratios were automatically calculated based on ratios of reporter ions originating from peptides that are distinct to each protein isoform. Relative occupancy

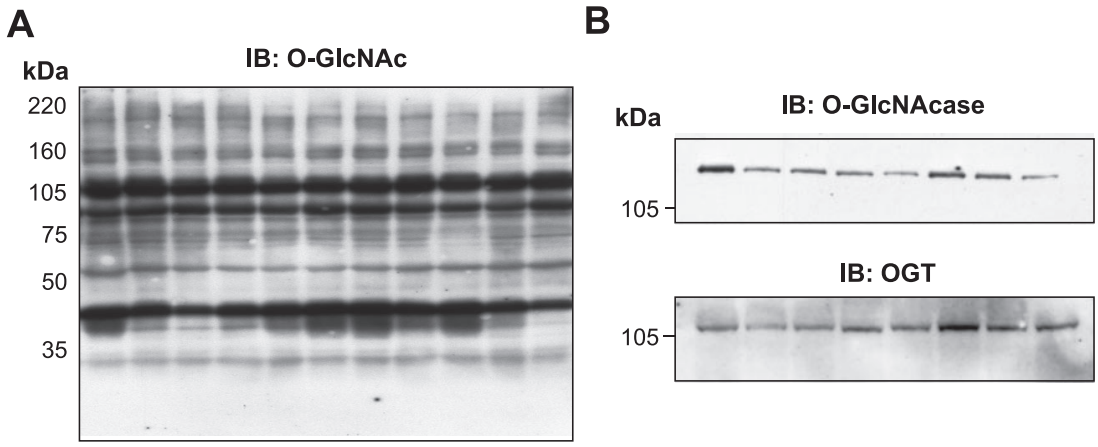

FIG. 1. Erythrocytic proteins are O-GlcNAc modified. Fifteen micrograms of erythrocyte proteins (hemoglobin depleted) were run on a 12.5% SDS-PAGE gel, transferred, and immunoblotted with antibodies against O-GlcNAc (*A*), O-GlcNAcase (*B*, *top*), or OGT (*B*, *bottom*). Different lanes represent samples from different individuals.

ratios (RORs) of O-GlcNAc between diabetic (D) and normal (N) samples were calculated using the following equation.

$$\text{ROR (D/N)} = \frac{[\text{O-GlcNAc}]}{[\text{Protein}]} \Big/ \frac{[\text{O-GlcNAc}]}{[\text{Protein}]} = \frac{[\text{O-GlcNAc}]}{[\text{O-GlcNAc}]} \times \frac{[\text{Protein}]}{[\text{Protein}]}$$

RESULTS

Erythrocytic proteins are O-GlcNAcylated. We initially tested the extent of GlcNAcylation in erythrocytic proteins by using a pan-specific O-GlcNAc antibody (CTD 110.6). Immunoblot data showed that multiple erythrocytic proteins are O-GlcNAc modified (Fig. 1*A*). Control immunoblotting with O-GlcNAc antibody but in the presence of excess free GlcNAc yielded little to no signal, suggesting the specificity of the blotting (data not shown). We next wanted to assess whether the GlcNAcylation cycles in erythrocytes. Both OGT and O-GlcNAcase, the enzymes responsible for O-GlcNAc cycling, are detected in erythrocytes (Fig. 1*B*). The donor substrate for OGT, UDP-GlcNAc, and O-GlcNAcase activity are also present in the erythrocytes (K.P., C.S., G.H., unpublished data). These data suggest that O-GlcNAc cycling may exist in human erythrocytes.

Selective enrichment and identification of putative GlcNAcylated proteins. Enrichment is key for mass spectrometric identification of GlcNAcylated proteins because of low stoichiometry and ion suppression by unmodified peptide ions in the mass spectrometer (21). Immunoisolation by pan-specific antibodies suffers from low efficiency due to relatively low binding affinity of the antibodies. Using lectins to enrich O-GlcNAc proteins suffers from low specificity because lectins may bind strongly to proteins with other forms of glycosylation (rev. in 21). In this study, a highly selective tagging method was used (18). This method takes advantage of the mutant UDP-galactose transferase (Y289L GalT1) (19), which has an enlarged donor-substrate binding pocket and can accommodate UDP-galactose analogs, in this case, UDP-Gal-ketone. GalT1 was used to enzymatically tag GlcNAc modifications on erythrocytic proteins with Gal-ketone. PNGase F was used to remove *N*-glycans. After enzymatic labeling, the ketone group was chemically tagged with an aminooxy biotin, which allowed capturing of O-GlcNAcylated proteins with streptavidin beads. Enriched proteins were then eluted, separated by SDS-PAGE, and identified by an ion trap mass spectrometer after in-gel digestion (Fig. 2*A*). By using this method, 25 erythrocyte proteins were identified as putatively GlcNAcylated (Table 1). A mock experiment with no UDP-Gal-ketone added yielded no signal when blotted with horseradish peroxidase–conjugated avidin, indicating the specificity of the approach (Fig. 2*B*). We further confirmed some of the putative GlcNAc proteins by first immunoprecipitating the proteins and then Western blotting with O-GlcNAc antibody (Fig. 2*C*). O-GlcNAc antibody competition with excess free GlcNAc was routinely performed and eliminated the signals in immunoblotting, documenting the antibody specificity (data not shown).

Mapping O-GlcNAc sites and site-specific quantitation. We first attempted to evaluate the overall dynamics of O-GlcNAc on whole protein levels by immunoprecipitation and imunoblotting, but no conclusive result was observed. This may not be surprising because the extent of GlcNAcylation is site-specifically regulated and because O-GlcNAc antibody may not be sensitive enough to recognize significant change of O-GlcNAc on one specific site when the protein is modified on multiple sites. Furthermore, one of the eventual goals of this study is to develop site specific O-GlcNAc antibodies. Thus, we wanted to map O-GlcNAc sites and quantitate O-GlcNAc site specifically. A recently developed chemoenzymatic tagging method was used to enrich O-GlcNAc at peptide levels (17,20). Y289L GalT1 was similarly used to label the GlcNAc moieties as described previously but on trypsin-digested peptides instead of whole proteins. With increased accessibility of the enzyme and donor substrate, the labeling with UDP-GalNAz (an analog of UDP-galactose with azide function group) is nearly 100% after overnight incubation at 4°C (21). Biotinylation was performed with the highly efficient copper-catalyzed cycloaddition reaction or "click" chemistry under mild conditions (22) (Fig. 3*A*). For this chemoenzymatic approach using UDP-GalNAz and biotin-PEG-alkyne, the tags added to the O-GlcNAc peptide exceed 772 Da in mass. Although the biotin-PEG-alkyne tags allow for highly selective enrichment of GlcNAcylated peptides, they are problematic; not only do they negatively affect the ionization efficiency, but the heavy tags also impose other challenges for mass spectrometric analysis. For example, fragmentation of the biotin moiety and the PEG linker arm on CAD makes the MS/MS spectra noisy and difficult to interpret. In addition, the tagging does not change the extremely labile nature of the β-O-linkage, which undergoes neutral loss before peptide backbone

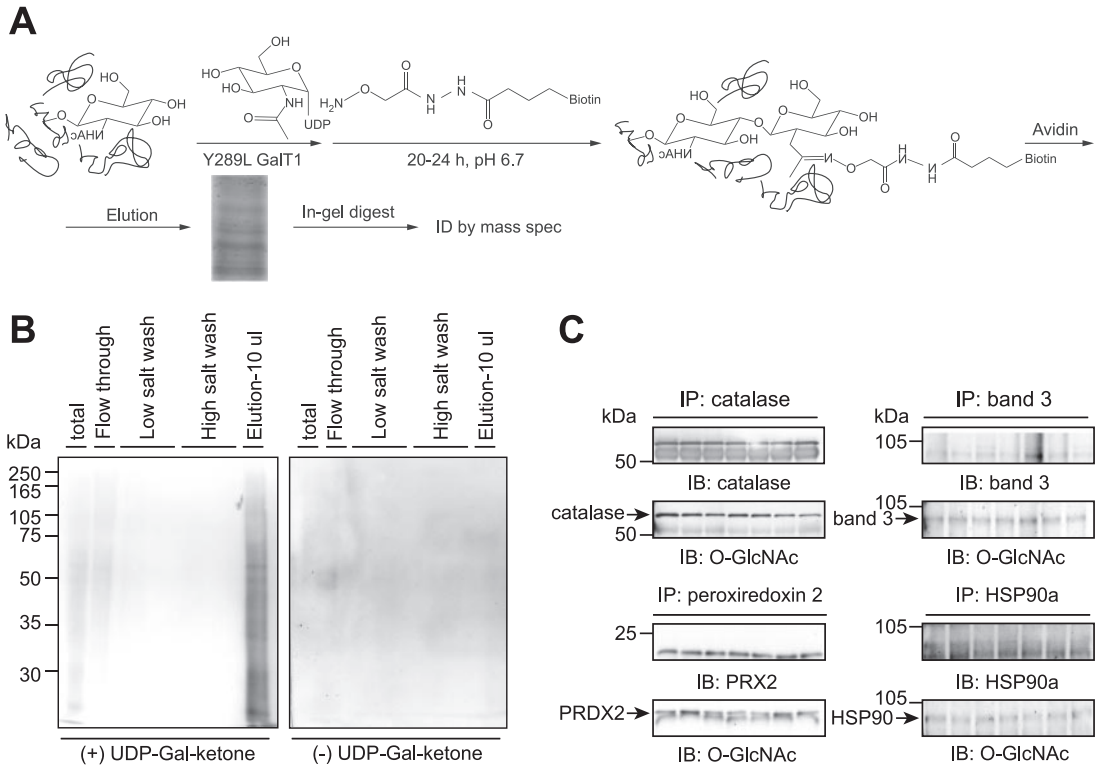

FIG. 2. Enrichment and identification of O-GlcNAc–modified erythrocytic proteins. *A*: Scheme for enriching O-GlcNAc proteins. *B*: Negative control of the approach. *C*: Confirmation of the O-GlcNAc states on several proteins.

fragmentation in CAD (Fig. 3*B*). To resolve these issues, we modified and combined a previously developed chemical derivatization method called BEMAD (β-elimination followed by Michael addition with DTT) (23) with the chemoenzymatic enrichment method. The BEMAD chemical derivatization was performed directly on the solid phase after the tagged peptides were captured by avidin beads (Fig. 3*C*, *inset*). The derivatized peptides were released from the solid phase with the O-GlcNAc, and tags were replaced by a DTT via Michael addition. The result-

TABLE 1
O-GlcNAcylated proteins enriched and identified from human erythrocytes

Protein description	Molecular weight (Da)	GenInfo identifier no.	Peptide
Catalase	59,947	Gi4557014	7
Aminolevulinic acid dehydrase isoform b	37,718	Gi34577066	3
Protease, serine 2, preprotein	26,927	Gi61097912	11
β-Globin	16,102	Gi4504349	19
Peroxiredoxin 2 isoform a	22,049	Gi32189392	10
Phosphatase and actin regulator 2	69,762	Gi7662248	5
Potassium channel tetramerization domain containing 18	47,223	Gi45387953	3
Organic cation transporter-like 3	61,435	Gi4758852	2
Peroxiredoxin 1	22,324	Gi4505591	8
α2-Globin	15,305	Gi4504345	12
Ubiquitin carrier protein	24,285	Gi7657046	6
Vacuolar protein sorting 13B isoform 2	161,849	Gi35493719	2
δ-Globin	16,159	Gi4504351	17
Spectrin-α, erythrocytic	282,024	Gi4507189	11
Hypothetical protein XP_378876	33,510	Gi51458600	2
Spectrin-β isoform a	268,630	Gi67782321	17
Spectrin-β isoform b	247,171	Gi67782319	15
Glycogen phosphorylase	97,487	Gi5032009	2
HSP90 α	98,622	Gi63029937	8
Band 3 anion transport protein	102,013	Gi4507021	9
N-acylaminocyl-peptide hydrolase	82,142	Gi23510451	5
Aldehyde dehydrogenase 1A1	55,454	Gi21361176	4
Attractin isoform 1	163,450	Gi21450861	3
Carbonic anhydrase II	29,246	Gi4557395	12
Glyceraldehyde-3-phosphate dehydrogenase	36,201	Gi7669492	7

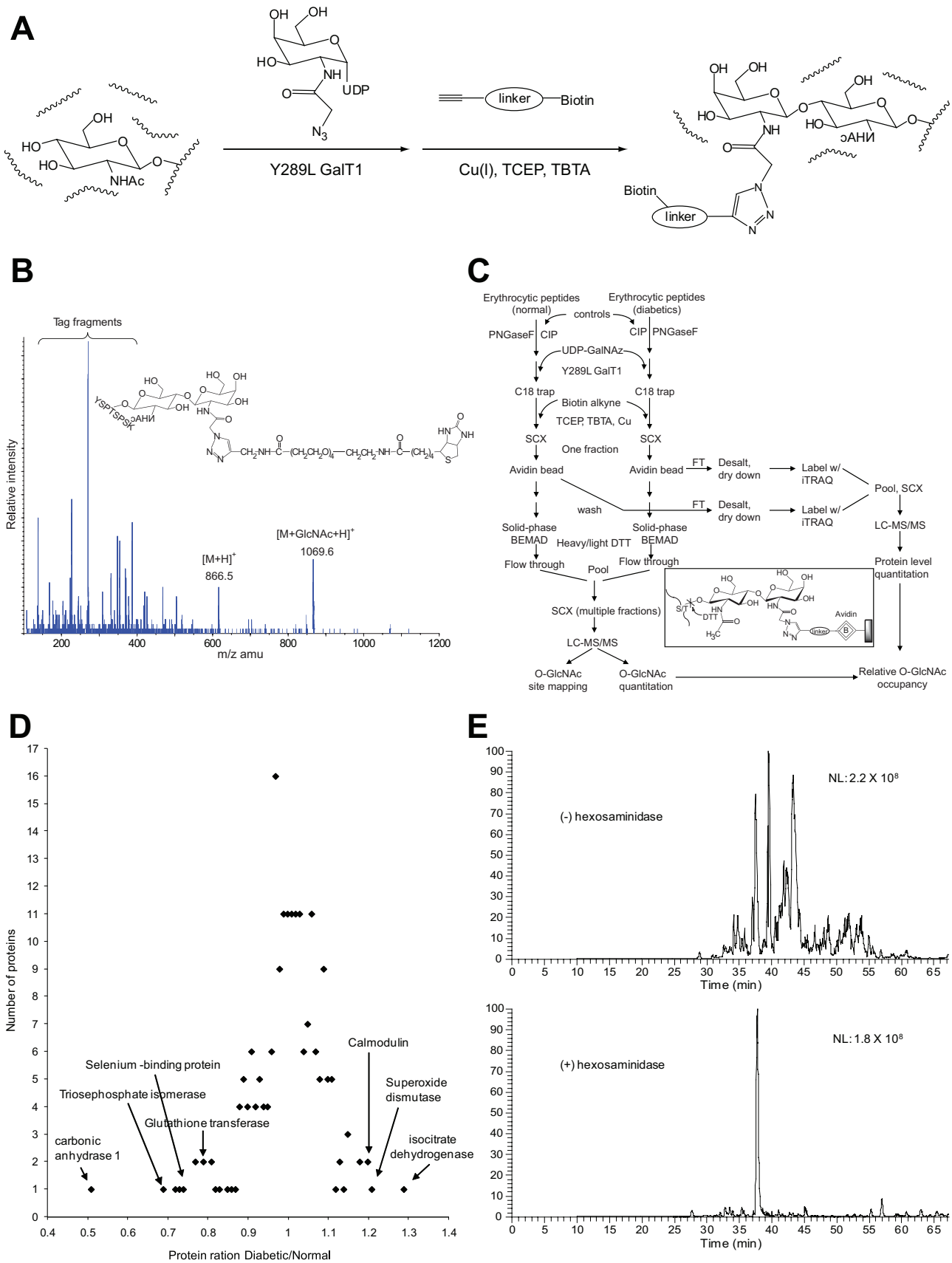

FIG. 3. Mapping O-GlcNAc sites and site-specific quantitation. *A*: Scheme for enrichment of O-GlcNAc peptides. *B*: Structure (*inset*) and CAD fragmentation of fully tagged O-GlcNAc peptide (YSPgTSPSK). $[M+GlcNAc+GalNAz+Biotin+3H]^{3+}$ = 614.6, $[M+H]^{+}$ = 866.5, $[M+GlcNAc+H]^{+}$ = 1069.6. *C*: Flow chart for comparing site-specific O-GlcNAc RORs. *Inset*: Scheme for solid-phase BEMAD. *D*: Protein expression level dynamics in diabetic erythrocytes compared with normal erythrocytes. *E*: Specificity control for the enrichment and site-mapping. Samples were untreated or treated with hexosaminidase at 37°C for 48 h before going through the work flow. Base peak chromatograms are shown. NL, intensity in counts normalized to 1 s; TBTA, Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine; TCEP, Tris(2-carboxyethyl)phosphine.

TABLE 2
Information on normal and diabetic blood donors

	n	Age	Sex (men/women)	A1C (%)	PG (mg/dl)	Diabetes duration (years)	Type of diabetes
Normal	10	27 ± 1.6	6/4	5.7 ± 0.1	84 ± 2.8	N/A	N/A
Diabetic	10	55 ± 4.5	5/5	9.5 ± 0.5	222 ± 28	14 ± 3.1	3 type 1, 7 type 2

Data are means ± SE. PG, highest recorded plasma glucose.

ing DTT modification is stable and can be easily identified by mass spectrometry. This approach also circumvents the need to break the strong biotin-avidin interaction with harsh conditions. Mass spectrometric quantitation of O-GlcNAc peptides is also readily enabled by isotopic labeling with deuterated DTT (DTT-d6), which introduces a 6-Da mass difference between the peptide pairs (e.g., normal vs. diabetic). The overall approach is shown as a flow chart in Fig. 3*C* and described in detail in RESEARCH DESIGN AND METHODS. Of course, it is possible that the apparent changes in GlcNAcylation may arise from different dynamics of protein expression or turnover. To address this factor, we labeled the flow-through of avidin chromatography, containing mostly unmodified peptides,

TABLE 3
O-GlcNAc site-mapping and comparison of site-specific O-GlcNAc RORs between normal and diabetic states

Protein name	Accession no.	O-GlcNAc peptides	Peptide score	Ratio D:N	ROR
Spectrin-β chain, erythrocytic	P11277	R.DV**S**SVELLMK.Y	56	1.0	0.97
		K.DLT**S**VLILQR.K	43	1.1	1.0
		K.LLT**S**QDVSYDEAR.N	87	1.2	1.2
		R.AQGLL**S**AGHPEGEQIIR.L	54	1.0	0.97
		R.LL**S**GEDVGQDEGATR.A	60	1.0	0.97
Carbonic anhydrase 1	P00915	K.Y**S**SLAEAASK.A	58	0.20	0.40
		K.ESI**S**VSSEQLAQFR.S	86	0.29	0.57
Hemoglobin subunit-β	P68871	R.FFESFGDL**S**TPDAVMGNPK.V	53	1.8	1.8
		K.VLGAF**S**DGLAHLDNLK.G	75	1.2	1.2
		K.G**T**FATLSELHCDK.L	67	1.1	1.1
Band 3 anion transport protein	P02730	K.A**S**TPGAAAQIQEVK.E	86	0.67	0.66
		K.H**S**HAGELEALGGVKPAVLTR.S	53	0.70	0.69
		K.IPPD**S**EATLVLVGR.A	69	1.1	1.1
Ankyrin-1	P16157	K.L**S**TPPPLAEEEGLASR.I	98	1.0	0.97
		K.VVTDET**S**FVLVSDK.H	52	1.5	1.5
		R.I**S**EILLDHGAPIQAK.T	49	1.1	1.1
		R.D**S**GEGDTTSLR.L	45	2.4	2.3
Spectrin-α chain	P02549	R.VS**S**QDYGR.D	57	1.0	1.0
		R.VILENIA**S**HEPR.I	42	1.0	1.0
		R.L**S**ESHPDATEDLQR.Q	50	1.0	1.0
Peroxiredoxin-2	P32119	K.A**S**AVVDGAFK.E	37	0.74	0.83
		R.L**S**EDYGVLK.T	51	0.82	0.92
Erythrocyte band 4.2	P16452	R.TQATFPI**S**SLGDR.K	37	0.90	0.86
GLUT1	P11166	R.TFDEIA**S**GFR.Q	59	1.1	1.3
Equilibrative nucleoside transporter 1	Q99808	K.DAQA**S**AAPAAPLPER.N	50	1.0	1.0
Protein 4.1	P11171	R.LT**S**TDTIPK.S	43	1.0	1.1
Glutathione transferase ω-1	P78417	K.G**S**APPGPVPEGSIR.I	56	2.5	3.2
Proteasome subunit-α type 5	P28066	K.S**S**LIILK.Q	54	1.3	1.4
Catalase	P04040	R.L**S**QEDPDYGIR.D	48	1.1	1.2
		R.F**S**TVAGESGSADTVR.D	78	2.1	3.4
α-Synuclein	P37840	K.TVEGAG**S**IAAATGFVK.K	50	1.3	1.6
Aquaporin-1	P29972	R.S**S**DLTDR.V	42	1.3	1.2
Hemoglobin subunit-α	P69905	K.FLASV**S**TVLTSK.Y	65	1.2	1.2
		R.MFL**S**FPTTK.T	35	2.3	2.2
		M.VL**S**PADK.T	43	1.0	0.96

Data are means of three experiments. Peptide scores listed are the highest scores in three independent experiments. Ratio D:N, ratio between diabetic and normal samples. ROR, *O*-GlcNAc ROR. Underlined residues show site of modification.

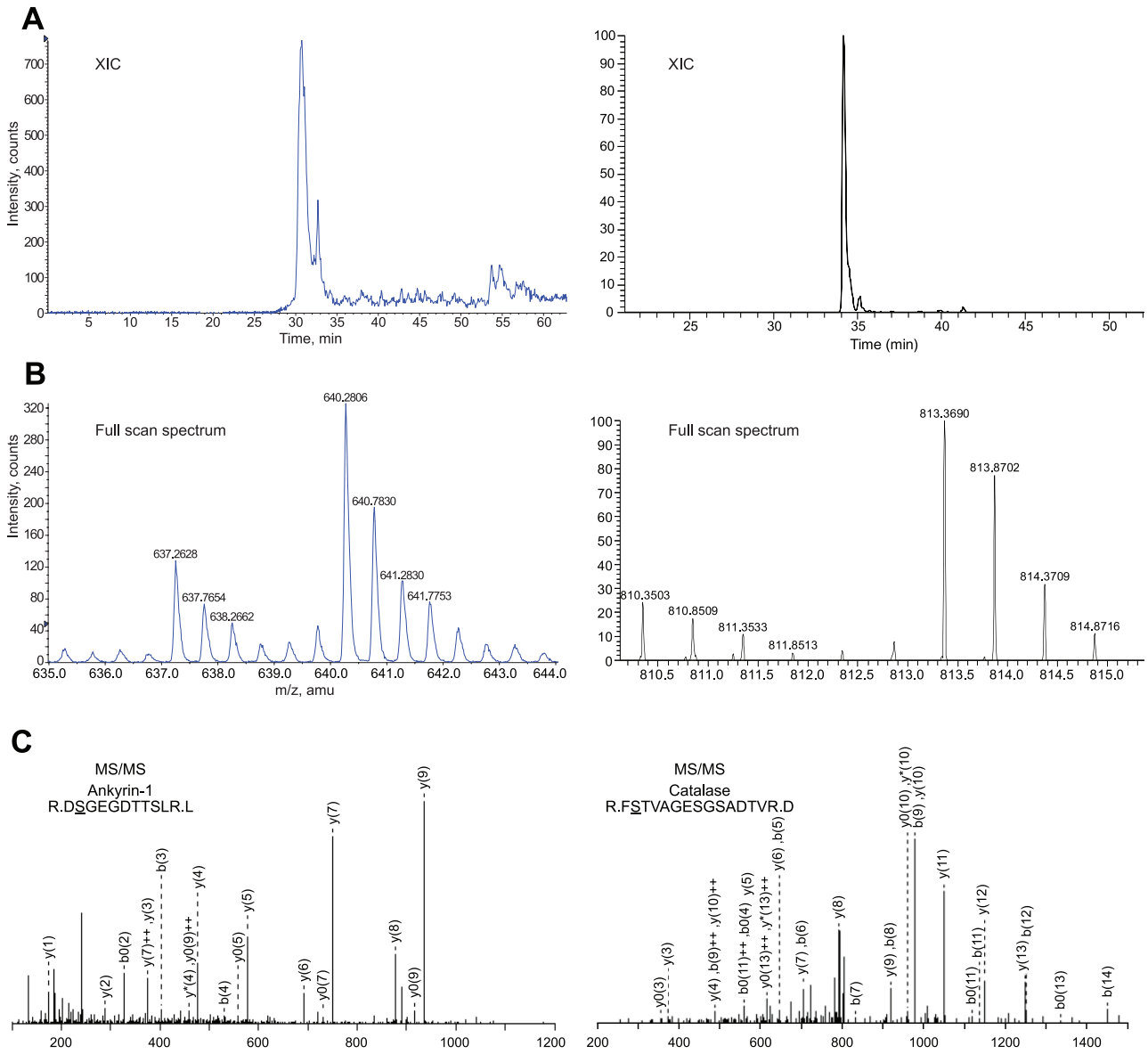

FIG. 4. O-GlcNAc as potential biomarkers for diabetes. Specific O-GlcNAc sites (underlined Ser) on ankyrin-1 (identified and quantified by QSTAR) and catalase (identified by LTQ-Orbitrap) were upregulated 2.7- and 3.9-fold, respectively. *A*: Extracted ion chromatogram (XIC). *B*: Averaged full-scan spectra during elution time of the ion pairs. *C*: MS/MS spectra that showed the peptide sequences and mapped DTT attachment sites.

with iTRAQ reagents and used it to quantitate relative changes of protein expression levels. With relative abundance of both O-GlcNAc peptides and corresponding protein levels, RORs of O-GlcNAc could then been calculated using a simple equation (see RESEARCH DESIGN AND METHODS).

Erythrocyte lysates from normal and diabetic blood donors (10 each; Table 2) were pooled separately and used as the starting materials after partial depletion of abundant hemoglobins. Three independent experiments were performed according to the flow chart shown in Fig. 3*C*. Using the standard of at least one unique peptide with a >99% confidence level, 206 erythrocyte proteins were identified and quantified (supplemental data, available in an online appendix at http://dx.doi.org/10.2337/db08-0994). Although most proteins were equally abundant, changes were observed for a few proteins between normal and diabetic samples (Fig. 3*D*). Thirty-five O-GlcNAc sites originating from 17 proteins were identified. The relative occupancy rates of O-GlcNAc at these sites between diabetic and normal states were calculated (Table 3). A negative control sample was first treated with hexosaminidase (an enzyme that removes GlcNAc) before enrichment and yielded no identification of a GlcNAcylated protein (Fig. 3*E*), indicating the specificity of the overall approach. Differentially regulated GlcNAcylation was observed on multiple sites originating from several proteins (Table 3; Fig. 4). This regulation is clearly site specific, as observed in the cases of ankyrin-1, hemoglobin α, and catalase (Table 3).

DISCUSSION

Erythrocytes are probably among the simplest of human cells. For a long time, erythrocytes had been regarded as a cytoplasm surrounded by a simplified membrane and consisting mainly of hemoglobins. A recent in-depth analysis of the erythrocyte proteome indicated that there are

likely far more complex cellular processes inside erythrocytes than previously known (24). Results presented here suggest that O-GlcNAc actively cycles on erythrocyte proteins.

Some of the earliest known GlcNAcylated proteins were detected in human erythrocytes (25). The challenges of studying O-GlcNAc by mass spectrometry come from its low stoichiometry, suppressed ionization efficiency in presence of unmodified peptides, and intrinsic lability in gas phase (21). In this study, highly efficient enrichment methods based on chemoenzymatic tagging addressed the first two challenges. Solid-phase chemical derivatization successfully circumvented the lability issue. As the exploratory phase of a project aimed at using O-GlcNAc as a potential biomarker for diagnostic of diabetes, we identified 25 O-GlcNAc modified proteins, mapped 35 O-GlcNAc sites, and compared the O-GlcNAc RORs between erythrocyte lysates obtained from normal and diabetic individuals. By using a rigorous mass spectrometric standard, we also identified 206 erythrocytic proteins and compared their abundance between normal and diabetic samples. A few proteins, such as carbonic anhydrase 1 (diabetic: normal 0.51), glutathione transferase ω 1 (0.79), GLUT1 (0.88), superoxide dismutase (1.21), and isocitrate dehydrogenase (1.29), were observed as differentially regulated in normal and diabetic samples. Although these protein level dynamics might not be conclusive because of relatively small sample size and inherent variation among individuals, these observations may reflect hyperglycemia and increased oxidative stress in diabetic patients.

Clinical diagnosis of diabetes has been evolving since the diagnostic criteria were first initiated in 1979 by the National Diabetes Data Group report (26). The glycemic criteria have been based on levels of glucose that associate with microvascular, specifically retinopathic, changes characteristic of diabetes. There are major limitations in the current criteria used for the diagnosis of diabetes. Fasting plasma glucose reflects only one aspect of glucose metabolism, which may be stated as the postabsorptive balance of hepatic glucose production and peripheral glucose uptake. It does not reflect the free-living, daily glycemic patterns, the prolonged fasted state, or the even postprandial state. The oral glucose tolerance test, in addition to being clinically cumbersome, is also nonphysiological (assuming most meal ingestion does not include 75 g concentrated sucrose). Assessing glucose tolerance with the single measure of plasma glucose 2 h after the oral glucose is therefore of limited usefulness. Another commonly used test is to assay for A1C. A1C values reflect an average glycemic status over several months' time (27). A1C assay has been recently proposed as a diagnostic criterion (28).

Perhaps the most apparent functional aspect of O-GlcNAc is its role in regulation of insulin signaling and as a mediator of glucose toxicity (2–15,29). Increasing global GlcNAcylation in adipocytes or muscle blocks insulin signaling at several points (12,29,30). Moderately increased UDP-GlcNAc levels in muscle induced insulin resistance (31), whereas overexpression of OGT in muscle or adipose causes insulin resistance and hyperleptinemia in transgenic mice (11). Diabetes is an extremely complicated syndrome. Although some controversies still exist about the roles of O-GlcNAc in diabetes (32), the results presented in this report along with the rapid cycling nature of GlcNAcylation and its sensitivity toward changes in glucose metabolism give site-specific GlcNAcylation on erythrocyte proteins great potential as biomarker(s) for detecting the early stages of diabetes.

Given the exploratory nature of the current study, quantitative measurements were based on relatively small sample sizes. Completion of the discovery phase will be followed by a validation phase, for which targeted high-throughput mass spectrometry will be adopted to determine the prevalence of O-GlcNAc dynamics on preselected sites among a large amount of samples. Polyclonal and monoclonal antibodies against O-GlcNAc on specific sites will be developed and used to screen a large number of samples from normal, pre-diabetic, and diabetic patients to further evaluate the feasibility of this approach. In any case, this study not only has identified important GlcNAcylated proteins and sites of modification in human erythrocytes but also suggests that O-GlcNAc cycling plays a role in erythrocyte biology.

ACKNOWLEDGMENTS

G.W.H. has received National Institutes of Health (NIH) Grant DK-71280 and NIH contract N01-HV-28180.

G.W.H. receives a percentage of royalties garnered by the university on sales of the CTD 110.6 antibody under a licensing agreement between Covance Research Products and The Johns Hopkins University School of Medicine. No other potential conflicts of interest relevant to this article were reported.

REFERENCES

1. Wang Z, Gucek M, Hart GW: Crosstalk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proc Natl Acad Sci U S A* 105:13793–13798, 2008
2. Copeland R, Bullen J, Hart GW: Cross-talk between GlcNAcylation and phosphorylation: roles in insulin resistance and glucose toxicity. *Am J Physiol Endocrinol Metab* 295:E17–E28, 2008
3. Love DC, Hanover JA: The hexosamine signaling pathway: deciphering the "O-GlcNAc code". *Sci STKE* 2005:re13, 2005
4. Hart GW, Housley MP, Slawson C: Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* 446:1017–1022, 2007
5. Wells L, Vosseller K, Hart GW: Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science* 291:2376–2378, 2001
6. Zachara NE, Hart GW: O-GlcNAc modification: a nutritional sensor that modulates proteasome function. *Trends Cell Biol* 14:218–221, 2004
7. McClain D, Crook E: Hexosamines and insulin resistance. *Diabetes* 45: 1003–1009, 1996
8. Housley MP, Rodgers JT, Udeshi ND, Kelly TJ, Shabanowitz J, Hunt DF, Puigserver P, Hart GW: O-GlcNAc regulates FoxO activation in response to glucose. *J Biol Chem* 283:16283–16292, 2008
9. Yang X, Ongusaha PP, Miles PD, Havstad JC, Zhang F, So WV, Kudlow JE, Michell RH, Olefsky JM, Field SJ, Evans RM: Phosphoinositide signalling links O-GlcNAc transferase to insulin resistance. *Nature* 451:964–969, 2008
10. Chia CW, Astor BC, Saudek CD: Continuous glucose excursions in everyday living: normal, pre-diabetes, and mild type 2 diabetes. *Diabetes* 52 (Suppl.):A100, 2004
11. Perreault L, Bergman BC, Playdon MC, Dalla Man C, Cobelli C, Eckel, RH: Impaired fasting glucose with or without impaired glucose tolerance: progressive or parallel states of prediabetes? *Am J Physiol Endocrinol Metab*. 295:E428–E435, 2008
12. Vosseller K, Wells L, Lane MD, Hart GW: Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in Akt activation in 3T3–L1 adipocytes. *Proc Natl Acad Sci U S A* 99:5313–5318, 2002
13. Parker G, Taylor R, Jones D, McClain D: Hyperglycemia and inhibition of glycogen synthase in streptozotocin-treated mice: role of *O*-linked *N*-acetylglucosamine. *J Biol Chem* 279:20636–20642, 2004
14. Parker GJ, Lund KC, Taylor RP, McClain DA: Insulin resistance of glycogen synthase mediated by *O*-linked *N*-acetylglucosamine. *J Biol Chem* 278: 10022–10027, 2003
15. McClain DA, Lubas WA, Cooksey RC, Hazel M, Parker GJ, Love DC,

Hanover JA: Altered glycan-dependent signaling induces insulin resistance and hyperleptinemia. *Proc Natl Acad Sci U S A* 99:10695–10699, 2002
16. Kreppel LK, Hart GW: Regulation of a cytosolic and nuclear O-GlcNAc transferase: role of the tetratricopeptide repeats. *J Biol Chem* 274:32015–32022, 1999
17. Wang Z, Pandey A, Hart GW: Dynamic interplay between *O*-linked *N*-acetylglucosaminylation and glycogen synthase kinase-3-dependent phosphorylation. *Mol Cell Proteomics* 6:1365–1379, 2007
18. Tai H, Khidekel N, Ficarro SB, Peters EC, Hsieh-Wilson LC: Parallel identification of O-GlcNAc-modified proteins from cell lysates. *J Am Chem Soc* 126:10500–10501, 2004
19. Ramakrishnan B, Qasba P: Structure-based design of β1,4-galactosyltransferase I (β4Gal-T1) with equally efficient *N*-acetylgalactosaminyltransferase activity: point mutation broadens β4gal-T1 donor specificity. *J Biol Chem* 277:20833–20839, 2002
20. Khidekel N, Ficarro SB, Peters EC, Hsieh-Wilson LC: Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain. *Proc Natl Acad Sci U S A* 101:13132–13137
21. Wang Z, Hart GW: Glycomic approaches to study GlcNAcylation: protein identification, site-mapping, and site-specific O-GlcNAc quantitation. *Clin Proteomics*. 4:5–13, 2008
22. Kolb H, Finn M, Sharpless K: Click chemistry: diverse chemical function from a few good reactions. *Angew Chem* 40:2004–2021, 2001
23. Wells L, Vosseller K, Cole R, Cronshaw J, Matunis M, Hart G: Mapping sites of *O*-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol Cell Proteomics* 1:791–804, 2002
24. Pasini EM, Kirkegaard M, Mortensen P, Lutz HU, Thomas AW, Mann M: In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood* 108:791–801, 2006
25. Holt GD, Haltiwanger RS, Torres CR, Hart GW: Erythrocytes contain cytoplasmic glycoproteins: O-linked GlcNAc on band 4.1. *J Biol Chem* 262:14847–14850, 1987
26. National Diabetes Data Group: Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 28:1039–1057, 1979
27. Saudek CD, Derr RL, Kalyani RR: Assessing glycemia in diabetes using self-monitoring blood glucose and hemoglobin A1c. *JAMA* 295:1688–1697, 2006
28. Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB: A new look at screening and diagnosing diabetes mellitus. *J Clin Endocrinol Metab* 93:2447–2453, 2008
29. Arias EB, Kim J, Cartee GD: Prolonged incubation in PUGNAc results in increased protein O-linked glycosylation and insulin resistance in rat skeletal muscle. *Diabetes* 53:921–930, 2004
30. Nelson BA, Robinson KA, Buse MG: Insulin acutely regulates Munc18-c subcellular trafficking: altered response in insulin-resistant 3T3–L1 adipocytes. *J Biol Chem* 277:3809–3812, 2002
31. Hawkins M, Barzilai N, Liu R, Hu M, Chen W, Rossetti L: Role of the glucosamine pathway in fat-induced insulin resistance. *J Clin Invest* 99:2173–2182, 1997
32. Robinson KA, Ball LE, Buse MG: Reduction of O-GlcNAc protein modification does not prevent insulin resistance in 3T3–L1 adipocytes. *Am J Physiol Endocrinol Metab* 292:E884–E890