9-1-2013

Targeted identification of glycosylated proteins in the gastric pathogen helicobacter pylori (Hp)

Kanokwan Champasa
Bowdoin College

Scott A. Longwell
Bowdoin College

Aimee M. Eldridge
Bowdoin College

Elizabeth A. Stemmler
Bowdoin College

Danielle H. Dube
Bowdoin College

⌘ Author's Choice

Targeted Identification of Glycosylated Proteins in the Gastric Pathogen *Helicobacter pylori* (*Hp*)*

Kanokwan Champasa‡¶, Scott A. Longwell‡¶, Aimee M. Eldridge‡, Elizabeth A. Stemmler‡, and Danielle H. Dube‡§

Virulence of the gastric pathogen *Helicobacter pylori (Hp)* is directly linked to the pathogen's ability to glycosylate proteins; for example, *Hp* flagellin proteins are heavily glycosylated with the unusual nine-carbon sugar pseudaminic acid, and this modification is absolutely essential for *Hp* to synthesize functional flagella and colonize the host's stomach. Although *Hp*'s glycans are linked to pathogenesis, *Hp*'s glycome remains poorly understood; only the two flagellin glycoproteins have been firmly characterized in *Hp*. Evidence from our laboratory suggests that *Hp* synthesizes a large number of as-yet unidentified glycoproteins. Here we set out to discover *Hp*'s glycoproteins by coupling glycan metabolic labeling with mass spectrometry analysis. An assessment of the subcellular distribution of azide-labeled proteins by Western blot analysis indicated that glycoproteins are present throughout *Hp* and may therefore serve diverse functions. To identify these species, *Hp's* azide-labeled glycoproteins were tagged via Staudinger ligation, enriched by tandem affinity chromatography, and analyzed by multidimensional protein identification technology. Direct comparison of enriched azide-labeled glycoproteins with a mock-enriched control by both SDS-PAGE and mass spectrometry-based analyses confirmed the selective enrichment of azide-labeled glycoproteins. We identified 125 candidate glycoproteins with diverse biological functions, including those linked with pathogenesis. Mass spectrometry analyses of enriched azide-labeled glycoproteins before and after cleavage of O-linked glycans revealed the presence of Staudinger ligation-glycan adducts in samples only after beta-elimination, confirming the synthesis of O-linked glycoproteins in *Hp*. Finally, the secreted colonization factors urease alpha and urease beta were biochemically validated as glycosylated proteins via Western blot analysis as well as by mass spectrometry analysis of cleaved glycan products. These data set the stage for the development of glycosylation-based therapeutic strategies, such as new vaccines based on natively glycosylated *Hp* proteins, to eradicate *Hp* infection. Broadly, this report validates metabolic labeling as an effective and efficient approach for the identification of bacterial glycoproteins. ***Molecular & Cellular Proteomics 12: 10.1074/mcp.M113.029561, 2568–2586, 2013.***

Helicobacter pylori (*Hp*)[1] infection poses a significant health risk to humans worldwide. The Gram-negative, pathogenic bacterium *Hp* colonizes the gastric tract of more than 50% of humans (1). Approximately 15% of infected individuals develop duodenal ulcers and 1% of infected individuals develop gastric cancer (2). Current treatment to clear infection requires "triple therapy" (3), a combination of multiple antibiotics that is often associated with negative side effects (4). Because of poor patient compliance and the evolution of antibiotic resistance, existing antibiotics are no longer effective at eradicating *Hp* infection (4). New treatment methods are needed to eliminate *Hp* from the human gastric tract.

Recent work has focused on gaining insights into the pathogenesis of *Hp* to aid the development of new treatments. The most recent findings in this area have conclusively revealed that glycosylation of proteins in *Hp* is required for pathogenesis. *Hp* use complex flagella, comprised of flagellin proteins, to navigate the host's gastric mucosa (5, 6). The flagellin proteins are heavily glycosylated with the unusual nine-carbon sugar pseudaminic acid, found exclusively in mucosal-associated pathogens (*Hp* (7), *Campylobacter jejuni*

From the ‡Department of Chemistry and Biochemistry, Bowdoin College, 6600 College Station, Brunswick, Maine 04011

⌘ *Author's Choice*—Final version full access.

Received March 27, 2013, and in revised form, May 24, 2013

Published, MCP Papers in Press, June 13, 2013, DOI 10.1074/mcp.M113.029561

[1] The abbreviations used are: *Hp*, *Helicobacter pylori*; MOE, metabolic oligosaccharide engineering; Ac_4GlcNAz, peracetylated *N*-azidoacetylglucosamine; Ac_4GlcNAc, peracetylated *N*-acetylglucosamine; GlcNAc, *N*-acetylglucosamine; mudPIT, multidimensional protein identification technology; PBS, phosphate buffered saline; ureA, urease alpha subunit; ureB, urease beta subunit; CAZy, carbohydrate active enzymes; Phos-FLAG, phosphine conjugated to a FLAG peptide (DYKDDDDK); Phos-FLAG-His_6, phosphine conjugated to a FLAG-His_6 peptide (DYKDDDDKHHHHHH); TCEP, tris(2-carboxyethyl)phosphine; HBFA, heptafluorobutyric acid; FBS, fetal bovine serum; HPLC-Chip/Q-TOFMS, High Performance Liquid Chromatographic-Chip Quadrupole Time-of-Flight Mass Spectrometer; nanoESI, nano-electrospray ionization; HexNAz, *N*-azidoacetylhexosamine; HexNAc, *N*-acetylhexosamine; CID, collision-induced dissociation.

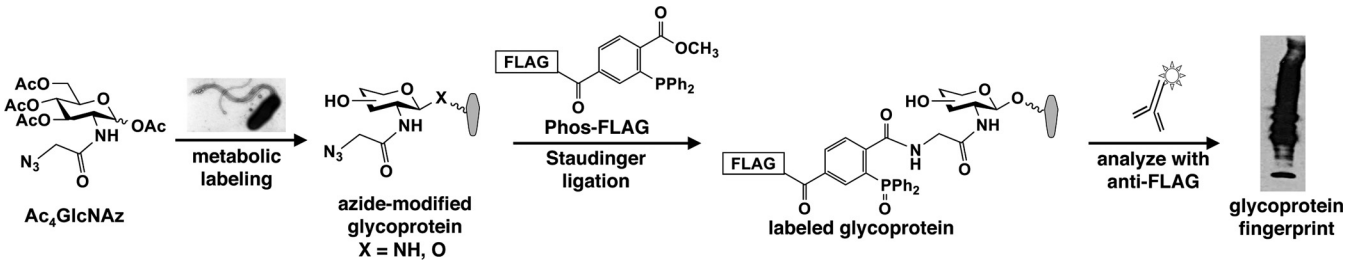

FIG. 1. **Metabolic oligosaccharide engineering facilitates labeling and detection of *Hp*'s glycoproteins.** Supplementation of *Hp* with Ac_4GlcNAz leads to metabolic labeling of *Hp*'s N-linked and O-linked glycoproteins with azides. Azide-modified glycoproteins covalently labeled with Phos-FLAG can be detected via Western blot analysis with anti-FLAG antibody to yield *Hp*'s glycoprotein fingerprint, which contains a large number of as-yet unidentified glycoproteins.

(8) and *Pseudomonas aeruginosa* (9)). This modification is absolutely essential for the formation of functional flagella on *Hp* (7, 10). Deletion of any one of the enzymes in the pseudaminic acid biosynthetic pathway results in *Hp* that lack flagella, are nonmotile, and are unable to colonize the host's stomach (7). Although pseudaminic acid is critical for *Hp* virulence, it is absent from humans (11, 12). Therefore, insights into *Hp*'s pathogenesis have revealed that *Hp's* glycan pseudaminic acid is a bona fide target of therapeutic intervention. This is one of a number of examples linking protein glycosylation to virulence in medically significant bacterial pathogens (13, 14).

Despite these findings, *Hp*'s glycome remains poorly understood overall. Only the two flagellin glycoproteins have been firmly characterized in *Hp* (7) to date. Nine other candidate glycoproteins have been identified in *Hp*, but their glycosylation status has not been biochemically confirmed (15). The relative paucity of information regarding *Hp's* glycoproteins is due in part to the previously held belief that protein glycosylation could not occur in bacteria (13, 16, 17). However, even after Szymanski (18, 19), Koomey (20), Guerry (21), Logan (7), Comstock and others (13, 16, 17) disproved this belief by firmly establishing the synthesis of glycoproteins in bacteria, the study of bacterial glycoproteins has presented unique challenges for analytical study (14, 22). For example, the unusual structures of bacterial glycans, which often contain amino- and deoxy-carbohydrates exclusively found in bacteria (12, 23–25), hampers their identification using existing tools. Though methods such as the use of glycan-binding reagents (20, 24, 26, 27) and periodic acid/hydrazide glycan labeling (15) have successfully detected glycoproteins in a range of bacteria, they present limitations. Glycan binding-based methods are often limited because of the unavailability of lectins or antibodies with binding specificity for glycosylated proteins in the bacteria of interest (14, 22). Periodic acid/hydrazide-based labeling is plagued by a lack of specificity for glycosylated proteins (15). Thus, an efficient and robust approach to discover *Hp's* glycoproteins is needed.

In previous work, we established that the chemical technique known as metabolic oligosaccharide engineering (MOE), which was developed by Bertozzi (28, 29), Reutter (30), and others for the study of mammalian glycoproteins, is a powerful approach to label and detect *Hp*'s glycoproteins (31). Briefly, *Hp* metabolically processes the unnatural, azide-containing sugar peracetylated *N*-azidoacetylglucosamine (Ac_4GlcNAz) (32), an analog of the common metabolic precursor *N*-acetylglucosamine (GlcNAc), into cellular glycoproteins (Fig. 1). Elaboration of azide-labeled glycoproteins via Staudinger ligation (33) with a phosphine probe conjugated to a FLAG peptide (Phos-FLAG) (34) followed by visualization with an anti-FLAG antibody (Fig. 1) revealed a glycoprotein fingerprint containing a large number of as-yet unidentified *Hp* glycoproteins that merit further investigation (31).

Here we describe a glycoproteomic identification strategy for the selective detection, isolation, and discovery of *Hp*'s glycoproteins. In particular, we demonstrate that glycan metabolic labeling coupled with mass spectrometry analysis is an efficient and robust chemical approach to identify novel glycoproteins in *Hp*. This work characterizes glycosylated virulence factors in *Hp*, thus opening the door to new vaccination and antibiotic therapies to eradicate *Hp* infection. Broadly, this work validates metabolic oligosaccharide engineering as a complementary method to discover bacterial glycoproteins.

EXPERIMENTAL PROCEDURES

Materials—Protease inhibitor mixture, antibiotics, anti-FLAG antibodies and anti-FLAG agarose were purchased from Sigma (St. Louis, MO). Sequencing grade trypsin was from Promega (Madison, WI), and fetal bovine serum (FBS) was from Invitrogen (Carlsbad, CA). Nitrocellulose paper, Ni-NTA resin, zinc stain, and 4–16% Ready Gel Tris-HCl gels were from BioRad (Hercules, CA). GlcNAz, Ac_4GlcNAc, Ac_4GlcNAz, Phos-FLAG (Phos-DYKDDDDK) and Phos-FLAG-His_6 (Phos-DYKDDDDKHHHHHH) were prepared as previously described (29). *Hp* strain 26695 was provided by Dr. Manuel Amieva (Stanford, CA). Azidoacetyl ethyl ester was a gift from Dr. Jennifer Prescher (Irvine, CA).

Metabolic Labeling of Hp—*Hp* strain 26695 was grown on horse blood agar plates for 3–4 days in a microaerophilic environment (14% CO_2) at 37 °C. The bacteria were then transferred to Brucella broth containing 10% FBS, 6 μg/ml vancomycin, and 1 mM peracetylated *N*-acetylglucosamine (Ac_4GlcNAc) or Ac_4GlcNAz, as previously described (31). Alternatively, the bacteria were transferred to Brucella broth containing 10% FBS and 6 μg/ml vancomycin supplemented with 5 mM sodium azidoacetate or 5 mM azidoacetyl ethyl ester. Liquid

cultures were grown for 3–5 days in a microaerophilic environment with gentle rocking at 37 °C, then harvested as described below.

Harvesting Hp to Probe for Total Cellular Azide-Labeled Glycoproteins—*Hp* grown in the presence of Ac_4GlcNAc, Ac_4GlcNAz, sodium azidoacetate, or azidoacetyl ethyl ester were centrifuged at 3500 rpm using a Sorvall Legend RT$^+$ centrifuge (Thermo Scientific) and washed two times with PBS. The cells were resuspended in ice cold lysis buffer (20 mM Tris-HCl, pH 7.4; 1% Igepal; 150 mM NaCl; 1 mM EDTA) containing protease inhibitors (protease inhibitor mixture, Sigma) for 5 mins at room temperature. Protein concentration of the samples was measured using BioRad's D_C protein concentration assay (BioRad) per manufacturer's instructions. All samples were standardized to 3.0 mg/ml. To probe for azide incorporation, the lysates were diluted 1:1 with 500 μM Phos-FLAG (34), reacted overnight at room temperature, and then analyzed via Western blot with anti-FLAG antibody (Sigma) or via SDS-PAGE with zinc stain (BioRad).

Subcellular Fractionation of Glycoproteins—*Hp* grown in the presence of Ac_4GlcNAc or Ac_4GlcNAz were harvested from Brucella broth by centrifugation at 6000 $\times$ *g*, and the conditioned medium was set aside to access secreted proteins. The harvested cells were fractionated as described by Hoshino *et al.* to obtain periplasmic, inner membrane-associated, and cytoplasmic protein fractions (35). Secreted proteins were isolated from the conditioned Brucella broth via TCA-precipitation as described by Bumann *et al.* (36). An acid-glycine extraction was performed as described previously to isolate *Hp's* surface-associated proteins (37). An *N*-lauroylsarcosine extraction was performed as reported by Hopf *et al.* to access outer-membrane *Hp* proteins (15, 38). The protein concentration of subcellular fractions was standardized to 3.0 mg/ml using a BioRad D_C Protein Assay (BioRad). Standardized samples were incubated 1:1 with 500 μM Phos-FLAG (34) overnight at room temperature to label azide-containing glycoproteins, and then analyzed by Western blot analysis with anti-FLAG antibody (Sigma) or via SDS-PAGE with zinc stain (BioRad). In addition, standardized periplasmic and cytoplasmic samples were assayed for malate dehydrogenase (35) activity alongside a positive control (malate dehydrogenase, Sigma) according to the manufacturer's (Sigma's) instructions to confirm the efficacy of subcellular fractionation.

Preparation and Analysis of Enriched Azide-Labeled Glycoproteins—*Hp* metabolically labeled with 1 mM Ac_4GlcNAc or Ac_4GlcNAz were centrifuged at 3000 $\times$ *g* and washed two times with phosphate buffered saline (PBS). The cells were lysed according to Koenigs *et al.* and centrifuged at 3000 $\times$ *g* for 10 min to remove insoluble debris (31). Following the protocol by Laughlin *et al.*, Phos-FLAG-His_6 was added as a solid directly to the lysate at a final concentration of 500 μM, and Staudinger ligation was run at room temperature for 24 h under argon (29, 39). The unreacted Phos-FLAG-His_6 was removed by Bio-Rad P-10 size-exclusion column according to the manufacturer's protocol. Flow-through fractions with $A_{280} > 0.050$ were combined and concentrated using a centrifugal filter device with 30 kDa molecular weight cutoff (Millipore). FLAG-His_6-tagged glycoproteins were enriched using 2 ml of α-FLAG agarose beads (Sigma) followed by 5 ml of nickel-nitrilotriacetic acid (Ni-NTA)-agarose resin (Bio-Rad), as previously described (29, 39). After rinsing away nonbinders, the bound glycoproteins were eluted from the nickel column with one column volume of a solution of 8 M urea, 0.1 M Na_2HPO_4, 10 mM Tris (pH 8.0), and 250 mM imidazole. Samples containing 15 μg of material from Ac_4GlcNAc and Ac_4GlcNAz-treated *Hp* both before purification (input) and after purification (eluent) were analyzed by SDS-PAGE to confirm successful enrichment. In addition, 150 μg of *Hp*'s enriched azide-labeled glycoproteins were suspended in 10 mM Tris (pH 8) and focused on 11 cm pH 3–10 nonlinear IPG strips (GE Healthcare) using GE Healthcare's Multiphor™ II system. These strips were reduced, alkylated and loaded onto 8–18% polyacrylamide gels (GE Healthcare), which were run at 10 °C and stained with zinc stain (BioRad).

Protein Identification by MudPIT—An azide-labeled enriched glycoprotein sample containing 150 μg of protein and a mock enriched control (which was labeled with Ac_4GlcNAc, reacted with Phos-FLAG-His_6, and enriched via anti-FLAG and Ni-NTA affinity chromatography, as described above) were concentrated using a Microcon centrifugal filter device (Millipore) to ~10 μl, suspended in 8 M urea and 100 mM Tris-HCl (pH 8.5), reduced with 5 mM *tris*(2-carboxyethyl)phosphine (TCEP), and reacted with 10 mM iodoacetamide. The samples were diluted by a factor of four with 100 mM Tris-HCl (pH 8.5) and 1 mM $CaCl_2$, and then digested with 0.25 μg of trypsin (Promega) overnight at 37 °C. The digests were mixed with formic acid to a final concentration of 5%, desalted using C_{18} Spec tips (Varian), fully dried by speed vacuum at room temperature, and analyzed by multidimensional protein identification technology (mudPIT) (40). The Vincent J. Coates Proteomics/Mass Spectrometry Laboratory at University of California Berkeley performed mass spectrometry analyses of the enriched, azide-labeled sample and mock-enriched control sample. A nano LC column was packed in a 100 μm inner diameter glass capillary with an emitter tip. The column consisted of 10 cm of Polaris C18 5 μm packing material (Varian), followed by 4 cm of Partisphere 5 SCX (Whatman). The column was loaded by use of a pressure bomb and washed extensively with buffer A (see below). The column was then directly coupled to an electrospray ionization source mounted on a Thermo LTQ XL linear ion trap mass spectrometer. An Agilent 1200 HPLC equipped with a split line so as to deliver a flow rate of 30 nl/min was used for chromatography. Peptides were eluted using an 8-step MudPIT procedure (40). Buffer A was 5% acetonitrile/0.02% heptaflurobutyric acid (HBFA); buffer B was 80% acetonitrile/0.02% HBFA. Buffer C was 250 mM ammonium acetate/5% acetonitrile/0.02% HBFA; buffer D was same as buffer C, but with 500 mM ammonium acetate.

The programs SEQUEST, as embodied in BIOWORKS BROWSER (version 3.3.1 SP1), (41) and DTASELECT (version 1.9) (42) were used to identify peptides and proteins from a database consisting of all proteins (1560 total) from *Hp* 26695 plus a collection of 160 common contaminants (1720 protein entries in the database actually searched). The *Hp* database was the newest available from NCBI RefSeq on October 7, 2010. Search parameters were altered to DTA generation 600.00–4000.00 with an absolute threshold of 500. The search was restricted to tryptic peptides and allowed three missed cleavages. Carboxyamidomethylation of cysteines as a fixed modification and oxidation of methionines as a variable modification were specified. The mass tolerance for precursor ions was 1.4 and for fragment ions was 1.0. Individual peptides were accepted with Xcorr of 1.8 for charge state +1, 2.2 for charge state +2 and 3.5 for charge state +3. These values have given < 1% false positives over many data sets. Proteins were further required to have at least two peptides meeting acceptance criteria. Data was considered high confidence if assigned spectra were present in the Ac_4GlcNAz samples but absent from the mock-enriched Ac_4GlcNAc control samples.

Prediction of Subcellular Localization—Initial predictions of the subcellular localization of *Hp* proteins identified by mudPIT were based on literature reports, where available. Because of the shortage of subcellular location information for hypothetical proteins and functionally uncharacterized proteins, each identified protein was also subjected to PSORTb v.3.0 to predict subcellular location (43).

Accessing Synthetic Staudinger Ligation-Glycan Standards—Staudinger ligation was performed overnight at 37 °C by mixing 1:1 solutions of GlcNAz (100 mM in H_2O) and Phos-FLAG-His_6 (100 mM in PBS) to produce GlcNAz-Phos-FLAG-His_6. Similarly, overnight incubation at 37 °C of solutions of GlcNAz (100 mM in H_2O) and Phos-FLAG-His_6 (500 μM in PBS) mixed 1:1 produced GlcNAz-Phos-FLAG-

His_6. Successful synthesis was confirmed by high-resolution mass spectrometry analysis as described below (see *HPLC-Chip/Q-TOF mass spectrometric analysis*).

Phos-FLAG Labeling to Produce "Tagged Lysates"—Hp metabolically labeled with 1 mM Ac_4GlcNAc or Ac_4GlcNAz were centrifuged at 3000 × *g* and washed two times with phosphate buffered saline (PBS). The cells were lysed according to Koenigs *et al.* (31). Lysates were rinsed extensively with PBS using a Microcon centrifugal filter device (Millipore, 10000 MWCO) to remove residual Ac_4GlcNAc or Ac_4GlcNAz. Lysates were reacted with 250 μM Phos-FLAG at room temperature overnight. The unreacted Phos-FLAG was removed by extensive rinsing with ddH_2O and passage of reacted lysates through a Microcon centrifugal filter device (Millipore, 10000 MWCO) to yield samples of "tagged lysates." These lysates were then subjected to beta-elimination, as described below.

Biochemical Purification of Azide-Labeled Urease—Urease was purified from *Hp* strain 26695 using ion exchange chromatography (IEC) and size exclusion chromatography (SEC) based on previously described purification schemes by Hu and Mobley (44, 45) and Rokita *et al.* (46). A 250 ml culture of *Hp* metabolically labeled with 1 mM Ac_4GlcNAz was inoculated and grown as described above. Cells were harvested by centrifugation at 10,000 × *g* and the resultant pellet was frozen at −20 °C. The cells were re-suspended in ice-cold buffer (20 mM sodium phosphate, pH 6.9) containing protease inhibitors (protease inhibitor mixture, Sigma) and 10 μg/ml DNase and then sonicated to achieve complete cell lysis. Concentrated stock solutions of EDTA and β-mercaptoethanol (βME) were added to the lysis solution to make it consistent with the initial IEC buffer (20 mM sodium phosphate, pH 6.9, 1 mM EDTA, 1 mM βME). Cellular debris was removed from the sample through centrifugation at 14,000 × *g*. The supernatant was loaded onto a Q-Sepharose Fast Flow column (25 ml, 3 cm diameter, GE Healthcare) that had been pre-equilibrated with the above IEC buffer. Urease was eluted from the column at ~150 mM NaCl using a linear gradient from 0 M to 0.5 M NaCl in IEC buffer. The presence of urease in the elution fractions was detected qualitatively using a phenol-red colorimetric assay (44). Fractions that exhibited urease activity were further analyzed by SDS-PAGE with zinc-stain (BioRad) and via Western with anti-ureA and anti-ureB antibodies (Santa Cruz Biotechnology) to verify the presence of ureA and ureB. Fractions with the most urease and fewest contaminants were pooled for subsequent purification by SEC. The pooled fractions were exchanged into the SEC buffer (20 mM sodium phosphate, pH 6.9, 150 mM NaCl, 1 mM EDTA, 1 mM βME) using a 9000 Da MWCO filter (Pierce) and concentrated to a final volume of 7 ml. This concentrated sample was loaded onto a pre-equilibrated Sepharose CL-6B (GE Healthcare) SEC column (200 ml total volume, 3 cm diameter) and eluted at 1 ml/min. Fractions containing urease were identified using the phenol-red assay and subsequently analyzed by SDS-PAGE. The purest urease-containing fractions (see supplemental Fig. S7) were pooled, concentrated to 0.5 mg/ml via ultrafiltration (Millipore, 100,000 MWCO), and mixed 1:1 with 500 μM Phos-FLAG at room temperature overnight. The unreacted Phos-FLAG was removed by extensive rinsing with 50% MeOH in ddH_2O, 15% acetonitrile in ddH_2O, and 100% ddH_2O over a Microcon centrifugal filter device (Millipore, 10000 MWCO) to yield samples of "tagged urease." This sample was then subjected to beta-elimination, as described below.

Beta-Elimination of Hp's O-Linked Glycans—Hp's enriched azide-labeled glycoproteins, Phos-FLAG-reacted Ac and Az lysates ("tagged lysates"), and purified azide-labeled urease ("tagged urease") were subjected to beta-elimination to remove glycans overnight at 4 °C using GlycoProfile Beta-elimination Kit (Sigma-Aldrich) according to manufacturer's instructions. Released glycans were then analyzed by mass spectrometry, as described below.

HPLC-Chip/Q-TOF Mass Spectrometric Analysis—Samples of *Hp's* enriched azide-labeled glycoproteins, Phos-FLAG-reacted Ac and Az lysates ("tagged lysates"), and purified azide-labeled urease ("tagged urease") were analyzed before and after beta-elimination of O-linked glycans using a 6530 High Performance Liquid Chromatographic-Chip Quadrupole Time-of-Flight Mass Spectrometer (HPLC-Chip/Q-TOFMS; Agilent Technologies, Santa Clara, CA). Chromatographic separation and nano-electrospray ionization (nanoESI) was performed with a 1260 Chip Cube system (Agilent Technologies) using a ProtID-chip with a 40 nL enrichment column and a 150 mm x 75 μm analytical column (Agilent Technologies). The enrichment and analytical columns were packed with 300 Å, 5 μm particles with C18 stationary phase. The mobile phases were 0.1% formic acid/H_2O (A) and 0.1% formic acid/acetonitrile (B). Samples (0.2–2 μl) were loaded on the enrichment column using 98:2 (A:B) at 4 μl/min and eluted through the analytical column using a gradient from 98:2 (A:B) to 80:20 (A:B) over 2 min to 2:98 (A:B) over a period of 8 min at 0.3 μl/min. Mass spectra (MS and MS/MS) were collected in positive ion mode; the ionization voltage ranged from 1850–1950 V and the ion source temperature was held at 350 °C. The fragmentor and collision cell voltages, as well as other tuning parameters, were optimized to minimize metastable water losses associated with ion sampling and transmission to the TOF analyzer. This was of particular importance for the analysis of HexNAz-Phos-FLAG-His_6 (**1**), for which metastable water losses became increasingly prominent for higher charge state ions. A similar effect, impacting metastable losses of sulfate from sulfated oligosaccharides, has been reported by Zaia and co-workers (47). Spectra were internally calibrated using methyl stearate ($C_{17}H_{35}CO_2CH_3$) or dibutyl phthalate ($C_{16}H_{22}O_4$) and hexakis(1H, 1H, 4H-hexafluorobutyloxy)phosphazine (HP-1221; $C_{24}H_{18}O_6N_3P_3F_{36}$), continuously introduced and detected as $[M+H]^+$. Collision-induced dissociation (CID)-MS/MS experiments were executed with precursor selection determined using a targeted or data-dependent approach. Precursor ions were subjected to CID using nitrogen as the target gas with collision energies ranging from 20–35 V.

Immunopurification of Hp's ureA and ureB—Lysate from metabolically labeled *Hp* was standardized to 4.4 mg/ml and incubated 1:1 with 500 μM Phos-FLAG overnight at room temperature to label azide-containing glycoproteins. Insoluble membrane material was separated from soluble proteins by centrifugation at 12,000 × *g* for 5 min. To the soluble protein samples, 1 μl of α-*Hp* ureA or ureB (Santa Cruz Biotechnology) was added per 100 μl of protein volume. The samples were incubated at 4 °C overnight with gentle shaking. After completion of the overnight incubation, 100 μl of a 50% Protein-G agarose suspension in HNTG washing buffer (20 mM HEPES buffer pH 7.5, containing 150 mM NaCl, 0.1% (w/v) Triton X-100, and 10% (w/v) glycerol) was added to *Hp* lysates and allowed to incubate for 90 min at 4 °C with gentle shaking. The immunoprecipated complexes were collected via centrifugation at 3000 × *g* for 5 min at 4 °C and the pellet was washed three times with ice-cold HNTG washing buffer. The final pellet was re-suspended in 40 μl of ice-cold HNTG washing buffer, diluted with an equal volume of 2 × SDS-loading buffer, and heated at 95 °C for 5 min. The boiled samples were centrifuged at 12,000 × *g* for 30 s to separate the agarose beads from the immunopreciptated proteins. Immunoprecipitated proteins were then analyzed by Western blot with anti-FLAG (Sigma) or anti-urease antibody (Santa Cruz Biotechnology).

Gel and Western Blot Analyses—All lanes were loaded with 10–15 μg of protein sample unless otherwise noted. Western blots were treated with HRP-conjugated anti-FLAG antibody (1:3000, Sigma), rabbit polyclonal anti-*Hp* ureA (1:3000, Santa Cruz Biotechnology), or rabbit polyclonal anti-*Hp* ureB antibody (1:3000, Santa Cruz Biotechnology), followed by goat α-rabbit IgG-HRP (1:10000, Santa Cruz

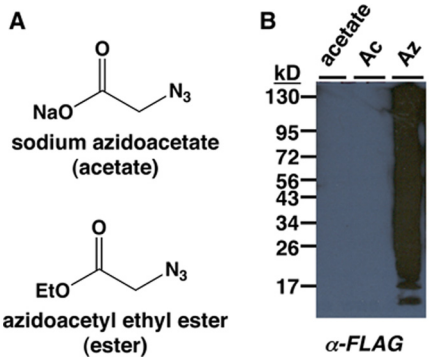

FIG. 2. **Metabolic labeling of *Hp's* proteins with azides.** *A*, Azidoacetyl compounds used in metabolic feeding experiments with *Hp*. *B*, Western blot analysis of lysate from *Hp* treated with 5 mM sodium azidoacetate (acetate), 1 mM Ac_4GlcNAz (Az), or with 1 mM of the azide-free control sugar Ac_4GlcNAc (Ac) for 4 days. After lysis, samples were reacted with Phos-FLAG and visualized by Western blot with anti-FLAG antibody. *C*, Western blot analysis of lysate from *Hp* treated with 5 mM azidoacetyl ethyl ester (ester), 1 mM Ac_4GlcNAz (Az), or with 1 mM of the azide-free control sugar Ac_4GlcNAc (Ac) for 4 days. After lysis, samples were reacted with Phos-FLAG and visualized by Western blot with anti-FLAG antibody.

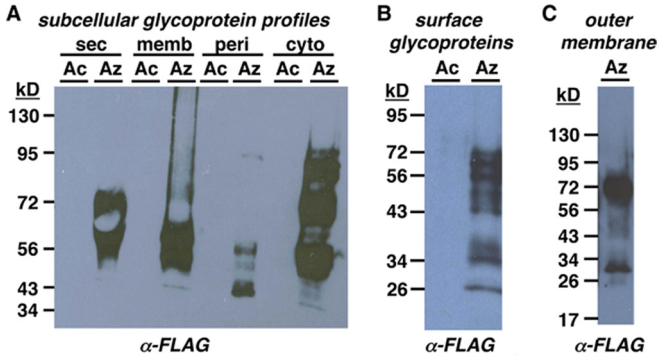

FIG. 3. **Subcellular distribution of *Hp's* azide-labeled glycoproteins.** Western blot analysis of subcellular fractions from *Hp* treated with 1 mM Ac_4GlcNAz (Az) or with the azide-free control sugar Ac_4GlcNAc (Ac) for 4 days. After subcellular fractionation, samples were reacted with Phos-FLAG and visualized by Western blot with anti-FLAG antibody. *A*, Glycoprotein profiles of four subcellular fractions are shown: secreted (sec), inner membrane (memb), periplasmic (peri), cytoplasmic (cyto). *B*, Glycoprotein profile of *Hp's* surface-associated proteins. *C*, Glycoprotein profile of *Hp's* outer membrane fraction.

Biotechnology), then developed with chemiluminescent substrate (Pierce). Polyacrylamide gels were visualized with zinc stain (BioRad).

RESULTS AND DISCUSSION

Metabolic Labeling of Hp's Proteins With Azides Does Not Appear to be Catabolic—Here we employed metabolic glycan labeling to characterize *Hp's* glycoproteins and unveil molecules that have the potential to serve as the basis of novel anti-*Hp* treatments. In previous work, our laboratory demonstrated that supplementation of *Hp's* media with 1 mM Ac_4GlcNAz for three to 5 days leads to incorporation of azide-dependent signal into a large number of glycoproteins (31). The majority, though not all, of this azide-dependent signal can be removed by glycosidases that catalyze the cleavage of certain *N*-linked and O-linked glycans (31). Thus, we sought to assess whether the glycosidase-resistant azide signal is because of an alternative metabolic fate of Ac_4GlcNAz in *Hp*; specifically, catabolism of the azidosugar to the azidoacetyl unit followed by subsequent activation and covalent addition to proteins. This type of azidosugar catabolism has been observed in some other organisms. If this process were to occur in *Hp*, then supplementation of *Hp's* media with the azidoacetyl moiety would likely lead to metabolic labeling of *Hp's* proteins with azides. To assess this possibility, *Hp* cells were grown in media supplemented with 5 mM sodium azidoacetate or with the more bioavailable azidoacetyl ethyl ester (Fig. 2*A*); Hang and coworkers demonstrated that structurally analogous compounds are cell permeable (48, 49). Ac_4GlcNAz or the azide-free sugar peracetylated GlcNAc (Ac_4GlcNAc) were employed as positive and negative controls of metabolic labeling, respectively. Ac_4GlcNAz treatment resulted in robust azide-dependent signal in lysates, whereas no azide-labeled proteins were observed in lysates from cells supplemented with Ac_4GlcNAc, with sodium azidoacetate (Fig. 2*B*), or with azidoacetyl ethyl ester (Fig. 2*C*). SDS-PAGE analysis confirmed the presence of equivalent protein levels in all lanes (supplemental Fig. S1). These data suggest that azide-dependent labeling of *Hp's* proteins is not because of catabolic incorporation of the azidoacetyl moiety into proteins. The remaining azide-dependent signal could be glycosidase resistant because of the presence of glycans that are not substrates of the enzymes, because similar glycosidase-resistance, and even resistance to chemical cleavage, has been observed for glycoproteins produced by *C. jejuni* (24). With the confidence that the vast majority of *Hp's* azide-labeled proteins are glycosylated, we undertook experiments to study and identify these labeled glycoproteins.

Assessment of the Subcellular Distribution of Hp's Azide-Labeled Glycoproteins Reveals Glycoproteins Throughout Hp, Including on the Cell Surface—To study *Hp*'s glycoproteins, we began by assessing the subcellular distribution of these species. In particular, we were curious to see whether any of *Hp's* secreted and cell surface proteins are glycosylated, as these proteins are likely to directly modulate interactions with the host and could be novel therapeutic targets. *Hp* cells supplemented with the azidosugar Ac_4GlcNAz or with the azide-free control Ac_4GlcNAc (31) were fractionated to acquire secreted (36), periplasmic, inner membrane-associated, cytoplasmic (35), and surface associated proteins (37). Further, fractions containing outer membrane proteins were also obtained (15, 38). Western blot visualization of azide-labeled glycoproteins (Fig. 3) and SDS-PAGE analysis of total protein (see supplemental Fig. S2*A*–C) demonstrated that protein composition varied between fractions, consistent with successful subcellular separation of proteins. Further, malate dehydrogenase assays (50) of fractions revealed activities con-

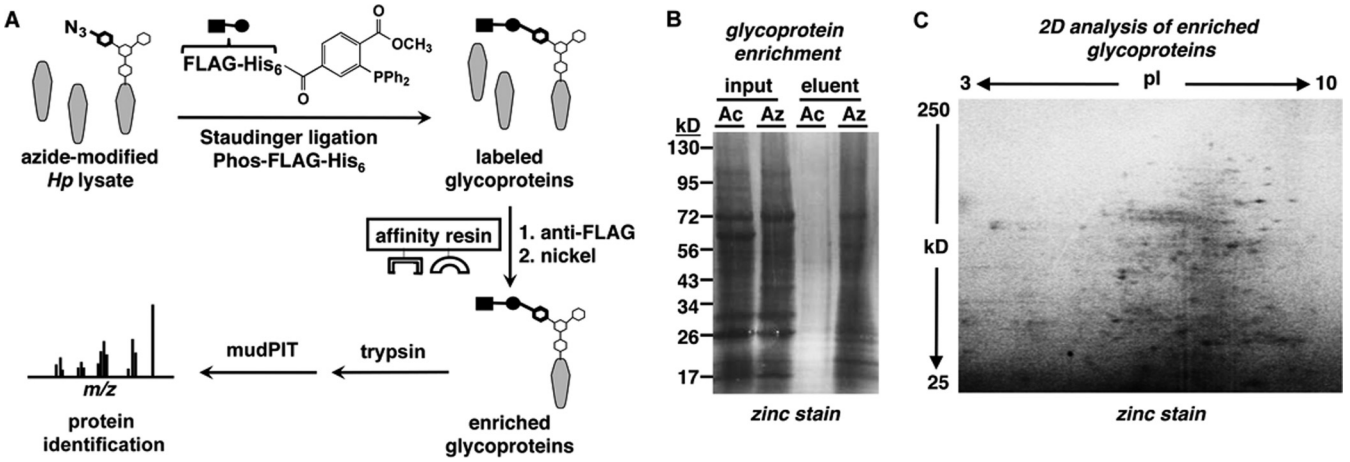

FIG. 4. **Enrichment of *Hp's* azide-labeled glycoproteins.** *A*, Targeted isolation of *Hp's* azide-labeled glycoproteins was facilitated by the Staudinger ligation. First, *Hp's* azide-labeled glycoproteins were reacted with Phos-FLAG-His_6 to yield glycoproteins containing a tandem affinity tag. Labeled glycoproteins were then enriched by anti-FLAG chromatography followed by Ni^{2+} affinity chromatography. Finally, enriched glycoproteins were trypsinized and analyzed by multidimensional protein identification technology (mudPIT). These experiments were performed along side an azide-free control. *B*, SDS-PAGE analysis of proteins from *Hp* treated with Ac_4GlcNAz (Az) or with the azide-free control sugar Ac_4GlcNAc (Ac) before enrichment (input) and after elution from the nickel column (eluent) indicates successful isolation of azide-labeled glycoproteins. Proteins were visualized with zinc stain. *C*, Two-dimensional gel analysis of *Hp*'s enriched, azide-modified glycoproteins. Proteins were visualized with zinc stain.

sistent with successful separation of cytoplasmic and periplasmic proteins (35) (supplemental Fig. S2*D*). No azide-labeled proteins were observed in fractions from cells supplemented with Ac_4GlcNAc, whereas Ac_4GlcNAz treatment resulted in robust azide-dependent signal in all subcellular fractions (Fig. 3). Our data demonstrate the remarkable selectivity of the Staudinger ligation between phosphines and azides, as these abiotic functional groups are nonreactive with endogenous functional groups yet display exquisitely selective reactivity with each other to form a covalent adduct (51, 52). Moreover, these data indicate that glycoproteins are present throughout *Hp* and may serve diverse functions. For example, modification of internal glycoproteins may be important for the physiology of *Hp*, whereas glycosylation of secreted and surface proteins may be crucial for *Hp* to survive and cause harm within the gastric mucosa.

The prominent display of glycans on a subset of *Hp's* cell surface proteins might be harnessed for novel anti-*Hp* therapeutic strategies. For example, these glycosylated proteins might provide the basis of a carbohydrate-based vaccine; analogous glycoconjugate vaccines have been remarkably successful in the clinic (14, 53). Moreover, if glycan modifications are important for the interactions of *Hp* with its host, small molecule inhibitors could target glycan synthesis to inactivate *Hp*. Finally, *Hp's* glycans could be selectively labeled with azides and targeted with phosphines to disrupt their function (14, 54, 55). Realizing these strategies would first require identification of *Hp's* glycoproteins.

Targeted Isolation of Hp's Azide-Labeled Glycoproteins was Facilitated by the Staudinger Ligation—To identify *Hp's* glycoproteins, we began by taking advantage of the metabolically incorporated azide, in conjunction with the Staudinger ligation, to enrich these species. Lysates from *Hp* supplemented with Ac_4GlcNAz or Ac_4GlcNAc were subjected to Staudinger ligation with a phosphine comprising a tandem FLAG-hexahistidine peptide (Phos-FLAG-His_6, Fig. 4*A*) (29, 56). Tagged glycoproteins containing both the FLAG and His_6 epitopes were enriched by purification via anti-FLAG chromatography followed by Ni^{2+} affinity chromatography (29, 56) (Fig. 4*A*). SDS-PAGE analysis enabled comparison of protein profiles from samples before purification (input) *versus* after enrichment of azide-labeled species (eluent). A large number of proteins were present in both the azide-labeled and acetyl-labeled samples before purification (Fig. 4*B*). In contrast, after purification, proteins were present in substantially higher abundance in the enriched, azide-labeled sample relative to the mock-enriched control (Fig. 4*B*). These data highlight the selectivity of the Staudinger ligation between phosphines and azides. The ladder of proteins observed in the enriched, azide-labeled fraction is consistent with the large number of azide-labeled proteins observed by Western blot analysis of lysate from Ac_4GlcNAz-treated cells (Figs. 2*B*, 2*C*). These data indicate a dramatic enrichment of *Hp's* azide-labeled glycoproteins using Phos-FLAG-His_6. Analysis of the enriched, azide-labeled glycoprotein sample by two-dimensional gel electrophoresis yielded greater than 50 detectable species (Fig. 4*C*), indicating that *Hp* synthesizes myriad glycoproteins. The large number of *Hp* glycoproteins detected is similar in magnitude to the number of glycoproteins identified in large-scale studies of the related species *C. jejuni* (57, 58) and in more distantly related bacteria, including *Mycobacterium tuberculosis* (26), *Neisseria sp.* (20)., and *Bacteroides sp.*

(27). *Hp* appears to dedicate considerable metabolic energy to glycoprotein synthesis, suggesting glycoproteins play an important role in *Hp's* physiology.

Identification of Hp's Azide-Labeled Glycoproteins by Mass Spectrometry Indicates Glycosylation of a Number of Virulence Factors—Motivated by these results, we next sought to identify *Hp's* glycoproteins by mass spectrometry analysis. *Hp's* enriched, azide-labeled glycoproteins and a mock-enriched control (which was labeled with Ac_4GlcNAc, reacted with Phos-FLAG-His_6, and enriched via anti-FLAG and Ni-NTA affinity chromatography, as described above) were digested with trypsin, then analyzed by multi-dimensional protein identification technology (mudPIT) (Fig. 4*A*) (40). Experimental detection of unique peptides enabled the unambiguous identification of 127 proteins in the azide-labeled sample. In contrast, the mock-enriched control that lacked the azide contained only two *Hp* proteins (HP0721, HP0570), both of which were also present in the azide-labeled sample. These proteins were subtracted from the list of putative glycoproteins identified in the azide-labeled sample, narrowing the pool of candidate glycoproteins to 125 (Table I and Table II). All proteins were identified based on two or more peptides meeting acceptance criteria. Detailed information on all of the proteins identified in the azide-labeled sample and the mock-enriched control, including detected peptides that made assignments possible, can be found in supplemental Table S1. These results demonstrate efficient enrichment and identification of azide-labeled proteins, suggesting this metabolic labeling strategy is a robust method to profile bacterial glycoproteins.

Of the putative glycoproteins identified, 70% have housekeeping functions, 21% are linked to *Hp*'s pathogenesis, and 9% have unknown function. From a basic biology perspective, the identification of a number of putative glycoproteins that have housekeeping functions suggests that glycosylation could play a role in proper folding and function (59), or perhaps that glycosylation provides a selective advantage for survival in the harsh conditions of the gastric tract. From a therapeutic perspective, it is intriguing that 26 of *Hp*'s glycoproteins have known roles in pathogenesis (Table I) that include enabling survival in the host stomach via iron acquisition (pfr) (60), pH resistance (ureA, ureB, tufB) (61), oxidative protection (tsaA, sodB, katA, tagD, msrA) (62, 63), motility (flgH) (64), adhesion (babC, babA, tlpA, dnaK, htrA) (65), eliciting an immune response (napA, msrA, icd, frdA) (66, 67), and causing harm to the host (cag7, groES, groEL, hopQ, fusA) (68, 69). The large percentage of virulence factors in *Hp's* glycoproteome, relative to the low percentage of virulence factors in *Hp's* proteome (70, 71), implicates *Hp's* glycans in pathogenesis. This observation suggests that the known link between protein glycosylation and virulence in *Hp*, which Logan firmly established with her studies of *Hp's* flagellin glycosylation (7), may be a widespread phenomenon within this bacterium.

We next considered the subcellular distribution of *Hp's* putative glycoproteins. Based on literature reports and *in silico* predictions, candidate glycoproteins appear to be expressed throughout *Hp* (Fig. 5), including intracellular locations such as the cytoplasm, inner membrane, and periplasm, and extracellular locations such as the outer membrane, cell surface (surface-associated), and the secreted milieu. Moreover, the predicted subcellular distribution of *Hp's* putative glycoproteins (Fig. 5) appears to reflect the subcellular distribution that we detected experimentally (Fig. 3). Follow-up experiments need to be conducted to establish whether these localization predictions are correct and to confirm that *Hp* synthesizes glycoproteins throughout the cell, as our data suggest.

We next compared the results of our analyses of *Hp's* glycoproteins to a recent study by Creuzenet and coworkers. In their study, they employed a periodic acid/hydrazide-based screen to identify nine candidate *Hp* glycoproteins (15). Five of the putative glycoprotein hits identified here (katA, gltA, tsf, atpA, atpD) were also identified as candidate glycoproteins in that screen, further validating our hit list. Moreover, their analyses indicated the presence of glycoproteins in soluble and membrane fractions of *Hp* (15), consistent with the subcellular distribution observed here. These findings underscore the value of our targeted strategy to detect and discover bacterial glycoproteins in a manner that is complementary to alternative techniques.

We note the absence of the flagellin proteins FlaA and FlaB, glycoproteins confirmed to be modified with pseudaminic acid, from our list of identified azide-labeled glycoproteins. This result is consistent with our previous studies, which demonstrated that Ac_4GlcNAz treatment does not lead to detectable incorporation of azide into *Hp's* flagellin proteins (31). Our previously reported data indicate that Ac_4GlcNAz is not converted to the complex sugar pseudaminic acid at detectable levels. Thus, though the studies described here identity a large number of putative *Hp* glycoproteins, they do not exhaustively identify all *Hp* glycoproteins.

Mass Spectrometry Analyses Reveal Staudinger Ligation-Glycan Adducts—We sought to characterize the nature of *Hp's* azide-modified glycans. A sample of *Hp's* enriched, azide-labeled glycoproteins was analyzed by HPLC-Chip/Q-TOFMS before and after being subjected to beta-elimination to cleave O-linked glycans. For these preliminary experiments, we focused on identifying the simplest Staudinger ligation-glycan adduct that would be formed as a result of metabolic labeling with Ac_4GlcNAz - addition of a single GlcNAz residue or, if the stereochemistry is changed, an *N*-azidoacetylhexosamine (HexNAz) residue onto a glycosylated protein. Thus, we directed our analysis at the identification of the Staudinger ligation-glycan adduct that would be formed by protein glycosylation with a single HexNAz residue followed by release via beta-elimination - HexNAz-Phos-FLAG-His_6 (**1**) (Fig. 6). Because this adduct has not been

TABLE I
Putative glycoproteins identified by mudPIT with links to pathogenesis

Gene	Acc. #	Protein name	Predicted MW (kDa)	Function/sequence features	Link to pathogenesis	Subcellular localization[a]	Spect-counts	Uni-pep	Seq-cov%
HP0073	P14916	ureA	27.8	Urease alpha subunit	Enables survival in acid	Cytoplasm, surface	96	4	34.4
HP0072	P69996	ureB	61.6	Urease beta subunit	Enables survival in acid	Cytoplasm, surface	188	7	22.7
HP0527	NP_207323	cag7	230.1	*cag* pathogenicity island, type IV secretion	Oncogene, gastric cancer	Outer membrane	2	2	1.4
HP1563	P21762	tsaA	23.5	Alkyl hydroperoxide reductase	Resists oxidative damage	Cytoplasm	428	12	65.7
HP0325	O25092	flgH	27.8	Flagellar export	Critical for motility	Outer membrane	4	2	7.2
HP0010	P42383	groEL	61.5	Chaperone	Gastric cancer-associated antigen, immunogen	Cytoplasm, surface	576	18	37.3
HP0011	P0A0R3	groES	13.7	Chaperone	Gastric cancer-associated antigen, immunogen	Cytoplasm, surface	546	7	41.9
HP1161	O25776	fldA	18.5	Flavodoxin	Critical metabolic gene	Unknown	2	2	20.8
HP0653	P52093	pfr	20.3	Nonheme iron-containing ferritin	Crucial for iron acquisition	Cytoplasm	35	3	25.6
HP0389	AAD07454	sodB	25.9	Superoxide dismutase	Resists oxidative damage	Periplasm	10	4	28.9
HP0875	P77872	katA	61.5	Catalase	Combats oxidative stress	Periplasm, surface	91	6	18.0
HP0390	O25151	tagD	19.3	Adhesin thiol peroxidase	Combats oxidative stress	Periplasm	3	3	32.6
HP0243	P43313	napA	17.9	Neutrophil activation	Immunomodulator, biofilm formation	Cytoplasm, surface	9	3	25.5
HP0224	O25011	msrA	43.2	Methionine sulfoxide reductase	Antioxidant, immunogen	Cytoplasm	42	4	18.6
HP0317	AAD07380	babC	85.4	Outer membrane protein	Adhesion, glycan binding	Outer membrane	4	2	5.4
HP1243	AAD08288	babA	83.3	Outer membrane protein	Adhesion, glycan binding	Outer membrane	2	1	2.8
HP1177	AAD08221	hopQ	73.3	Outer membrane protein	Present on disease-related strains	Outer membrane	14	2	5.5
HP1205	P56003	tufB	45.9	Translation elongation factor	Resistance to prolonged acid exposure	Cytoplasm	10	4	17.1
HP1123	O25748	slyD	21.3	Peptidyl-prolyl cis-trans isomerase	Important for urease assembly	Cytoplasm	2	2	16.2
HP1195	P56002	fusA	80.9	Translation elongation factor	Duodenal ulcer-related antigen	Cytoplasm, surface	36	8	21.0
HP0027	P56063	icd	50.1	Isocitrate dehydrogenase	Immunogen, induces humoral immune system	Cytoplasm	21	4	15.6
HP0192	O06913	frdA	84.0	Fumarate reductase	Immunogen, crucial for survival in gastric mucosa	Inner membrane	28	5	13.2
HP0109	P55994	dnaK	70.6	Chaperone and heat shock protein 70	Stress induced surface adhesion, immunogen	Cytoplasm, surface	6	3	8.7
HP0082	AAD07152	tlpC	79.1	Methyl-accepting chemotaxis transducer	Assist colonization of the stomach	Inner membrane	10	3	7.6
HP1019	AAD08063	htrA	50.6	Serine protease, chaperone	Disrupts intracellular adhesion	Periplasm, secreted	5	2	6.6
HP0099	AAD07167	tlpA	78.4	Methyl-accepting chemotaxis protein	Chemotaxis receptor	Outer membrane	3	1	3.2

[a] Subcellular localization was based on literature reports or predicted using PSORTb v.3.0.

TABLE II
Additional putative glycoproteins identified by mudPIT

Gene	Acc. #	Protein name	Predicted MW (kDa)	Function/sequence features	Subcellular localization[a]	Spect-counts	Uni-pep	Seq-cov%
HP0025	AAD07106	omp2	81.6	Outer membrane protein	Outer membrane	5	3	7.9
HP0191	O06914	frdB	29.3	Fumarate reductase beta	Inner membrane	15	4	26.2
HP0599	AAD07662	hylB	51.0	Hemolysin secretion	Outer membrane	31	8	29.5
HP0371	AAD07435	fabE	18.1	Biotin carboxyl carrier	Unknown	35	4	26.7
HP0110	P55970	grpE	23.3	Heat shock protein	Cytoplasm	14	5	29.9
HP0631	AAD07691	hydA	44.6	Quinine reactive hydrogenase	Inner membrane	4	3	9.1
HP1198	O25806	rpoBC	339.5	RNA polymerase	Cytoplasm	9	6	3.6
HP1496	P56078	ctc	20.9	General stress protein	Cytoplasm	6	3	33.7
HP0824	P66928	trxA	12.5	Thioredoxin A	Cytoplasm	20	2	26.8
HP0825	3ISH_B	trxB	35.5	Thioredoxin reductase	Cytoplasm	22	6	33.1
HP1458	AAD08500	—	12.4	Thioredoxin	Cytoplasm	47	2	26.4
HP1325	O25883	fumC	53.6	Fumarase	Cytoplasm	23	5	25.5
HP1555	P55975	tsf	41.6	Translation elongation factor	Cytoplasm	17	3	9.1
HP0266	O25045	pyrC	44.6	Dihydroorotase	Unknown	19	5	25.1
HP1135	AAD08177	atpH	21.3	ATP synthase F1 delta	Membrane	10	2	18.0
HP1134	P55987	atpA	58.1	ATP synthase F1 alpha	Membrane	28	5	13.2
HP1132	P55988	atpD	54.1	ATP synthase F1 beta	Membrane	26	6	22.3
HP1099	AAD08142	eda	23.9	Aldolase	Cytoplasm	5	3	23.2
HP0154	P48285	eno	49.1	Enolase	Cytoplasm	12	4	21.6
HP0512	P94845	glnA	57.4	Glutamate synthase	Cytoplasm	32	4	15.4
HP0779	AAD07828	acnB	97.7	Aconitase	Cytoplasm	72	10	19.9
HP0163	P56074	hemB	38.1	Delta-aminolevulinic acid dehyratase	Cytoplasm	5	3	15.2
HP0974	P56196	pgm	57.7	Phosphoglycerate mutase	Cytoplasm	5	4	18.1
HP0294	O25067	aimE	39.7	Aliphatic amidase	Cytoplasm	31	3	17.9
HP0649	P56149	aspA	54.6	Aspartate ammonia lyase	Cytoplasm	17	5	17.5
HP0865	O25536	dut	16.8	Nucleotidohydrolase	Unknown	3	1	17.5
HP0020	AAD07088	nspC	48.2	Carboxynorpermidine decarboxylase	Unknown	5	4	16.9
HP1540	AAD08580	fbcF	19.2	Ubiquinol cytochrome c oxidase	Inner membrane	18	2	15.9
HP0480	O25225	yihK	69.9	GTP-binding	Inner membrane	10	5	14.1
HP0954	O25608	—	25.4	NAD(P)H reductase	Cytoplasm	3	2	14.0
HP0176	P56109	tsr	35.7	Fructose-bisphosphate aldolase	Cytoplasm	3	2	13.8
HP1110	AAD08154	porA	47.0	Pyruvate ferredoxin oxidoreductase	Unknown	47	4	18.9
HP1111	AAD08155	pfor	36.9	Pyruvate ferredoxin oxidoreductase beta	Unknown	37	3	16.9
HP0632	AAD07692	hydB	67.6	Quinone-reactive Ni-Fe hydrogenase	Inner membrane	7	4	13.5
HP0026	P56062	gltA	50.9	Citrate synthase	Cytoplasm	71	6	13.2
HP0757	AAD07805	—	34.9	Beta-alanine synthase	Cytoplasm	6	2	12.4
HP0485	AAD07551	—	37.7	Catalase-like protein	Periplasm	4	3	12.3
HP0201	AAD07269	plsX	38.4	Fatty acid phospholipid synthesis	Cytoplasm	5	2	12.1
HP0183	P56089	glyA	48.0	Serine hydroxyl-methyltransferase	Cytoplasm	17	3	11.9
HP0690	AAD07742	fadA	43.4	Acetyl coenzyme acetyltransferase	Cytoplasm	5	2	11.9

TABLE II—*continued*

Gene	Acc. #	Protein name	Predicted MW (kDa)	Function/sequence features	Subcellular localization[a]	Spect-counts	Uni-pep	Seq-cov%
HP0402	P56145	pheT	89.4	phenylalanyl-tRNA synthase	cytoplasm	20	5	11.7
HP0096	AAD07165	—	36.7	phosphoglycerate dehydrogenase	cytoplasm	6	2	11.7
HP0576	O25300	lepB	35.3	signal peptidase I	inner membrane	11	3	11.1
HP1013	O25657	dapA	34.8	dihydrodipicolinate synthetase	cytoplasm	8	2	10.8
HP0397	AAD07461	serA	60.8	phosphoglycerate dehydrogenase	cytoplasm	12	3	10.7
HP1576	O26096	abc	38.6	ABC transporter	inner membrane	4	2	10.7
HP0006	P56061	panC	32.8	pantoate-beta-alanine ligase	cytoplasm	2	2	10.7
HP0558	NP_207353	fabF	45.8	beta ketoacyl-acyl carrier protein synthase	inner membrane	8	2	10.6
HP0330	O25097	ilvC	38.5	ketol-acid reductoisomerase	cytoplasm	2	2	10.3
HP0672	AAD07733	aspB	45.1	mitochondrial signature protein	cytoplasm	2	2	10.0
HP1103	O25731	glk	38.6	glucokinase	cytoplasm	11	2	9.9
HP0221	AAD07289	—	38.3	nifU-like	cytoplasm	2	2	9.9
HP0604	AAD07669	hemE	40.4	uroporphyrinogen decarboxylase	cytoplasm	3	2	9.8
HP0306	P56115	hemL	49.3	aminomutase	cytoplasm	3	2	9.7
HP1266	AAD08310	nqo3	99.1	NADH-ubiquinone oxidoreductase	unknown	8	4	9.3
HP1366	AAD08410	mboIIR	52.6	type IIS restriction enzyme R	unknown	4	2	8.5
HP0490	AAD07558	—	45.2	putative potassium channel	unknown	2	2	8.5
HP1181	AAD08227	mfs	51.4	multidrug-efflux transporter	inner membrane	4	2	8.4
HP0777	P56106	pyrH	27.4	uridine 5′ monophosphate kinase	cytoplasm	13	2	8.3
HP1422	P56456	ileS	111.3	Isoleucyl-tRNA synthetase	cytoplasm	5	4	8.1
HP0322	AAD07390	—	61.4	poly E-rich protein	cytoplasm	5	3	7.9
HP0382	AAD07451	—	48.5	zinc metallo protease	unknown	3	2	7.9
HP0859	AAD07905	rfaD	39.4	mannoheptose epimerase	cytoplasm	5	2	7.8
HP1100	P56111	—	70.2	6-phophogluconate dehydratase	cytoplasm	55	2	6.4
HP0210	P56116	htpG	74.8	chaperone and heat shock protein	cytoplasm	2	2	7.2
HP0056	AAD07126	—	141.6	pyrroline-5-carboxylate dehydrogenase	cytoplasm	14	5	7.1
HP0728	O25428	tilS	41.7	tRNA(Ile)-lysidine synthetase	unknown	3	2	7.1
HP1045	O25686	acoE	78.8	acetyl-CoA synthetase	cytoplasm	7	3	6.7
HP0169	P56113	prtC	49.9	collagenase	unknown	2	2	6.7
HP1104	AAD08150	cad	40.6	cinnamyl-alcohol dehydrogenase	cytoplasm	7	2	6.6
HP0470	AAD07532	pepF	70.9	oligoendopeptidase F	cytoplasm	5	3	6.4
HP0422	AAD07486	speA	74.0	arginine decarboxylase	periplasm	7	3	6.3
HP0510	P94844	dapB	29.4	dihydrodipicolinate reductase	cytoplasm	4	2	6.3

TABLE II—*continued*

Gene	Acc. #	Protein name	Predicted MW (kDa)	Function/sequence features	Subcellular localization[a]	Spect-counts	Uni-pep	Seq-cov%
HP0680	P55982	nrdA	94.7	Ribonucleoside-diphosphate reductase	Cytoplasm	2	2	5.5
HP1430	P56185	—	81.3	Conserved hypothetical ATP-binding protein	Unknown	4	2	5.5
HP1012	AAD08056	pqqE	52.9	Protease	Unknown	2	2	5.3
HP1213	O25812	pnp	80.8	Polynucleotide phosphorylase	Cytoplasm	2	2	4.8
HP1278	P56142	trpB	45.0	Tryptophan synthase beta	Cytoplasm	16	1	4.6
HP1072	P55989	copA	86.1	Copper-transporting ATPase	Inner membrane	3	2	4.3
HP1450	O25989	yidC	65.8	Protein translocase	Inner membrane	6	2	4.3
HP0121	P56070	ppsA	95.8	Phosphoenolpyruvate synthase	Cytoplasm	3	2	4.1
HP0607	NP_207402	acrB	119.6	Acriflavine resistance	Inner membrane	2	2	4.0
HP1402	AAD08445	hsdR	121.7	Type I restriction enzyme	Unknown	2	2	3.4
HP1112	P56468	purB	52.6	Adenylosuccinate lyase	Cytoplasm	5	1	3.4
HP1478	AAD08516	uvrD	81.7	DNA helicase II	Cytoplasm	2	2	3.3
HP1547	P56457	leuS	97.7	Leucyl-tRNA synthase	Cytoplasm	3	2	2.9
HP0791	Q59465	cadA	78.9	Cadmium-transporting ATPase	Inner membrane	2	2	2.9
HP0600	AAD07663	spaB	71.8	Multidrug resistance	Inner membrane	3	1	2.7
HP0696	AAD07747	—	90.7	N-methylhydantoinase	Unknown	3	2	2.6
HP0958		—	31.3	Hypothetical protein	Unknown	2	2	12.3
HP0787		—	48.2	Conserved integral membrane protein	Inner membrane	21	3	9.3
HP0231		—	31.1	*Hypothetical protein*	Unknown	10	4	22.9
HP0318		—	30.1	*Hypothetical protein*	Unknown	35	3	18.4
HP0086		—	53.3	*Hypothetical protein*	Unknown	35	3	9.1
HP0773		—	41.9	*Hypothetical protein*	Unknown	32	2	8.9
HP0754		—	9.7	*Hypothetical protein*	Unknown	7	2	34.1
HP0396		—	74.4	*Hypothetical protein*	Unknown	3	3	8.8
HP1079		—	45.2	*Hypothetical protein*	Unknown	34	2	6.9
HP1143		–	53.2	Hypothetical protein	Unknown	4	1	5.0

[a] Subcellular localization was based on literature reports or predicted using PSORTb v.3.0.

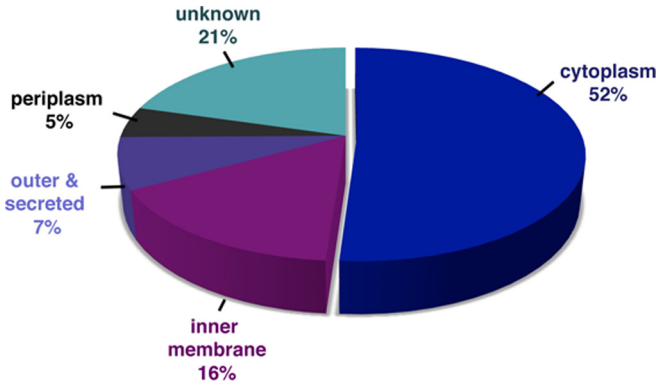

FIG. 5. **Predicted subcellular location of candidate glycoproteins.** The cellular locations of each identified protein was determined by a combination of literature reports and predictions by PSORTb v.3.0. Secondary surface-associated localization (n = 8) was not included.

previously characterized by MS, we first synthesized GlcNAz-Phos-FLAG-His_6 and analyzed its chromatographic and mass spectrometric properties using LC-nanoESI and collision-induced dissociation (CID). The nanoESI mass spectrum of the synthetic standard GlcNAz-Phos-FLAG-His_6 (M = 2400.90 Da) is dominated by $[M+3H]^{3+}$, $[M+4H]^{4+}$, and $[M+5H]^{5+}$ ions (supplemental Fig. S3); additionally, we observed metastable losses of water that were most apparent for the $[M+5H]^{5+}$ charge state. The intensities of these peaks were highly dependent on voltages controlling ion transmission, so we attempted to minimize these contributions through careful instrument tuning. CID-MS/MS analysis of the $[M+4H]^{4+}$ ion at m/z 601.24 revealed fragmentation rationalized by initial loss of the neutral glycan via pathway (a) to form product ion Ia^{4+} (m/z 556.46) (Figs. 6 and 7*A*), which proceeds, via net loss of CO and CH_2NH, to form the base peak in the spectrum

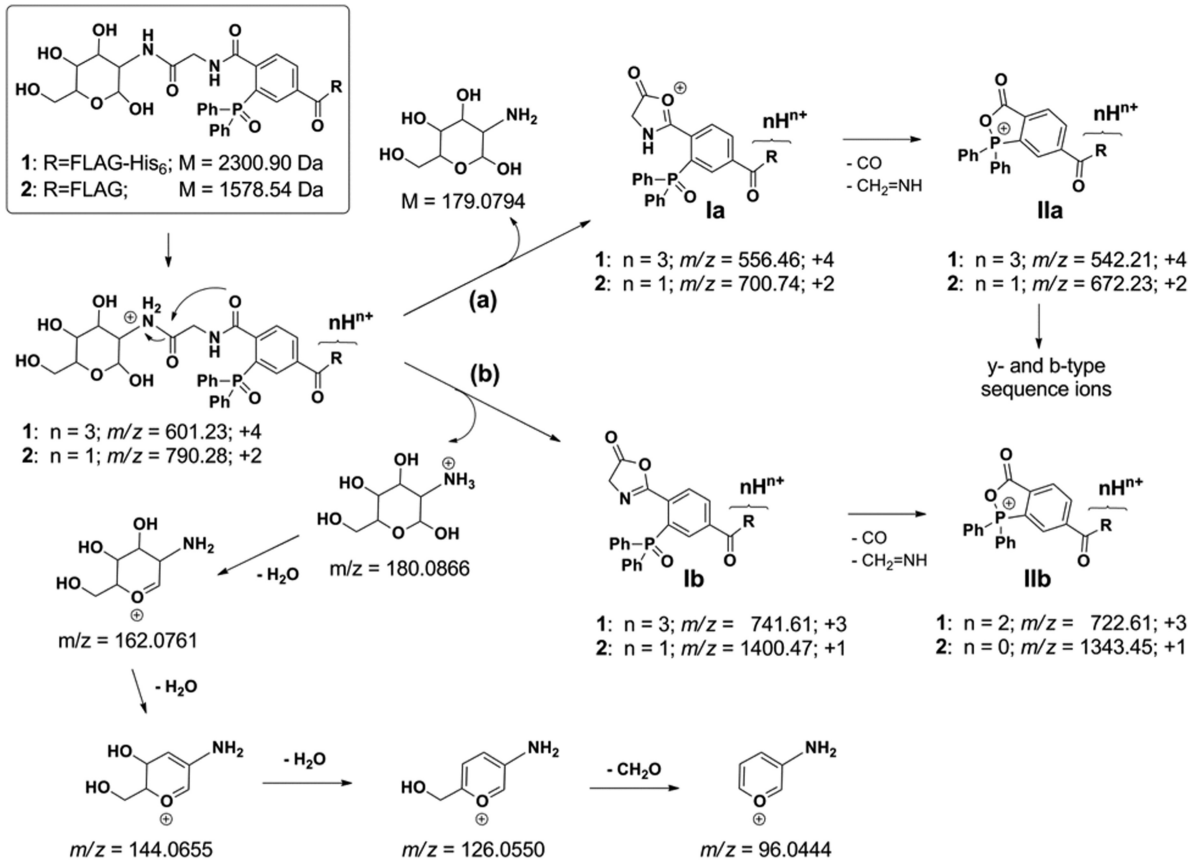

FIG. 6. **Proposed mass spectrometric fragmentation pathways for Staudinger ligation-glycan adducts 1 (HexNAz-Phos-FLAG-His$_6$) and 2 (HexNAz-Phos-FLAG).** CID-MS/MS of nanoESI-generated $[M+4H]^{4+}$ (adduct **1**) and $[M+2H]^{2+}$ (adduct **2**) ions yield fragment ions Ia and IIa, formed via pathway (a), through which loss of the neutral glycan is hypothesized to generate cycylized Ia, which is followed by sequential losses of CO and CH_2NH to generate IIa. Charge-retention on the glycan and subsequent fragmentation via pathway (b) rationalizes formation of glycan-derived putative product ions at *m/z* 162.08, *m/z* 144.07, *m/z* 126.06, and *m/z* 96.04. Production of y- and b-type fragment ions is attributed to subsequent fragmentation of IIa yielding C- and N-(Phos-containing)-terminus fragments, respectively (see Figs. 7*E* and 8*E*).

(product ion IIa^{4+} at *m/z* 542.21) (Figs. 6 and 7*A*; supplemental Table S3). We also detect triply charged versions of product ions at *m/z* 741.63 and 722.61 (ions Ib^{3+} and IIb^{3+}, respectively) (Figs. 6 and 7*A*), which may be formed by loss of the positively charged glycan via pathway (b) (Fig. 6). This pathway may also be responsible for formation of the glycan-derived product ions that appear in the spectrum at *m/z* 162.08, *m/z* 144.07, *m/z* 126.06, and *m/z* 96.04 (see proposed structures in Figs. 6 and 7*B*). We also detected FLAG-His$_6$-derived product ions. These include abundant immonium ions at *m/z* 110.07 (His) and *m/z* 129.10 (Lys), as well as y-type (72) fragments containing the FLAG-His$_6$ C terminus (see Fig. 7*E*) and b-type (72) fragments; the latter are denoted by b_n^* and contain the glycan-eliminated N terminus (see Fig. 7*E*) that is attributed to subsequent fragmentations from ion IIa^{4+}. To determine if the Staudinger ligation-glycan adduct HexNAz-Phos-FLAG-His$_6$ (**1**) is released from *Hp's* enriched, azide-labeled glycoproteins on beta-elimination, a sample was analyzed by HPLC-Chip/Q-TOFMS before and after being subjected to beta-elimination to cleave O-linked glycans. Although analysis of the pre-beta-eliminated sample showed no evidence for product **1**, LC-nanoESI analysis of the beta-eliminated sample showed a chromatographic peak appearing at the retention time characteristic of synthetic **1**, and producing a mass spectrum with *m/z* values consistent with the mass of product **1** (M = 2,300.90 Da, mass measurement error < 5 ppm) (supplemental Fig. S4 and supplemental Table S2). Furthermore, when the $[M+4H]^{4+}$ ion at *m/z* 601.24 from putative product **1** was subjected to MS/MS analysis, the measured mass spectrum showed excellent agreement with the synthetic standard (Figs. 7*C*, 7*D*; supplemental Table S4). These results suggest that one metabolic fate of Ac_4GlcNAz in *Hp* is conversion to HexNAz (presumably GlcNAz) and incorporation into *Hp's* O-linked glycoproteins.

Although the above data indicate the presence of a Staudinger ligation-glycan adduct in an azide-labeled glycoprotein sample, we sought to confirm that these adducts are absent from acetyl-labeled controls. Because of the metastable water losses observed with the Phos-FLAG-His$_6$ tag and the multiplicity of charge states, we chose to conduct fol-

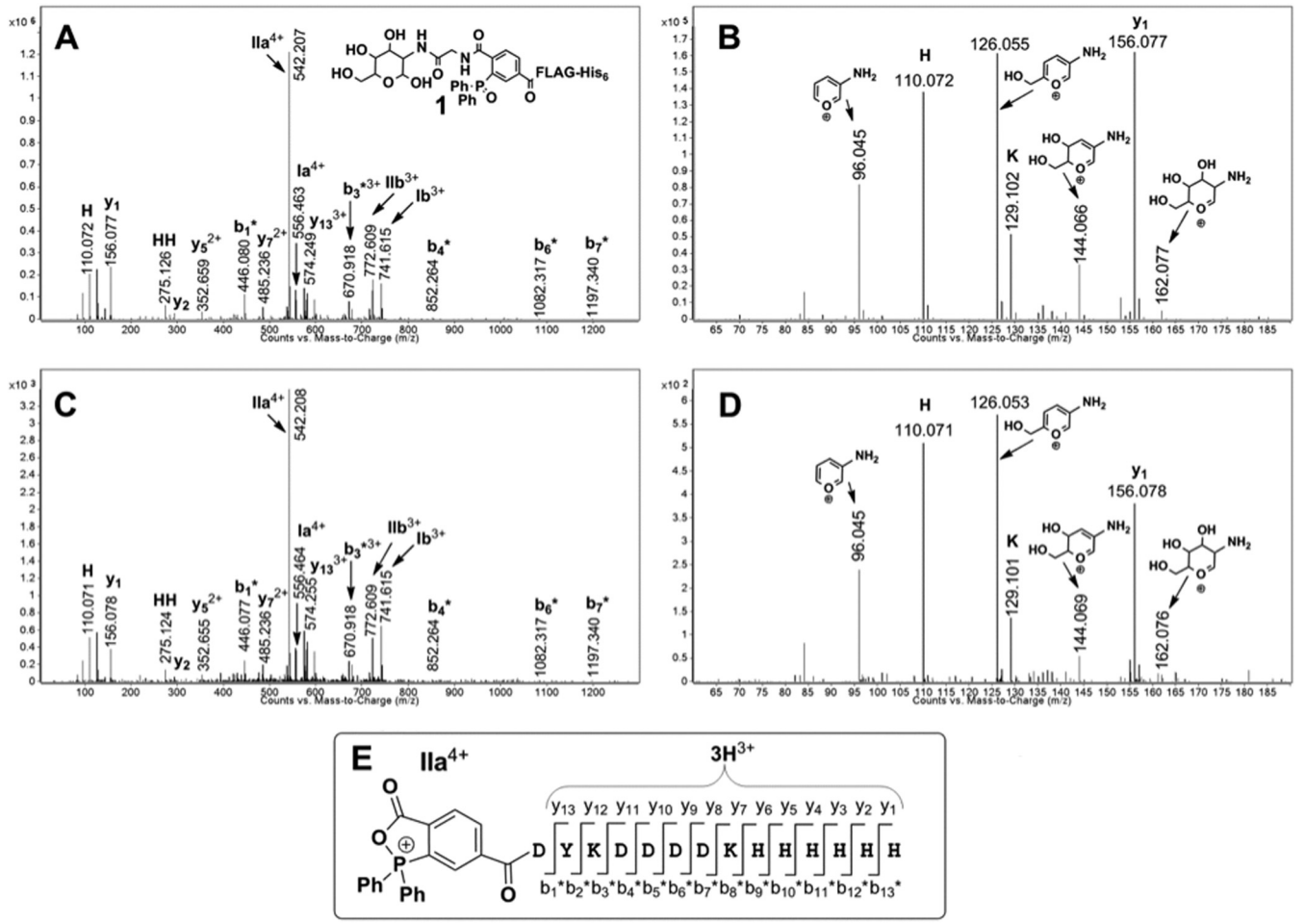

FIG. 7. **LC-nanoESI-CID-MS/MS spectra of (*A*, *B*) synthetic standard GlcNAz-Phos-FLAG-His$_6$ and (*C*, *D*) beta-eliminated glycans from enriched azide-labeled glycoprotein sample.** *A*, CID-MS/MS spectrum of GlcNAz-Phos-FLAG-His$_6$ synthetic standard **1** showing ions Ia^{4+} and Ib^{3+} (see Fig. 6) resulting from glycan loss, and ions IIa^{4+} and IIb^{3+} following net loss of CO, NH = CH2 (Fig. 6). As summarized in (*E*), subsequent fragmentation of ion IIa^{4+} is proposed to yield y-type fragment ions that include the C terminus of the FLAG-His$_6$ peptide and b-type ions, labeled with an asterisk, that include the Phos-modified N terminus. *B*, Expansion (magnified 5x along the *y* axis) of the low-mass region from (*A*) displays a variety of peaks that include glycan-derived fragments with proposed structures represented. Immonium ions (H and K) are also detected in this mass range. *C*, CID-MS/MS spectra of beta-eliminated glycans from enriched azide-labeled glycoprotein sample showing mass spectral features that agree with those of the synthetic standard (*A*). *D*, Expansion (magnified 5x along the *y* axis) of the low-mass region from (*C*), showing the production of glycan-derived peaks from the beta-eliminated putative HexNAz-Phos-FLAG-His$_6$ (**1**). *E*, Schematic representation of ions (y- and b-type) derived from further IIa^{4+} fragmentation. The asterisk associated with the b-type ions denotes a Phos-modified N terminus. Charge states are labeled when greater than +1. All CID-MS/MS spectra were measured using nitrogen as the collision gas, a 4 Da window for ion isolation, and collision energy of 20 V.

low-up experiments with the less complex Phos-FLAG tag. We first characterized synthetic GlcNAz-Phos-FLAG (**2**) (M = 1578.54 Da) by LC-nanoESI and CID. With the basic His$_6$ chain eliminated, **2** yielded a nanoESI mass spectrum dominated by the lower charge state [M+2H]$^{2+}$ ion at *m/z* 790.28 (supplemental Fig. S5; supplemental Table S5). As was observed for **1**, the CID-MS/MS spectrum for GlcNAz-Phos-FLAG (**2**) (Fig. 8*A* and 8*B*) was dominated by peaks resulting from neutral glycan loss via pathway (a) to form product ions Ia^{2+} and IIa^{2+} (Figs. 6 and 8*A*); however, glycan-derived fragments (*m/z* = 96.04, 126.05, 144.07 and 162.08) were weaker, whereas y- and b-type sequence ions (Figs. 8*A* and 8*E*) were more abundant in the spectrum of GlcNAz-Phos-FLAG (**2**), a change that may result from the lower charge state of this precursor.

With characterization of the targeted analyte in hand, lysates from *Hp* supplemented with Ac$_4$GlcNAz or Ac$_4$GlcNAc were rinsed extensively to remove residual free sugar, and then subjected to Staudinger ligation with Phos-FLAG. After extensive washing to remove unreacted Phos-FLAG from labeled lysates via ultrafiltration, samples were analyzed by HPLC-Chip/Q-TOFMS before and after being subjected to beta-eliminations to cleave O-linked glycans. LC-nanoESI analysis of the beta-eliminated glycans from the azide-labeled

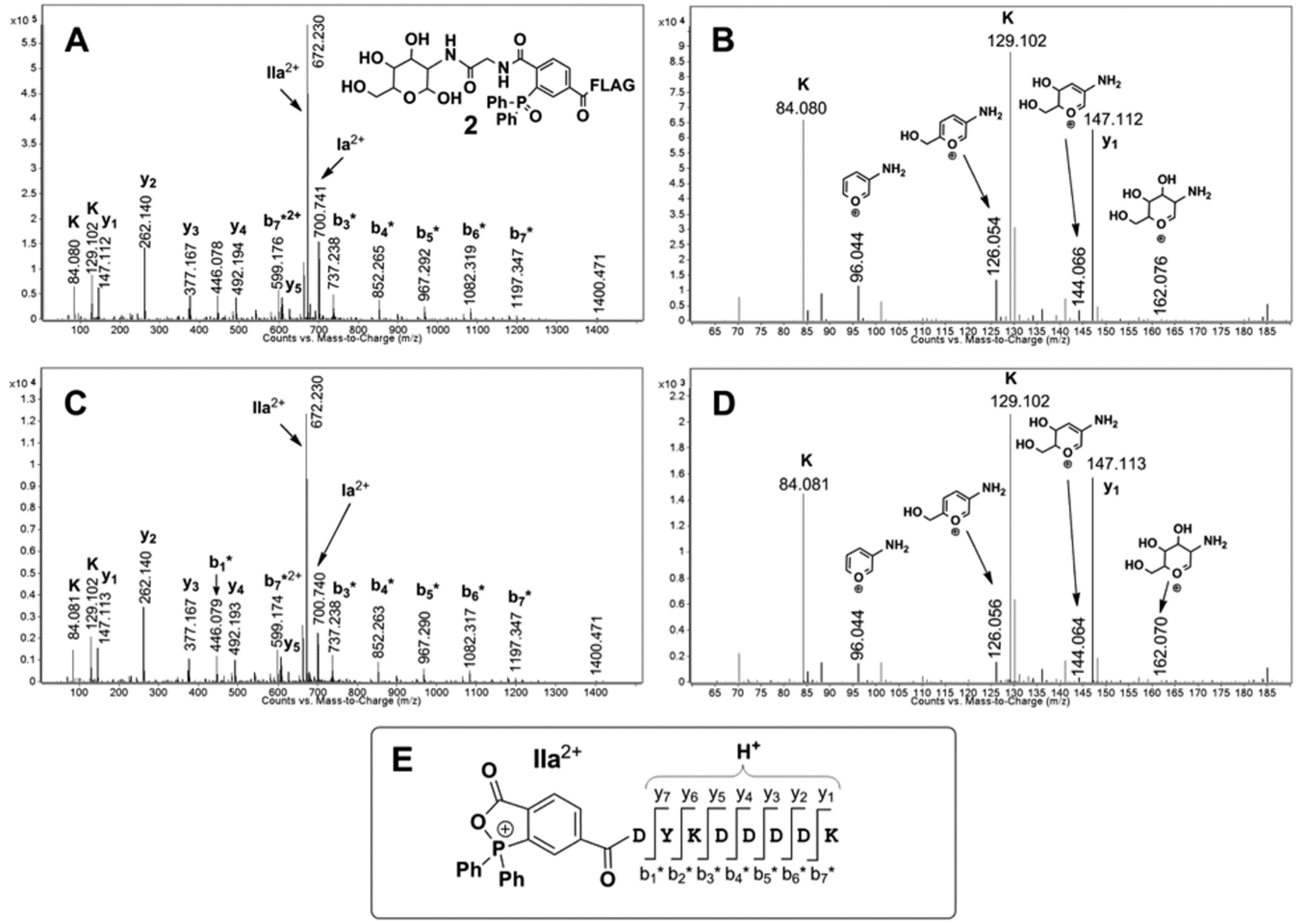

FIG. 8. **LC-nanoESI-CID-MS/MS spectra of (*A*, *B*) synthetic standard GlcNAz-Phos-FLAG (2) and (*C*, *D*) beta-eliminated glycans from an azide-labeled, Phos-FLAG-ligated, *Hp* cell lysate.** *A*, CID-MS/MS spectrum of GlcNAz-Phos-FLAG synthetic standard **2** showing ions Ia^{2+} (see Fig. 6) resulting from glycan loss, and ions IIa^{2+} following net loss of CO, NH = CH2 (Fig. 6). As summarized in (*E*), subsequent fragmentation of ion IIa^{2+} is proposed to yield y-type fragment ions that include the C terminus of the FLAG peptide and b-type ions, labeled with an asterisk, that include the Phos-modified N terminus. *B*, Expansion of the low-mass region from (*A*) displays a variety of peaks that include glycan-derived fragments with proposed structures represented. Immonium ions (K) are also detected in this mass range. *C*, CID-MS/MS spectra of beta-eliminated glycans from an azide-labeled, Phos-FLAG-ligated, *Hp* cell lysate showing mass spectral features that agree with those of the synthetic standard (*A*). *D*, Expansion of the low-mass region from (*C*), showing the production of glycan-derived peaks from the beta-eliminated putative HexNAz-Phos-FLAG (**2**). *E*, Schematic representation of ions (y- and b-type) derived from further IIa^{4+} fragmentation. The asterisk associated with the b-type ions denotes a Phos-modified N terminus. Charge states are labeled when greater than +1. All CID-MS/MS spectra were measured using nitrogen as the collision gas, a 4 Da window for ion isolation, and collision energy of 35 V.

sample revealed a doubly-charged peak at *m/z* 790.28 that corresponds to the mass of HexNAz-Phos-FLAG (**2**) (M = 1578.54 Da, mass measurement error < 5 ppm; supplemental Fig. 6*E* and supplemental Table S2). The CID-MS/MS spectrum of the peak at *m/z* 790.28 (Fig. 8*C* and 8*D*; supplemental Table S6) is consistent with the Staudinger ligation product between GlcNAz and Phos-FLAG - HexNAz-Phos-FLAG (**2**). This HexNAz-Phos-FLAG adduct is not detectable in the beta-eliminated sample from the corresponding acetyl-labeled control (supplemental Fig. S6*C*), nor is it present at detectable levels in tagged azide- or acetyl-labeled lysates before beta-elimination (supplemental Fig. S6*A* and S6*B*). The exclusive detection of this adduct in beta-eliminated azide-labeled samples indicates that it is derived from glycosylated *Hp* proteins and that beta-elimination conditions released this glycan adduct. The absence of this adduct in samples before beta-elimination further supports the conclusion that this adduct is derived from *Hp* glycoprotein conjugates and is released by the beta-elimination reaction. These data provide strong evidence that some *Hp* glycoproteins are covalently modified with O-linked HexNAz.

N-acetylhexosamine (HexNAc) residues are found in a number of characterized bacterial glycoproteins, including those synthesized by *Campylobacter jejuni* (73) and *Helicobacter pullorum* (74). Indeed, *H. pullorum* has an N-linked protein glycosylation system that makes use of a pentasaccharide

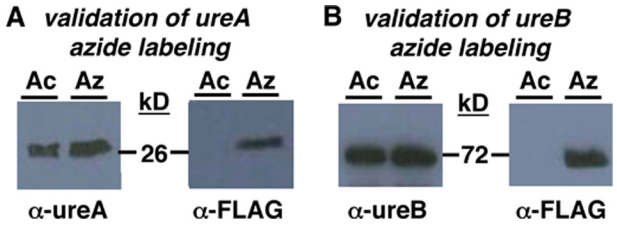

FIG. 9. **Secondary screens to assess the glycosylation status of urease alpha (ureA) and urease beta (ureB) subunits.** *A*, Western blot analysis of immunopurified ureA from *Hp* treated with Ac_4GlcNAz (Az) or Ac_4GlcNAc (Ac). These samples were reacted with Phos-FLAG, electrophoresed, and then probed with anti-ureA or anti-FLAG antibody. *B*, Western blot analysis of immunopurified ureB from *Hp* treated with Ac_4GlcNAz (Az) or Ac_4GlcNAc (Ac). These samples were reacted with Phos-FLAG, electrophoresed, and then probed with anti-ureB or anti-FLAG antibody.

with a core HexNAz. Structural studies are underway in our laboratory to assess whether *Hp's* O-linked HexNAz is elaborated on to form a higher order glycan.

In our experiments, detection of HexNAz-Phos-FLAG-His_6 and HexNAz-Phos-FLAG adducts by HPLC-Chip/Q-TOFMS analysis was greatly aided by the presence of the peptide moiety, which has favorable chromatographic and ionization characteristics. Thus, MOE facilitated detection of *Hp's* labeled glycans in unanticipated ways. If chromatographic resolution and signal enhancement are generalizable, metabolic glycan labeling coupled to reaction with a readily ionizable tag has the potential to amplify glycan detection in myriad bacteria and overcome challenges associated with the poor resolution and ionizablility of glycans.

ureA and ureB are Biochemically Validated Glycoproteins—Of the putative glycoproteins identified, two of them - urease alpha (ureA) and urease beta (ureB) - are surface-associated colonization and persistance factors (75). We sought to evaluate their glycosylation status, because this has important implications for host-pathogen interactions and the development of novel anti-*Hp* therapeutics. To assess the glycosylation status of ureA and ureB, these proteins were immunopurified from *Hp* treated with 1 mM Ac_4GlcNAz or Ac_4GlcNAc, reacted with Phos-FLAG, and then analyzed by Western blot. Probing immunopurified ureA with anti-ureA revealed the presence of a band at ~27 kDa (Fig. 9*A*), indicating successful enrichment of ureA. Further, probing with anti-FLAG antibody revealed the presence of azide-dependent signal in ureA (Fig. 9*A*). Similar analyses of immunopurified ureB indicated successful enrichment of ureB (Fig. 9*B*) and the presence of azide-dependent signal in ureB (Fig. 9*B*). These results suggest that ureA and ureB are glycosylated with azide-modified glycans.

Based on our glycan mass spectrometry results, we were curious to see whether urease (composed of alpha and beta subunits) is among the glycoproteins modified with O-linked HexNAz in *Hp*. To assess this possibility, urease was biochemically purified from Ac_4GlcNAz-labeled *Hp.* Purified urease (supplemental Fig. S7) was reacted with Phos-FLAG and rinsed extensively to remove residual Phos-FLAG. *Hp's* labeled urease was analyzed by HPLC-Chip/Q-TOFMS before and after being subjected to beta-elimination to cleave O-linked glycans. Although analysis of the pre-beta-eliminated sample showed no evidence for product **2** (supplemental Fig. S8*A*), LC-nanoESI analysis of the beta-eliminated sample showed a chromatographic peak appearing at the retention time characteristic of synthetic HexNAz-Phos-FLAG (**2**) (supplemental Fig. S8*B*) and producing a mass spectrum (supplemental Fig. S8*C*) with *m/z* values consistent with the mass of product **2** ($M = 1578.54$ Da, mass measurement error < 5 ppm; supplemental Table S2). Furthermore, when the $[M+2H]^{2+}$ ion at *m/z* 790.28 from putative product **2** was subjected to MS/MS analysis, the measured mass spectrum showed excellent agreement with the synthetic standard (supplemental Figs. S8*D* and S8*E*; supplemental Table S7). These results suggest that urease is glycosylated with O-linked HexNAz and have implications for the development of urease-based therapeutic strategies to prevent and treat *Hp* infection. Taken together, these data validate urease as a glycosylated protein and indicate that MOE is a robust approach to profiling bacterial glycoproteins.

CONCLUDING REMARKS

Glycosylation of *Hp's* flagellin proteins has been directly linked to *Hp's* pathogenesis, suggesting that *Hp's* glycoproteins are potential targets of therapeutic intervention. Despite their potential therapeutic importance, only the two flagellin glycoproteins have been firmly characterized in *Hp* (7). Mounting evidence from our lab (76) and others (15) suggests protein glycosylation is common in *Hp*. Employing experimental approaches to selectively study and identify *Hp's* glycoproteins should facilitate the discovery of new therapeutic targets and shed light on the role of protein glycosylation in *Hp's* physiology. In this article, we investigated the production of glycosylated proteins in *Hp* using a combination of metabolic glycan labeling and mass spectrometry analysis. Metabolic labeling of glycans with unnatural sugars and subsequent enrichment has been applied to identify glycoproteins in eukaryotic systems (29, 56, 77). Though metabolic glycan labeling has been used to study bacterial glycolipids (78), to our knowledge this study is the first to demonstrate the applicability of this chemical approach to bacterial glycoproteomics (79).

Metabolic incorporation of azide-labeled sugars, followed by reaction with phosphine probes via Staudinger ligation, facilitates detection, visualization, and enrichment of cellular glycoproteins (29). We began our studies by assessing the subcellular distribution of *Hp's* azide-labeled glycoproteins to gain insight into their potential functions. Western blot data indicated that glycoproteins are present throughout *Hp,* including on the cell surface, embedded in both the inner and outer membranes, and within the periplasm and cytoplasm.

The observed subcellular distribution suggests that protein glycosylation is abundant, widespread, and critical for *Hp's* physiology. Broadly, glycans modulate protein function in one of two ways: by providing epitopes for binding and recognition by other proteins (80, 81), and by stabilizing proteins (*e.g.* folding, solubility, protease susceptibility) (82–84). *Hp's* intracellular glycans are likely involved in stabilizing proteins, whereas those exposed to the extracellular milieu (surface and secreted glycoproteins) may stabilize proteins as well as modulate interactions with the host.

Based on our results, the subcellular distribution of glycoproteins in *Hp* appears to be broader than that observed in other bacteria to date. For example, in *C. jejuni* (85, 86), *Neisseria sp.* (20, 87), and *Bacteroides sp.* (27)., general protein glycosylation systems are present within the periplasm and modify proteins that traffic through this compartment (*e.g.* periplasmic, membrane-bound, and secreted proteins). In *Actinobacillus*, in contrast, there is a general protein glycosylation system within the cytosol that modifies cytosolic proteins (88). Finally, in *Haemophilus influenzae*, monosaccharides are added to proteins in the cytoplasm and then transported to the cell envelope (89). Future research will have to be conducted to confirm the localization of our candidate glycoproteins. If the observed, broad distribution is confirmed by follow-up experiments, this result suggests that *Hp* may have two general protein glycosylation systems (one within the cytosol and one within the periplasm) or an exclusively cytoplasmic glycosylation pathway that modifies proteins before sorting. Based on predictions by the carbohydrate-active enzymes (CAZy) database (90), there are 22 glycosyltransferases within *Hp* strain 26695, and half of these have unknown functions. Therefore, there are numerous glycosyltransferases within *Hp* that may be responsible for the synthesis of the observed glycoproteins. Characterizing which of these genes are responsible for carbohydrate assembly pathways is challenging because, in contrast to most bacteria, the genes for *Hp* carbohydrate biosynthesis pathways (*e.g.* O-antigen biosynthesis) are spread throughout the chromosome (91). Efforts are underway in our laboratory to identify the glycosyltransferases responsible for protein glycosylation in *Hp*.

We were excited to observe that a subset of *Hp's* azide-modified glycoproteins are present on the cell surface, as these surface-exposed glycoproteins and the glycans that modify them are potential therapeutic targets. For example, these glycosylated proteins could provide the basis for a carbohydrate-based vaccine (14). Unusual bacterial monosaccharides, such as di-*N*-acetylbacillosamine produced by *Neisseria sp.* (92), can themselves be immunogenic. These glycans, in the context of a glycosylated protein, may serve as particularly effective conjugate vaccines. Alternatively, glycosylation pathways could be targeted with small molecule inhibitors to interrupt interactions between *Hp* and its host. Finally, our demonstration that azide-labeled glycoproteins are present on the cell surface suggests that azide-covered *Hp* could be targeted with therapeutic phosphines. Covalent modification with properly designed phosphines could disrupt glycan function or render *Hp* innocuous within the host (93). Indeed, Kaewsapsak *et al.* designed phosphine therapeutics conjugated to the immune stimulant 2,4-dinitrophenyl to selectively kill *Hp* covered with azide-labeled glycans (55).

Staudinger ligation with Phos-FLAG-His_6 allowed for selective and efficient isolation of *Hp's* azide-labeled glycoproteins. MudPIT analyses of enriched azide-labeled proteins and a mock-enriched control enabled the unambiguous identification of 125 putative glycoproteins. We identified a large number of glycoprotein hits in this study relative to the small number (nine) identified in Creuzenet and coworkers' recent periodic acid/hydrazide-based screen of *Hp's* glycoproteins (15). We attribute much of this difference to the high incidence of background signal associated with periodic acid/hydrazide chemistry (15), rendering the identification of glycoprotein signal relative to nonglycoprotein noise difficult. In contrast, metabolic glycan labeling has a very high signal-to-noise ratio, enabling the facile detection and selective enrichment of glycoproteins. Thus, metabolic glycan labeling offers a complementary and efficient alternative to existing approaches for bacterial glycoprotein detection and identification.

In our studies, we detected Staudinger ligation-HexNAz adducts in beta-eliminated glycan samples from enriched azide-labeled glycoproteins, azide-labeled *Hp*, and purified azide-labeled urease. The exclusive presence of these adducts in samples from azide-labeled *Hp* provides evidence that the Staudinger ligation is exquisitely selective for azide-bearing molecules. Moreover, these results indicate that one metabolic fate of Ac_4GlcNAz in *Hp* is conversion to HexNAz. The addition of the Phos-FLAG moiety onto the HexNAz epitope greatly enhanced the chromatographic resolution, ionization, and detection of this glycan species. Optimization of the phosphine probe to contain isotopic tags or characteristic fragment ions will further facilitate glycoprotein identification and glycan characterization. Thus, MOE is a robust approach for glycoprotein discovery.

A large number of proteins with known links to colonization, persistence and virulence were identified as glycoprotein hits (Table I), and two of these colonization factors, ureA and ureB, were biochemically validated as glycoproteins in our studies. This finding is consistent both with previous studies indicating that flagellin glycosylation is important to *Hp's* ability to survive within the host (94), and more broadly, with the known links between protein glycosylation and virulence in a number of medically significant bacterial pathogens (13, 14). Further, this observed link underscores the importance of *Hp's* glycoproteins as potential therapeutic targets.

The discovery that the urease subunits are glycosylated has potential implications for the development of urease-based therapeutic strategies to prevent and treat *Hp* infection. Vaccines based on recombinant urease have had some success

in preclinical (95, 96) and clinical studies (97, 98). However, these vaccines are based on recombinant proteins produced in *Escherichia coli* or *Salmonella* in their nonglycosylated state. New vaccines based on *Hp's* natively glycosylated urease have the potential to stimulate a robust immune response and thus protect patients more effectively than those based on recombinant urease.

In summary, we were able to use metabolic glycan labeling to enrich and identify 125 putative *Hp* glycoproteins. Further, we validated the glycosylation status of two of these hits. Finally, we characterized the resulting glycan adducts. These results reveal that glycosylated proteins are abundant in *Hp,* indicate that *Hp*'s glycoproteins have a wide range of functions, and reveal new links between *Hp's* pathogenesis and glycosylation. Armed with this information, the stage is set for the development of glycosylation-based therapeutic strategies to eradicate *Hp* infection. In particular, this work opens the door to phosphine-based therapeutics that target *Hp's* azide-labeled glycans.

Broadly, this work validates MOE as a viable approach to discover and characterize bacterial glycoproteins. In addition to enabling the discovery of *Hp*'s glycoproteins, MOE has the potential to facilitate the study of glycoproteins in other bacteria. Indeed, reports of metabolic labeling of glycoproteins with azide-modified pseudaminic acid in *C. jejuni* (54) and with alkyne-bearing fucosylated sugars in *Bacteroidales sp.* (99) suggest that this glycoproteomic approach should transfer readily to these organisms. The ability to use a general metabolic precursor rather than a specific glycan-binding reagent will facilitate the detection and discovery of glycoproteins in a broad range of bacteria. Thus, MOE will propel bacterial glycoproteomics forward and, in the process, unveil novel therapeutic targets.

Acknowledgments—We thank J. Precher (UC Irvine), M. Amieva (Stanford), B. Gorske (Bowdoin), and M. Feldman (U Alberta) for helpful conversations. L. Kohlstaedt (UC Berkeley Proteomics Facility) performed MudPIT analysis. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences, the National Institutes of Health, or the National Science Foundation.

* K.C. was supported by James Stacy Coles and Kibbe Science Research Fellowships. S.A.L. was supported by the Arnold and Mabel Beckman Foundation and by grants from the National Center for Research Resources (5P20RR016463) and the National Institute of General Medical Sciences (8 P20 GM103423) from the National Institutes of Health. D.H.D. acknowledges support from the NIH (award R15GM093867), Research Corporation, and the Camille and Henry Dreyfus Foundation. E.A.S. and D.H.D. acknowledge support from the NSF (award MRI 1126657).

[S] This article contains supplemental Figs. S1 to S8 and Tables S1 to S7.

§ To whom correspondence should be addressed: Department of Chemistry and Biochemistry, Bowdoin College, Brunswick, ME 04011. Tel.: 207-798-4326; Fax: 207-725-3017. E-mail: ddube@bowdoin.edu.

¶ These authors contributed equally to this work.

REFERENCES

1. Graham, D., Malaty, H., Evans, D., Evans, D., Kelin, P., and Adam, E. (1991) Epidemiology of *Helicobacter pylori* in an asymptomatic population in the United States. Effect of age, race, and socioeconomic status. *Gastroenterology* **100,** 1495–1501
2. Marshall, B. J. (1994) *Helicobacter pylori. Am. J. Gastroenterol.* **89,** S116–S128
3. Suerbaum, S., and Michetti, P. (2002) *Helicobacter pylori* infection. *N. Eng. J. Med.* **347,** 1175–1186
4. Megraud, F., and Marshall, B. (2000) How to treat *Helicobacter pylori.* First-line, second-line, and future therapies. *Am. J. Gastroenterol.* **29,** 759–773
5. Ottemann, K. M., and Lowenthal, A. C. (2002) *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infect. Immun.* **70,** 1984–1990
6. McGee, D. J., Langford, M. L., Watson, E. L., Carter, J. E., Chen, Y. T., and Ottemann, K. M. (2005) Colonization and inflammation deficiencies in Mongolian gerbils infected by *Helicobacter pylori* chemotaxis mutants. *Infect. Immun.* **73,** 1820–1827
7. Schirm, M., Soo, E. C., Aubry, A. J., Austin, J., Thibault, P., and Logan, S. M. (2003) Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori. Mol. Microbiol.* **48,** 1579–1592
8. Goon, S., Kelly, J. F., Logan, S. M., Ewing, C. P., and Guerry, P. (2003) Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized via the Cj1293 gene. *Mol. Microbiol.* **50,** 659–671
9. Knirel, Y. A., Vinogradov, E. V., Shashkov, A. S., Dmitriev, B. A., Kochetkov, N. K., Stanislavsky, E. S., and Mashilova, G. M. (1986) Somatic antigens of *Pseudomonas aeruginosa. Eur. J. Biochem.* **157,** 129–138
10. Josenhans, C., Vossebein, L., Friedrich, S., and Suerbaum, S. (2002) The neuA/flmD gene cluster of *Helicobacter pylori* is involved in flagellar biosynthesis and flagellin glycosylation. *FEMS Microbiol. Lett.* **210,** 165–172
11. Schoenhofen, I. C., Lunin, V. V., Julien, J. P., Li, Y., Ajamian, E., Matte, A., Cygler, M., Brisson, J. R., Aubry, A., Logan, S. M., Bhatia, S., Wakarchuk, W. W., and Young, N. M. (2006) Structural and functional characterization of PseC, an aminotransferase involved in the biosynthesis of pseudaminic acid, an essential flagellar modification in *Helicobacter pylori. J. Biol. Chem.* **281,** 8907–8916
12. Obhi, R. K., and Creuzenet, C. (2005) Biochemical characterization of the *Campylobacter jejuni* Cj1294, a novel UDP-4-keto-6-deoxy-GlcNAc aminotransferase that generates UDP-4-amino-4,6-dideoxy-GalNAc. *J. Biol. Chem.* **280,** 20902–20908
13. Benz, I., and Schmidt, M. A. (2002) Never say never again: protein glycosylation in pathogenic bacteria. *Mol. Microbiol.* **45,** 267–276
14. Dube, D. H., Champasa, K., and Wang, B. (2011) Chemical tools to discover and target bacterial glycoproteins. *Chem. Commun.* **47,** 87–101
15. Hopf, P. S., Ford, R. S., Zebian, N., Merkx-Jacques, A., Vijayakumar, S., Ratnayake, D., Hayworth, J., and Creuzenet, C. (2011) Protein glycosylation in *Helicobacter pylori*: Beyond the flagellins? *PLoS ONE* **6,** e25722
16. Messner, P. (2004) Prokaryotic glycoproteins: unexplored but important. *J. Bacteriol.* **186,** 2517–2519
17. Schmidt, M. A., Riley, L. W., and Benz, I. (2003) Sweet new world: glycoproteins in bacterial pathogens. *Trends Microbiol.* **11,** 554–561
18. Szymanski, C. M., and Wren, B. W. (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat. Rev. Micro.* **3,** 225–237
19. Kelly, J., Jarrell, H., Millar, L., Tessier, L., Fiori, L. M., Lau, P. C., Allan, B., and Szymanski, C. M. (2006) Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J. Bacteriol.* **188,** 2427–2434
20. Vik, A., Aas, F. E., Anonsen, J. H., Bilsborough, S., Schneider, A., Egge-Jacobsen, W., and Koomey, M. (2009) Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae. Proc. Natl. Acad. Sci. U.S.A.* **106,** 4447–4452
21. Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., and Guerry, P. (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni. Mol. Microbiol.* **32,** 1022–1030
22. Balonova, L., Hernychova, L., and Bilkova, Z. (2009) Bioanalytical tools for the discovery of eukaryotic glycoproteins applied to the analysis of bacterial glycoproteins. *Expert Rev. Proteom.* **6,** 75–85
23. Stimson, E., Virji, M., Makepeace, K., Dell, A., Morris, H. R., Payne, G.,

Saunder, J. R., Jennings, M. P., Barker, S., and Panico, M. (1995) Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Mol. Microbiol.* **17,** 1201–1214
24. Young, N. M., Brisson, J. R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., Jarrell, H. C., Cadotte, N., St. Michael, F., Aberg, E., and Szymanski, C. M. (2002) Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* **277,** 42530–42539
25. Schirm, M., Schoenhofen, I. C., Logan, S. M., Waldron, K. C., and Thibault, P. (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Anal. Chem.* **77,** 7774–7782
26. Gonzalez-Zamnorano, M., Mendoza-Hernandez, G., Xolalpa, W., Parada, C., Vallecillo, A. J., Bigi, F., and Espitia, C. (2009) *Mycobacterium tuberculosis* glycoproteomics based on ConA-lectin affinity capture of mannosylated proteins. *J. Proteom. Res.* **8,** 721–733
27. Fletcher, C. M., Coyne, M. J., Villa, O. F., Chatzidaki-Livanis, M., and Comstock, L. E. (2009) A general O-glycosylation system important to the physiology of a major human intestinal symbiont. *Cell* **137,** 321–331
28. Dube, D. H., and Bertozzi, C. R. (2003) Metabolic oligosaccharide engineering as a tool for glycobiology. *Curr. Opin. Chem. Biol.* **7,** 616–625
29. Laughlin, S. T., and Bertozzi, C. R. (2007) Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nat. Prot.* **2,** 2930–2944
30. Keppler, O. T., Horstkorte, R., Pawlita, M., Schmidt, C., and Reutter, W. (2001) Biochemical engineering of the N-acyl side chain of sialic acid: biological implications. *Glycobiology.* **11,** 11R–18R
31. Koenigs, M. B., Richardson, E. A., and Dube, D. H. (2009) Metabolic profiling of *Helicobacter pylori* glycosylation. *Mol. Biosyst.* **5,** 909–912
32. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A., and Bertozzi, C. R. (2003) A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 9116–9121
33. Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction. *Science* **287,** 2007–2010
34. Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 19–24
35. Hoshino, H., Tsuchida, A., Kametani, K., Mori, M., Nishizawa, T., Suzuki, T., Nakamura, H., Lee, H., Ito, Y., Kobayashi, M., Masumoto, J., Fujita, M., Fukuda, M., and Nakayama, J. (2011) Membrane-associated activation of cholesterol alpha-glucosyltransferase, an enzyme responsible for biosynthesis of cholesteryl-alpha-D-glucopyranoside in *Helicobacter pylori* critical for its survival. *J. Histochem. Cytochem.* **59,** 98–105
36. Bumann, D., Aksu, S., Wendland, M., Janek, K., Zimny-Arndt, U., Sabarth, N., Meyer, T. F., and Jungblut, P. R. (2002) Proteome analysis of secreted proteins of the gastric pathogen. *Helicobacter pylori*. *Infect. Immun.* **70,** 3396–3403
37. Utt, M., Nilsson, I., Ljungh, A., and Wadstrom, T. (2002) Identification of novel immunogenic proteins of *Helicobacter pylori* by proteome technology. *J. Immunol. Meth.* **259,** 1–10
38. Baik, S. C., Kim, K. M., Song, S. M., Kim, D. S., Jun, J. S., Lee, S. G., Song, J. Y., Park, J. U., Kang, H. L., Lee, W. K., Cho, M. J., Youn, H. S., Ko, G. H., and Rhee, K. H. (2004) Proteomic analysis of the sarcosine-insoluble outer membrane fraction of *Helicobacter pylori* strain 26695. *J. Bacteriol.* **186,** 949–955
39. Laughlin, S. T., Baskin, J. M., Amacher, S. L., and Bertozzi, C. R. (2008) In vivo imaging of membrane-associated glycans in developing zebrafish. *Science* **320,** 664–667
40. Washburn, M. P., Wolters, D., and Yates, J. R. 3rd. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotech.* **19,** 242–247
41. Eng, J. K., McCormack, A. L., and Yates, J. R. 3rd. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5,** 976–989
42. Tabb, D. L., McDonald, W. H., and Yates, J. R. 3rd. (2002) DTASelect and contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. *J. Prot. Res.* **1,** 21–26
43. Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., Foster, L. J., and Brinkman, F. S. L. (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26,** 1608–1615
44. Hu, L. T., and Mobley, H. L. (1990) Purification and N-terminal analysis of urease from *Helicobacter pylori*. *Infect. Immun.* **58,** 992–998
45. Hu, L. T., Foxall, P. A., Russell, R., and Mobley, H. L. (1992) Purification of recombinant *Helicobacter pylori* urease apoenzymes encoded by ureA and ureB. *Infect. Immun.* **60,** 2657–2666
46. Rokita, E., Makristathis, A., Hirschl, A. M., and Rotter, M. L. (2000) Purification of surface-associated urease from *Helicobacter pylori*. *J. Chromatography B* **737,** 203–212
47. Staples, G. O., Shi, X., and Zaia, J. (2010) Extended N-sulfated domains reside at the nonreducing end of heparan sulfate chains. *J. Biol. Chem.* **285,** 18336–18343
48. Yang, Y. Y., Ascano, J. M., and Hang, H. C. (2010) Bioorthogonal chemical reporters for monitoring protein acetylation. *J. Am. Chem. Soc.* **132,** 3640–3641
49. Yang, Y. Y., Yu-Ying, Y., Grammel, M., Markus, G., Hang, H. C., and Howard, H. C. (2011) Identification of lysine acetyltransferase p300 substrates using 4-pentynoyl-coenzyme A and bioorthogonal proteomics. *Bioorg. Med. Chem. Lett.* **21,** 4976–4979
50. Luo, C., Wang, X., Long, J., and Liu, J. (2006) An NADH-tetrazolium-coupled sensitive assay for malate dehydrogenase in mitochondria and crude tissue homogenates. *J. Biochem. Biophys. Methods* **68,** 101–111
51. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems. *Nat. Chem. Biol.* **1,** 13–21
52. Lin, F. L., Hoyt, H. M., van Halbeek, H., Bergman, R. G., and Bertozzi, C. R. (2005) Mechanistic investigation of the Staudinger ligation. *J. Am. Chem. Soc.* **127,** 2686–2695
53. Verez-Bencomo, V., Fernández-Santana, V., Hardy, E., Toledo, M. E., Rodriguez, M. C., Heynngnezz, L., Rodriguez, A., Baly, A., Herrera, L., Izquierdo, M., Villar, A., Valdés, Y., Cosme, K., Deler, M. L., Montane, M., Garcia, E., Ramos, A., Aguilar, A., Medina, E., Torano, G., Sosa, I., Hernandez, I., Martinez, R., Muzachio, A., Carmenates, A., Costa, L., Cardoso, F., Campa, C., Diaz, M., and Roy, R. (2004) A synthetic conjugate polysaccharide vaccine against *Haemophilus influenzae* type b. *Science* **305,** 522–525
54. Liu, F., Aubry, A. J., Schoenhofen, I. C., Logan, S. M., and Tanner, M. E. (2009) The engineering of bacteria bearing azido-pseudaminic acid-modified flagella. *ChemBioChem* **10,** 1317–1320
55. Kaewsapsak, P., Esonu, O., and Dube, D. H. (2013) Recruiting the host's immune system to target *Helicobacter pylori's* surface glycans. *ChemBioChem* **14,** 721–726
56. Laughlin, S. T., and Bertozzi, C. R. (2009) In vivo imaging of *Caenorhabditis elegans* glycans. *ACS Chem. Biol.* **4,** 1068–1072
57. Young, N. M., Brisson, J. R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., Jarrell, H. C., Cadotte, N., St. Michael, F., Aberg, E., and Szymanski, C. M. (2002) Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* **277,** 42530–42539
58. Scott, N. E., Parker, B. L., Connolly, A. M., Paulech, J., Edwards, A. V. G., Crossett, B., Falconer, L., Kolarich, D., Djordjevic, S. P., Hojrup, P., Packer, N. H., Larsen, M. R., and Cordwell, S. J. (2011) Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS Applied to the N-linked glycoproteome of *Campylobacter jejuni*. *Mol. Cell. Proteom* 10, M000031-MCP201
59. Helenius, A., and Aebi, M. (2001) Intracellular functions of N-linked glycans. *Science* **291,** 2364–2369
60. Waidner, B., Greiner, S., Odenbreit, S., Kavermann, H., Velayudhan, J., Stähler, F., Guhl, J., Bissé, E., van Vliet, A. H. M., Andrews, S. C., Kusters, J. G., Kelly, D. J., Haas, R., Kist, M., and Bereswill, S. (2002) Essential role of ferritin Pfr in *Helicobacter pylori* iron metabolism and gastric colonization. *Infect. Immun.* **70,** 3923–3929
61. Eaton, K. A., Brooks, C. L., Morgan, D. R., and Krakowka, S. (1991) Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect. Immun.* **59,** 2470–2475
62. Harris, A. G., Wilson, J. E., Danon, S. J., Dixon, M. F., Donegan, K., and Hazell, S. L. (2003) Catalase (KatA) and KatA-associated protein (KapA) are essential to persistent colonization in the *Helicobacter pylori* SS1 mouse model. *Microbiology* **149,** 665–672
63. Momynaliev, K. T., Kashin, S. V., Chelysheva, V. V., Selezneva, O. V.,

Demina, I. A., Serebryakova, M. V., Alexeev, D., Ivanisenko, V. A., Aman, E., and Govorun, V. M. (2009) Functional divergence of *Helicobacter pylori* related to early gastric cancer. *J. Proteome Res.* **9,** 254–267
64. Eaton, K. A., Morgan, D. R., and Krakowka, S. (1992) Motility as a factor in the colonisation of gnotobiotic piglets by *Helicobacter pylori. J. Med. Microbiol.* **37,** 123–127
65. Odenbreit, S. (2005) Adherence properties of *Helicobacter pylori*: Impact on pathogenesis and adaptation to the host. *Int. J. Med. Microbiol.* **295,** 317–324
66. Ge, Z., Feng, Y., Dangler, C. A., Xu, S., Taylor, N. S., and Fox, J. G. (2000) Fumarate reductase is essential for *Helicobacter pylori* colonization of the mouse stomach. *Micro. Pathogen.* **29,** 279–287
67. Wang, G., Alamuri, P., and Maier, R. J. (2006) The diverse antioxidant systems of *Helicobacter pylori. Mol. Microbiol.* **61,** 847–860
68. Rohde, M., Püls, J., Buhrdorf, R., Fischer, W., and Haas, R. (2003) A novel sheathed surface organelle of the *Helicobacter pylori* cag type IV secretion system. *Mol. Microbiol.* **49,** 219–234
69. Lin, Y. F., Chen, C. Y., Tsai, M. H., Wu, M. S., Wang, Y. C., Chuang, E. Y., Lin, J. T., Yang, P. C., and Chow, L. P. (2007) Duodenal ulcer-related antigens from *Helicobacter pylori. Mol. Cell. Proteom.* **6,** 1018–1026
70. Kavermann, H., Burns, B. P., Angermuller, K., Odenbreit, S., Fischer, W., Melchers, K., and Haas, R. (2003) Identification and characterization of *Helicobacter pylori* genes essential for gastric colonization. *J. Exp. Med.* **197,** 813–822
71. Costa, A. C., Figueiredo, C., and Touati, E. (2009) Pathogenesis of *Helicobacter pylori i*nfection. *Helicobacter* **14,** 15–20
72. Roepstorff, P. (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biochem. Enviro. Mass Spec.* **11,** 601
73. Linton, D., Allan, E., Karlyshev, A. V., Cronshaw, A. D., and Wren, B. W. (2002) Identification of *N*-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in. *Campylobacter jejuni. Mol. Microbiol.* **43,** 497–508
74. Jervis, A. J., Langdon, R., Hitchen, P. G., Lawson, A. J., Wood, A., Fothergill, J. L., Morris, H. R., Dell, A., Wren, B. W., and Linton, D. (2010) Characterizatin of N-linked protein glycosylation in *Helicobacter pullorum. J. Bacteriol.* **192,** 5228–5236
75. Salama, N. R., Hartung, M. L., and Muller, A. Life in the human stomach: persistence strategies of the bacterial pathogen *Helicobacter pylori. Nat. Rev. Micro.* **11,** 1181–1188
76. Koenigs, M. B., Richardson, E. A., and Dube, D. H. (2009) Metabolic profiling of *Helicobacter pylori* glycosylation. *Mol. Biosyst.* **5,** 909–912
77. Yang, L. F., Nyalwidhe, J. O., Guo, S. Q., Drake, R. R., and Semmes, O. J. (2011) Targeted identification of metastasis-associated cell-surface sialoglycoproteins in prostate cancer. *Mol. Cell. Proteom.* 10, M110.007294
78. Swarts, B. M., Holsclaw, C. M., Jewett, J. C., Alber, M., Fox, D. M., Siegrist, M. S., Leary, J. A., Kalscheuer, R., and Bertozzi, C. R. Probing the *Mycobacterial* trehalome with bioorthogonal chemistry. *J. Am. Chem. Soc.* **134,** 16123–16126
79. Longwell, S. A., and Dube, D. H. Deciphering the bacterial glycocode: recent advances in bacterial glycoproteomics. *Curr. Opin. Chem. Biol.* **17,** 41–48
80. Haltiwanger, R. S., and Lowe, J. B. (2004) Role of glycosylation indevelopment. *Annu. Rev. Biochem.* **73,** 491–537
81. Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Ann. Rev. Biochem.* **73,** 1019–1049
82. Langsford, M. L., Gilkes, N. R., Singh, B., Moser, B., Miller, R. C., Warren, R. A., and Kilburn, D. G. (1987) Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. *FEBS Lett.* **225,** 163–167
83. Meldgaard, M., and Svendsen, I. (1994) Different effects of N-glycosylation on the thermostability of highly nomologous bacterial (1,3–1,4)-beta-glucanases secreted from yeast. *Microbiology* **140,** 159–166
84. Nakatsukasa, K., Okada, S., Umebayashi, K., Fukuda, R., Nishikawa, S., and Endo, T. (2004) Roles of O-mannosylation of aberrant proteins in reduction of the load for endoplasmic reticulum chaperones in yeast. *J. Biol. Chem.* **279,** 49762–49772
85. Kelly, J., Jarrell, H., Millar, L., Tessier, L., Fiori, L. M., Lau, P. C., Allan, B., and Szymanski, C. M. (2006) Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J. Bacteriol.* **188,** 2427–2434
86. Nita-Lazar, M., Wacker, M., Schegg, B., Amber, S., and Aebi, M. (2005) The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. *Glycobiology* **15,** 361–367
87. Hartley, M. D., Morrison, M. J., Aas, F. E., Børud, B., Koomey, M., and Imperiali, B. (2011) Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N′-diacetylbacillosamine. *Biochemistry* **50,** 4936–4958
88. Schwarz, F., Fan, Y. Y., Schubert, M., and Aebi, M. (2011) Cytoplasmic N-glycosyltransferase of *Actinobacillus pleuropneumoniae* is an inverting enzyme and recognizes the NX(S/T) consensus sequence. *J. Biol. Chem.* **286,** 35267–35274
89. Grass, S., Buscher, A. Z., Swords, W. E., Apicella, M. A., Barenkamp, S. J., Ozchlewski, N., and St. Geme, J. W. 3rd (2003) The *Haemophilus influenzae* HMW1 adhesin is glycosylated in a process that requires HMW1C and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. *Mol. Microbiol.* **48,** 737–751
90. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233–D238
91. Hug, I., Couturier, M. R., Rooker, M. M., Taylor, D. E., Stein, M., and Feldman, M. F. (2010) *Helicobacter pylori* lipopolysaccharide is synthesized via a novel pathway with an evolutionary connection to protein N-glycosylation. *PLoS Pathog.* 6, e1000819
92. Børud, B., Viburiene, R., Hartley, M. D., Paulsen, B. S., Egge-Jacobsen, W., Imperiali, B., and Koomey, M. (2011) Genetic and molecular analyses reveal an evolutionary trajectory for glycan synthesis in a bacterial protein glycosylation system. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 9643–9648
93. Dube, D. H. (2012) Metabolic labeling of bacterial glycans with chemical reporters. In: Reid, C. W., Twine, S. M., and Reid, A. N., eds. *Bacterial glycomics: current research, technology and applications*, pp. 229–242, Caister Academic Press, Norfolk, UK
94. Schirm, M., Soo, E. C., Aubry, A. J., Austin, J., Thibault, P., and Logan, S. M. (2003) Structural, genetic and functional characterization of the flagellin glycosylation process in *Helicobacter pylori. Mol. Microbiol.* **48,** 1579–1592
95. Dubois, A., Lee, C. K., Fiala, N., Kleanthous, H., Mehlman, P. T., and Monath, T. (1998) Immunization against natural *Helicobacter pylori* infection in nonhuman primates. *Infect. Immun.* **66,** 4340–4346
96. Londoño-Arcila, P., Freeman, D., Kleanthous, H., O'Dowd, A. M., Lewis, S., Turner, A. K., Rees, E. L., Tibbitts, T. J., Greenwood, J., Monath, T. P., and Darsley, M. J. (2002) Attenuated *Salmonella enterica* serovar typhi expressing urease effectively immunizes mice against *Helicobacter pylori* challenge as part of a heterologous mucosal priming-parenteral boosting vaccination regimen. *Infect. Immun.* **70,** 5096–5106
97. Michetti, P., Kreiss, C., Kotloff, K. L., Porta, N., Blanco, J. L., Bachmann, D., Herranz, M., Saldinger, P. F., Corthesy-Theulaz, I., Losonsky, G., Nichols, R., Simon, J., Stolte, M., Ackerman, S., Monath, T. P., and Blum, A. L. (1999) Oral immunization with urease and *Escherichia coli* heat-labile enterotoxin is safe and immunogenic in *Helicobacter pylori*-infected adults. *Gastroenterology* **116,** 804–812
98. Velin, D., and Michetti, P. (2010) Advances in vaccination against *Helicobacter pylori. Exp. Rev. Gastroenterol. Hepatol.* **4,** 157–166
99. Besanceney-Webler, C., Jiang, H., Wang, W., Baughn, A. D., and Wu, P. (2011) Metabolic labeling of fucosylated glycoproteins in Bacteroidales species. *Bioorg. Med. Chem. Lett.* **21,** 4989–4992