DOI 10.1002/pmic.201300144 *Proteomics* 2014, *14*, 1020–1030

RESEARCH ARTICLE

Global identification of *O*-GlcNAc transferase (OGT) interactors by a human proteome microarray and the construction of an OGT interactome

Rui-Ping Deng[1,2*], *Xiang He*[1,2*], *Shu-Juan Guo*[1,2], *Wei-Feng Liu*[3], *Yong Tao*[3] *and Sheng-Ce Tao*[1,2,4]

[1] Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, P. R. China
[2] State Key Laboratory of Oncogenes and Related Genes, Shanghai, P. R. China
[3] CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Chinese Academy of Sciences, Beijing, P. R. China
[4] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, P. R. China

O-Linked β-*N*-acetylglucosamine (*O*-GlcNAcylation) is an important protein PTM, which is very abundant in mammalian cells. *O*-GlcNAcylation is catalyzed by *O*-GlcNAc transferase (OGT), whose substrate specificity is believed to be regulated through interactions with other proteins. There are a handful of known human OGT interactors, which is far from enough for fully elucidating the substrate specificity of OGT. To address this challenge, we used a human proteome microarray containing ~17 000 affinity-purified human proteins to globally identify OGT interactors and identified 25 OGT-binding proteins. Bioinformatics analysis showed that these interacting proteins play a variety of roles in a wide range of cellular functions and are highly enriched in intra-Golgi vesicle-mediated transport and vitamin biosynthetic processes. Combining newly identified OGT interactors with the interactors identified prior to this study, we have constructed the first OGT interactome. Bioinformatics analysis suggests that the OGT interactome plays important roles in protein transportation/localization and transcriptional regulation. The novel OGT interactors that we identified in this study could serve as a starting point for further functional analysis. Because of its high-throughput and parallel analysis capability, we strongly believe that protein microarrays could be easily applied for the global identification of regulators for other key enzymes.

Received: April 11, 2013
Revised: January 23, 2014
Accepted: January 29, 2014

Keywords:
Glycoproteomics / Human proteome microarray / *O*-GlcNAcylation / OGT / OGT interactome / Protein–protein interaction

Additional supporting information may be found in the online version of this article at the publisher's web-site

Correspondence: Professor Sheng-Ce Tao, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai 200240, P. R. China
E-mail: taosc@sjtu.edu.cn

Abbreviations: **BP**, biological process; **CC**, cellular component; **HAAO**, 3-hydroxyanthranilate 3,4-dioxygenase; **MCODE**, Molecular Complex Detection; **OGT**, *O*-GlcNAc transferase; **PSAT1**, phosphoserine aminotransferase 1; **TET**, ten eleven translocation

1 Introduction

Glycosylation is very abundant and functionally important in both eukaryotes and prokaryotes [1]. Among a special type of protein glycosylation, the addition of *O*-linked β-*N*-acetylglucosamine (*O*-GlcNAcylation) [2], exhibits

*These two authors contributed equally to this study.

Colour Online: See the article online to view Figs. 2, 3 and 5 in colour.

significantly different properties from the "classical" glycosylations [3]. Most notably, *O*-GlcNAcylation is a dynamic and inducible PTM [4] in which GlcNAc is (reversibly) added to serine and threonine residues of nuclear and cytosolic proteins [5]. The addition/removal modifications are catalyzed by a pair of enzymes, namely *O*-GlcNAc transferase (OGT) [3] and β-D-*N*-acetylglucosaminidase [6]. OGT uses uridine diphosphate *N*-acetyl glucosamine from the hexosamine biosynthetic pathway as the direct sugar donor [6], whereas β-D-*N*-acetylglucosaminidase is responsible for the removal of the *O*-GlcNAc from the proteins.

It is known that OGT could glycosylate numerous proteins, including nuclear pore proteins, tumor suppressors, transcription factors, kinases, phosphatases, and histones [7–9]. Interestingly, most of the known serine/threonine sites for *O*-GlcNAcylation are also sites for phosphorylation or adjacent to phosphorylation sites. As two of the most dynamic modifications, *O*-GlcNAcylation and phosphorylation show a complex interplay: they can either competitively occupy a single site or proximal sites, or noncompetitively occupy different sites on a substrate [6]. *O*-GlcNAcylation is functionally important and is involved in many biological processes (BPs), such as transcription, translation, cytoskeletal assembly, and signal transduction [10]. Functional studies also show that *O*-GlcNAcylation is highly related to a variety of challenging diseases, including diabetes [10], cancer [10], and degenerative diseases [11]. However, how *O*-GlcNAcylation is related to these diseases and the underlying molecular mechanisms are still elusive.

OGT is highly conserved from metazoans to human [3,12]. This protein is composed of a C-terminal catalytic domain and an N-terminal tetratricopeptide repeat domain, which mediates multimerization and protein–protein interactions [6]. There are three human OGT isoforms that can be transcribed from a single gene, namely ncOGT (p110, 110 kDa), mOGT (p103, 103 kDa), and sOGT (p78, 78 kDa) [3]. The ncOGT isoform can form either a homodimer or heterotrimer with sOGT in the cytoplasm and nucleus depending on cell type, while mOGT is localized in the mitochondria [3]. The N-terminal domain of human ncOGT contains the longest tetratricopeptide repeats of 13.5 repeats. ncOGT is the major form of OGT, and functionally, the most important. The catalytic activity of OGT is regulated by intracellular (uridine diphosphate *N*-acetyl glucosamine) [5, 6] that fluctuates proportionally in response to the flow of nutrients, such as glucose and glutamine [6,13,14]. Thus, the *O*-GlcNAcylation substrates may represent "metabolic sensors" that adjust their functions according to the cellular status of nutrition [6].

Since the discovery of *O*-GlcNAcylation, there are more than 3000 proteins that can be *O*-GlcNAcylated that have been identified [15], and it is believed that many more remain to be discovered. Although *O*-GlcNAcylation has many properties similar to that of phosphorylation, its regulation is expected to be drastically different since there are 518 protein kinases [16] and only one OGT coding gene in human [3]. How a single OGT could specifically modify more than 3000 protein substrates is a fundamental problem in the field.

One possibility is that the specificity is effected through a set of binding partners of OGT, forming putative substrate-specific holoenzyme complexes, each with unique protein substrate specificity [3]. However, to date, there are only a handful of OGT-interacting proteins that have been identified (Supporting Information Table 1) in humans [17]. To fully understand how the activity of OGT is carried out through the interaction with its protein interactors, it is essential that novel interacting partners of OGT be identified.

Proteome microarrays [18], known as a powerful technology for proteomics study, are fabricated by spotting down thousands to even ten thousands of proteins of a given species on chemically modified glass slides [18, 19]. The protein microarray technology provides a versatile platform for the characterization of hundreds of thousands of proteins in a highly parallel and high-throughput manner [20, 21]. Proteome microarrays have been applied to study a variety of types of biochemical properties of proteins and have been successfully employed in a wide range of studies that related to protein expression profiling, biomarker identification, cell surface marker/glycosylation profiling, clinical diagnosis, and environmental/food safety analysis and PTMs [18, 20–23]. Recently, a proteome microarray carrying ~17 000 human proteins was constructed [24], which could serves as an ideal platform for globally identifying OGT-interacting proteins.

Herein, we interrogated this human proteome microarray with ncOGT to globally identify OGT-interacting proteins and identified 25 novel OGT interactors. Bioinformatics analysis showed that these interacting proteins play a variety of roles in a wide range of cellular functions and are highly enriched in intra-Golgi vesicle-mediated transport and water-soluble vitamin biosynthetic process. The interactions of two of these novel interactors, phosphoserine aminotransferase 1 (PSAT1) and 3-hydroxyanthranilate 3,4-dioxygenase (HAAO), were validated by reciprocal co-IP, and the binding parameters were also measured by kinetics analysis. Interestingly, Western blotting showed that PSAT1 and HAAO were also *O*-GlcNAcylated in vivo. Combining these newly identified OGT interactors with the interactors identified prior to this study, we construct the first OGT interactome in humans. Bioinformatics analysis suggests that the OGT interactome play important roles in protein transportation/localization and transcriptional regulation. The novel OGT interactors that we discovered in this work can serve as a valuable resource for more detailed global analysis of OGT specificity and function in humans.

2 Materials and methods

2.1 Chemicals

All chemicals were obtained from Sigma unless otherwise stated. ProteinG-Agarose and Attractene Transfection

Reagent were purchased from Roche (Basel, Switzerland) and QIAGEN (Hilden, Germany), respectively. The Co-Immuno-precipitation kit and EZ-Link® Sulfo-NHS-LC-Biotinylation Kit were provided by Pierce (Rockford, USA). The GenePix 4200A fluorescence microarray scanner and analysis software GenePix Pro 6.0 were both from Axon Instruments (IL, USA). The anti-c-Myc antibody was from Cell Signaling Technology, Inc. (Danvers, USA). All the IRDye®-conjugated secondary antibodies and the Odyssey Infrared Imaging System were purchased from LI-COR (NE, USA). The ncOGT entry clone was purchased from FulenGen Co. (Guangzhou, China). The *O*-GlcNAc-specific antibody that we used in this study was purchased from Covance Inc. (Shanghai, China). The sample dilute (SD) buffer (10 mM PBS + 0.1% BSA + 0.02% Tween 20 pH 7.4) was provided by ForteBio (Menlo Park, USA). PNGase F was obtained from New England Biolabs (Beijing, China).

2.2 Fabrication of the human proteome microarray

The human proteome microarrays were fabricated as described by Jeong et al. [21]. In short, ~17 000 human proteins, which represent the majority of the human proteome were affinity purified using GST tag and Glutathione Agarose Beads. After elution, the quality of these proteins was monitored by Western blotting using an anti-GST antibody. Each protein was printed duplicate on a full moon slide from Full Moon BioSystems, Inc. (Sunnyvale, CA, USA) using a SmartArrayer 48 microarrayer from CapitalBio Co. (Beijing, China). The printed microarrays were stored at −80°C prior to use.

2.3 Identification of OGT interactors using the human proteome microarrays

Proteome microarrays were preincubated with blocking buffer (TBS with 0.1% Tween 20, and 1% BSA) at room temperature for 1 h with gentle agitation. The microarrays were then incubated with an affinity-purified human ncOGT (25 ng/μL) tagged by both GST and V5 for 1 h at room temperature. The microarrays were washed 3 × 5 min with TBS buffer with 0.1% Tween 20 (TBST), incubated with 1:1000 diluted V5 antibody in blocking buffer for 1 h at room temperature, followed by 3 × 5 min washes in TBST. The microarrays were further incubated with a Cy3-conjugated second antibody at 1:1000 dilution for 1 h at room temperature, followed by 3 × 5 min washes. The microarrays were spun dry at 1500 rpm for 3 min and subjected for scanning with a GenePix 4200A to visualize and record the results. GenePix Pro 6.0 was applied for data analysis.

2.4 Proteome microarray data processing

The proteome microarray data were extracted with GenePix Pro 6.0 and processed as previously described [25]. In brief, the signal-to-noise ratio (S/N) of each spot was set as F532 Median/B532 Median. The S/N of each protein was averaged from the two repeated spots of the same protein. The signals of the microarrays incubated with and without ncOGT were defined as S/N(+) and S/N(−). The ratio of S/N(+) and S/N(−) was set as Calling_score. To call the final potential OGT interactors, the cutoff was set as S/N(+) ≥ 3 and Calling_score ≥ 2.

2.5 Bioinformatics analysis of the ncOGT interactome

The 25 interacting protein candidates (Table 1) and the OGT interactome (Supporting Information Table 1 and Table 1) were analyzed using the PANTHER classification system with default settings [26]. GO analysis [27] was performed using DAVID Bioinformatics Resources 6.7 [28, 29] to check the enrichment of any special type of proteins of the OGT-interacting proteins. The human proteome was set as the background for DAVID analysis. The significance of the enrichments was further judged using a modified Fisher's exact test, A Benjamini–Hochberg false discovery rate correction was applied for calculating the *p* value [28, 29]. The BPs, cellular components (CCs), and molecular functions of the ncOGT interactome were analyzed. Protein interaction networks for the OGT interactome were generated by STRING and Cytoscape [30, 31]. This interaction network was further analyzed for densely connected regions using a graph theoretic clustering algorithm—"Molecular Complex Detection" (MCODE) [32]. MCODE is a plug-in in Cytoscape. The top two high-ranking modules were selected for further analysis.

2.6 Cell culture and plasmids construction

The 293T cells were routinely maintained in DMEM with 10% FBS and 1% penicillin/streptomycin at 37°C in a 5% CO_2-humidified atmosphere. For transfecting, full-length complementary DNAs of human OIP106, PSAT1, HAAO were PCR amplified and cloned into the PCMV-Myc vector [33], while ncOGT was PCR amplified and cloned into the Pflag-CMV-4 vector [34].

2.7 Co-immunoprecipitation

The 293T cells were co-transfected with ncOGT and OIP106, PSAT1, HAAO using Attractene Transfection Reagent according to the manufacturer's instruction. After 48 h post-transfection, cells were harvested, washed with PBS, and transferred to a clean 1.5 mL Eppendorf tube. For the specificity control, 40 μM of PUGNAc was added to the cell culture after 24-h posttransfection, the cells were treated for another 24 h before harvesting. Cells were lysed with lysis buffer for 10 min on the ice and cellular debris was removed by centrifugation for 20 min at 5000 × *g* and 4°C. Supernatants were

Table 1. The list of the OGT interactors identified on the human proteome microarray

No.	UniProt	Name	S/N(+)	Calling_score	Alias/description
1	P61106	RAB14	14.55	29.10	Ras-related protein Rab-14
2	P46952	HAAO	12.85	35.94	3-Hydroxyanthranilate oxygenase
3	Q9Y617	PSAT1	11.84	71.09	Phosphohydroxythreonine aminotransferase
4	Q96DG6	CMBL	9.52	8.57	Carboxymethylenebutenolidase homolog
5	Q92796	DLG3	8.20	24.15	Disks large homolog 3
6	Q9NWV4	C1orf123	8.03	64.27	UPF0587 protein C1orf123
7	P30047	GCHFR	7.68	24.57	GTP cyclohydrolase I feedback regulatory protein
8	P50995	ANXA11	7.62	4.52	56 kDa autoantigen
9	Q14353	GAMT	7.37	8.29	Guanidinoacetate *N*-methyltransferase
10	Q99829	CPNE1	6.96	3.71	Copine I
11	Q8N4Q0	ZADH2	6.84	23.93	Zinc-binding alcohol dehydrogenase domain-containing protein 2
12	Q86V88	MDP-1	6.71	60.42	Magnesium-dependent phosphatase 1
13	P51570	GALK1	6.56	14.78	Galactose kinase
14	Q92917	GPKOW	6.54	4.53	G patch domain-containing protein 5
15	O14579	COPE	5.32	47.96	Coatomer subunit epsilon
16	Q13057	COASY	5.29	3.84	Bifunctional coenzyme A synthase
17	Q99747	NAPG	5.02	7.81	*N*-Ethylmaleimide-sensitive factor attachment protein gamma
18	P16455	MGMT	4.78	4.18	6-*O*-Methylguanine-DNA methyltransferase
19	Q9H2H8	PPIL3	4.53	42.98	Cyclophilin-like protein PPIL3
20	Q8WUD1	RAB2B	4.40	63.24	Ras-related protein Rab-2B
21	Q9NZL9	MAT2B	4.29	4.06	Methionine adenosyltransferase II beta
22	Q96E17	RAB3C	4.02	32.18	Ras-related protein Rab-3C
23	Q6PH85	DCUN1D2	3.77	6.46	DCUN1 domain-containing protein 2
24	Q9UBQ0	VPS29	3.75	26.22	Vesicle protein sorting 29
25	Q9NQZ5	STARD7	3.06	8.57	Gestational trophoblastic tumor protein 1

collected and protein concentrations were measured in duplicate for antibody coupling. For each protein tested, one vial of 1 mg cell lysate was immunoprecipitated by 2 μg mouse monoclonal anti-FLAG antibody and another vial of 1 mg cell lysate was immunoprecipitated by 2 μg mouse anti-c-Myc antibody, for overnight at 4°C with gentle shaking, followed by incubation with 20 μL ProteinG Sepharose for 2 h at 4°C with gentle shaking. The beads were washed six times with 200 μL of TBS each. The supernatant was completely removed. For the specificity control, the immunoprecipitated proteins were treated with PNGase according to the manufacturer's instruction for 1 h at 37°C. Thirty microliters 1 × SDS buffer was added and boiled for 10 min at 99°C. Proteins were then resolved on a 10% SDS-PAGE gel followed by Western blotting with the reciprocal antibodies.

2.8 Western blotting

Proteins were separated on 10% SDS-PAGE gel, transferred to a NC membrane. After blocked at room temperature for 1 h with 5% nonfat milk, the membranes were probed overnight at 4°C with a rabbit monoclonal anti-c-Myc antibody and a mouse monoclonal anti-FLAG antibody. Three 5 min TBST washes with agitation were then performed. The membrane was reacted at room temperature for 2 h with an IRDye® 680CW Donkey anti-rabbit antibody and a IRDye® 800CW Donkey anti-mouse antibody. The results were then recorded by the Odyssey Infrared Imaging System.

2.9 ForteBio Octet kinetic assays

The ForteBio Octet system was applied for measuring the binding kinetics of ncOGT and its interactors. The affinity-purified ncOGT was biotinylated using the protein biotinylation kit from Pierce according to the manual. The biotinylated ncOGT was tethered on the tip surface of a streptavidin-coated sensor. The binding partner (OIP106,PSAT1,HAAO) in SD buffer is then exposed to the tethered one, and binding is measured by coincident change in the interference pattern [35, 36].

3 Results

3.1 Identified ncOGT-interacting protein candidates

To globally identify the interactors for ncOGT, a human proteome microarray with ~17 000 individually affinity-purified N-terminal GST-tagged human proteins was incubated with affinity-purified N-terminal GST- and V5-dual-tagged human ncOGT. A parallelly processed microarray with the addition of ncOGT was included as a negative control. The microarrays were further incubated with a mouse anti-V5 antibody and a Cy3-labeled anti-mouse antibody (Fig. 1A and B). This experiment has been repeated twice. Twenty-five candidates (Table 1) with an S/N(+) ≥ 3 and Calling_score ≥ 2 were determined as potential ncOGT protein interactors. Four representatives of the potential interactors are shown in Fig. 1C.

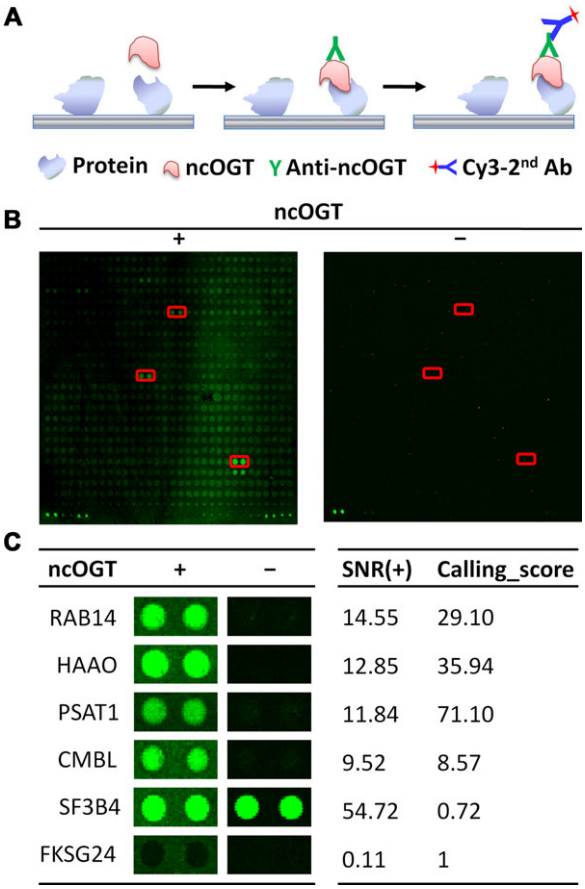


Figure 1. Identification of ncOGT-interacting proteins on a human proteome microarray. (A) The schematic of the proteome microarray strategy for identification of ncOGT-interacting proteins. (B) Two representative blocks from the same position of the microarray incubated with ncOGT and the negative control microarray without ncOGT are shown. (C) By comparing the signals from the two microarrays, 25 novel ncOGT interactors were identified. Four representative interacting proteins are shown. Two other proteins, i.e. SF3B4 and FKSG24, with nonspecific binding and without binding, respectively, are also shown. S/N(+) and Calling_score of these representatives are given.

3.2 Categorization and GO analysis of the potential ncOGT interactors

To understand the biological relevance of the potential ncOGT interactors (Table 1), we applied the online protein classification tool, PANTHER, to the 25 newly identified ncOGT interactors to identify enrichment for specific BPs, molecular functions, or protein classes. We found that the ncOGT interactors could be classified as ten groups of BP (Fig. 2A), the top one of which was metabolic processes group (GO: 0008152; 33%). Many of the interactors are also involved in the transport (GO: 0006810; 26%). The candidates could also be classified into two groups of molecular function (Fig. 2B), the larger of which was catalytic activity (GO: 0003824; 78%), and the second one is binding (GO: 0005488; 22%). Finally, the candidates could be classified into ten groups of protein class (Fig. 2C), the most significant two of which were membrane traffic protein (PC00150; 22%), transferase (PC00220; 17%). These data indicate that ncOGT is a critical regulator of diverse cellular functions involving widespread BPs.

To gain insight into possible functional roles of the newly identified ncOGT-binding proteins, the enrichment of pathways and ontology terms of the ncOGT inteactors was analyzed using DAVID in comparison to their occurrence in the human proteome. The candidate proteins were examined for enrichment in BP, CC, and molecular function. For BP, the top ten GO terms of $p < 0.05$ are shown in Fig. 3A. The candidate list is substantially enriched in intra-Golgi vesicle-mediated transport ($p = 0.026$), water-soluble vitamin biosynthetic process ($p = 0.032$), and Golgi vesicle transport ($p = 0.014$). The candidate list is also significantly enriched for carboxylic acid biosynthetic process, and establishment of protein localization. For CC, the GO terms of $p < 0.05$ are shown in Fig. 3B. There is a significant amount of the candidates localizing in coated vesicles ($p = 0.016$). Except the internal side of plasma membrane, other significantly enriched CC terms were mostly vesicle related, such as vesicle ($p = 0.008$), cytoplasmic membrane-bounded vesicle ($p = 0.004$), membrane-bounded vesicle ($p = 0.005$), and cytoplasmic vesicle ($p = 0.007$).

3.3 Reciprocal immunoprecipitation validation of the ncOGT-interacting proteins

To confirm the protein–protein interactions identified on the proteome microarray, we performed co-immunoprecipitation and Western blotting with a subset of these putative interactors. Four proteins with the highest S/N(+), namely PSAT1, HAAO, RAB14 and CMBL, and other two proteins, MDP-1 with a median S/N(+) value and RAB3C with a low S/N(+) value, were selected for further analysis. In addition, we also included OIP106 in these experiments as a positive control since it has been previously reported to physically interact with ncOGT [8]. To facilitate the reciprocal co-IP, an N-terminal FLAG-tagged ncOGT construct and N-terminal c-Myc-tagged constructs of the two putative ncOGT interactors were evaluated (Fig. 4A). The ncOGT construct, paired with the candidate constructs, was co-transfected in 293T cells, and as shown in Fig. 4B, OIP106, PSAT1, and HAAO could be readily detected in the ncOGT complex immunoprecipitated with an anti-FLAG antibody, while CMBL is negative. Furthermore, reciprocal IP-Western blotting confirmed the specific interactions of OIP106, PSAT1, and HAAO with ncOGT in cells. Thus, there is compelling evidence showing that PSAT1 and HAAO are *bona fide* ncOGT-interacting proteins.

The interaction of ncOGT with its binding partner could have two biological functions: first, the binding partner could regulate the function of ncOGT, such as determining its

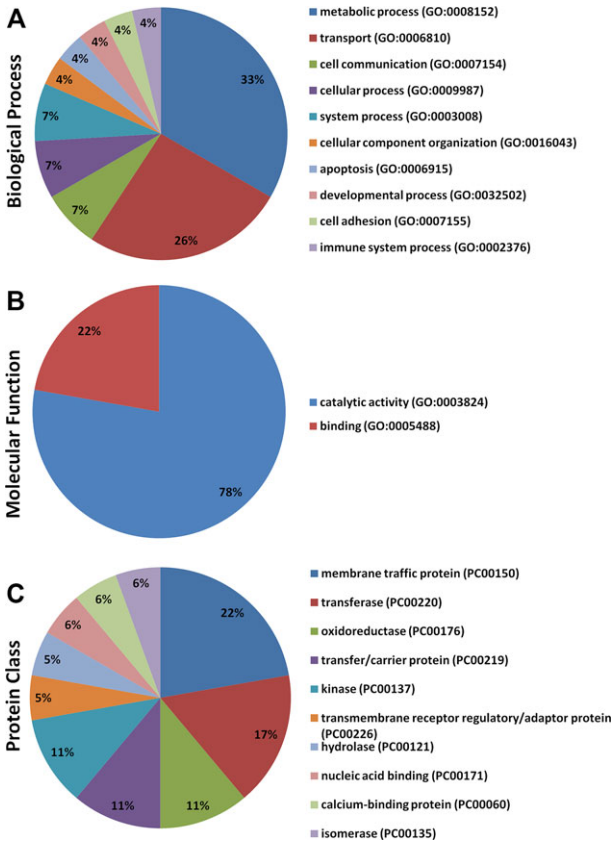


Figure 2. Functional distribution of the 25 novel ncOGT-interacting proteins according to their (A) biological process, (B) molecular function, and (C) protein class. Categorizations are based on information provided by the online resource PANTHER classification system.

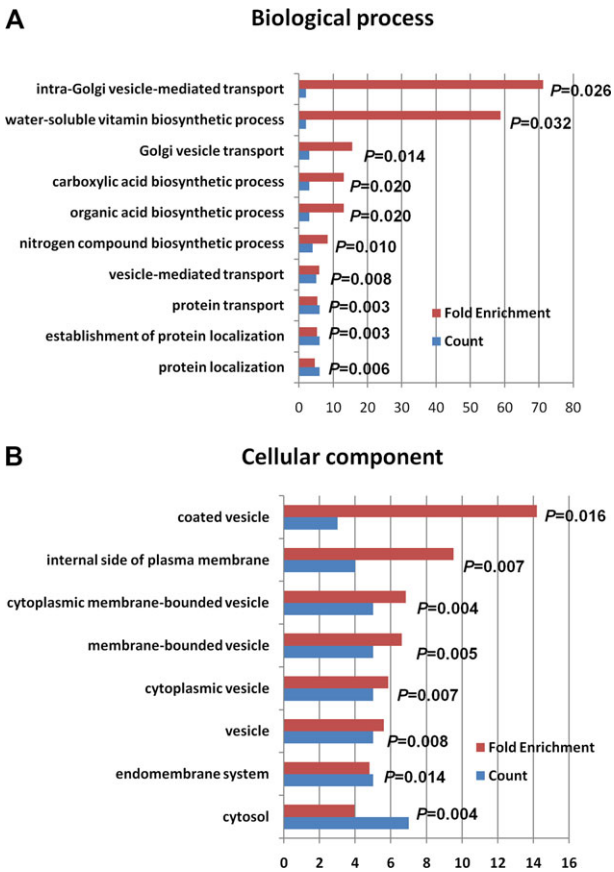
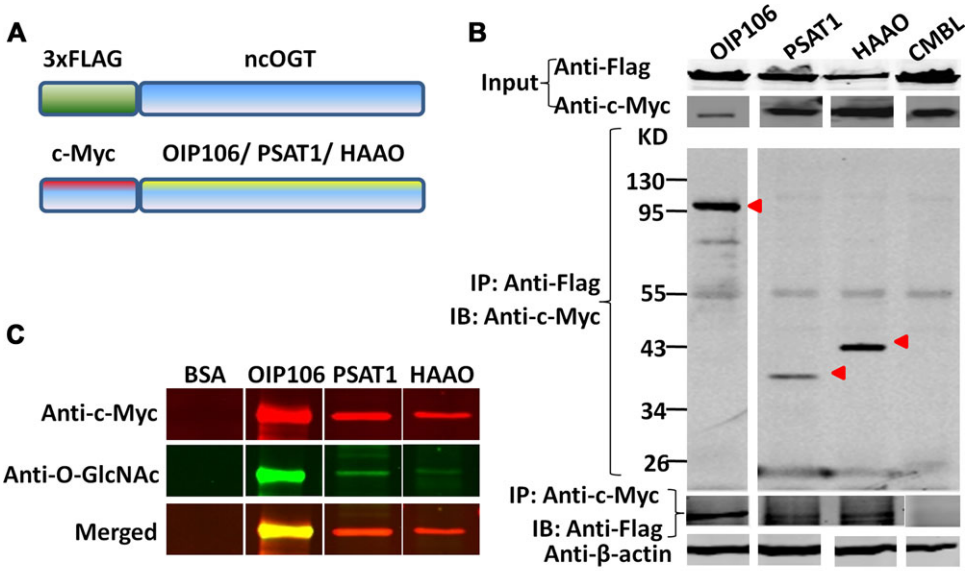


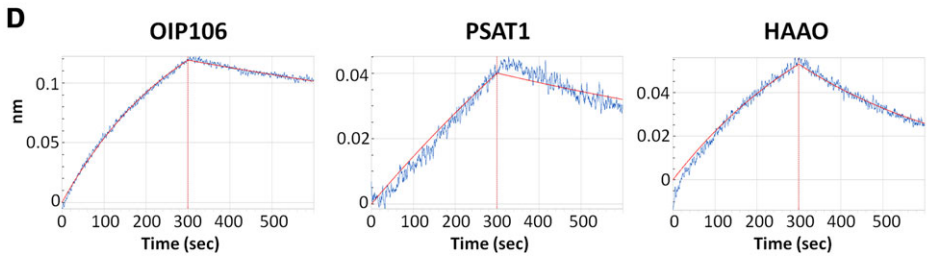
Figure 3. Functional distribution of 25 novel ncOGT-interacting proteins based on their (A) biological process and (B) cellular component. Categorizations are based on information provided by the online tool DAVID classification system.

substrate specificity, and second, the binding partner itself could be regulated through interaction with ncOGT, perhaps as a consequence of *O*-GlcNAcylation. To test the second possibility, the 293T cells transfected with c-Myc-tagged OIP106, HSAT1, or HAAO were lysed and probed with anti-c-Myc and anti-*O*-GlcNAc antibodies simultaneously, and probed with appropriate fluorescent secondary antibodies. We found that PSAT1, HAAO, and OIP106 were all positively stained by the anti-*O*-GlcNAc antibody, while there was no signal for the negative control BSA (Fig. 4C). It is possible that the anti-*O*-GlcNAc antibody may nonspecifically binds to proteins or to N-linked GlcNAc-GlcNAc. To rule out this possibility, the immunoprecipitated proteins were treated PNGase F to remove the possible *N*-glycans on the proteins. The results clearly showed that the *O*-GlcNAcylation signal was not affected by the PNGase F treatment (Supporting Information Fig. 1A). To further confirm the specificity of the anti-*O*-GlcNAc antibody, the cells were treated with an *O*-GlcNase inhibitor PUGNAc. Western blotting showed that the protein level was not affected while the *O*-GlcNAcylation signals of PSAT1, HAAO, and OIP106 were significantly enhanced (Supporting Information Fig. 1B). This result clearly indicates that PSAT1, HAAO, and OIP106 are *O*-GlcNAcylated in vivo.

3.4 Validation of ncOGT-interacting proteins by biolayer interferometry

To further validate the microarray experiments, we used biolayer interferometry to analyze the interaction and quantify the kinetic parameters of putative ncOGT interactors. In this work, we evaluated the interaction between ncOGT and PSAT1 and HAAO, and also between ncOGT and OIP106 and BSA as positive and negative controls, respectively. This experiment clearly showed strong interactions between ncOGT and PSAT1, HAAO, and OIP106, but none between ncOGT and BSA under the same conditions. From this measurement, we determined the equilibrium dissociation constants (K_D) for the interaction between ncOGT and OIP16, PSAT1, and HAAO to be 94, 3380, 18 700 nM, respectively (Fig. 4D). Hence, as might be expected from their identification in the





Name	K_D (M)	K_{on} (1/Ms)	K_{dis} (1/s)
OIP106	9.39×10^{-8}	5.71×10^{3}	5.36×10^{-4}
PSAT1	3.38×10^{-6}	2.22×10^{2}	7.51×10^{-4}
HAAO	1.87×10^{-5}	2.40×10^{2}	2.39×10^{-3}

Figure 4. Validation of the interaction between ncOGT and the novel interactors by reciprocal co-IP and biolayer interferometry. (A). The constructs for the co-IP validation. (B) Reciprocal co-IP of ncOGT and the interactors (PSAT1 and HAAO), OIP106 was set as positive control. (C) The ncOGT interactors (PSAT1, HAAO, and OIP106) are *O*-GlcNAcylated in vivo. PSAT1 and HAAO are immunoprecipitated using an anti-c-Myc-antibody. OIP106 and BSA are included as positive and negative control, respectively. (D) Binding kinetics of ncOGT immobilized on SA Bio-sensors and the ncOGT interactors (PSAT1, HAAO, and OIP106).

microarray experiment, these interactions are relatively high affinity.

3.5 Construction and bioinformatic analysis of the OGT interactome

Previous to our work, there were 57 OGT interactors identified, somewhat sporadic studies (Supporting Information Table 1) [17]. To provide a more system-wide level of understanding these interactors, we combined the human OGT interactors discovered from previous studies and our study to produce a single set of the OGT interactome.

To understand the biological relevance of the OGT interactome, protein classification analysis was performed using PANTHER. As shown in Supporting Information Table 2, the top three BP groups are metabolic processes (GO: 0008152; 25.1%), cellular process (GO: 0009987; 12.6%), and transport (GO: 0006810; 12.6%). The top CC group is intracellular (GO: 0005622; 63.6%), which indicates that a significant portion of the OGT interactome is cytosolic. In addition, the top three molecular function groups are catalytic activity (GO: 0003824; 35.7%), binding (GO: 0005488; 27.1%), and transcription regulator activity (GO: 0030528; 14.3%). The top three pathways are Wnt signaling pathway (P00057; 10.5%), Parkinson disease (P00049; 8.8%), and gonadotropin-releasing hormone receptor pathway (P06664; 8.8%). Finally, the top three protein classification categories are transferase (PC00220; 17%), nucleic acid binding (PC00171; 12.8%), and transcription factor (PC00218; 10.6%).

To understand the functional roles of the OGT interactome, the enrichment of GO terms and pathways was analyzed using DAVID. As shown in Supporting Information Table 3, the most enriched GO terms of BP include positive

regulation of protein complex assembly (Fold Enrichment >21 and p value $< 1 \times 10^{-3}$) and protein transport (Fold Enrichment >3 and p value of 0.0012). The top enriched CC term is nuclear lumen (Fold Enrichment of 2.48 and p value of 0.0005). The most enriched molecular function terms are transcription factor binding (Fold Enrichment of 4.3 and p value of 0.0001) and *N*-acetyltransferase activity (Fold Enrichment of 10.6 and p value of 0.006), while the most significant enriched pathways are Hedgehog signaling pathway (Fold Enrichment of 12.52 and p value of 0.0034) and Wnt signaling pathway (Fold Enrichment of 5.8 and p value of 0.0088).

To create significance out of the protein list of the OGT interactome, we constructed biological interaction networks of these proteins. The ncOGT-interacting proteins that we identified (Table 1) and the previously known ones (Supporting Information Table 1) were combined and imported into STRING and Cytoscape to build the network. After removing isolated nodes, a network with 67 protein nodes was constructed (Fig. 5A). To better understand the BPs, it is very important to map interactions between protein complexes. To characterize protein complexes which associate with OGT, we therefore analyzed the OGT interactome for highly connected regions using MCODE, and determined the interconnectivity between these complexes and OGT by protein network analysis. As shown in Fig. 5B and C, two highly connected clusters were identified, both of them are associated with transcriptional regulation. These results suggest that *O*-GlcNAcylation play important roles in regulating gene expression.

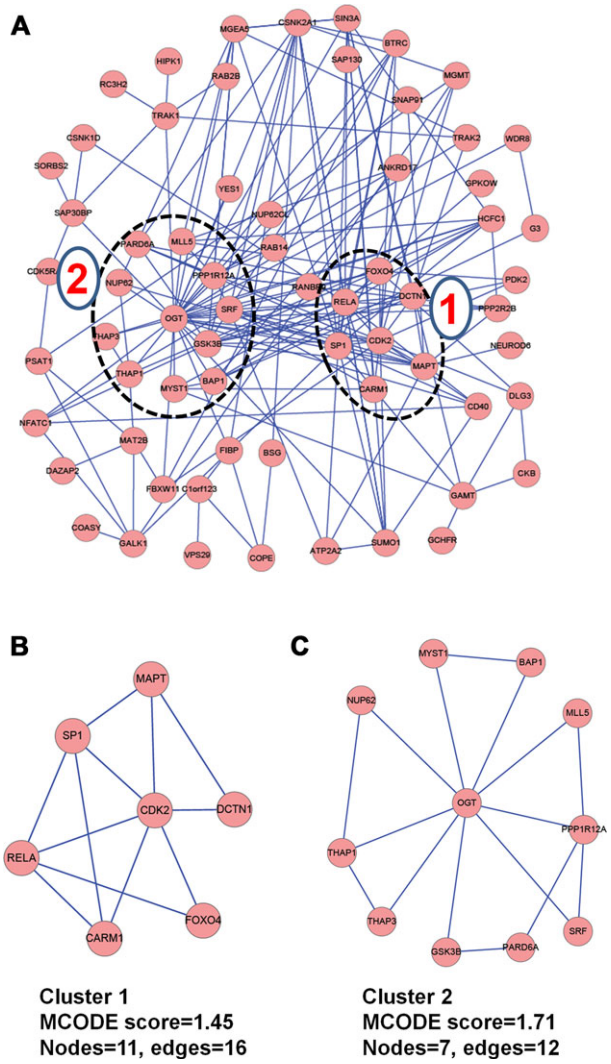


Figure 5. The protein–protein interaction network of the OGT interactome. (A) The network. (B, C) The two most significant subnetworks. The network was generated by the online tool STRING and Cytoscape 2.8. The significant subnetworks were identified by the Cytoscape plug-in MCODE.

4 Discussion

O-GlcNAcylation on serine/threonine is one of the most abundant PTMs in mammalian cells, where it plays key roles in numerous BPs and is involved in many challenging diseases. There are more than 1000 proteins that could be *O*-GlcNAcylated in human cells [15] and many more remain to be discovered. However, there is only one enzyme, OGT, that is responsible for *O*-GlcNAcylation in human cells. It is known that the substrate specificity of OGT is regulated by its interacting proteins [3]. However, prior to our work, there were only 57 known human OGT interactors, which is not enough to fully elucidate the substrate specificity of OGT. A key step in understanding the substrate specificity and physiological function of OGT is assigning OGT to appropriate biochemical pathways through the identification of its interacting partners. In this study, we have performed a global identification of OGT interactors using a human proteome microarray containing ~17 000 affinity-purified human proteins. Twenty-five OGT interactors were successfully identified. Bioinformatics analysis showed that these interacting proteins play roles in a wide range of cellular functions and are highly enriched in intra-Golgi vesicle-mediated transport and vitamin biosynthetic process. The interactions between OGT and several of the newly identified interactors, e.g. PSAT1 and HAAO were validated by reciprocal co-IP and further characterized by kinetics analysis. Western blotting showed that PSAT1 and HAAO were also *O*-GlcNAcylated in vivo. Combining newly identified OGT interactors with the interactors identified prior to this study, we have constructed the first OGT interactome. Bioinformatics analysis suggests that the OGT interactome play important roles in protein transportation/localization and transcriptional regulation.

There are several reasons that the human proteome microarray is a suitable platform for global identification of protein–protein interactions for OGT: (1) all the human proteins on the proteome microarray were overexpressed and purified from an eukaryotic system, *Saccharomyces cerevisiae*, presumably preserving native PTMs, such as acetylation and

phoshporylation, which should thus maintain their designated biological functions and binding activities; (2) every protein on the proteome microarray is immobilized on an addressable location, which is very convenient for the identification of novel protein–protein interactions; (3) the amount of the overexpressed proteins in *S. cerevisiae* is generally higher than that from mammalian cells. Thus, the local concentrations of proteins on the proteome microarray are relatively high, which could greatly facilitate the identification of OGT interactors. Naturally, the human proteome microarray is not limited to studies with OGT: it is generally applicable for global protein-protein interaction (PPI) studies of many other proteins of interest.

Both protein classification analysis (Fig. 2) and GO analysis (Fig. 3) showed that the OGT interactors (Table 1) from the human proteome microarray experiment were significant enriched for vesicle-mediated transport, specifically, intra-Golgi vesicle-mediated transport. By GO definition, intra-Golgi vesicle-mediated transport is the directed movement of substances, usually proteins, within the Golgi, mediated by small transport vesicles. The specific targeting of different classes of transport vesicles to their distinct membrane destinations is essential to maintain the unique composition of Golgi. It is shown that different targeting reactions involve distinct protein complexes that act to mark the target organelle for incoming vesicles [37]. Thus, it is possible that OGT regulates the intra-Golgi vesicle-mediated transport through direct interactions with the components of specific protein complexes, such as NAPG and COPE. Besides direct interaction, OGT may also regulate vesicle-mediated transport through *O*-GlcNAcylates the component of some of the key protein complex; indeed, a component of COPII complex-Sec24p is *O*-GlcNAcylated during interphase [38]. When cells enter mitosis, the *O*-GlcNAc modification is replaced by phosphorylation. This could be used to explain why ER-to-Golgi transport is blocked in mammalian cells during mitosis.

The interactions of OGT with its interactors may have multiple functional roles. Besides regulating the substrate specificity of OGT, the interactors may also regulate the activity of OGT, and vice versa. For example, Capotosti et al. [39] showed that OGT recognizes the human epigenetic cell cycle regulator HCF-1 at the six centrally located HCF-1-repeat sequences. Besides the designated *O*-GlcNAcylaion activity, OGT also showed specific protease activity on the HCF-1-repeat sequences, which mediated the activation of HCF-1(C)-subunit in M phase progression. Chen et al. [40] also found that ten eleven translocation (TET)2 and TET3 of the TET enzymes, which are responsible for converting 5-methylcytosine to 5-hydroxymethylcytosine and regulate gene transcription, bind to OGT. Functional study showed that this interaction facilitates OGT-dependent histone *O*-GlcNAcylation with functional consequences in transcription regulation.

This point could be further strengthened by the observation found here that the two reciprocal Co-IP assays validated that the OGT interactors PSAT1 and HAAO were also *O*-GlcNAcylated in vivo. PSAT1, is an enzyme involved in serine biosynthesis and has been linked with cell proliferation in vitro [41]. Studies also have shown that PSAT1 is involved in schizophrenia, breast cancer, and colon cancer [41, 42]. HAAO is an enzyme responsible for the conversion of 3-hydroxyanthranilic acid (2-amino-3-hydroxybenzoic acid; 3HANA) to quinolinic acid (2,3-pyridinedicarboxylic acid; QUIN) [43]. HAAO is widely distributed in peripheral organs, such as liver and kidney, and is also present in low amounts in the CNS. HAAO is also a potent excitotoxin that selectively activates *N*-methyl-D-aspartate receptors and it has been linked to neurological disorders such as epilepsy, Huntington's disease, hepatic encephalopathy, and AIDS-related dementia [44–46]. Recent research showed that hypermethylation of HAAO was associated with microsatellite instability in endometrial carcinomas [47] and HAAO may be a novel candidate epigenetic biomarker for ovarian cancer detection [48]. The interaction of OGT with PSAT1 and HAAO, and the in vivo *O*-GlcNAcylation of these two proteins suggest OGT may regulate their activity through either binding and/or glycosylation.

Somewhat surprisingly, there is no overlap between our newly discovered OGT interactors (Table 1) and the known OGT interactors (Supporting Information Table 1). However, as many other studies have shown different methods, especially proteomic methods possess distinct strengths and weaknesses and provide complementary types of information [49]. For example, the proteome microarrays detect binary interactions, while the MS-based techniques identify protein complexes. The MS-based techniques identify interactions at approximately physiological protein concentrations, however, the protein concentrations are more homogenous among proteins on the proteome microarray and this may facilitate identification of interacting proteins of low level under physiological conditions. No overlap between the two lists may also indicate that there are possibly many more human OGT-interacting proteins to be discovered. Therefore, to take advantage of the complementary capability and to systematically understand the substrate specificity and the function of OGT, we combined the human OGT interactors discovered from previous studies (Supporting Information Table 1) and our study (Table 1) into a single set as the OGT interactome. To prove the reliability of the proteome microarray-based strategy, we have already validated a few newly discovered candidates as representatives in this study; however, further validation may be helpful to make full use of the OGT interactome. Both protein classification analysis (Supporting Information Table 2) and GO analysis (Supporting Information Table 3) showed that the OGT interactome was significantly enriched for protein transport/localization, transcription factor binding, and the Wnt signaling pathway. The enrichment of protein transport/localization is consistent with the enrichment of vesicle-mediated transport of the OGT interactors that we identified on the human proteome microarray. The enrichment in transcription factor binding and the Wnt signaling pathway indicates that OGT may be a key regulator of gene expression and signaling pathways,

which is consistent with previous recent study showing that TET2 interacted with OGT and facilitates OGT-dependent histone *O*-GlcNAcylation, which are associated with gene transcription regulation [40].

Taken together, we have performed global Protein-Protein Interaction analysis of human OGT using the human proteome microarray. Combining the OGT interactors identified in this study and those from literatures, we have constructed, to our knowledge, the first OGT interactome. Our results expand the number of proteins that could directly interact with OGT. We believe that the novel OGT interactors identified in this study and the OGT interactome that we have constructed could serve as a valuable resource for a further study of the role of *O*-GlcNAcylation. It is now important to further characterize the interactions of OGT with different binding partners in different physiological and pathophysiological contexts. We believe that the outcome of such studies will facilitate both basic study and clinical research.

We thank Dr. Daniel Czajkowsky for critical reading. This work was supported by the State Key Development Program for Basic Research of China (grant no. 2010CB529205), the National High Technology Research and Development Program of China (nos. 2012AA020103 and 2012AA020203), and the National Natural Science Foundation of China (grant nos. 31370813 and 31000388).

The authors have declared no conflict of interest.

5 References

[1] Moremen, K. W., Tiemeyer, M., Nairn, A. V., Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* 2012, *13*, 448–462.

[2] Torres, C. R., Hart, G. W., Topography and polypeptide distribution of terminal N-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for O-linked GlcNAc. *J. Biol. Chem.* 1984, *259*, 3308–3317.

[3] Hu, P., Shimoji, S., Hart, G. W., Site-specific interplay between O-GlcNAcylation and phosphorylation in cellular regulation. *FEBS Lett.* 2010, *584*, 2526–2538.

[4] Wells, L., Whalen, S. A., Hart, G. W., O-GlcNAc: a regulatory post-translational modification. *Biochem. Bioph. Res. Co.* 2003, *302*, 435–441.

[5] Kreppel, L. K., Hart, G. W., Regulation of a cytosolic and nuclear O-GlcNAc transferase—role of the tetratricopeptide repeats. *J. Biol. Chem.* 1999, *274*, 32015–32022.

[6] Zeidan, Q., Hart, G. W., The intersections between O-GlcNAcylation and phosphorylation: implications for multiple signaling pathways. *J. Cell. Sci.* 2010, *123*, 13–22.

[7] Zachara, N. E., Hart, G. W., The emerging significance of O-GlcNAc in cellular regulation. *Chem. Rev.* 2002, *102*, 431–438.

[8] Iyer, S. P. N., Akimoto, Y., Hart, G. W., Identification and cloning of a novel family of coiled-coil domain proteins that interact with O-GlcNAc transferase. *J. Biol. Chem.* 2003, *278*, 5399–5409.

[9] Zhang, S. S., Roche, K., Nasheuer, H. P., Lowndes, N. F., Modification of histones by sugar beta-N-acetylglucosamine (GlcNAc) occurs on multiple residues, including histone H3 serine 10, and is cell cycle-regulated. *J. Biol. Chem.* 2011, *286*, 37483–37495.

[10] Slawson, C., Copeland, R. J., Hart, G. W., O-GlcNAc signaling: a metabolic link between diabetes and cancer? *Trends Biochem. Sci.* 2010, *35*, 547–555.

[11] Issad, T., Masson, E., Pagesy, P., O-GlcNAc modification, insulin signaling and diabetic complications. *Diabetes Metab.* 2010, *36*, 423–435.

[12] Marz, P., Stetefeld, J., Bendfeldt, K., Nitsch, C. et al., Ataxin-10 interacts with O-linked beta-N-acetylglucosamine transferase in the brain. *J. Biol. Chem.* 2006, *281*, 20263–20270.

[13] Marshall, S., Nadeau, O., Yamasaki, K., Dynamic actions of glucose and glucosamine on hexosamine biosynthesis in isolated adipocytes—differential effects on glucosamine 6-phosphate, UDP-N-acetylglucosamine, and ATP levels. *J. Biol. Chem.* 2004, *279*, 35313–35319.

[14] McClain, D. A., Crook, E. D., Hexosamines and insulin resistance. *Diabetes* 1996, *45*, 1003–1009.

[15] Zachara, N. E., Vosseller, K., Hart, G. W., Detection and analysis of proteins modified by O-linked N-acetylglucosamine. *Current Protocols in Protein Science* 2011, *66*, 12.8.1–12.8.33.

[16] Manning, G., Whyte, D. B., Martinez, R., Hunter, T., Sudarsanam, S., The protein kinase complement of the human genome. *Science* 2002, *298*, 1912–1934.

[17] Cheung, W. D., Sakabe, K., Housley, M. P., Dias, W. B., Hart, G. W., O-Linked beta-N-acetylglucosaminyltransferase substrate specificity is regulated by myosin phosphatase targeting and other interacting proteins. *J. Biol. Chem.* 2008, *283*, 33935–33941.

[18] Zhu, H., Qian, J., Applications of functional protein microarrays in basic and clinical research. *Adv. Genet.* 2012, *79*, 123–155.

[19] Zhu, H., Bilgin, M., Bangham, R., Hall, D. et al., Global analysis of protein activities using proteome chips. *Science* 2001, *293*, 2101–2105.

[20] Tao, S. C., Chen, C. S., Zhu, H., Applications of protein microarray technology. *Comb. Chem. High. T. Scr.* 2007, *10*, 706–718.

[21] Jeong, J. S., Rho, H. S., Zhu, H., A functional protein microarray approach to characterizing posttranslational modifications on lysine residues. *Methods Mol. Biol.* 2011, *723*, 213–223.

[22] Yang, L. N., Guo, S. J., Li, Y., Zhou, S. M., Tao, S. C., Protein microarrays for systems biology. *Acta Bioch. Bioph. Sin.* 2011, *43*, 161–171.

[23] Chen, C. S., Korobkova, E., Chen, H., Zhu, J. et al., A proteome chip approach reveals new DNA damage recognition activities in *Escherichia coli*. *Nat. Methods* 2008, *5*, 69–74.

[24] Jeong, J. S., Jiang, L. Z., Albino, E., Marrero, J. et al., Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. *Mol. Cell. Proteomics* 2012, *11*, O111.016253.

[25] Wang, Z. X., Deng, R. P., Jiang, H. W., Guo, S. J. et al., Global identification of prokaryotic glycoproteins based on

an *Escherichia coli* proteome microarray. *PloS One* 2012, *7*, e49080.

[26] Mi, H. Y., Guo, N., Kejariwal, A., Thomas, P. D., PANTHER version 6: protein sequence and function evolution data with expanded representation of biological pathways. *Nucleic Acids Res.* 2007, *35*, D247–D252.

[27] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D. et al., Gene Ontology: tool for the unification of biology. *Nat. Genet.* 2000, *25*, 25–29.

[28] Huang, D. W., Sherman, B. T., Lempicki, R. A., Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 2009, *4*, 44–57.

[29] Riedel, T., Held, B., Nolan, M., Lucas, S. et al., Genome sequence of the orange-pigmented seawater bacterium Owenweeksia hongkongensis type strain (UST20020801(T)). *Stand. Genomic Sci.* 2012, *7*, 120–130.

[30] Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M. et al., The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* 2011, *39*, D561–D568.

[31] Maere, S., Heymans, K., Kuiper, M., BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* 2005, *21*, 3448–3449.

[32] Bader, G. D., Hogue, C. W., An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 2003, *4*, 2.

[33] Li, C. Y., Peng, J., Ren, L. P., Gan, L. X. et al., Roles of histone hypoacetylation in LAT expression on T cells and Th2 polarization in allergic asthma. *J. Transl. Med.* 2013, *11*, 26.

[34] Nyaga, S. G., Lohani, A., Jaruga, P., Trzeciak, A. R. et al., Reduced repair of 8-hydroxyguanine in the human breast cancer cell line, HCC1937. *BMC Cancer* 2006, *6*, 297.

[35] Wilson, J. L., Scott, I. M., McMurry, J. L., Optical biosensing. *Biochem. Mol. Biol. Edu.* 2010, *38*, 400–407.

[36] Abdiche, Y., Malashock, D., Pinkerton, A., Pons, J., Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet. *Anal. Biochem.* 2008, *377*, 209–217.

[37] Guo, W., Sacher, M., Barrowman, J., Ferro-Novick, S., Novick, P., Protein complexes in transport vesicle targeting. *Trends Cell. Biol.* 2000, *10*, 251–255.

[38] Dudognon, P., Maeder-Garavaglia, C., Carpentier, J. L., Paccaud, J. P., Regulation of a COPII component by cytosolic O-glycosylation during mitosis. *FEBS Lett.* 2004, *561*, 44–50.

[39] Capotosti, F., Guernier, S., Lammers, F., Waridel, P. et al., O-GlcNAc Transferase catalyzes site-specific proteolysis of HCF-1. *Cell* 2011, *144*, 376–388.

[40] Chen, Q., Chen, Y., Bian, C., Fujiki, R., Yu, X., TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* 2013, *493*, 561–564.

[41] Ozeki, Y., Pickard, B. S., Kano, S., Malloy, M. P. et al., A novel balanced chromosomal translocation found in subjects with schizophrenia and schizotypal personality disorder: altered L-serine level associated with disruption of PSAT1 gene expression. *Neurosci. Res.* 2011, *69*, 154–160.

[42] Vie, N., Copois, V., Bascoul-Mollevi, C., Denis, V. et al., Overexpression of phosphoserine aminotransferase PSAT1 stimulates cell growth and increases chemoresistance of colon cancer cells. *Mol. Cancer* 2008, *7*, 14.

[43] Dilovic, I., Gliubich, F., Malpeli, G., Zanotti, G., Matkovic-Calogovic, D., Crystal structure of bovine 3-hydroxyanthranilate 3,4-dioxygenase. *Biopolymers* 2009, *91*, 1189–1195.

[44] Heyes, M. P., Metabolism and neuropathologic significance of quinolinic acid and kynurenic acid. *Biochem. Soc. Trans.* 1993, *21*, 83–89.

[45] Schwarcz, R., Tamminga, C. A., Kurlan, R., Shoulson, I., Cerebrospinal fluid levels of quinolinic acid in Huntington's disease and schizophrenia. *Ann. Neurol.* 1988, *24*, 580–582.

[46] Achim, C. L., Heyes, M. P., Wiley, C. A., Quantitation of human immunodeficiency virus, immune activation factors, and quinolinic acid in AIDS brains. *J. Clin. Invest.* 1993, *91*, 2769–2775.

[47] Huang, Y. W., Luo, J., Weng, Y. I., Mutch, D. G. et al., Promoter hypermethylation of CIDEA, HAAO and RXFP3 associated with microsatellite instability in endometrial carcinomas. *Gynecol. Oncol.* 2010, *117*, 239–247.

[48] Huang, Y. W., Jansen, R. A., Fabbri, E., Potter, D. et al., Identification of candidate epigenetic biomarkers for ovarian cancer detection. *Oncol. Rep.* 2009, *22*, 853–861.

[49] Yu, X., Ivanic, J., Memisevic, V., Wallqvist, A., Reifman, J., Categorizing biases in high-confidence high-throughput protein-protein interaction data sets. *Mol. Cell. Proteomics* 2011, *10*, M111 012500.

