ACS Editors' Choice

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Terms of Use

Microarray Discovery of New OGT Substrates: The Medulloblastoma Oncogene OTX2 Is *O*-GlcNAcylated

Rodrigo F. Ortiz-Meoz,[†] Yifat Merbl,[‡] Marc W. Kirschner,[‡] and Suzanne Walker*[,†]

[†]Department of Microbiology and Immunobiology, and [‡]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, United States

Supporting Information

ABSTRACT: *O*-GlcNAc transferase (OGT) is a serine/threonine glycosyltransferase that is essential for development and continues to be critically important throughout life. Understanding OGT's complex biology requires identifying its substrates. Here we demonstrate the utility of a microarray approach for discovering novel OGT substrates. We also report a rapid method to validate OGT substrates that combines in vitro transcription-translation with *O*-GlcNAc mass tagging. Among the validated new OGT targets is Orthodenticle homeobox 2 (OTX2), a transcription factor critical for brain development, which is primarily expressed only during early embryogenesis and in medulloblastomas, where it functions as an oncogene. We show that endogenous OTX2 from a medulloblastoma cell line is *O*-GlcNAcylated at several sites. Our results demonstrate that protein microarray technology, combined with the target validation strategy we report, is useful for identifying biologically important OGT substrates, including substrates not present in most tissue types or cell lines.

Protein post-translational modifications (PTMs) are crucial for the proper metabolic control of protein activity. One of the most common PTMs in eukaryotes is *O*-GlcNAcylation, the attachment of *N*-acetyl glucosamine (GlcNAc) to serine and threonine side chains of intracellular proteins. This modification is effected by *O*-GlcNAc transferase (OGT), an enzyme essential for development that has been implicated in a variety of different signaling pathways.[1] OGT targets include kinases, phosphatases, transcription factors, and metabolic enzymes, and OGT has been implicated in insulin signaling, stress response pathways, and cell-cycle regulation, among other processes. Dysregulated *O*-GlcNAcylation has been linked to cancer, diabetic complications, and other pathologies.[2] Deconvoluting OGT's myriad functions requires identifying its cellular substrates and considerable effort has been devoted to this endeavor. The vast majority of known substrates were identified using proteomic approaches based on liquid chromatography tandem mass spectrometry (LC-MS/MS). These methods require that *O*-GlcNAcylated cellular proteins be enriched from the general protein population prior to LC-MS/MS analysis.[3] A number of enrichment approaches have been developed based on lectin affinity capture or bioorthogonal labeling with biotin followed by streptavidin affinity capture.[4] While proteomic approaches are powerful, they have limitations. Enrichment methods can isolate proteins that are not *O*-GlcNAcylated as well as ones that are, resulting in false positives, and some low abundance OGT substrates may still be missed. Furthermore, *O*-GlcNAcylated proteins expressed only during development or in specialized cell types will not be identified in standard cell lines.

Human protein microarrays can circumvent common pitfalls of methods that rely on substrate enrichment from complex mixtures. Thousands of purified human proteins, including rare and disease-relevant proteins, are present in known, roughly equivalent concentrations on these arrays with their identities spatially encoded.[5] Large microarrays have been used to identify substrates for phosphorylation, acetylation and ubiquitylation, but have not been applied to *O*-GlcNAcylation.[6] Here we demonstrate the utility of large microarrays for discovering novel OGT substrates. Among a dozen new substrates confirmed using an efficient, orthogonal validation method is OTX2, a master regulator of neural development and a known oncogene for medulloblastoma, the most common malignant brain tumor in children.[7]

Human OGT contains a C-terminal catalytic domain fused to an N-terminal tetratricopeptide repeat (TPR) domain, which is involved in protein−protein interactions.[8] OGT is expressed as three different isoforms that contain identical catalytic domains but different numbers of TPRs: nucleocytoplasmic OGT (ncOGT) contains 12.5 TPRs, mitochondrially targeted OGT (mOGT) contains 9.5 TPRs, and short OGT (sOGT) contains 2.5 TPRs.[9] We purified each isoform from a bacterial overexpression system and tested them for glycosylation of protein substrates in HeLa cell extracts.[10] ncOGT and mOGT, but not sOGT, showed robust glycosylation activity in these extracts (Figure S1), consistent with previous reports.[9] Therefore, we used ncOGT and mOGT to probe a commercially available microarray containing ∼8000 unique human proteins. Microarrays were incubated with 40 μM UDP-GlcNAc and 3 μM purified ncOGT or mOGT at room temperature for one hour, and *O*-GlcNAc-modified proteins were then detected using a monoclonal anti-*O*-GlcNAc primary antibody (CTD110.6, Sigma) and a fluorescent secondary antibody (anti-mouse IgG, Alexa-fluor 646 conjugate, Cellsignal) (Figure 1a).[11] Because the anti-*O*-GlcNAc antibody CTD110.6 can react with pre-existing *O*-GlcNAc residues on the arrayed proteins, and has also been shown to bind to both O- and N-linked glycans,[12] we incubated a control microarray

Received: January 15, 2014
Published: February 28, 2014

   dx.doi.org/10.1021/ja500451w | *J. Am. Chem. Soc.* 2014, 136, 4845−4848

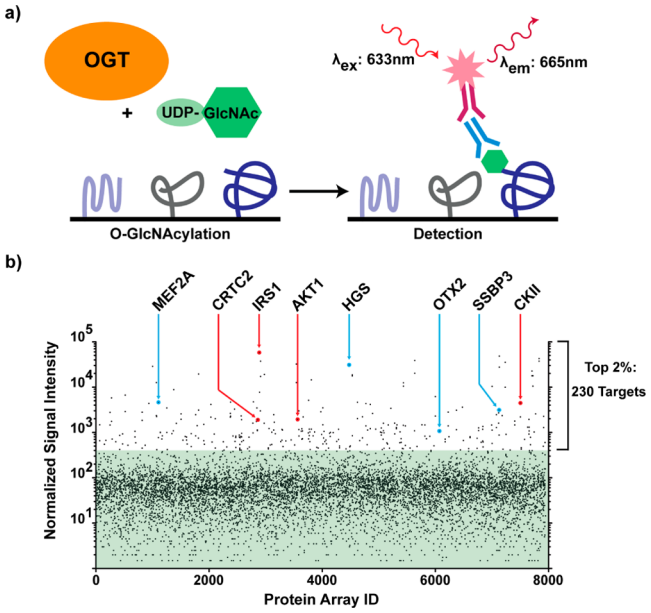

Figure 1. Known and novel OGT targets can be identified on microarrays. (a) Schematic of the microarray approach. A fluorescent antibody reports on *O*-GlcNAcylated proteins. (b) Normalized signal intensities plotted against Array ID. Selected known (red) and novel (blue) targets are indicated. Complete list of 230 targets can be found in Tables S1 and S2.

with UDP-GlcNAc, but without exogenous OGT, and subsequently treated it with the antibodies. The signal intensities from this control array were subtracted from the data for the test arrays to ensure that any observed signal increase in the test arrays was due to glycosylation by exogenously added OGT. The top 2% of normalized signal intensities for each OGT isoform, encompassing a total of 230 proteins, were identified as possible targets of OGT (Figure 1b). This is a stringent cutoff and there are likely many other OGT targets in the array, our first goal was not to comprehensively assess all substrates, but to establish whether a microarray approach would be useful for identifying new substrates. Among the 230 proteins were several known targets of human OGT, including IRS1 and AKT1, which are involved in insulin signaling, CRTC2, a central regulator of gluconeogenic gene expression, and CKII, a well-studied kinase involved in numerous cellular processes.[13] However, the majority of the top hits were not previously identified human OGT targets. Ingenuity Pathway Analysis showed that the hits fall into several functional classes, with the major categories including kinases, transcription factors and apoptosis-related proteins (Table S3).

Validation is a major bottleneck for all OGT substrate discovery efforts. Putative OGT substrates identified by proteomic methods are typically validated by transfecting mammalian cell lines with an overexpression vector, immunoprecipitating the expressed protein with a specific antibody, and immunoblotting for the *O*-GlcNAc modification using an anti-*O*-GlcNAc antibody.[6b] Alternatively, chemoenzymatic methods can be used to install a biotin handle on all *O*-GlcNAcylated proteins in a cell lysate, and specific proteins of interest can be identified after streptavidin immunoprecipitation by Western blotting using protein-specific antibodies.[14] These approaches work, but are time-intensive, and screening multiple targets requires a collection of protein-specific antibodies or mammalian overexpression vectors. To accelerate substrate validation, we developed an approach based on a method used to confirm ubiquitylation substrates.[15] In this approach, a radiolabeled protein is expressed in vitro from a commercially available cDNA clone and the *O*-GlcNAc post-translational modification is then detected following gel electrophoresis (Figure 2a). Unlike ubiquitylation, *O*-GlcNAc modifications are

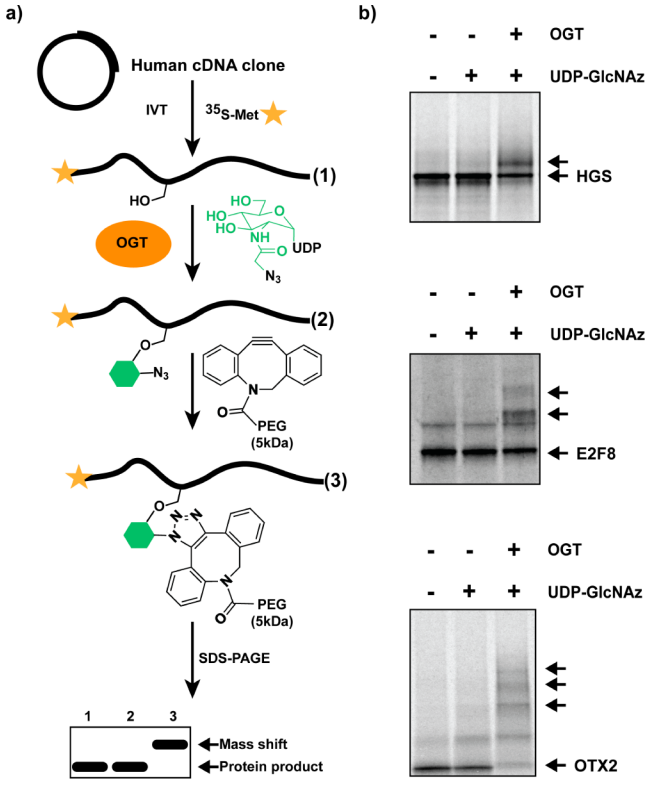

Figure 2. A rapid, orthogonal validation method confirms new OGT substrates. (a) Schematic of the validation method. (b) Radiolabeled HGS, E2F8, and OTX2 bands shift to higher molecular weight bands. Nine other OGT substrates were validated (Figure S4 and Table S4).

usually not detectable by mass shifts unless many sites are modified. Therefore, we employed a mass-tagging strategy to overcome this limitation.[16] Candidate proteins were expressed from their respective cDNA clones in an in vitro transcription/translation system (IVT, TnT SP6, Promega) supplemented with ^{35}S-methionine (^{35}S-Met), which radiolabels the expressed protein. Reactions were then incubated with OGT and a UDP-GlcNAc analogue containing an N-acyl azide (UDP-GlcNAz; Figure 2),[17] followed by azadibenzocyclooctyne-PEG, which adds a 5 kDa tag to each GlcNAz.[18] Prior to using the method to validate new substrates, we confirmed that OGT efficiently transfers GlcNAz to IVT-expressed nucleoporin62 (Nup62), a well-characterized OGT target that has ten glycosylation sites and thus undergoes a significant mass shift even without a PEG tag (Figure S2).[19] We then verified the mass-tagging method using IVT-expressed Nup62 and CRTC2, which also has multiple *O*-GlcNAcylation sites (Figure S3).[13c] For both proteins, gel analysis of IVT reactions showed a single band that shifted to a set of higher molecular weight bands only in reactions containing OGT as well as UDP-GlcNAz and azadibenzocyclooctyne-PEG. Twenty IVT-expressed proteins were then analyzed using the method outlined in Figure 2a, and 12 of these were mass shifted, confirming the microarray identification of them as substrates (Figure 2, Table S4, Figure

dx.doi.org/10.1021/ja500451w | J. Am. Chem. Soc. 2014, 136, 4845−4848

S4). The remainder either are not OGT substrates or failed to confirm for technical reasons; i.e., the IVT-expressed protein is not correctly folded or is not identical to the form on the microarray, which was expressed in cells and may contain other PTMs. To distinguish between these possibilities, other validation methods can be used.[6b] Nevertheless, the confirmed substrates highlight the utility of the approach. The IVT, OGT, and mass-tagging reactions can be accomplished in under four hours, making it possible to validate a large number of possible substrates in a day once in vitro expression from a commercially available cDNA clone has been verified. This validation procedure should be generally useful for confirming new OGT substrates. Several of the twelve validated OGT substrates are involved in gene transcription,[20] including SSBP2, SSBP3, the glucocorticoid receptor, MEF2A, E2F8, and OTX2. Some of these proteins showed a single 5 kDa shift, consistent with one glycosylation event per protein (HGS, Figure 2b), while others showed multiple bands, indicating several *O*-GlcNAc modifications per protein (OTX2 and E2F8, Figure 2b).[21]

OTX2, a previously unknown OGT substrate, is a transcription factor that acts as a master regulator of brain development.[7] Although *otx2* expression is switched off in most tissues after early embryogenesis, it is highly expressed in ∼75% of medulloblastomas, responsible for the majority of fatal childhood brain cancers.[22] Along with c-Myc, OTX2 has been identified as an important oncogene in these tumors.[23] After verifying that OTX2 is expressed in the medulloblastoma cell line D283 Med (Figure S5), we adapted a previously developed mass-tagging approach to assess its endogenous *O*-GlcNAcylation status (Figure 3a).[21] In the adaptation of this approach, a

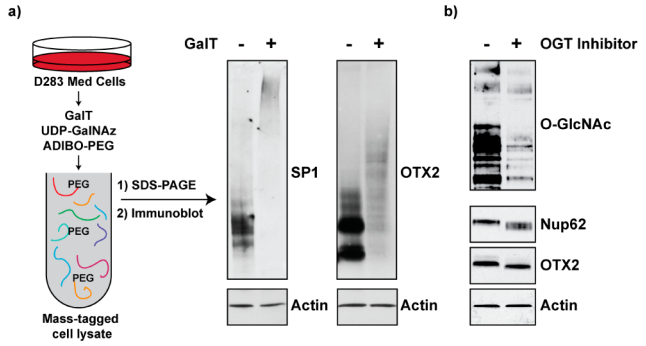

Figure 3. Endogenous OTX2 is *O*-GlcNAcylated in a medulloblastoma cell line. (a) Schematic of a validation method adapted from ref 21 with data shown for SP1 (positive control) and OTX2. (b) Immunoblot analysis of inhibitor-treated D283 Med cells. *O*-GlcNAc antibody shows decreased global *O*-GlcNAc, Nup62 and OTX2 show mass shifts when OGT is inhibited.

β-1,4-galactosyltransferase variant, $GalT_{Y289}$, is used to transfer azido-GalNAc (GalNAz) to existing *O*-GlcNAcs on proteins obtained from a cell lysate. Lysates are then treated with ADIBO-PEG to install a mass tag on each GalNAz. Following SDS-PAGE of total cell lysate, proteins of interest are detected by immunoblotting. SP1, a ubiquitous human transcription factor known to be highly *O*-GlcNAcylated,[24] shifted to higher molecular weight upon mass-tagging while actin, which is not *O*-GlcNAcylated, did not. As with SP1, the OTX2 bands shifted to higher molecular weight (Figure 3a), confirming that this protein is *O*-GlcNAcylated. Several shifted bands were observed, implying that OTX2 from medulloblastoma cells is *O*-GlcNAcylated at multiple sites, as it is when expressed via IVT (Figure 2b). We confirmed this finding by treating D283 Med cells with 50 μM Ac-5S-GlcNAc, which is metabolized to the validated OGT inhibitor UDP-5S-GlcNAc.[25] Using anti-Nup62 and anti-OTX2 antibodies, we found that inhibitor treatment resulted in a detectable shift to lower molecular weight for both proteins, consistent with the loss of multiple *O*-GlcNAc modifications due to OGT inhibition (Figure 3b). The roles of *O*-GlcNAc modification in the oncogenic functions of OTX2 are under investigation.

We have shown that human protein microarrays can detect novel OGT substrates. We note that antibody bias and subtraction of the control array mean that we are not detecting all of the *O*-GlcNAcylated substrates on the array. The development of alternative detection methods should overcome this limitation. While these microarrays do not yet contain all human proteins, they contain many thousands, including substrates of low abundance, temporally limited expression or glycosylation patterns, or highly restricted cell type specificity. Hence, microarray substrate profiling, combined with the IVT validation method described, can powerfully complement existing proteomic approaches. Moreover, microarray methods are uniquely well suited to exploring the effects of changes in OGT structure on substrate selection.

■ ASSOCIATED CONTENT

Supporting Information

Experimental procedures, microarray protocol and data analysis, IVT validation data, and list of OGT targets. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

suzanne_walker@hms.harvard.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The OGT inhibitor was a gift from David Vocadlo (Simon Fraser University). This research was supported by the National Institutes of Health (GM094263, GM26875, and GM100539).

■ REFERENCES

(1) (a) Hart, G. W.; Slawson, C.; Ramirez-Correa, G.; Lagerlof, O. *Annu. Rev. Biochem.* **2011**, *80*, 825. (b) Hanover, J. A.; Krause, M. W.; Love, D. C. *Nat. Rev. Mol. Cell. Biol.* **2012**, *13* (5), 312.

(2) (a) Wells, L.; Kreppel, L. K.; Comer, F. I.; Wadzinski, B. E.; Hart, G. W. *J. Biol. Chem.* **2004**, *279* (37), 38466. (b) Yang, X. Y.; Ongusaha, P. P.; Miles, P. D.; Havstad, J. C.; Zhang, F. X.; So, W. V.; Kudlow, J. E.; Michell, R. H.; Olefsky, J. M.; Field, S. J.; Evans, R. M. *Nature* **2008**, *451* (7181), 964. (c) Lehman, D. M.; Fu, D. J.; Freeman, A. B.; Hunt, K. J.; Leach, R. J.; Johnson-Pais, T.; Hamlington, J.; Dyer, T. D.; Arya, R.; Abboud, H.; Goring, H. H. H.; Duggirala, R.; Blangero, J.; Konrad, R. J.; Stern, M. P. *Diabetes* **2005**, *54* (4), 1214. (d) Zachara, N. E.; O'Donnell, N.; Cheung, W. D.; Mercer, J. J.; Marth, J. D.; Hart, G. W. *J. Biol. Chem.* **2004**, *279* (29), 30133. (e) Slawson, C.; Zachara, N. E.; Vosseller, K.; Cheung, W. D.; Lane, M. D.; Hart, G. W. *J. Biol. Chem.* **2005**, *280* (38), 32944. (f) Lazarus, B. D.; Love, D. C.; Hanover, J. A. *Int. J. Biochem. Cell Biol.* **2009**, *41* (11), 2134. (g) Caldwell, S. A.; Jackson, S. R.; Shahriari, K. S.; Lynch, T. P.; Sethi, G.; Walker, S.; Vosseller, K.; Reginato, M. J. *Oncogene* **2010**, *29* (19), 2831. (h) Ma, Z.; Vocadlo, D. J.; Vosseller, K. *J. Biol. Chem.* **2013**, *288*

(21), 15121. (i) Lynch, T. P.; Ferrer, C. M.; Jackson, S. R.; Shahriari, K. S.; Vosseller, K.; Reginato, M. J. *J. Biol. Chem.* **2012**, *287* (14), 11070.
(3) Hahne, H.; Sobotzki, N.; Nyberg, T.; Helm, D.; Borodkin, V. S.; van Aalten, D. M.; Agnew, B.; Kuster, B. *J. Proteome. Res.* **2013**, *12* (2), 927.
(4) (a) Sprung, R.; Nandi, A.; Chen, Y.; Kim, S. C.; Barma, D.; Falck, J. R.; Zhao, Y. *J. Proteome. Res.* **2005**, *4* (3), 950. (b) Nandi, A.; Sprung, R.; Barma, D. K.; Zhao, Y.; Kim, S. C.; Falck, J. R.; Zhao, Y. *Anal. Chem.* **2006**, *78* (2), 452. (c) Gurcel, C.; Vercoutter-Edouart, A. S.; Fonbonne, C.; Mortuaire, M.; Salvador, A.; Michalski, J. C.; Lemoine, J. *Anal. Bioanal. Chem.* **2008**, *390* (8), 2089. (d) Zaro, B. W.; Yang, Y. Y.; Hang, H. C.; Pratt, M. R. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (20), 8146. (e) Alfaro, J. F.; Gong, C. X.; Monroe, M. E.; Aldrich, J. T.; Clauss, T. R.; Purvine, S. O.; Wang, Z.; Camp, D. G., II; Shabanowitz, J.; Stanley, P.; Hart, G. W.; Hunt, D. F.; Yang, F.; Smith, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (19), 7280. (f) Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (36), 13132.
(5) MacBeath, G.; Schreiber, S. L. *Science* **2000**, *289* (5485), 1760.
(6) (a) A small kinase microarray in which 3H-GlcNAc was detected was used to identify kinase substrates of OGT. See: (b) Dias, W. B.; Cheung, W. D.; Hart, G. W. *Biochem. Biophys. Res. Commun.* **2012**, *422* (2), 224.
(7) Beby, F.; Lamonerie, T. *Exp. Eye. Res.* **2013**, *111*, 9.
(8) Jinek, M.; Rehwinkel, J.; Lazarus, B. D.; Izaurralde, E.; Hanover, J. A.; Conti, E. *Nat. Struct. Mol. Biol.* **2004**, *11* (10), 1001.
(9) Lazarus, B. D.; Love, D. C.; Hanover, J. A. *Glycobiology* **2006**, *16* (5), 415.
(10) (a) Merbl, Y.; Kirschner, M. W. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (8), 2543. (b) Gross, B. J.; Kraybill, B. C.; Walker, S. *J. Am. Chem. Soc.* **2005**, *127* (42), 14588.
(11) Comer, F. I.; Vosseller, K.; Wells, L.; Accavitti, M. A.; Hart, G. W. *Anal. Biochem.* **2001**, *293* (2), 169.
(12) Isono, T. *PLoS One* **2011**, *6* (4), e18959.
(13) (a) Whelan, S. A.; Dias, W. B.; Thiruneelakantapillai, L.; Lane, M. D.; Hart, G. W. *J. Biol. Chem.* **2010**, *285* (8), 5204. (b) Wang, S.; Huang, X.; Sun, D.; Xin, X.; Pan, Q.; Peng, S.; Liang, Z.; Luo, C.; Yang, Y.; Jiang, H.; Huang, M.; Chai, W.; Ding, J.; Geng, M. *PLoS One* **2012**, *7* (5), e37427. (c) Dentin, R.; Hedrick, S.; Xie, J.; Yates, J., III; Montminy, M. *Science* **2008**, *319* (5868), 1402. (d) Tarrant, M. K.; Rho, H. S.; Xie, Z.; Jiang, Y. L.; Gross, C.; Culhane, J. C.; Yan, G.; Qian, J.; Ichikawa, Y.; Matsuoka, T.; Zachara, N.; Etzkorn, F. A.; Hart, G. W.; Jeong, J. S.; Blackshaw, S.; Zhu, H.; Cole, P. A. *Nat. Chem. Biol.* **2012**, *8* (3), 262. (e) 11 out of 80 validated *O*-GlcNAc proteins discovered in ref 3 were also present on the protein microarray; 5 were in our top 2% of hits, and 7 were in the top 6%.
(14) Tai, H. C.; Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2004**, *126* (34), 10500.
(15) Merbl, Y.; Refour, P.; Patel, H.; Springer, M.; Kirschner, M. W. *Cell* **2013**, *152* (5), 1160.
(16) Boyce, M.; Bertozzi, C. R. *Nat. Methods* **2011**, *8* (8), 638.
(17) Vocadlo, D. J.; Hang, H. C.; Kim, E. J.; Hanover, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (16), 9116.
(18) Yao, J. Z.; Uttamapinant, C.; Poloukhtine, A.; Baskin, J. M.; Codelli, J. A.; Sletten, E. M.; Bertozzi, C. R.; Popik, V. V.; Ting, A. Y. *J. Am. Chem. Soc.* **2012**, *134* (8), 3720.
(19) Holt, G. D.; Snow, C. M.; Senior, A.; Haltiwanger, R. S.; Gerace, L.; Hart, G. W. *J. Cell. Biol.* **1987**, *104* (5), 1157.
(20) Slawson, C.; Hart, G. W. *Nat. Rev. Cancer* **2011**, *11* (9), 678.
(21) Rexach, J. E.; Rogers, C. J.; Yu, S. H.; Tao, J.; Sun, Y. E.; Hsieh-Wilson, L. C. *Nat. Chem. Biol.* **2010**, *6* (9), 645.
(22) Adamson, D. C.; Shi, Q.; Wortham, M.; Northcott, P. A.; Di, C.; Duncan, C. G.; Li, J.; McLendon, R. E.; Bigner, D. D.; Taylor, M. D.; Yan, H. *Cancer Res.* **2010**, *70* (1), 181.
(23) Bai, R. Y.; Staedtke, V.; Lidov, H. G.; Eberhart, C. G.; Riggins, G. J. *Cancer Res.* **2012**, *72* (22), 5988.
(24) Jackson, S. P.; Tjian, R. *Cell* **1988**, *55* (1), 125.
(25) Gloster, T. M.; Zandberg, W. F.; Heinonen, J. E.; Shen, D. L.; Deng, L.; Vocadlo, D. J. *Nat. Chem. Biol.* **2011**, *7* (3), 174.