THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 16, pp. 11186–11192, April 21, 2006

The Group B Streptococcal Sialic Acid *O*-Acetyltransferase Is Encoded by *neuD*, a Conserved Component of Bacterial Sialic Acid Biosynthetic Gene Clusters*

Received for publication, December 27, 2005, and in revised form, February 6, 2006 Published, JBC Papers in Press, February 20, 2006, DOI 10.1074/jbc.M513772200

Amanda L. Lewis[‡§1], **Mary E. Hensler**[¶], **Ajit Varki**[§‖**], **and Victor Nizet**[§¶2]

From the [‡]*Division of Biological Sciences and the Departments of* [¶]*Pediatrics,* [‖]*Medicine, and* **Cellular and Molecular Medicine, *and the* [§]*Glycobiology Research and Training Center, School of Medicine, University of California San Diego, California 92093-0687*

Nearly two dozen microbial pathogens have surface polysaccharides or lipo-oligosaccharides that contain sialic acid (Sia), and several Sia-dependent virulence mechanisms are known to enhance bacterial survival or result in host tissue injury. Some pathogens are also known to *O*-acetylate their Sias, although the role of this modification in pathogenesis remains unclear. We report that *neuD*, a gene located within the Group B *Streptococcus* (GBS) Sia biosynthetic gene cluster, encodes a Sia *O*-acetyltransferase that is itself required for capsular polysaccharide (CPS) sialylation. Homology modeling and site-directed mutagenesis identified Lys-123 as a critical residue for Sia *O*-acetyltransferase activity. Moreover, a single nucleotide polymorphism in *neuD* can determine whether GBS displays a "high" or "low" Sia *O*-acetylation phenotype. Complementation analysis revealed that *Escherichia coli* K1 NeuD also functions as a Sia *O*-acetyltransferase in GBS. In fact, NeuD homologs are commonly found within Sia biosynthetic gene clusters. A bioinformatic approach identified 18 bacterial species with a Sia biosynthetic gene cluster that included *neuD*. Included in this list are the sialylated human pathogens *Legionella pneumophila, Vibrio parahemeolyticus, Pseudomonas aeruginosa*, and *Campylobacter jejuni,* as well as an additional 12 bacterial species never before analyzed for Sia expression. Phylogenetic analysis shows that NeuD homologs of sialylated pathogens share a common evolutionary lineage distinct from the poly-Sia *O*-acetyltransferase of *E. coli* K1. These studies define a molecular genetic approach for the selective elimination of GBS Sia *O*-acetylation without concurrent loss of sialylation, a key to further studies addressing the role(s) of this modification in bacterial virulence.

Sialic acids (Sias)[3] are a family of related nine-carbon acidic sugars (nonulosonic acids) that are found at the outer ends of glycan chains on all vertebrate cells. Over 20 microbial pathogens are known to use surface Sia decoration as a form of molecular mimicry that is often crucial to virulence (1). Several convergent strategies for microbial Sia decoration have been identified, including endogenous biosynthesis, scavenging, or direct transfer of host Sias to their own surface glycans (1). Heavily sialylated bacteria often cause sepsis and meningitis by exhibiting Sia-dependent serum resistance due to deactivation of the alternative complement pathway (2–4). In addition, bacterial Sias can enhance intracellular survival (5), participate in biofilm formation (6), or mask underlying antibody epitopes (7). Antibodies elicited against sialylated bacterial glycolipids can also cross-react with host peripheral nerve axons, leading to the debilitating autoimmune disorder called Guillain-Barré syndrome (8, 9). In short, there are multiple Sia-dependent virulence mechanisms that enhance bacterial survival and/or result in host tissue injury.

Several pathogenic bacteria are known to modify Sia residues by *O*-acetylation (10–18). In other contexts, *O*-acetylation alters immunogenicity of polysaccharide epitopes (19–21), affects Sia-dependent interactions with host leukocyte receptors and complement (22), and can have intracellular consequences for apoptotic regulation (36). Despite the potential of Sia *O*-acetylation to affect pathologic processes, little is known of its role at the host-pathogen interface. Elucidation of the molecular genetic basis of this modification is critical for mechanistic understanding of Sia as a microbial virulence determinant.

Group B *Streptococcus* (GBS) is the leading cause of sepsis and meningitis in human newborn infants and is increasingly associated with invasive infections in adult patient populations (23). The surface capsular polysaccharides (CPS) of all nine identified GBS serotypes invariably display a terminal α2–3-linked *N*-acetylneuraminic acid (Neu5Ac), the predominant Sia on the surface of human cells. Recently, we discovered that these Neu5Ac residues are modified with various levels of *O*-acetylation (17). Due to the chemical lability of *O*-acetyl esters, nearly three decades of previously published investigations on the biochemistry of the GBS CPS addressed neither the presence of Sia *O*-acetylation nor its potential implications for pathogenesis and immunogenicity. We demonstrate the genetic basis for Sia *O*-acetylation in GBS, including a polymorphism associated with variation in *O*-acetylation levels among clinical GBS strains. We show that the GBS Sia *O*-acetyltransferase, NeuD, is itself required for CPS sialylation, a role that can be separated from its *O*-acetyltransferase activity. A bioinformatic approach further revealed that *neuD* homologs are present within many bacterial Sia biosynthetic gene clusters.

EXPERIMENTAL PROCEDURES

Strains and Culture Conditions, DNA Isolation, and PCR—Wild-type (WT) serotype III GBS strain COH1 was the background for all isogenic modifications. GBS were propagated on Todd-Hewitt agar plates or liquid broth at 37 °C without shaking unless specified, with 10 μg/ml erythromycin or 2 μg/ml chloramphenicol selection when indicated. *Escherichia coli* K1 strain ATCC 23511 was propagated in Luria broth (or agar) at 37 °C.

* This work was funded by National Institutes of Health Grants R01HD051796 (to V. N.), R01GM32373 (to A. V.), and P01HL57345 (to A. V.) and an American Heart Association Established Investigator Award (to V. N.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by University of California, San Diego Institutional National Institutes of Health Training Grant 5 T32 DK07202.

[2] To whom correspondence should be addressed. Tel.: 858-534-7408; Fax: 858-534-5611; E-mail: vnizet@ucsd.edu.

[3] The abbreviations used are: Sia, sialic acid; GBS, Group B *Streptococcus*; Neu5Ac, *N*-acetylneuraminic acid; CPS, capsular polysaccharide; GAT, galactoside acetyltransferase; CAT, chloramphenicol acetyltransferase; WT, wild type; HPLC, high pressure liquid chromatography; Leg, legionaminic acid; Pse, pseudaminic acids.

Allelic Replacement Mutagenesis—Precise in-frame allelic replacement of the GBS *neu*D gene with the chloramphenicol acetyltransferase (*cat*) gene was performed as described previously for allelic replacement of the GBS *iagA* gene (24). The primers used to amplify *neu*D upstream and downstream flanking sequences were: NeuDF 5′-gctacatggagaaagtgtag-3′ + NeuDIOF 5′-cattttagattattctcctttaaatca-3′, and NeuDR 5′-caatgtcgtaaacagaccag-3′ + NeuDIOR 5′-taaaggatatgttatgaagccaatttg-3′, respectively. The temperature-sensitive targeting plasmid pNeuD-KO derived by this technique was used to transform WT COH1. Differential temperature and antibiotic selection (17) facilitated precise allelic replacement of *neuD* with *cat* to yield mutant COH1ΔNeuD, as confirmed by direct sequence analysis of chromosomal DNA. Slight modifications of the same protocol were used to generate a derivative of COH1 containing a single base pair mutation in *neuD* changing residue 88 of the encoded protein from phenylalanine to cysteine (F88C). The targeting vector was generated from a NeuDF + NeuDR PCR amplicon of type II GBS strain DK23 bearing the identified *neuD* polymorphism. This vector was used to transform COH1ΔNeuD with temperature shifts and differential antibiotic selection identifying a chloramphenicol-sensitive mutant in which *cat* was precisely replaced with the DK23 *neu*D gene. The resulting mutant (COH1-NeuDF88C) was confirmed by direct sequence analysis to be identical to WT COH1 except bearing the DK23 version of *neuD*.

Complementation Studies—WT GBS COH1 *neuD* was subcloned from the pCR2.1 vector above into streptococcal expression vector pDCerm (25), yielding pGBS-NeuD. *E. coli* K1 strain ATCC23511 *neu*D was amplified using primers 5′-cgtctagacgatgagtaaaaagttaataatatttggtgcggg-3′ and 5′-cgggatcccgttcattccccctaattaatcttgttgg-3′ based on published sequence (26) and cloned into pDCerm to produce pEcK1-NeuD. Sequencing of *neuD* from strain ATCC23511 revealed 99.5% identity with the published *neuD* sequence of *E. coli* K1 RS218 (NCBI accession number AAC43301), differing only at residue 164, where cysteine replaces valine. Each NeuD expression construct was used to transform electrocompetent COH1ΔNeuD, and the resulting erythromycin-resistant transformants were analyzed by 1,2-diamino-4,5-methylene dioxybenzene derivatization and HPLC as below.

Site-directed Mutagenesis—Site-directed mutagenesis of NeuD lysine 123 to alanine in pGBS-NeuD plasmid was accomplished by QuikChange (Stratagene) using the PAGE-purified primer 5′-ggggctttataggttctaaagtggcgctgtttgataacaacg-3′ and its reverse complement followed by DpnI digestion to remove residual template. DNA isolated by ethanol precipitation was transformed into electrocompetent *E. coli* MC1061, and plasmid pGBS-NeuD-K123A was confirmed by sequencing to harbor only the single residue change in the *neuD* coding region.

Release, 1,2-Diamino-4,5-methylene Dioxybenzene Derivatization, and HPLC Analysis and Quantitation of Sialic Acids—Sias were released from phosphate-buffered saline-washed GBS using 2 N acetic acid for 3 h at 80 °C (27). Acid release is known to induce some migration of *O*-acetyl esters from the 7–9-carbon positions (27); this behavior has also been documented for GBS (17). Sias released from GBS were fluorescently labeled with 1,2-diamino-4,5-methylene dioxybenzene and resolved by reverse phase HPLC as described (17). Peaks were assigned based on retention times of standard sialic acids, as well as NaOH lability of *O*-acetyl esters (17).

Homology Modeling—Homology modeling of NeuD was based on the Protein Data Bank coordinates (1KRR) of the *E. coli* galactoside acetyltransferase (GAT) in complex with acetyl-coenzyme A. Modeling was accomplished using the program Modeler 7v7 (Andrej Sali) and an alignment of the two sequences produced by clustalX. The root mean square difference between GAT coordinates and the modeled NeuD structure was determined using the Unix-based program O using residues 83–99 and 105–144 of GAT and residues 91–107 and 111–160 of NeuD. Energy parameters as determined by Modeler were acceptable with no error messages during execution of the script. PROCHECK was used to generate Ramachandran plots of both the experimentally determined (GAT) and the modeled (NeuD) coordinates, each showing two residues near the C terminus in disallowed regions of the plot. Reconstruction of the modeled NeuD trimer complete with the coordinates of acetyl-coenzyme A from 1KRR was performed using the program O.

Bioinformatic Approach for Locating neuD Homologs in Sia Biosynthetic Clusters—We compiled lists of organisms that encode homologs of each of the GBS gene products (NeuA–D) with characterized roles in Sia biosynthesis. BLAST searches on the public domain NCBI data base were performed using sequences of the four GBS genes. Alternatively, the "Clusters of Orthologous Groups of proteins," or the COG system (28), available through The Institute for Genomic Research (TIGR) microbial genome data base, was used to generate lists of organisms with NeuA–D homologs. The four lists were compared and compiled manually to determine which organisms contained all four biosynthetic genes. Although there are a few cases in which the UDP-GlcNAc epimerase (NeuC) is encoded separately, we only included organisms that had a clearly defined clustering of Sia biosynthetic genes (see accession numbers). Note that the term "Sia biosynthetic clusters" is used to describe gene arrangements and their homologies, without necessarily implying the production of Sias.

Phylogenetic Analysis of Bacterial Hexapeptide Repeat Family O-Acetyltransferases—ClustalX was used to produce a multiple alignment comparing NeuD amino acid sequences with Sia *O*-acetyltransferases (NeuO and OatWY) from *E. coli* (AAX11174) and *Neisseria meningitidis* (CAC38867), serine acetyltransferases from *E. coli* (NP_312512) and *Zea mays* (corn; AAN76864), and *E. coli* GAT (NP414876) and MAT (maltose *O*-acetyltransferase, P77791). A Nexus file format of the multiple alignment was used in PAUP 4.0b10 to create an unrooted parsimony tree using the "heuristic search" command with default settings.

RESULTS

NeuD Is Involved in Sia Biosynthesis—We previously speculated that the GBS *neuD* gene may encode a Sia *O*-acetyltransferase (17). This hypothesis was based on its physical location (clustered with other GBS Sia biosynthesis genes) and its homology to hexapeptide repeat acetyltransferases of the CysE/LacA/LpxA/NodL family. Non-polar allelic replacement mutagenesis of *neuD* in WT GBS COH1 resulted in severe attenuation of overall Sia biosynthesis and *O*-acetylation (Fig. 1*A*). The COH1ΔNeuD mutant lost buoyancy and precipitated readily from static culture, consistent with the requirement of sialylation for maximal CPS expression in GBS (29). Both Sia biosynthesis and *O*-acetylation could be restored to the COH1ΔNeuD mutant by complementation in *trans* with pGBS-NeuD (Fig. 1*A*). Higher levels of *O*-acetylation upon overexpression of NeuD on the high copy number complementation vector likely reflect stoichiometric constraints for proper functions that are common among biosynthetic complexes. The simple mutagenesis and complementation analysis proved insufficient to define the hypothesized role of NeuD as an *O*-acetyltransferase. We thus undertook further studies to dissect the potential dual functions of NeuD in Sia biosynthesis and *O*-acetylation.

NeuD Is the GBS Sia O-Acetyltransferase—A strategy of site-directed mutagenesis was employed to discriminate between hypothetical dual functions of NeuD in GBS Sia biosynthesis and Sia *O*-acetylation. Homology modeling of NeuD was used to predict a residue in the pro-

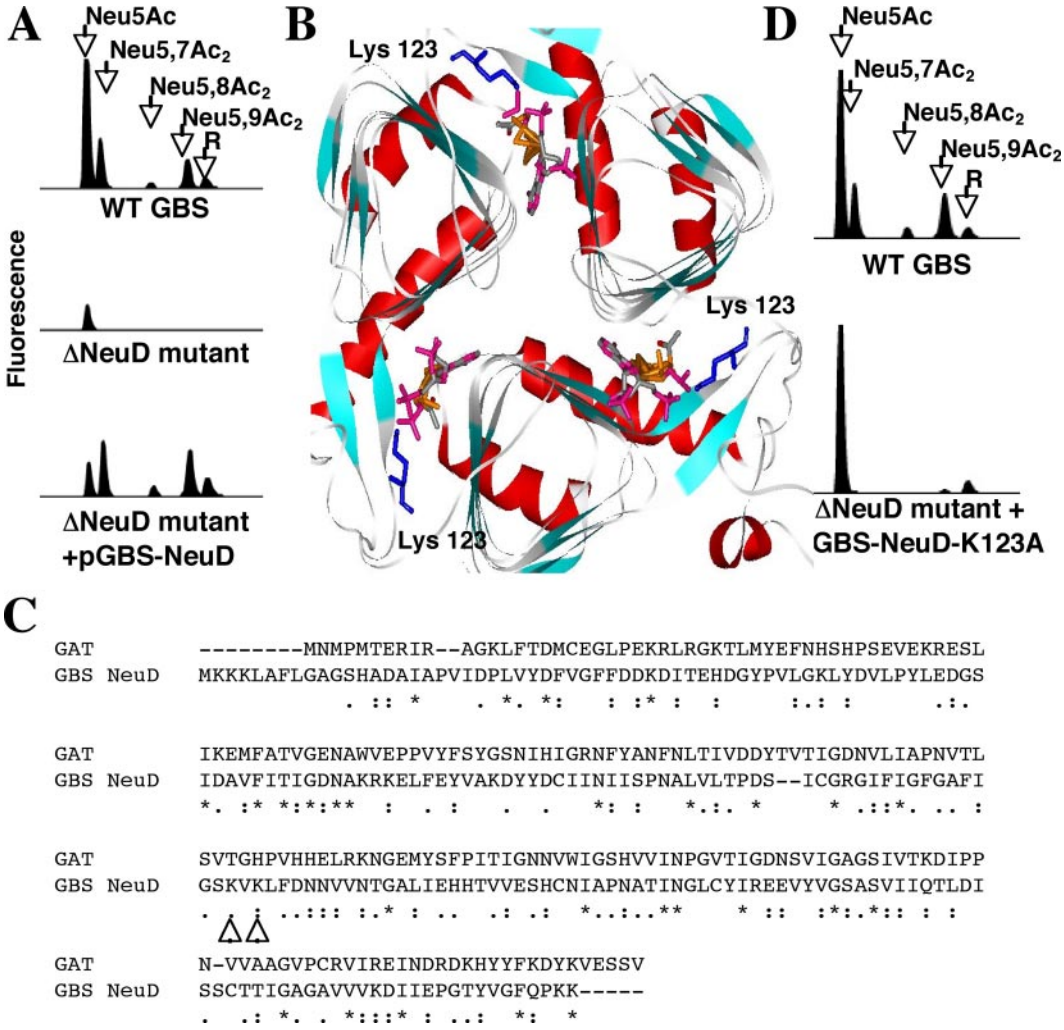

FIGURE 1. **NeuD is the GBS sialic acid *O*-acetyltransferase.** *A*, HPLC resolution of fluorescently labeled Sias isolated from WT GBS strain COH1, *neuD* allelic replacement mutant COH1ΔNeuD (ΔNeuD), and mutant complemented with *neu*D on an expression plasmid. Loss of *neu*D markedly reduces GBS sialylation. *B*, homology model of the NeuD trimer depicting lysine 123 in putative interaction with acetyl-coenzyme A based upon the known structure of *E. coli* galactoside acetyltransferase. *C*, Clustalx alignment of GBS NeuD with the *E. coli* GAT used for homology modeling. *Arrows* indicate lysine residues 121 and 123; the latter aligns with His-115 of GAT. *D*, complementation of the COH1ΔNeuD mutant with the K123A mutant of NeuD restores sialylation but not *O*-acetylation. Neu5Ac is *N*-acetylneuraminic acid, and Neu5,7Ac$_2$, Neu5,8Ac$_2$, and Neu5,9Ac$_2$, refer to the 7-, 8-, and 9-*O*-acetylated versions of Neu5Ac. *R* refers to a reagent peak of unknown identity.

posed GBS NeuD active site that would be specifically involved in *O*-acetyl transfer to Sia residues. The aim was to complement the asialo-COH1ΔNeuD mutant with a modified version of GBS *neuD* that could still perform the presumed Sia biosynthetic function but not the *O*-acetyltransferase function. The GBS NeuD model in Fig. 1*B* is based on the crystal structure of the *E. coli* GAT (Protein Data Bank number 1KRR). Overlay of these molecules reveals a striking similarity in three-dimensional structure, despite only 16% identity between their primary sequences. When compared with the 2.5 Å resolution structure of GAT (30), the GBS NeuD model had a root mean square difference of 1.04 over 55 amino acids involved in the left-handed β-helix domain of this enzyme family. Examination of the GBS NeuD model predicted lysine residue 123 to be involved in the critical acetyl-coenzyme A binding function (Fig. 1*B*). GBS NeuD also contains a lysine residue at position 121 that may be involved in the acetylation reaction. We chose Lys-123 for mutagenesis because it aligns with His-115 of the *E. coli* galactoside acetyltransferase (Fig. 1*C*), which is predicted to abstract a proton from the acceptor hydroxyl group, allowing attack of the resulting carbonyl by acetyl CoA (31). Site-directed mutagenesis was used to change NeuD lysine 123 to alanine. The resulting plasmid (pGBS-NeuD-K123A) was introduced by electroporation into GBS mutant COH1ΔNeuD. As predicted, the NeuD-K123A plasmid restored Sia biosynthesis to the ΔNeuD mutant. In contrast, *O*-acetylation of Neu5Ac was barely detectable (Fig. 1*D*), indicating that NeuD is indeed the GBS Sia *O*-acetyltransferase and that Lys-123 is required for efficient acetyl transfer. These data suggest that despite their ancient relationship, GAT and NeuD likely share a similar mechanism of action.

A neuD Polymorphism Associated with Varied O-Acetylation Levels among GBS Strains—During our initial discovery of Neu5Ac *O*-acetylation in GBS CPS, we noted that clinical isolates expressed either higher (40–55%) or lower (2–15%) levels of this Sia modification (17). To probe the basis for this GBS strain variation in Sia *O*-acetylation, we sequenced the complete *neu*D gene from three strains belonging to each group ("high OAc" *versus* "low OAc"). We found that a single non-synonymous nucleotide change correlated with *O*-acetylation status, wherein high OAc strains (including COH1) had a thymine at position 264, and low OAc strains had a guanine at that position. This polymorphism leads to an amino acid coding change at position 88, with phenylalanine in high OAc strains and cysteine in low OAc strains (Fig. 2*A*). To ascertain whether a functional relationship existed between the identified nucleotide polymorphism and the observed levels of Sia *O*-acetylation, the *neu*D gene of low OAc serotype II strain DK23 was used to precisely

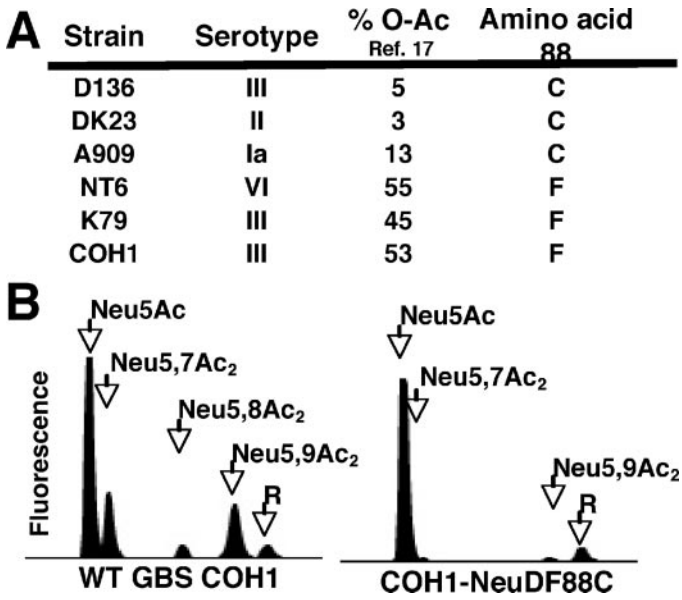

FIGURE 2. **A single nucleotide polymorphism in NeuD associated with variant GBS *O*-acetylation (*O-Ac*) phenotype.** *A*, sequence polymorphism in NeuD from GBS clinical strains exhibiting high or low *O*-acetylation phenotypes. *B*, precise chromosomal replacement of WT COH1 *neuD* with the F88C polymorphic version converts the strain from a high OAc to low OAc phenotype. See the legend for Fig. 1 for a description of the HPLC peaks.

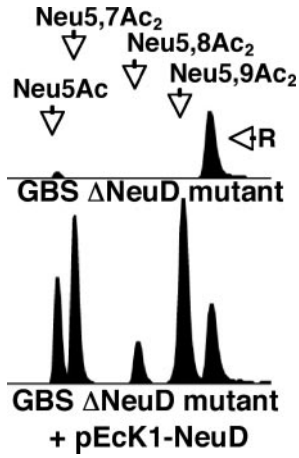

FIGURE 3. ***E. coli* K1 NeuD is a functional *O*-acetyltransferase.** HPLC analysis of fluorescently labeled total sialic acids from the GBS ΔNeuD mutant before and after transformation with a plasmid expressing NeuD from *E. coli* K1. See the legend for Fig. 1 for a description of the HPLC peaks.

replace the native *neu*D gene within the chromosome of high OAc WT strain COH1 (see "Experimental Procedures"). The resulting "knock-in" mutant COH1-NeuDF88C, differing from WT COH1 by just the single nucleotide polymorphism in *neuD*, was found to exhibit a low OAc phenotype (Fig. 2*B*). This result identifies the existence of at least one specific genetic polymorphism that contributes functionally to variation in total *O*-acetylation levels among GBS strains.

E. coli K1 neuD Encodes an Active Sia O-Acetyltransferase—Homologs of the GBS Sia biosynthetic enzymes (NeuA, B, and C) can be identified in a number of other organisms including sialylated pathogens and their human hosts. Like GBS, *E. coli* K1 is a leading cause of neonatal sepsis and meningitis, expresses a surface CPS containing Neu5Ac, and further modifies this Sia by *O*-acetylation (10, 32). *E. coli* K1 also has a homolog of the GBS *neuD* gene alongside the Sia biosynthetic cluster. Earlier studies demonstrated that disruption of the *E. coli* K1 *neu*D gene leads to total loss of Sia expression and that Sia expression could be restored to the resulting mutant by introduction of plasmids expressing either the *E. coli* K1 or the GBS versions of NeuD (33). Discovery of the NeuO poly-Sia *O*-acetyltransferase (34), coupled with an experiment in which the complete *E. coli* K1 Sia gene cluster (including *neuD*) failed to confer CPS Sia *O*-acetylation to a non-K1 *E. coli* host (35), led the investigators to a logical conclusion that a role for NeuD in Sia *O*-acetylation was unlikely for either GBS or *E. coli* K1 (34). Despite these considerations, the current studies allowed us to conclude that NeuD is in fact the GBS Sia *O*-acetyltransferase. Thus, we next asked whether the homologous *E. coli* K1 NeuD also has demonstrable Sia *O*-acetyltransferase activity. Expression vector pNeuD-EcK1, containing the *neu*D gene of *E. coli* K1 strain ATCC23511, was introduced into the asialo-GBS mutant COH1ΔNeuD and was observed to restore both Sia biosynthesis and high levels of *O*-acetylation (Fig. 3). This result indicates that despite only 30% sequence identity between the GBS and *E. coli* K1 NeuD enzymes, they are both able to carry out Sia biosynthetic and *O*-acetyltransferase functions in GBS.

NeuD Homologs Are Broadly Represented among Bacterial Species—Given the orthologous relationship between GBS and *E. coli* K1 *neuD* genes, we explored whether *neuD* is a common component of Sia biosynthetic clusters. Using a bioinformatic approach, we found *neu*D homologs that are physically associated with Sia biosynthetic gene clusters in the genomes of 18 bacterial species (Fig. 4, *A* and *B*), many of which have not to date been analyzed for Sia expression (Fig. 4*B*) Phylogenetic analysis indicates that NeuD homologs of sialylated pathogens (in Fig. 4*A*) have a common evolutionary lineage, distinct from chloramphenicol, maltose, galactoside, and serine acetyltransferases (Fig. 4*C*). The tree also shows that the Sia *O*-acetyltransferases from *E. coli* (NeuO) and *N. meningitidis* (OatWY), which likely both act on polysialic acid, evolved independently from the NeuD subfamily, appearing more closely related to maltose and galactoside acetyltransferases. A second predicted *O*-acetyltransferase of this family in the GBS genome (AAM99894) clusters phylogenetically with the maltose and galactoside acetyltransferases (Fig. 4*C*). Single crossover (Campbell-type) plasmid integrational disruption of this gene in WT GBS strain COH1 resulted in no change in overall sialylation or *O*-acetylation (data not shown), indicating that it is not involved in either of these processes.

DISCUSSION

In the present work, we show that NeuD has dual roles in GBS sialylation and Sia *O*-acetylation that can be distinguished using site-directed mutagenesis. We identified Lys-123 as a critical residue for efficient acetyl transfer and demonstrate that a naturally occurring single residue polymorphism in the *neuD* coding region allows variation in the GBS *O*-acetylation phenotype. Complementation analysis shows that *E. coli* K1 *neuD* also encodes an active Sia *O*-acetyltransferase within the context of the GBS biosynthetic machinery. Indeed, *neuD* homologs are found within Sia biosynthetic gene clusters in at least 18 bacterial species, 12 of which are not currently known to synthesize Sias. On mammalian cells, Sia *O*-acetylation can affect apoptotic regulation (36), susceptibility to alternative complement, tissue tropism, and binding to leukocyte-expressed Sia-binding proteins (22). Due to past issues of chemical lability and genetic intractability, understanding the role(s) of bacterial Sia *O*-acetylation has been very difficult. These studies offer a specific approach to experimentally assess the importance of this modification in bacterial pathogenesis.

Pathogenic strains of *E. coli* and *C. jejuni* that cause bloody diarrhea in humans often express Neu5Ac as part of their lipo-oligosaccharides (15, 37). Although the *E. coli* O104 antigen is known to be *O*-acetylated

A

Organism	Sia Biosynthetic Gene Cluster	Sia type	OAc position	Sia Structures
Group B Streptococcus	B SAG1161 · C 1160 · D 1159 · A 1158	Neu	7,8,9-OAc[17]	Neu
Escherichia coli K1	D AAC43301 · B AAC43302 · A AAA24210 · C Q47400	Neu	7,9-OAc[10]	
Escherichia coli O104	D AAK64367 · B 64368 · A 64369 · C 64370	Neu	7,9-OAc[15]	
Escherichia coli O145	D AAX58757 · B 58758 · A 58759 · C 58760	Neu[18]	?	Leg
Legionella pneumophilia sg-1	A lpg0751 · B 0752 · C 0753 · D 0754	Leg	8-OAc[16]	
Campylobacter jejuni 11168	D cj1321 · B 1327 · C 1328 · A 1331	Neu + Pse	?	Pse
	D cj1123 · B 1141 · C 1142 · A 1143		8-OAc[13,14]	
Pseudomonas aeruginosa (O12)	C AAM27588 · B 27589 · D 27590 · A 27591	Pse	4-OAc[12]	
Vibrio parahaemolyticus	C VP0199 · B 0200 · D 0201 · A 0203	Leg[50]	?	

B

Organism	"NeuA"	"NeuB"	"NeuC"	"NeuD"
Leptospira interrogans **56601**	LA1605	1609	1610	1608
Leptospira interrogans **Copenhageni Fiocruz L1-130**	LIC12177	12173	12172	12174
Vibrio vulnificus **YJ016**	VV0316	312	311	313
Chromobacterium violaceum **ATCC12472**	CV4028	4031	4020	4024
Vibrio fischeri **ES114**	VF0147	143	142	144
Geobacillus kaustophilus **HTA426**	GK3120	3123	3126	3122
Idiomarina loihiensis **L2TR**	IL0554	551	550	552
Synechococcus **sp. WH8102**	SYNW0444	448	449	447
Bradyrhizobium japonicum **USDA110**	blr5996	5992	5991	5993
Bdellovibrio bacteriovorus	Bd1686	1683	1680	1684
Clostridium thermocellum **ATCC27405**	ZP_00060397	60395	60396	60394
Geobacter sulfurreducens	GSU1972	1970	2245	1973
Magnetospirillum magnetotacticum	ZP_00056632	56630	56629	56631
Homo sapiens	Q8NFW8	NP_061156	AAS59259	
Mus musculus	AAH31500	NP_444409	Q91WG8	

C

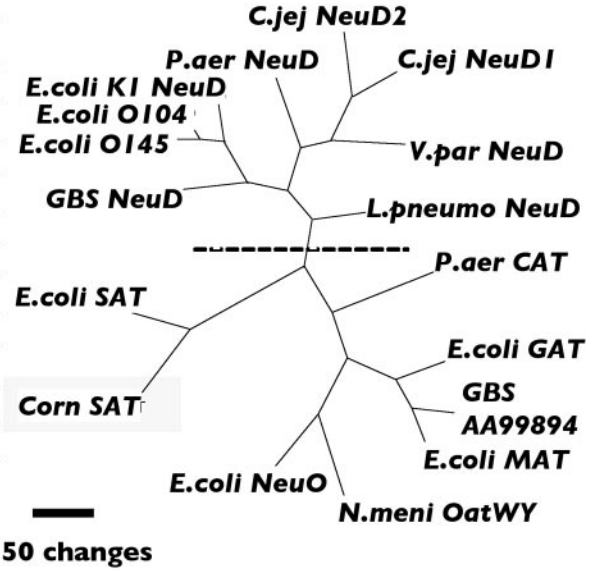

FIGURE 4. **NeuD is conserved among bacterial Sia biosynthetic gene clusters.** *A*, schematic representations of Sia biosynthetic gene clusters of bacteria known to synthesize *N*-acetylneuraminic (*Neu*) acid, Leg, or Pse. Letter designations refer to homologs of known GBS enzymes: NeuC, which catalyzes the epimerization of UDP-GlcNAc to *N*-acetyl mannosamine (49); NeuB, which carries out the condensation of *N*-acetyl mannosamine and phosphoenolpyruvate to form Neu5Ac (50); and NeuA, which activates Neu5Ac with CTP to generate CMP-Neu5Ac (51). Core Sia structures are marked with *arrows* to indicate reported *O*-acetylation (*OAc*) sites. Accession numbers are from NCBI or TIGR. *B*, additional bacterial species possessing Sia biosynthetic gene clusters (NeuA–D) but as yet not known to synthesize Sia. *C*, parsimony-based phylogenetic tree indicating the relationship of putative NeuD orthologs to other hexapeptide acetyltransferases including GAT, maltose *O*-acetyltransferase (*MAT*), CAT, serine acetyltransferase (*SAT*), or polysialic acid acetyltransferases of *E. coli* (NeuO) or *N. meningitides* (*N.meni*) (OatWY). AAM99894 is a second *O*-acetyltransferase of this family in the GBS genome. *C.jej*, *C. jejuni*; *P.aer*, *P. aeruginosa*; *V.par*, *V. parahemeolyticus*; *L.pneumo*, *L. pneumophila*.

(15), this modification has not been reported for *E. coli* O145 or lipooligosaccharides of any *Campylobacter* species. *O*-acetylation of lipooligosaccharide Sia is especially relevant because mimicry of host gangliosides is felt to be a pivotal factor in the autoimmune damage to motor axons that characterizes *Campylobacter*-associated Guillain-Barré syndrome (8). Serum antibodies recognizing *O*-acetyl forms of gangliosides GD1b, GD3, or GT3 have been identified in subsets of patients with this nervous system disorder (38, 39). We identify NeuD

homologs in the Sia biosynthetic clusters of *C. jejuni*, *E. coli* O104, and *E. coli* O145 (Fig. *4*).

Some bacterial agents of pneumonia synthesize nine-carbon acidic sugars that are structurally similar to Neu5Ac using homologous biosynthetic pathways (40). These sugars, known as legionaminic acids (Leg) and pseudaminic acids (Pse) (Fig. *4A*), have been called "Sia-like" molecules (41) and are found in over a dozen bacterial species including *L. pneumophila*, *P. aeruginosa*, and pathogenic *Vibrio spp* (42). Biochemical analysis has shown that Leg and Pse molecules on the surfaces of *L. pneumophila*, *P. aeruginosa,* and *C. jejuni* are modified by *O*-acetylation (12, 14, 16). We now show that like the Neu5Ac biosynthetic clusters, both Leg and Pse biosynthetic clusters also include a *neuD* homolog (Fig. 4*A*), further supporting the evolutionary relationship between these biosynthetic pathways and their inclusion in the larger category of Sias. The presence of NeuD in Leg and Pse biosynthetic clusters predicts a genetic basis for *O*-acetylation of these related sugars.

Unlike the poly-Sia *O*-acetyltransferases of *E. coli* K1 and *N. meningitidis*, we could not reliably align the legionaminic acid *O*-acetyltransferase of *L. pneumophila* to analyze its relationship with other enzymes in this family. Initial studies of a *L. pneumophila* Lag-1 mutant strongly suggested its role as a Leg *O*-acetyltransferase (43). It was later recognized, however, that *L. pneumophila* often expresses Lag-1-independent *O*-acetylation of Leg residues in the LPS core region (44). Thus, the *L. pneumophila neuD* homolog represents a candidate gene for the remaining *O*-acetylation in the *lag-1* mutant. It is theoretically possible that, like *L. pneumophila*, *E. coli* K1 strains express NeuO-independent *O*-acetylation of Sia moieties that are not accessible to specific polyclonal antisera used for *O*-acetyl analysis. Alternatively, NeuD may only be functional as an *O*-acetyltransferase in *E. coli* strains that produce sialylated lipo-oligosaccharide antigens (O104 and O145). If the latter were the case, K1 neuD would be conserved in this biosynthetic gene cluster solely for its role in sialylation. Further studies are needed to determine the actual roles of *neuD* homologs in these bacteria.

Of the 18 bacterial species we identified using our bioinformatic approach, 12 are not currently known to express Sias (Fig. *4B*). Among these are the human pathogens *Vibrio vulnificans,* causing wound infections and septicemia (45), *Chromobacterium violaceum,* an agent of serious invasive disease in pediatric age groups (46), and *Leptospira interrogans*, a leading zoonotic pathogen producing systemic illnesses manifest by hepatic and renal dysfunction (47). For investigators working with each of these species, our identification of Sia gene clusters containing *neuD* may stimulate a search for the presence of Sias with *O*-acetyl modifications that likely play a role in environmental niche ecology and/or interactions with mammalian hosts.

Although we have not precisely defined the duality of the role of NeuD in bacterial Sia biosynthesis, others have presented evidence that the *E. coli* K1 NeuD protein interacts with NeuB, the sialic acid synthetase (48). Our data are consistent with previous speculation that NeuD may play a structural role in stabilization of a Sia biosynthetic complex of enzymes (48), distinct from its *O*-acetyltransferase activity. Further investigations of the genetic basis and functional role of *O*-acetylation in *Legionella*, *Pseudomonas*, *Campylobacter, Leptospira, Vibrio*, etc. may require a similar approach to the present one, involving site-directed active site mutagenesis to discriminate between these dual biosynthetic roles. Some pathogens including *N. meningitidis* may have lost the requirement for NeuD in Sia biosynthesis since a corresponding gene is not easily discernable near the Sia biosynthetic cluster. Notably, unlike the other components of the Sia biosynthesis pathway (NeuA–C), which share homology between humans and bacteria (Fig. *4B*), there are no mammalian homologs of any hexapeptide repeat acetyltransferase, including NeuD. Thus, NeuD may represent a unique bacterial target for rational drug design aimed at eliminating surface expression of sialic acid, a virulence phenotype of several human pathogens.

Acknowledgments—We thank Russell Doolittle for guidance in homology modeling and Warren Lewis for assistance with ChemDraw.

REFERENCES

1. Vimr, E., and Lichtensteiger, C. (2002) *Trends Microbiol.* **10,** 254–257
2. Marques, M. B., Kasper, D. L., Pangburn, M. K., and Wessels, M. R. (1992) *Infect. Immun.* **60,** 3986–3993
3. Vogel, U., Weinberger, A., Frank, R., Muller, A., Kohl, J., Atkinson, J. P., and Frosch, M. (1997) *Infect. Immun.* **65,** 4022–4029
4. Ram, S., Sharma, A. K., Simpson, S. D., Gulati, S., McQuillen, D. P., Pangburn, M. K., and Rice, P. A. (1998) *J. Exp. Med.* **187,** 743–752
5. Charland, N., Kobisch, M., Martineau-Doize, B., Jacques, M., and Gottschalk, M. (1996) *FEMS Immunol. Med. Microbiol.* **14,** 195–203
6. Jurcisek, J., Greiner, L., Watanabe, H., Zaleski, A., Apicella, M. A., and Bakaletz, L. O. (2005) *Infect. Immun.* **73,** 3210–3218
7. Guerry, P., Ewing, C. P., Hickey, T. E., Prendergast, M. M., and Moran, A. P. (2000) *Infect. Immun.* **68,** 6656–6662
8. Hughes, R. A., and Cornblath, D. R. (2005) *Lancet* **366,** 1653–1666
9. Koga, M., Takahashi, M., Masuda, M., Hirata, K., and Yuki, N. (2005) *Neurology* **65,** 1376–1381
10. Orskov, F., Orskov, I., Sutton, A., Schneerson, R., Lin, W., Egan, W., Hoff, G. E., and Robbins, J. B. (1979) *J. Exp. Med.* **149,** 669–685
11. Bhattacharjee, A. K., Jennings, H. J., Kenny, C. P., Martin, A., and Smith, I. C. (1976) *Can. J. Biochem.* **54,** 1–8
12. 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
13. 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
14. 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
15. Gamian, A., Romanowska, E., Ulrich, J., and Defaye, J. (1992) *Carbohydr. Res.* **236,** 195–208
16. Knirel, Y. A., Rietschel, E. T., Marre, R., and Zahringer, U. (1994) *Eur. J. Biochem.* **221,** 239–245
17. Lewis, A. L., Nizet, V., and Varki, A. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 11123–11128
18. Feng, L., Senchenkova, S. N., Tao, J., Shashkov, A. S., Liu, B., Shevelev, S. D., Reeves, P. R., Xu, J., Knirel, Y. A., and Wang, L. (2005) *J. Bacteriol.* **187,** 758–764
19. Berry, D. S., Lynn, F., Lee, C. H., Frasch, C. E., and Bash, M. C. (2002) *Infect. Immun.* **70,** 3707–3713
20. Fattom, A. I., Sarwar, J., Basham, L., Ennifar, S., and Naso, R. (1998) *Infect. Immun.* **66,** 4588–4592
21. McNeely, T. B., Staub, J. M., Rusk, C. M., Blum, M. J., and Donnelly, J. J. (1998) *Infect. Immun.* **66,** 3705–3710
22. Shi, W. X., Chammas, R., Varki, N. M., Powell, L., and Varki, A. (1996) *J. Biol. Chem.* **271,** 31526–31532
23. Edwards, M. S., Nizet, V., and Baker, C. J. (2006) in *Infectious Diseases of the Fetus and Newborn Infant* (Remington, J., Klein, J., Baker, C. J., and Wilson, C. B., eds) 6th Ed., pp. 403–464, Elsevier Saunders, Philadelphia, PA
24. Doran, K. S., Engelson, E. J., Khosravi, A., Maisey, H. C., Fedtke, I., Equils, O., Michelsen, K. S., Arditi, M., Peschel, A., and Nizet, V. (2005) *J. Clin. Investig.* **115,** 2499–2507
25. Jeng, A., Sakota, V., Li, Z., Datta, V., Beall, B., and Nizet, V. (2003) *J. Bacteriol.* **185,** 1208–1217
26. Annunziato, P. W., Wright, L. F., Vann, W. F., and Silver, R. P. (1995) *J. Bacteriol.* **177,** 312–319
27. Varki, A., and Diaz, S. (1984) *Anal. Biochem.* **137,** 236–247
28. Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J., and Natale, D. A. (2003) *BMC Bioinformatics* **4,** 41
29. Chaffin, D. O., Mentele, L. M., and Rubens, C. E. (2005) *J. Bacteriol.* **187,** 4615–4626
30. Wang, X. G., Olsen, L. R., and Roderick, S. L. (2002) *Structure* (*Camb.*) **10,** 581–588
31. Lewendon, A., Ellis, J., and Shaw, W. V. (1995) *J. Biol. Chem.* **270,** 26326–26331
32. Higa, H. H., and Varki, A. (1988) *J. Biol. Chem.* **263,** 8872–8878
33. Daines, D. A., Wright, L. F., Chaffin, D. O., Rubens, C. E., and Silver, R. P. (2000) *FEMS Microbiol. Lett.* **189,** 281–284
34. Deszo, E. L., Steenbergen, S. M., Freedberg, D. I., and Vimr, E. R. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 5564–5569

35. Silver, R. P., Finn, C. W., Vann, W. F., Aaronson, W., Schneerson, R., Kretschmer, P. J., and Garon, C. F. (1981) *Nature* **289,** 696–698
36. Malisan, F., Franchi, L., Tomassini, B., Ventura, N., Condo, I., Rippo, M. R., Rufini, A., Liberati, L., Nachtigall, C., Kniep, B., and Testi, R. (2002) *J. Exp. Med.* **196,** 1535–1541
37. Moran, A. P., Rietschel, E. T., Kosunen, T. U., and Zahringer, U. (1991) *J. Bacteriol.* **173,** 618–626
38. Hitoshi, S., Kusunoki, S., Kon, K., Chiba, A., Waki, H., Ando, S., and Kanazawa, I. (1996) *J. Neuroimmunol.* **66,** 95–101
39. Koga, M., Yuki, N., Ariga, T., and Hirata, K. (1999) *J. Neurol. Sci.* **164,** 50–55
40. Luneberg, E., Zetzmann, N., Alber, D., Knirel, Y. A., Kooistra, O., Zahringer, U., and Frosch, M. (2000) *Int. J. Med. Microbiol.* **290,** 37–49
41. Angata, T., and Varki, A. (2002) *Chem. Rev.* **102,** 439–469
42. Knirel, Y. A., Shashkov, A. S., Tsvetkov, Y. E., Jansson, P. E., and Zahringer, U. (2003) *Adv. Carbohydr. Chem. Biochem.* **58,** 371–417
43. Zou, C. H., Knirel, Y. A., Helbig, J. H., Zahringer, U., and Mintz, C. S. (1999) *J. Bacteriol.* **181,** 4137–4141
44. Kooistra, O., Luneberg, E., Lindner, B., Knirel, Y. A., Frosch, M., and Zahringer, U. (2001) *Biochemistry* **40,** 7630–7640
45. Strom, M. S., and Paranjpye, R. N. (2000) *Microbes Infect.* **2,** 177–188
46. Chattopadhyay, A., Kumar, V., Bhat, N., and Rao, P. (2002) *J. Pediatr. Surg.* **37,** 108–110
47. Bharti, A. R., Nally, J. E., Ricaldi, J. N., Matthias, M. A., Diaz, M. M., Lovett, M. A., Levett, P. N., Gilman, R. H., Willig, M. R., Gotuzzo, E., and Vinetz, J. M. (2003) *Lancet Infect. Dis.* **3,** 757–771
48. Daines, D. A., and Silver, R. P. (2000) *J. Bacteriol.* **182,** 5267–5270
49. Vann, W. F., Daines, D. A., Murkin, A. S., Tanner, M. E., Chaffin, D. O., Rubens, C. E., Vionnet, J., and Silver, R. P. (2004) *J. Bacteriol.* **186,** 706–712
50. Suryanti, V., Nelson, A., and Berry, A. (2003) *Protein Expression Purif.* **27,** 346–356
51. Haft, R. F., and Wessels, M. R. (1994) *J. Bacteriol.* **176,** 7372–7374