Bacterial genome sequencing has provided a wealth of genetic data. However, the definitive functional characterization of hypothetical open reading frames and novel biosynthetic genes remains challenging. This is particularly true for genes involved in protein glycosylation because the isolation of their glycan moieties is often problematic. We have developed a focused metabolomics approach to define the function of flagellar glycosylation genes in *Campylobacter jejuni* 81–176. A capillary electrophoresis-electrospray mass spectrometry and precursor ion scanning method was used to examine cell lysates of *C. jejuni* 81–176 for sugar nucleotides. Novel nucleotide-activated intermediates of the pseudaminic acid (Pse5NAc7NAc) pathway and its acetamidino derivative (PseAm) were found to accumulate within the metabolome. Therefore, to facilitate the elucidation of unknown intermediates involved in novel biosynthetic pathways, we felt it might be more useful to employ a focused approach to simplify the metabolite pool and target relevant compounds.

Glycosylation of prokaryotic proteins in both N- and O-linkage with novel glycan components is now well established (3–5). *Campylobacter jejuni* is unique in having both a general protein glycosylation pathway that is responsible for the N-linked addition of a heptasaccharide containing N-acetylgalactosamine, glucose, and 2,4-diacetamido-2,4,6-trideoxy-α-D-glucopyranose, which is a bacillosamine derivative found in the N-linked protein glycan. Using this focused metabolomics approach, PseB, pseC, pseE, pseF, and pseI were found to be directly involved in either the biosynthesis of CMP-Pse5NAc7NAc or CMP-Pse5NAc7Am. In contrast, it was shown that pseD, pseG, pseJ, pseK, pseL, and pseP have no role in the CM-Pse5NAc7NAc or CM-Pse5NAc7Am pathways. These results demonstrate the usefulness of this approach for targeting compounds within the bacterial metabolome to assign function to genes, identify metabolomic intermediates, and elucidate novel biosynthetic pathways.

The availability of bacterial genomic sequences has provided unprecedented opportunity for comparative studies. The functional analysis of each respective genome is currently under way and often involves an integrative, multidisciplinary approach that combines bioinformatics, mutagenesis, proteomics, and microarray technologies. Most recently, metabolomics-based analyses, although limited in their number, are becoming established as an additional tool to elucidate gene function (1, 2). Metabolomics is the characterization of all low molecular weight compounds in a defined biological system and differs from classical metabolism studies by its greater breadth and speed of metabolite analysis. Although recent advances in mass spectrometry (MS),2 such as the ultimate resolving capability of the Fourier-transform ion cyclotron resonance MS, and also in nuclear magnetic resonance spectroscopy (NMR) technologies, such as the development of higher magnetic field spectrometers and cryogenically cooled probes (cold probes), has greatly facilitated metabolomic analysis (2), such an undertaking still presents numerous technological challenges due to the complexity of the metabolome. Therefore, to facilitate the elucidation of unknown intermediates involved in novel biosynthetic pathways, we felt it might be more useful to employ a focused approach to simplify the metabolite pool and target relevant compounds.

Glycosylation of prokaryotic proteins in both N- and O-linkage with novel glycan components is now well established (3–5).

2 The abbreviations used are: MS, mass spectrometry; Bac, 2,4-diamino-2,4,6-trideoxy-α-D-glucopyranose (also known as bacillosamine); CE, capillary electrophoresis; ESM, electrospray ionization-MS; HILIC, hydrophilic interaction liquid chromatography; HMBC, heteronuclear multiple bond correlation; HSQC, heteronuclear single quantum correlation; NOESY, nuclear Overhauser effect spectroscopy; Pse5NAc7NAc, 3,5,7,9-tetraacyclod-α-D-glucopyranose; Pse5NAc7Am, 5-acetamido-7-acetamidino-3,5,7,9-tetradeoxy-α-D-glucopyranose, which is a bacillosamine derivative found in the N-linked protein glycan. Using this focused metabolomics approach, PseB, pseC, pseE, pseF, and pseI were found to be directly involved in either the biosynthesis of CMP-Pse5NAc7NAc or CMP-Pse5NAc7Am. In contrast, it was shown that pseD, pseG, pseJ, pseK, pseL, and pseP have no role in the CM-Pse5NAc7NAc or CM-Pse5NAc7Am pathways. These results demonstrate the usefulness of this approach for targeting compounds within the bacterial metabolome to assign function to genes, identify metabolomic intermediates, and elucidate novel biosynthetic pathways.

Received for publication, April 19, 2006, and in revised form, May 8, 2006 Published, JBC Papers in Press, May 9, 2006, DOI 10.1074/jbc.M603777200

From the †National Research Council, Institute for Biological Sciences, Ottawa, Ontario K1A 0R6, Canada, ‡National Research Council, Institute for Marine Biosciences, Halifax, Nova Scotia B3H 3Z1, Canada, ¶Department of Chemistry, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada, and §Naval Medical Research Center, Silver Spring, Maryland 20910

**Functional Characterization of the Flagellar Glycosylation Locus in *Campylobacter jejuni* 81–176 Using a Focused Metabolomics Approach**

David J. McNally†, Joseph P. M. Hui‡, Annie J. Aubry†, Kenneth K. K. Mui†, Patricia Guerry‡, Jean-Robert Brisson‡, Susan M. Logan†, and Evelyn C. Soo†

JULY 7, 2006 • VOLUME 281 • NUMBER 27 • JOURNAL OF BIOLOGICAL CHEMISTRY 18489
erogeneous (involving between 25–50 genes depending on the strain) and contains a large number of hypothetical genes as well as genes implicated in glycan biosynthesis (10–12). Although the specific mechanism of the flagellar O-linked process is currently poorly described, the functional characterization of the biosynthetic pathway enzymes responsible for the production of Pse5NAc7NAc and related derivatives has recently received attention (13, 14).

Importantly, Pse5NAc7NAc and PseAm were shown to be essential for flagellar assembly and consequent motility (7, 15–17) and more recently the presence of these novel carbohydrate moieties on the flagellin was shown to play a role in pathogenesis (10, 15, 18). It is believed that agents which interfere with the flagellin glycosylation process and inhibit infection would offer tremendous therapeutic potential even if these agents were not detrimental to the bacteria (19, 20). Thus, the detailed characterization of the genetic locus as well as the elucidation of flagellar glycan biosynthetic pathways is of great significance. Recently, the entire flagellin glycosylation locus of 26 genes from C. jejuni 81–176 was characterized (10). Mutational analysis revealed that a total of nine genes from the locus were involved in flagellin glycosylation, seven of which resulted in a non motile phenotype (pseB, pseC, pseE, pseF, pseG, pseH, and pseL). In addition, mutation of two other genes (pseA and pseD) resulted in flagellin glycosylated with Pse5NAc7NAc but lacking PseAm. Although this study demonstrated a role for several genes in the biosynthesis or transfer of the Pse5NAc7NAc/PseAm moiety, the precise function of these genes remains to be elucidated.

In earlier work we had developed a focused metabolomics approach to examine sugar-nucleotide metabolites related to pseudaminic acid biosynthesis within the metabolome of Helicobacter pylori and C. jejuni (15, 21). Using CE-ESMS and precursor ion scanning for fragment ions characteristic of CMP, we were able to identify CMP-Pse5NAc7NAc and CMP-PseAm as nucleotide-activated precursors related to the biosynthesis of Pse5NAc7NAc and PseAm in C. jejuni 81–176 (21). In addition, during precursor ion scanning for fragment ions characteristic of the UDP carrier, the accumulation of UDP-linked mono- and diacetamido triideoxyhexose intermediates was observed in two defined mutants, pseF (Cj1311) and pseL (Cj1317), which clearly implicated these two proteins in the Pse5NAc7NAc biosynthetic pathway. In this study we expand this work by including all the mutants described by Guerry et al. (10) to provide new data on the role of these genes in the Pse5NAc7NAc and PseAm pathways. The metabolomes of the C. jejuni 81–176 wild-type strain and isogenic mutants were probed for nucleotide-activated monosaccharides relevant to Pse5NAc7NAc/PseAm biosynthesis using the CE-ESMS and precursor ion scanning approach. Next, nucleotide-activated metabolites of interest were purified from cell lysates using a novel hydrophilic interaction liquid chromatography (HILIC)-MS method. The identities of HILIC-MS-purified metabolites were then determined with NMR at 600 MHz (1H) equipped with a (HILIC)-MS method. The identities of HILIC-MS-purified metabolites were then determined with NMR at 600 MHz (1H) equipped with a (HILIC)-MS method. The identities of HILIC-MS-purified metabolites were then determined with NMR at 600 MHz (1H) equipped with a (HILIC)-MS method. The identities of HILIC-MS-purified metabolites were then determined with NMR at 600 MHz (1H) equipped with a (HILIC)-MS method.
NMR Spectroscopy of Purified Metabolites—Metabolites were lyophilized, resuspended in 200 μl of 99% D₂O (Cambridge Isotopes Laboratories Inc., Andover, MA), and analyzed by NMR spectroscopy. To observe exchangeable N-H protons, all MS, MS/MS, and MS³ acquisitions were performed using the linear ion trap (LIT)-mode operating at a scan rate of 1000 a.m.u./s in negative mode and using a typical fixed LIT fill time of 20 ms. A collision energy in the range of 35–45 eV was used for MS3. Analyst 1.4.1 software (AB/MDS Sciex) was used for data acquisition and processing.

RESULTS

CE-ESMS Analysis of Cell Lysates for Metabolites Relevant to CMP-PseSNac7NAc/PseAm Biosynthesis

Lysates from 15 mutants in the flagellin glycosylation locus of strain 81–176 were compared with the wild-type by CE-ESMS for sugar nucleotide intermediates of the PseSNac7NAc/PseAm biosynthetic pathway. This included seven non-motile mutants, three of which encode enzymes in the Pse biosynthetic pathway (pseF, pseC, and pseD) (13, 14), one that has been tentatively assigned as CMP-PseSNac7NAc synthase (pseE) (15), and two of unknown function (pseG and pseH) (15). Two mutants were included that were motile but produced flagellin modified by PseSNac7NAc/PseAm but lacking PseAm (pseA and pseD), and the remaining six mutants were fully motile with no changes in flagellin glycosylation as detected by isoelectric focusing gels (10). CE-ESMS results indicated that no detectable intermediates of the PseSNac7NAc/PseAm biosynthetic pathway were detected in wild-type 81–176, as previously reported (21) (Fig. 1a; Table 1). The first two enzymatic steps of PseSNac7NAc/PseAm synthesis encoded by pseB and pseC (13), as summarized in Fig. 2. A novel UDP-sugar intermediate was observed in the pseC mutant (m/z 631) (Fig. 1b), and tandem mass spectrometry experiments performed as in previous studies (21) revealed an oxonium ion corresponding to the sugar moiety at m/z 229 (data not shown), but no intermediates were observed in the pseB mutant. Mutants in the genes encoding the two remaining known enzymes in the PseSNac7NAc biosynthetic pathway (PseD and PseF) accumulated UDP-187 intermediates, as did mutants in pseD and pseH, suggesting that they may encode the two missing enzymatic steps (Fig. 2). The final non-motile mutant (pseE) did not accumulate any detectable novel intermediates, and in comparable fashion to the parent strain both CMP-PseSNac7NAc and CMP-PseAm were present. This suggests that the pseE gene product is not involved in biosynthesis of either

| FIGURE 1. CE-ESMS and precursor ion scanning for fragments related to CMP (m/z 322)- or UDP (m/z 385)-activated sugars in cell lysates. a, wild-type C. jejuni 81–176. b, isogenic mutant pseC. c, isogenic mutant pseA. CE-MS conditions were an Agilent CE system coupled to a 4000 QTRAP (AB/Sciex) mass spectrometer via a sheath flow interface and operated in the negative ion mode. The buffer was morpholine formate (30 mM; pH 9.0). The injection volume was 30 nl, and the voltage was 25 kV. The sheath buffer was 2-propanol/methanol (2:1, v/v); GS1 = 12; CUR = 18; IS = –5000 V. |
**TABLE 1**

| Strain/Cj number | Gene annotation | Phenotype of mutant | Intracellular sugar-nucleotide |
|------------------|-----------------|---------------------|-------------------------------|
| 81–176           | pseB, pseC      | mot−                | CMP-Pse5NAc7NAc, CMP-PseAm    |
| Cj1293           | pseC            | mot−                | Nothing detected              |
| Cj1294           | pseF            | mot−                | CMP-Pse5NAc7NAc, CMP-PseAm    |
| Cj1311           | pseG            | mot−                | UDP-229, UDP-187              |
| Cj1313           | pseH            | mot−                | UDP-229, UDP-187              |
| Cj1314c          | probable cyclase| mot+ / wt IEF       | CMP-Pse5NAc7NAc, CMP-PseAm    |
| Cj1315c          | Amidotransferase| mot+ / wt IEF       | CMP-Pse5NAc7NAc, CMP-PseAm, UDP-229, UDP-187 |
| Cj1316c          | pseA            | mot+/IEF change/ PseAm−, 486 Da− | CMP-Pse5NAc7NAc, CMP-PseAm |
| Cj1317           | pseI            | mot−                | UDP-229, UDP-187              |
| Cj1333           | mafs3           | mot+ / wt IEF       | CMP-Pse5NAc7NAc, CMP-PseAm    |
| Cj1337           | pseD            | mot+ / IEF change/ PseAm−; 487 Da+ | CMP-Pse5NAc7NAc, CMP-PseAm, UDP-229, UDP-187 |
| Cj1341c          | maf6            | mot+ / wt IEF       | CMP-Pse5NAc7NAc, CMP-PseAm    |
| Cj1342c          | maf7 (37)       | mot+ / wt IEF       | CMP-Pse5NAc7NAc, CMP-PseAm    |

* Cj number refers to the gene showing highest homology in *C. jejuni* 1168 and in which the mutation was made in *C. jejuni* 81–176 (10).

* Gene annotations are from Guerry et al. (10) or Karlyshev et al. (37).

* The phenotype of the mutant was determined on motility agar/isolectric focusing (IEF) gels/MS structural characterization of the flagellin protein (10). A change in IEF indicates a change in glycosylation pattern. wt, wild type.

---

**FIGURE 2. Hypothetical biosynthetic pathways for CMP-pseudaminic acid (I) and UDP-6-deoxy-α-D-GlcNAc (II, UDP-2,4-diacetamido-bac).** The enzymatic activities of PseB, PseC, PgfI and PgfE have been described by Schoenhofen et al. (13) and that of Pse by Chou et al. (14). Broken arrows indicate biosynthetic steps for which the respective enzymatic activities have still to be assigned. The presumed activity of PseF is based on homology to NeuA from the sialic acid biosynthetic pathway.

---

of the CMP-activated sugars but may be involved in either transfer of the glycan moieties to the flagellin protein or at a later stage of the assembly process.

The two mutants that lacked PseAm on flagellin, pseA and pseD, displayed a different metabolic profile. Only CMP-Pse5NAc7NAc was observed in the pseA mutant (Fig. 1c), consistent with a role for the gene product in the biosynthesis of CMP-PseAm from CMP-Pse5NAc7NAc. In contrast, the pseD mutant accumulated both CMP-Pse5NAc7NAc and CMP-PseAm in identical fashion to the parent strain, indicating that PseD is not involved in the biosynthesis of CMP-PseAm. In an analogous fashion to PseE above, the gene product may, however, be involved in glycan transfer to flagellin. Two mutants, Cj1314c and Cj1315c, which had a detectable phenotype in Guerry et al. (10), accumulated, in addition to CMP-Pse5NAc7NAc and CMP-PseAm, UDP-229 and UDP-187 intermediates. This suggests that these two genes play a more subtle role in the Pse5NAc7NAc biosynthetic pathway, which remains to be elucidated. The remaining mutants, CjB1301, Cj1334, Cj1341c, and Cj1342c, all of which were fully motile (10), did not accumulate any novel UDP-linked intermediates, and the metabolomes contained both CMP-Pse5NAc7NAc and CMP-PseAm in a similar fashion to the parent strain and consistent with the phenotypes originally described.

There is a particular interest in defining the precise structure of the CMP-PseAm (compound I) and the UDP-diacetamido-trideoxyhexose (compound II) precursors for a number of reasons. In earlier work, Thibault et al. (7) identified Pse5NAc7NAc and PseAm as the major carbohydrate modifications on the flagellin of two *Campylobacter* species, namely *C. jejuni* 81–176 and *Campylobacter coli* VC167. Although the structure of Pse5NAc7NAc from the flagellin of VC167 was determined by NMR, that of PseAm remains unknown. Subsequent efforts to obtain sufficient amounts of PseAm from purified flagellin protein for NMR analysis have been unsuccessful and so the metabolome provides an opportunity for the structural analysis of the novel PseAm monosaccharide made by *C. jejuni*.

In this study pseC was found to be rather unique in terms of the accumulation of a sugar-nucleotide metabolite suspected to be directly related to Pse5NAc7NAc biosynthesis (UDP-229, UDP-diacetamido-trideoxyhexose, II). Recent functional characterization of PseC had demonstrated that this enzyme is a 4-amino-transferase responsible for producing UDP-6-deoxy-β-1-AltNAc4N using the product of PseB, UDP-4-keto-4,6-dideoxy-β-1-AltNAc (Fig. 2) (13). Because this biosynthetic step would occur before the production of UDP-6-deoxy-β-1-AltNAc4N, the accumulation of a UDP-diacetamido-trideoxyhexose metabolite within pseC prompted further investigation. Accordingly, CMP-PseAm (I) and the unknown UDP-diacetamido-trideoxyhexose metabolite (II) were purified from the cell lysates of the wild-type and pseC, respectively, for unambiguous identification by NMR.

**HILIC-MS Purification of CMP-PseAm (I) and UDP-diacetamido-trideoxyhexose (II) from Cell Lysates**

By using CE-ESMS and morpholine/formate as the separation buffer, a recent study that examined intracellular sugar-nucleotides in *C. jejuni* 81–176 reported impressive resolution of sugar-nucleotides, including separation of CMP-Pse5NAc7NAc and CMP-PseAm (21). CE-ESMS is not amenable to the larger scale isolations of individual metabolites due to the
nanoliter sample volumes that are involved in the technique. LC would, therefore, be more suitable for larger scale purifications of intracellular metabolites. However, because sugar-nucleotides are not well retained on typical reverse-phase LC stationary phases, it was necessary to employ an alternate mode of LC. High performance anion-exchange chromatography and pulsed amperometric detection is an established technique for carbohydrate analysis. In addition, the use of ion-pair liquid chromatography with triethylammonium acetate as the ion-pairing reagent was recently shown to offer good selectivity toward sugar-nucleotides. Although it is possible to use either of these techniques for the purification of the sugar-nucleotide precursors, pure standards for CMP-PseAm (I) and UDP-diacetamido-trideoxyhexose (II) would be necessary for the identification of these novel metabolites in a cell lysate extract. Because these standards are unavailable, mass spectrometry is required to identify novel metabolites. HILIC displays unique selectivity toward polar compounds and employs mobile phase conditions that are amenable to mass spectrometry. Therefore, it was felt more desirable to employ HILIC in this study for the separation and purification of the intracellular sugar-nucleotides.

To optimize HILIC-MS conditions for the isolation of CMP-PseAm (I) and the UDP-diacetamido-trideoxyhexose (II) from cell lysates, initial efforts were focused on separating a mixture of sugar-nucleotide standards representative of those commonly involved in glycoconjugate biosynthesis pathways. A mixture of commercially available sugar-nucleotides was prepared that included CMP-β-D-Neu5Ac, UDP-α-D-Glc, UDP-α-D-GalNAc, GDP-α-D-Man, and ADP-Glc. HILIC separations were performed on the TSKgel Amide-80 (Tosoh Bioscience, Montgomeryville, PA) using mobile phases based on acetonitrile (A) and ammonium acetate (6.5 mM; pH 5.5) (B). Although ADP-Glc and UDP-α-D-GalNAc co-eluted on the HILIC column, base-line resolution of the remaining sugar-nucleotides was achieved (Fig. 3a). Detection of the sugar-nucleotides as their characteristic [M-H]− ions was performed on the 4000 QTRAP mass spectrometer, and confirmation of their identities was achieved by performing tandem mass spectrometry experiments on the individual sugar-nucleotides. Because the molecular weight of the individual sugar-nucleotides in the standard mixture was known, it was possible to readily identify sugar-nucleotides on the basis of their characteristic m/z values. However, when studying novel glycoconjugate biosynthesis pathways, the m/z values of the sugar-nucleotide precursors may not be known. Therefore, it would be highly useful to selectively detect sugar-nucleotides without prior knowledge of the identity. This was achieved in earlier CE-ESMS studies using precursor ion scanning for a fragment ion (s) characteristic of the nucleotide carriers CMP, UDP, ADP, and GDP. As expected, this precursor ion scanning approach with HILIC-MS separation of sugar-nucleotides provided the means to detect selectively for sugar-nucleotides (Fig. 3b and c).

Cell lysate extracts from the wild-type C. jejuni 81–176 and the isogenic mutant pseC were then analyzed specifically for the presence of CMP-PseAm (I) and UDP-diacetamido-trideoxyhexose (II) using the optimized conditions. HILIC-MS analysis of cell lysates obtained from C. jejuni 81–176 and pseC revealed a number of different unknown polar metabolites including CMP-PseAm (I) and UDP-diacetamido-trideoxyhexose (II). Although it was possible to extract ions corresponding to CMP-PseAm and UDP-diacetamido-trideoxyhexose from the total ion chromatograms, a precursor ion scanning approach was employed to ascertain where these metabolites eluted under the selected HILIC-MS conditions. HILIC-MS and precursor ions scanning for fragment ions related to CMP (m/z 322) revealed the presence of metabolites within the cell lysates of C. jejuni 81–176 that were suspected to be CMP-linked (Fig. 4). The mass spectra corresponding to the ions detected at 15.77 and 20.32 min revealed the presence of CMP-Pse5NAc7NAc and CMP-PseAm (I), which was confirmed by tandem mass spectrometry experiments. With regard to HILIC-MS of pseC cell lysates, co-elution of the unknown UDP-diacetamido-trideoxyhexose (II) was observed with another metabolite. It was, therefore, necessary to optimize the gradient elution conditions for isolation of the UDP-diacetamido-trideoxyhexose intermediate. A shallower gradient and incorporation of methanol in the mobile phase was necessary to achieve adequate resolution of the metabolite on the HILIC phase. The optimized conditions for the HILIC-MS analysis of pseC was a linear gradient of 80 A to 70% A in 20 min before returning to 80% A at 25 min, where A was acetonitrile, and B was 90−10 v/v ammonium acetate (6.5 mM; pH 5.5)/methanol. Using the selected mobile phase conditions, large scale isolations of CMP-PseAm (I) and UDP-diacetamido-trideoxyhexose (II) from cell lysates of C. jejuni 81–176 and pseC, respectively, were achieved. The fractions containing CMP-PseAm or UDP-diacetamido-trideoxyhexose were pooled, dried on a SpeedVac concentrator and analyzed by CE-ESMS as a rapid means to determine the purity of the metabolites before NMR analysis.
Metabolomic Analysis of the Flagellar Glycosylation Locus

Structural Elucidation of I and II by NMR Spectroscopy with Cold Probe Technology

CMP-5-acetamido-7-acetamido-3,5,7,9-tetraeoyoxy-1-glycero-α-L-manno-nonulosonic Acid (CMP-Pse5NAc7Am, I)—Based on HILIC-MS analysis that indicated a CMP-PseAm metabolite, a COSY experiment was used initially to target resonances that are unique to Pse, such as H9, H3ax, and H3eq proton signals (not shown). The complete spin system for CMP-Pse5NAc7Am (I) was then meticulously unraveled using one- and two-dimensional NMR experiments (Fig. 5, Table 2). By comparing integrals to an internal TSP (3-(trimethylsilyl)propionic acid-d4 sodium salt) standard of known concentration (27), the amount of CMP-Pse5NAc7NAc present was determined to be ~30 μg (0.25 ms). Because of the low concentration of the sample and the high degree of spectral overlap, the heightened sensitivity of the cold probe was pivotal in characterizing the structure of CMP-Pse5NAc7Am.

The proton spectrum revealed broad resonances for the H3ax proton and distinct sharp signals originating from the CH3 protons of the NAc and Am groups at δH 2.04 and 2.24 ppm, respectively. Signals corresponding to minor unknown contaminants were also observed within the proton spectrum (Fig. 5a). Selective one-dimensional TOCSY experiments were used to assign the protons and to measure J1H,H coupling constants. One-dimensional TOCSY of H3ax revealed signals for H3eq, H4, and H5 (Fig. 5b), whereas that of H7 was used to assign H5, H6, H8, and H9 (Fig. 5c). One-dimensional TOCSY of H9 showed signals for H6, H7, and H8 (Fig. 5d). Interestingly, an additional coupling of 4.5 Hz was observed for H3ax that was difficult to explain. A two-dimensional TOCSY correlation of the NAc NH showed a correlation to H7 (Fig. 5f), indicating the location of the NAc group at C5 and the Am at C7. These findings were corroborated by the results of two-dimensional COSY and TOCSY (90 ms) experiments that revealed J-coupled correlations between the NAc NH and H3ax, H3eq, H4, and H5 as well as between the Am NH and H6, H7, H8, and H9. A long range TOCSY correlation (90 ms) was also observed between the Am NH at δH 9.25 ppm and one of the Am N7H7 protons at δH 8.82 ppm (not shown). A 13C HSQC experiment revealed clear 1H,13C correlations that permitted the complete assignment of carbon chemical shifts (Fig. 5g).

Importantly, the chemical shifts of H7 and C7 differed by ΔδH 0.33 and ΔδC 5.90 ppm, respectively, compared with the monosaccharide of Pse5NAc7NAc (32). These differences are in agreement with the effects of an acetamido group, which was reported to cause a local shift of ΔδH 0.53 ppm and ΔδC 5.87 ppm (29). These results established the location of the acetamido group at C7 of the nonulosonic acid metabolite.

By comparing chemical shifts and coupling constants to those reported for Pse in the literature, the absolute configuration of CMP-Pse5NAc7Am (I) was determined to be D-glycero-L-manno (14, 29, 30–34). The relatively small difference in proton chemical shifts for H3ax and H3eq (ΔδH 0.59 ppm) indicated that I had the α configuration with the C1 carbonyl group occupying the equatorial position. In β-Pse, the difference for H3ax and H3eq tends to be larger and ranges from ΔδH 0.77 to 1.10 ppm (29, 32, 34). The chemical shift for C6 at δC 71.9 ppm also supported that C1 occupies the equatorial position since when C1 is axial, C6 typically resonates downfield at δC 74.2 ppm (34). These findings for the absolute and anomic configurations for I were, therefore, in good agreement with previous studies that examined Pse5NAc7NAc in C. jejuni and C. coli (7, 14). Together, these results provided conclusive evidence showing the identity of compound I to be CMP-Pse5NAc7Am.

UDP-2,4-diacetamido-2,4,6-trideoxy-a-D-glucopyranosyl (UDP-6-deoxy-a-D-GlcNAc4NAc, II)—By comparing chemical shifts (δ ppm) and coupling constants (J1H,H) to those reported in the literature (35, 36), the UDP-diacetamido-trideoxyhexose metabolite (II) was identified as UDP-2,4-diacetamido-2,4,6-trideoxy-a-D-glucopyranosyl (UDP-6-deoxy-a-D-GlcNAc4NAc). The 1H NMR spectrum acquired at 600 MHz (1H) with the cold probe revealed sharp peaks characteristic of UDP-6-deoxy-a-D-GlcNAc4NAc (Fig. 6a, Table 2). Despite the low amount of metabolite present (~10 μg, 0.08 μM), the ratio of signal-to-noise (S/N) was determined to be 100:1 (S/N was determined with respect to the anomeric signal of II at δH 5.49 ppm). Few signals originating from degradation products or contaminants were observed, thereby demonstrating the effectiveness of HILIC chromatography for purifying nucleotide-activated sugar metabolites. In contrast, signals within the 1H NMR spectrum of the same sample acquired at 500 MHz (1H) with a standard 3-mm probe were not as clearly discernible, and the signal/noise ratio was less than half that obtained using the cold probe (signal/noise ratio = 39:1) (not shown). Selective one-dimensional TOCSY experiments were used to assign proton resonances and measure J1H,H coupling constants. One-dimensional-TOCSY of H1 revealed J-coupled peaks representing H2, H3, and H4 (Fig. 6b), whereas that of H6 revealed peaks for H2, H3, H4, and H5 (Fig. 6c). Vicinal coupling constants for H1, H2, H3, and H4 indicated a pyranose sugar with an α-D-gluco configuration. Nuclear Overhauser effects observed for H2 and H4 as well as for H3 and H5 using a two-

FIGURE 4. HILIC-MS and precursor ion scanning for fragment ions related to CMP (m/z 322) of cell lysates from C. jejuni B1-176. a, total ion chromatogram (m/z 100–1000); b, mass spectra at 15.77 min corresponding to CMP-Neu5Ac and 16.25 min corresponding to CMP-Pse5NAc7NAc. c, mass spectrum at 20.32 showing the presence of CMP-PseAm.
The 31P HSQC spectrum revealed 1H31P correlations for H1 and H5 of ribose at 12.5 ppm and for H2 of ribose at 10.6 ppm (Fig. 6d). The 13C HSQC spectrum (160 transients, 128 increments) of 1H9 (80 ms) of the same sample acquired at 500 MHz (1H) with a standard 3-mm probe was substantially noisier, and 13C1H correlations for H1 and H4 were not observed. The UDP-6-deoxy-α-D-GlcNAc4NAc metabolite was concluded to have the D-configuration based on a recent study that showed 2-acetamido-bacillosamine is produced from UDP-α-D-GlcNAc in C. jejuni (13). Together, these results established without ambiguity the identity of compound II as UDP-6-deoxy-α-D-GlcNAc4NAc.

**DISCUSSION**

It is generally accepted that a non-targeted study of all low molecular weight molecules is the ultimate way to gain comprehensive understanding of metabolism. However, our studies on flagellin glycosylation clearly show that there is in fact much to be gained from using a focused metabolomics approach to target a particular metabolic pathway. Utilizing a novel targeted metabolomics approach, we have uncovered three novel biosynthetic gene functions, completed the structural assignment of two unique glycan-associated intracellular metabolites and demonstrated for the first time in vivo a unique interaction between two distinct protein glycosylation pathways.

The approach described here provides a specific functional assignment in the flagellar glycosylation pathway for a number of genes.
from the *Campylobacter* flagellar glycosylation locus. The recent complete genetic characterization of the flagellar glycosylation locus identified all the genes from the locus that were involved in motility or flagellar glycosylation (10). This work complements and extends these findings to provide specific functional assignments for a number of the gene products. We have now been able to demonstrate that in addition to pseB, pseC, pseF, and pseI, the products of three other genes, pseA, pseD, and pseH, are integral members of the pathway. Insertional inactivation of pseD, which has a motility-related phenotype, had no effect on the production of the Pse biosynthetic pathway was interrupted at later stages of the pathway. In this study we expand upon this observation and show that when the Pse biosynthetic pathway is interrupted *in vivo* through inactivation of pseC, the biosynthetic product of Pse is diverted to the 2,4-diacetamido-bacillosamine pathway. 2,4-Diacetamidobacillosamine is an important component of the N-linked heptasaccharide that modifies many proteins in *Campylobacter* (6). It was speculated that this back-epimerization reaction that converts UDP-4-keto-4,6-dideoxy-α-D-GlcNAc to UDP-4-keto-4,6-dideoxy-β-L-altanNAc might be a mechanism that shunts metabolites from the Pse pathway to the Bac pathway. In this study we expand upon this observation and show that when the Pse biosynthetic pathway is interrupted *in vivo* through inactivation of pseC, the biosynthetic product of Pse is diverted to the 2,4-diacetamido-bacillosamine pathway, resulting in the accumulation of the UDP-diacetamido-trideoxyhexose metabolite (UDP-6-deoxy-α-D-GlcNAc4NAc, II) in the metabolome.

In contrast, a different metabolic profile was obtained when the Pse biosynthetic pathway was interrupted at later stages of the pathway. Inactivation of either pseI, which has been shown to encode the Pse synthase (14), and pseF which, based on homology to neuA is likely to be the CMP-Pse5NAc7NAc synthase, produces a metabolite profile reflecting regulation or feedback inhibition of the pathway. It should be noted that the putative substrates of these enzymatic reactions, 6-deoxy-β-L-altanNAc4NAc (pseI) and Pse5NAc7NAc (pseF), are not nucleotide-activated and so neither would be detected using the current precursor ion-scanning analysis. However, metabolomic analysis of pseI and pseF revealed an accumulation of UDP-acetamido-amino-trideoxyhexose.

---

**Metabolomic Analysis of the Flagellar Glycosylation Locus**

**TABLE 2**

NMR chemical shifts (δ ppm) and coupling constants (J Hz) for CMP-5-acetamido-7-acetamidino-3,5,7,9-tetraoxo-L-manno-nonulosonic acid (CMP-Pse5NAc7Am, I) and UDP-2,4-diacetamido-2,4,6-trideoxy-α-D-glucopyranose (UDP-6-deoxy-α-D-GlcNAc4NAc, II)

| Compound | Proton (1H) Chemical shift δH (ppm) | Carbon (13C) Chemical shift δC (ppm) | Coupling constant J (Hz) |
|----------|------------------------------------|-------------------------------------|-------------------------|
| CMP-Pse5NAc7Am (I) | | | |
| H3ax | 1.66 | C1 | 92.1 |
| H3eq | 2.25 | C3 | 36.5 |
| H4 | 4.28 | C5 | 49.2 |
| H5 | 4.33 | C6 | 71.9 |
| H6 | 4.36 | C7 | 58.6 |
| H7 | 3.88 | C8 | 68.1 |
| H9 | 1.27 | C9 | 22.6 |
| NAc CH3 | 2.04 | NAc CH3 | 175.1 |
| UDP-6-deoxy-α-D-GlcNAc4NAc (II) | | | |
| H1 | 5.49 | C1 | 95.2 |
| H2 | 4.05 | C2 | 54.8 |
| H3 | 3.79 | C3 | 69.8 |
| H4 | 3.69 | C4 | 57.8 |
| H5 | 4.06 | C5 | 69.2 |
| H6 | 1.20 | C6 | 18.0 |
| NAc CH3 | 2.03 | NAc CH3 | 23.0 |

*Not resolved. The chemical shifts (δH) for NH and N1H4 protons were determined in 95% H2O (5% D2O) at pH 3.5.*

---

## References

1. Guerry et al. (2010).

2. Recently the precise functional characterization of dehydratase/aminotransferase enzyme pairs PseB and PseC was completed (13). Utilizing recombinant purified enzymes, it was demonstrated *in vitro* that PseB, a C6 dehydratase/C5 epimerase, was capable of synthesizing UDP-4-keto-4,6-dideoxy-α-D-GlcNAc only after considerable amounts of UDP-4-keto-4,6-dideoxy-β-L-altanNAc had accumulated (13). This study had also shown that the UDP-4-keto-4,6-dideoxy-α-D-GlcNAc is an important metabolite within the closely related 2,4-diacetamido-bacillosamine biosynthetic pathway. 2,4-Diacetamidobacillosamine is an important component of the N-linked heptasaccharide that modifies many proteins in *Campylobacter*.

---

## Notes

- **Metabolomic Analysis of the Flagellar Glycosylation Locus**

- **TABLE 2**

- **Metabolomic Analysis of the Flagellar Glycosylation Locus**

- **TABLE 2**

- **Metabolomic Analysis of the Flagellar Glycosylation Locus**

- **Metabolomic Analysis of the Flagellar Glycosylation Locus**
and UDP-diacetamido-trideoxyhexose metabolites in both isogenic mutants. Based on these observations, it appears that interruption of the Pse5NAc7NAc pathway at these later biosynthetic nodes results in the probable accumulation of intermediates from earlier steps of the pathway. Although the precise form of this inhibition remains to be established, two possible alternatives are a typical feedback inhibition that results from the accumulation of biosynthetic intermediates or the perturbation of a higher order biosynthetic enzyme protein complex. The loss of a single biosynthetic enzyme component after mutation of a genetic target may destabilize the complex. We are currently addressing this issue. The accumulation of UDP-acetamido-amino-trideoxyhexose and UDP-diacetamido-trideoxyhexose compounds within the metabolomes of \textit{pseG} and \textit{pseH} provides the first definitive evidence that the proteins encoded by these genes are involved in the Pse5NAc7NAc biosynthetic pathway in \textit{C. jejuni}. An earlier metabolomics study that examined an isogenic mutant of the \textit{pseG} homolog in \textit{H. pylori} (HP0326B) also revealed an accumulation of nucleotide-activated intermediates of identical mass within the metabolome, although the precise structural configuration was not determined (15). Although the biosynthetic function for the gene products of \textit{pseG} and \textit{pseH} are not known, the only steps that do not have a functional assignment within the Pse5NAc7NAc pathway are UDP-6-deoxy-\(\beta\)-L-altNAc4N 4-acetyltransferase and the UDP-6-deoxy-\(\beta\)-L-altNAc4NAc nucleotidase.

Most significantly, this study presents the precise structural characterization of CMP-Pse5NAc7Am found on the flagellar filament of \textit{C. jejuni} 81–176. Based on mass spectroscopy fragmentation patterns, the acetamidino functional group had been tentatively assigned to C5 of Pse in an earlier study (7). We now show that it is actually located at the C7 position. In addition, although we clearly demonstrated the loss of CMP-Pse5NAc7Am in a \textit{pseA} mutant, we did not observe an accumulation of UDP-activated intermediates with an acetamidino group. These observations indicate that for this strain, the branching point of the Pse5NAc7NAc pathway leading to Pse5NAc7Am production likely occurs at a later stage (see Fig. 2). Antigenic differences between the flagellin of \textit{C. coli} and \textit{C. jejuni} have previously been shown to likely reside in the PseAm residues decorating the respective flagellin proteins (16). The findings of this current work now present a means to further explore the basis of the antigenic differences between the two strains.

REFERENCES

1. Weckwerth, W., and Morgenthal, K. (2005) Drug Discov. Today 10, 1551–1558
2. Weckwerth, W. (2003) Annu. Rev. Plant Biol. 54, 669–689
3. Benz, I., and Schmidt, M. A. (2002) Mol. Microbiol. 45, 267–276
4. Upreti, R. K., Kumar, M., and Shankar, V. (2003) Proteomics 3, 363–379
5. Symanski, C. M., and Wren, B. W. (2005) Nat. Rev. Microbiol. 3, 225–237
6. 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 Symanski, C. M. (2002) J. Biol. Chem. 277, 42530–42539
7. Thibault, P., Logan, S. M., Kelly, J. F., Brisson, J. R., Ewing, C. P., Trust, T. J., and Guerry, P. (2001) J. Biol. Chem. 276, 34862–34870
Metabolomic Analysis of the Flagellar Glycosylation Locus

8. Glover, K. J., Weerapan, E., Numao, S., and Imperiali, B. (2005) *Chem. Biol. (Lond.)* **12**, 1311–1315

9. Linton, D., Dorrell, N., Hitchen, P. G., Amber, S., Karlyshev, A. V., Morris, H. R., Dell, A., Valvano, M. A., Aebi, M., and Wren, B. W. (2005) *Mol. Microbiol.** 55**, 1695–1703

10. Guerry, P., Ewing, C. P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P., and Logan, S. (2006) *Mol. Microbiol.** 60**, 299–311

11. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Penn, C. W., Moule, S., Floyd, M. J., Quail, M. A., Rajandream, M. A., Rutherford, K. M., van Vliet, A. H., Whitehead, S., and Barrell, B. G. (2000) *Nature** 403**, 665–668

12. Champion, O. L., Gaunt, M. W., Gundogdu, O., Elmi, A., Witney, A. A., Dorrell, N., and Wren, B. W. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16043–16048

13. Schoenhofen, I. C., McNally, D. J., Vinogradov, E., Whitfield, D., Young, N. M., Dick, S., Wakarchuk, W. W., Brissou, J. R., and Logan, S. M. (2006) *J. Biol. Chem.* **281**, 723–732

14. Guerry, P., Ewing, C. P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P., and Logan, S. (2006) *Mol. Microbiol.** 60**, 299–311

15. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Penn, C. W., Moule, S., Floyd, M. J., Quail, M. A., Rajandream, M. A., Rutherford, K. M., van Vliet, A. H., Whitehead, S., and Barrell, B. G. (2000) *Nature** 403**, 665–668

16. Champion, O. L., Gaunt, M. W., Gundogdu, O., Elmi, A., Witney, A. A., Dorrell, N., and Wren, B. W. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16043–16048

17. Schirm, M., Soo, E. C., Aubry, A. J., Austin, J., Thibault, P., Ewing, C. P., and Guerry, P. (2002) *Mol. Microbiol.** 46**, 587–597

18. Arora, S. K., Neely, A. N., Blair, B., Lory, S., and Ramphal, R. (2005) *Infect. Immun.* **73**, 4395–4398

19. Allbone, L. E. (2002) *Expert Opin. Investig. Drugs** 11**, 1149–1159

20. Kauppi, A. M., Nordfelth, R., Uvell, H., Wolf-Watz, H., and Elofsson, M. (2003) *Chem. Biol. (Lond.)** 10**, 241–249

21. Soo, E. C., Aubry, A. J., Logan, S. M., Guerry, P., Kelly, J. F., Young, N. M., and Thibault, P. (2004) *Anal. Chem.* **76**, 619–626

22. Rabini, J., Maki, M., Savilahti, E. M., Jarvinen, N., Penttila, L., and Renkonen, R. (2001) Glycoconj. J. **18**, 799–805

23. Brisson, J. R., Sue, S. C., Wu, W. G., McManus, G., Nghi, P. T., and Uhrin, D. (2002) in *NMR Spectroscopy of Glycoconjugates* (Jimenez-Barbero, J., and Peters, T., eds) pp. 59–93, Wiley-VCH, Weinheim, Germany

24. Uhrin, D., and Brissou, J. R. (2000) in *NMR in Microbiology: Theory and Applications* (Barbotin, J. N., and Portais, J. C., eds) pp. 165–190, Horizon Scientific Press, Wymondham, UK

25. Wang, X., Li, W., and Rasmussen, H. T. (2005) *J. Chromatogr. A** 1083**, 58–62

26. Guo, Y., and Gaiki, S. (2005) *J. Chromatogr. A** 1074**, 71–80

27. Burton, I. W., Quilliam, M. A., and Walter, J. A. (2005) *Anal. Chem.* **77**, 3123–3131

28. Hermansson, K., Perry, M. B., Altman, E., Brissou, J. R., and Garcia, M. M. (1993) *Eur. J. Biochem.* **212**, 801–809

29. Kenne, L., and Lindberg, B. (1988) *Carbohydr. Res.* **184**, 288–291

30. Edebrink, P., Jansson, P. E., Bogwald, J., and Hoffman, J. (1996) *Carbohydr. Res.* **287**, 225–245

31. Sadowskaya, I., Brissou, J. R., Khieu, N. H., Mutharia, L. M., and Altman, E. (1998) *Eur. J. Biochem.* **253**, 319–327

32. Knirel, Y. A., Kocharova, N. A., Shashkov, A. S., Dmitriev, B. A., Kochetkov, N. K., Stanislavsky, E. S., and Mashilova, G. M. (1987) *Eur. J. Biochem.* **163**, 639–652

33. Knirel, Y. A., Vinogradov, E. V., Shashkov, A. S., Dmitriev, B. A., Kochetkov, N. K., Stanislavsky, E. S., and Mashilova, G. M. (1986) *Eur. J. Biochem.* **157**, 129–138

34. Kondakova, A. N., Perepelov, A. V., Bartodziejska, B., Shashkov, A. S., Senchenkova, S. N., Wykrota, M., Knirel, Y. A., and Rozalski, A. (2001) *Carbohydr. Res.* **333**, 241–249

35. Hanniffy, O. M., Shashkov, A. S., Senchenkova, S. N., Tomshich, S. V., Komandrov, N. A., Romanenko, I. A., Knirel, Y. A., and Savage, A. V. (1998) *Carbohydr. Res.* **307**, 291–298

36. Kocharova, N. A., Perepelov, A. V., Zatonsky, G. V., Shashkov, A. S., Knirel, Y. A., Jansson, P. E., and Weintraub, A. (2001) *Carbohydr. Res.* **330**, 83–92

37. Karlyshev, A. V., Linton, D., Gregson, N. A., and Wren, B. W. (2002) *Microbiology* **148**, 473–480