THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 2, pp. 855–865, January 11, 2008

Characterization of the *N*-Acetyl-5-neuraminic Acid-binding Site of the Extracytoplasmic Solute Receptor (SiaP) of Nontypeable *Haemophilus influenzae* Strain 2019*[S]

Received for publication, August 9, 2007, and in revised form, October 15, 2007 Published, JBC Papers in Press, October 18, 2007, DOI 10.1074/jbc.M706603200

Jason W. Johnston[‡1], Nathan P. Coussens[§1,2], Simon Allen[¶1], Jon C. D. Houtman[‡], Keith H. Turner[§], Anthony Zaleski[‡], S. Ramaswamy[§], Bradford W. Gibson[¶], and Michael A. Apicella[‡3]

From the [‡]Department of Microbiology and [§]Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242 and [¶]The Buck Institute for Age Research, Novato, California 94945

Nontypeable *Haemophilus influenzae* is an opportunistic human pathogen causing otitis media in children and chronic bronchitis and pneumonia in patients with chronic obstructive pulmonary disease. The outer membrane of nontypeable *H. influenzae* is dominated by lipooligosaccharides (LOS), many of which incorporate sialic acid as a terminal nonreducing sugar. Sialic acid has been demonstrated to be an important factor in the survival of the bacteria within the host environment. *H. influenzae* is incapable of synthesizing sialic acid and is dependent on scavenging free sialic acid from the host environment. To achieve this, *H. influenzae* utilizes a tripartite ATP-independent periplasmic transporter. In this study, we characterize the binding site of the extracytoplasmic solute receptor (SiaP) from nontypeable *H. influenzae* strain 2019. A crystal structure of *N*-acetyl-5-neuraminic acid (Neu5Ac)-bound SiaP was determined to 1.4 Å resolution. Thermodynamic characterization of Neu5Ac binding shows this interaction is enthalpically driven with a substantial unfavorable contribution from entropy. This is expected because the binding of SiaP to Neu5Ac is mediated by numerous hydrogen bonds and has several buried water molecules. Point mutations targeting specific amino acids were introduced in the putative binding site. Complementation with the mutated *siaP* constructs resulted either in full, partial, or no complementation, depending on the role of specific residues. Mass spectrometry analysis of the *O*-deacylated LOS of the R127K point mutation confirmed the observation of reduced incorporation of Neu5Ac into the LOS. The decreased ability of *H. influenzae* to import sialic acid had negative effects on resistance to complement-mediated killing and viability of biofilms *in vitro*, confirming the importance of sialic acid transport to the bacterium.

Haemophilus influenzae is a Gram-negative bacterium and a mucosal pathogen in humans. Strains can be encapsulated or nonencapsulated (nontypeable; NTHi). Encapsulated strains can cause meningitis and septicemia, whereas nontypeable strains commonly cause various local mucosal infections, including otitis media and bronchitis. The outer membrane of NTHi is exposed to a challenging environment in humans, and thus components of the outer membrane have undergone extensive study in an attempt to understand how this bacterium evades the host immune response. A major outer membrane component of NTHi is lipooligosaccharide (LOS).[4] These complex glycolipids were demonstrated to play a role in microbial virulence and pathogenicity (1). LOS contains carbohydrate epitopes that mimic human glycosphingolipids, allowing the bacteria to evade the host immune response (2). The LOS of NTHi is a heterogeneous mixture of glycoforms (3–5). Sialic acid (*N*-acetyl-5-neuraminic acid (Neu5Ac)) can be incorporated into the LOS as a terminal nonreducing sugar, protecting the bacterium from complement-mediated killing by normal human serum (6–8). The acceptors for sialic acid are lactose, *N*-acetyllactosamine, and possibly *N*-acetylgalactosamine (9). The precise structures of many of these sialylated LOS glycoforms have yet to be fully elucidated (10).

LOS is sialylated prior to reaching the outer membrane by one of the following three sialyltransferases: SiaA, Lic3A, and LsgB (7, 10). The donor for this transfer is CMP-sialic acid, which is synthesized from CTP, and sialic acid by SiaB, the CMP-sialic acid synthetase (6). Sialic acid is also incorporated into the biofilm matrix (11) and is important for the long term viability of cells within the biofilm (12). Aside from incorporation into LOS, sialic acid can be utilized as a carbon and nitrogen source by the bacteria, the first step of which is mediated by neuraminyl lyase (NanA) (13).

As *H. influenzae* is not capable of synthesizing sialic acid; the bacteria depend on transport of exogenous sources. It was recently elucidated that NTHi uses a tripartite ATP-independent periplasmic (TRAP) transporter to import sialic acid. This

* This work was supported in part by NIAID Grants AI024616 and AI30040 from the National Institutes of Health. 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.

The atomic coordinates and structure factors (code 3b50) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1 and S2 and Figs. S1 and S2.

[1] These authors contributed equally to this work.

[2] Recipeint of support from the National Institutes of Health Interdisciplinary Immunology Postdoctoral Training Program Grant NRSA AI07260-21.

[3] To whom correspondence should be addressed: 51 Newton Rd., IA City, IA 52242. Tel.: 319-335-7807; Fax: 319-335-9006; E-mail: michael-apicella@uiowa.edu.

[4] The abbreviations used are: LOS, lipooligosaccharide; *O*-LOS, *O*-deacylated LOS; TRAP, tripartite ATP-independent periplasmic; ELISA, enzyme-linked immunosorbent assay; MALDI-LIT-MS, matrix-assisted laser desorption ionization-linear ion trap instrument-mass spectrometry; CFU, colony-forming units; PBS, phosphate-buffered saline; ITC, isothermal titration calorimetry; MES, 4-morpholineethanesulfonic acid; ESR, extracytoplasmic solute-binding receptor; KDO, ketodeoxyoctanoate; Neu5Ac, *N*-acetyl-5-neuraminic acid.

TABLE 1
Strains and plasmids used in this study

Strain	Relevant phenotype or genotype	Source or Ref.
E. coli DH5α	F⁻ ϕ*dlacZ*ΔM15 Δ(*lacZ*YA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17($r_K^-m_K^+$)*pho*A *sup*E44λ-*thi*-1 *gyr*A96 *rel*A1	Invitrogen
E. coli TOP10	F⁻*mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen
H. influenzae		
NTHi 2019	Clinical respiratory isolate	32
NTHi 2019*siaT*	*siaT* (HI0147) mutant, erythromycin resistant	12
NTHi 2019*siaP*	*siaP* (HI0146) mutant, kanamycin (ribostamycin)	This study
NTHi 2019*siaP* complemented	NTHi 2019*siaP* mutant with an intact copy of *siaP* within HI0601.1, kanamycin (ribostamycin), Spectinomycin	This study
NTHI 2019*siaPsiaT*	*siaPT* double mutant, kanamycin (ribostamycin) and erythromycin-resistant	This study
Plasmids	**Description and selection marker**	**Source or Ref.**
pCR2.1-TOPO	TA cloning vector, ampicillin, kanamycin	Invitrogen
pBSL86	Source of kanamycin resistance gene *nptII*, ampicillin, kanamycin	16
pCR146	pCR2.1-TOPO containing region around HI0146 in NTHi 2019 genome, ampicillin	This study
pCR146Kn	pCR146 with *nptII* inserted in the PacI site of HI0146, ampicillin, kanamycin	This study
p601.1-Sp2	pCR601.1 with the EcoRI/BsrG fragment from pCRSp2 in pCR601.1 digested with BsgI and BsrGI, ampicillin, spectinomycin	17
p601.1Sp-146	p601.1-Sp2 with intact HI0146 gene inserted into the SmaI site 3′ to the spectinomycin gene, ampicillin, spectinomycin	This study
pBAD102D-TOPO	Protein expression vector, ampicillin	Invitrogen
pKHT101	pBAD102 containing *siaP*, ampicillin	This study

transporter consists of two distinct integral membrane components that are fused in NTHi, SiaT (12) (also known as SiaQM (14)) and an extracytoplasmic solute receptor, SiaP, which binds sialic acid with high affinity. A previous study of SiaP was carried out with the unencapsulated serotype d laboratory strain Rd (14), and its structure was determined by x-ray crystallography (15). SiaP belongs to the periplasmic binding protein type II superfamily of α/β proteins. This family includes several of the sugar transport-binding proteins. The overall structure is similar to the one reported by Muller *et al.* (15). It is a two-domain protein, and the ligand binds inbetween the two domains. In this study, the binding site of SiaP from the pathogenic NTHi strain 2019 is characterized, and the mechanism of Neu5Ac binding is elucidated.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions—Stains and plasmids used in this study are listed in Table 1. Unless otherwise stated, *Escherichia coli* was grown at 37 °C in Luria-Bertani (LB) medium with or without agar (2.0%) and supplemented with antibiotics as needed. Wild-type *H. influenzae* was grown on brain heart infusion agar (Difco) supplemented with 10 μg/ml hemin and 10 μg/ml NAD (sBHI) at 37 °C in 5% CO_2. sBHI agar was supplemented with Neu5Ac or antibiotics as indicated.

Cloning and Mutagenesis of HI0146 (siaP)—PCR was used to amplify a 4,474-bp fragment from NTHi 2019 genomic DNA containing the region from the 5′-half of HI0143 to the 5′-half of HI0147 (see supplemental Table S2 for primer sequences). This fragment was cloned into pCR2.1-TOPO (Invitrogen), creating pCR146. Sequencing and comparison with the *H. influenzae* Rd KW-20 genome confirmed the fragment contained HI0146. The 1,268-bp SmaI fragment from pBSL86 (16) containing the kanamycin resistance gene *nptII* was ligated into the PacI site of pCR146 that had been blunt-ended and phosphatase-treated. This plasmid was designated pCR146Kn. The correct orientation and structure of the construct was confirmed by sequencing. NTHi 2019 and NTHi 2019*siaT* were transformed with uncut plasmid pCR146Kn. Mutants were confirmed by PCR and Southern blot analysis. The resulting mutants, NTHi 2019*siaP* and 2019*siaPsiaT*, were used for subsequent analysis.

Complementation of the NTHi 2019siaP Mutant—The open reading frame HI0601.1 contains an authentic frameshift in the NTHi 2019 genome (17). We used this region to insert a copy of *siaP* to complement the 2019*siaP* mutant as described previously (17). PCR was used to amplify a 1,023-bp fragment containing the *siaP* open reading frame from NTHi 2019 genomic DNA using primers 146C-1 and 146C-2. The PCR fragment was digested with SmaI and ligated into the SmaI site of p601.1-Sp2, creating p601.1Sp-146. The construct was confirmed by sequencing. NTHi 2019*siaP* was transformed with plasmid p601.1-Sp-146, and transformants were selected on sBHI agar with 15 μg/ml ribostamycin and 25 μg/ml spectinomycin. Southern analysis was used to confirm the correct construction of the mutants.

Site-directed Mutagenesis—Point mutations were introduced into *siaP* using the GeneTailor site-directed mutagenesis system (Invitrogen). Primers were designed from the sequence of *siaP* to produce the mutations (supplemental Table S2). Template DNA (p601.1Sp146Comp) was first methylated using DNA methylase and protocols from Invitrogen. The methylated template DNA was amplified by PCR using Platinum *Taq* DNA polymerase, High Fidelity (Invitrogen). Reactions were transformed into DH5α-T1^R *E. coli*, and transformants were selected on LB agar containing 125 μg/ml ampicillin and 50 μg/ml spectinomycin. Transformants were passed three times on selective media, and the presence of point mutations was confirmed by DNA sequencing.

NTHi strain 2019*siaP* was transformed by incubation with the mutated p601.1Sp146Comp plasmid, and selection was performed using sBHI containing spectinomycin (25 μg/ml) and ribostamycin (15 μg/ml). Transformants were passed three times on selective media, and the presence of the mutation in the genome was confirmed by sequencing a cloned PCR fragment containing the *siaP* gene.

Expression and Purification of SiaP—Plasmid pKHT101 was created by the directional ligation of a blunt-end PCR fragment amplified from NTHi strain 2019 genomic DNA with primers HI0146F1 and HI0146R1 to linearized pBAD102/D-TOPO (Invitrogen) following the manufacturer's protocol. Cloning into and expression from pBAD102/D-TOPO results in a protein with an N-terminal thioredoxin tag and a C-terminal 6× histidine tag. The ligated DNA was transformed into *E. coli* TOP10 chemically competent cells (Invitrogen). Transformants were screened, and DNA sequencing was used to confirm the construct. Transformed cells were grown at 37 °C and 225 rpm in LB broth containing 100 μg/ml ampicillin. At mid-log phase (A_{600} ~0.5), expression of SiaP was induced with 0.002% arabinose. Four hours after induction, the culture was harvested by centrifugation for 30 min at 7,280 × *g* and 4 °C.

Cell pellets were resuspended in 45 ml of buffer A (50 mM sodium phosphate, 150 mM NaCl, 5 mM imidazole, pH 7.0) per 1 liter of culture and lysed using a French press at 16,000 p.s.i. The lysate was centrifuged for 20 min at 186,010 × *g* and 4 °C to separate the soluble and insoluble fractions. The resulting supernatant was mixed with TALON metal affinity resin (Clontech) pre-equilibrated in buffer A. The slurry was allowed to incubate on a rotating platform at 4 °C for 5 h and then it was transferred to a gravity column. The resin was washed with 10 column volumes of buffer A followed by 10 column volumes of buffer A with 1 M NaCl. Finally, elution buffer (50 mM sodium phosphate, 150 mM NaCl, 500 mM imidazole, pH 7.0) was used to elute the protein. The predominant form of the purified protein lacked the thioredoxin fusion, as confirmed by N-terminal sequencing.

The eluted protein was dialyzed against buffer B (20 mM Tris, 10 mM NaCl, pH 8.0) and loaded onto a Q-Sepharose Fast Flow column (XK 16, Amersham Biosciences) equilibrated with buffer B. After the protein was loaded, the column was washed with 55 ml of buffer B. For elution, a linear gradient of buffer B containing 10–500 mM NaCl, with a total volume of 130 ml, was used. The flow rate was 0.5 ml/min. SiaP eluted between 200 and 250 mM NaCl. This purification step allowed for the removal of the thioredoxin-fused protein form.

Crystallization and Structure Determination—Diffraction quality crystals were grown at 4 °C using the hanging-drop, vapor-diffusion method from drops consisting of a mixture of equal volumes of protein solution (37 mg/ml SiaP in 20 mM HEPES, 10 mM NaCl, 10 mM Neu5Ac, pH 8.0) and reservoir solution (100 mM MES, 30% w/v PEG 6000, pH 6.0; Qiagen).

SiaP crystals were flash-frozen in mother liquor containing 10% (v/v) glycerol. The crystals were mounted in loops for data collection at 100 K from the GM/CA CAT beamline at the Advanced Photon Source, Argonne, IL. The data were collected on a MAR-CCD detector with a crystal to detector distance of 250 mm. The data were processed and scaled using D*trek (18).

Molecular replacement was carried out using the Neu5Ac2en-bound SiaP (Protein Data Bank ID 2CEX). Model building was carried out using the program Coot (19), and refinement was carried out using the program REFMAC5 from the CCP4 package (20). The quality of the structure was assessed using the programs PROCHECK (21) and WHATIF (22). All indicators suggest that the quality of the structure is good. The details of the crystallographic data collection and refinement are presented in Table 2. The structure has been deposited in the RCSB (Protein Data Bank ID 3B50).

TABLE 2
Data collection and refinement statistics

Data collection statistics	
Resolution (Å)	56 to 1.4
Wavelength (Å)	0.979
Space group	$P2_12_12_1$
Cell (Å)	$a = 47.5$; $b = 74.6$; $c = 86.5$
Completeness (%) (last shell)	99.6 (1.45–1.4:96.8)[a]
Redundancy	7.0 (5.3)
I/σ_I	10.1 (2.5)
R_{merge} (%)[b]	7.9 (49.4)
Refinement statistics	
Resolution range (Å)	56.5–1.4
R (%)[c]	16.1
R_{free} (%)[d]	20.0
rms bonds (Å)	0.012
rms angles (°)	1.46
No. of water molecules	481
No. of protein atoms	4958 (310 amino acid residues)
No. of sugar atoms	21
Ramachandran analysis (%)	
Most favored	99.4
Allowed	0.6

[a] Values in parentheses relate to the highest resolution shell.
[b] $R_{merge} = \Sigma|I| - <I>/\Sigma I$, where I is the observed intensity, and $<I>$ is the average intensity obtained from multiple observations of symmetry-related reflections after rejections.
[c] $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.
[d] R_{free} is the R-factor calculated with 5% of the reflections chosen at random and omitted from refinement.

Isothermal Titration Calorimetry—Isothermal titration calorimetry (ITC) measurements were performed using a VP-ITC microcalorimeter (Microcal, Northhampton, MA). Protein concentration was determined in 6.0 M guanidine hydrochloride by measuring the A_{280} and using an extinction coefficient of 20,480 M^{-1} cm^{-1}. Titrations were performed at 25 °C in 20 mM HEPES, pH 8.0, with 10 mM NaCl. The concentrations of SiaP and Neu5Ac varied among titrations, and each binding experiment was repeated 1–2 times. The nonspecific heat released by the dilution of both the sugar and the protein was calculated by averaging the heat liberated during the last 3–5 injections. This value was then subtracted from the raw heat released during each injections. To derive the value of stoichiometry, the data sets were analyzed using a single-site binding model from the ORIGIN ITC analysis software (Microcal). In this analysis the values for stoichiometry, change in enthalpy (ΔH), and affinity were fit using nonlinear least squares analysis. To derive the values for affinity and ΔH, data sets were globally analyzed using the single-site binding model in the program SEDPHAT (23). In this analysis, the value for stoichiometry was fixed at 1:1, whereas ΔH and affinity were fit using nonlinear least squares analysis. Errors for affinity and ΔH derived from the global analysis were calculated using the Monte-Carlo analysis software built into SEDPHAT. The values for the change in entropy ($T\Delta S$) were calculated from the globally derived values of affinity and ΔH using Gibb's free energy equation $\Delta G^0 = -RT\ln K_a = \Delta H - T\Delta S$, where ΔG^0 is Gibb's free energy; R is the gas constant, and T is the temperature of the binding reaction.

Enzyme-linked Immunosorbent Assay (ELISA)—Bacteria were grown on sBHI agar plates containing 100 μM Neu5Ac. Cells were harvested and resuspended in PBS (20 mM

Na_2HPO_4, 150 mM NaCl, pH 7.0) at an A_{600} of 0.15. The wells of Costar 3590 96-well EIA/RIA plates (Corning Glass) were filled with 100 μl of the bacterial suspension, and six wells for each strain were tested. The plates were dried at 40 °C overnight. The plates were washed four times with 72 mM sodium acetate, 165 mM sodium chloride, 0.55% Tween 20 using an ELx50 Auto Strip Washer (BioTek Instruments, Winooski, VT). A neuraminidase digest was performed on three of the wells from each strain by adding 100 μl of 0.5 unit/ml neuraminidase (Roche Diagnostics) diluted in neuraminidase buffer (50 mM sodium acetate, 154 mM NaCl, 9 mM $CaCl_2$, 25 μg/ml human serum albumin, pH 5.6). The other three wells received 100 μl of the buffer alone. The plate was incubated for 6 h at 37 °C, and the wells were washed. The monoclonal antibody 3F11, diluted 1:250 in TBST (10 mM Tris, 165 mM NaCl, 0.3% Tween 20, pH 7.4), was added to each well. The plates were incubated at room temperature overnight (~18 h) followed by a wash. The secondary antibody, phosphatase-labeled goat anti-mouse IgM (Kirkegaard & Perry Laboratories, Gaithersburg, MD), diluted 1:2000 in TBST, was added. Plates were incubated for 1 h at room temperature and washed. *p*-Nitrophenyl phosphate (1.0 mg/ml; Sigma) in 8.2% diethanolamine, 0.5 mM $MgCl_2$, pH 9.8 was added. When yellow color had developed, the A_{405} was read using a Microplate Autoreader EL311 (BioTek Instruments). The mean for each strain was converted to a ratio of A_{405} with neuraminidase treatment divided by A_{405} without neuraminidase.

Purification and Preparation of O-Deacylated LOS and Neuraminidase Treatment—LOS was purified from plate-grown bacteria as described previously (12). To make the LOS more amenable for mass spectrometric analysis, *O*-linked fatty acids were removed from the lipid A moiety as described previously (24). The highly purified LOS (~0.1 mg) was incubated in anhydrous hydrazine (50 μl; Sigma) at 37 °C for 35 min with mixing every 10 min. Samples were cooled on ice prior to and after the addition of ice-cold acetone (250 μl; Sigma) then transferred to −20 °C for 2 h. The quenched reaction mixture was centrifuged at 12,000 × *g* for 45 min at 4 °C. The supernatant was removed, and the pelleted *O*-deacylated LOS (*O*-LOS) was dissolved in MilliQ H_2O (50 μl) and evaporated with a speed vacuum system. To remove salts and other low molecular weight contaminants, the *O*-LOS (~20–30 μg) was suspended on a nitrocellulose membrane (type VS, 0.025 μm; Millipore Corp., Billerica, MA) over water for ~1 h. The desalted *O*-LOS was removed from the membrane, concentrated with a speed vacuum system, and analyzed by mass spectrometry on a matrix-assisted laser desorption ionization-linear ion trap instrument-mass spectrometry (MALDI-LIT-MS). For removal of Neu5Ac, the *O*-LOS (~20–30 μg) was digested in 10 mM ammonium acetate, pH 6.0, containing immobilized neuraminidase from *Clostridium perfringens* (type VI; Sigma) for 20 h at 30 °C with shaking. The enzyme was pelleted by centrifugation, and the supernatant (~15 μl) was transferred to a nitrocellulose membrane for drop dialysis. The desialylated *O*-LOS was concentrated and analyzed by MALDI-LIT-MS.

Mass Spectrometry of O-LOS—Aliquots of *O*-LOS were desalted by mixing with Dowex 50 beads (100–200 mesh, NH_4^+ form; Bio-Rad) prior to mass spectrometry. Mass spectrometric data were collected on a Finnigan™ LTQ™ linear ion trap mass spectrometer coupled to a vMALDI ion source (ThermoElectron, San Jose, CA). Aliquots of the *O*-LOS (0.5 μl) were spotted onto a 96-well stainless steel sample plate and allowed to dry. Dried samples were overlaid with 0.5 μl of 50 mg/ml 2,5-dihydroxybenzoic acid) in 50% acetonitrile. Data were collected in negative ion mode over two mass ranges: *m/z* 800–4000 and 2000–4000. Tandem MS data were collected in negative mode to confirm sialylation, using a precursor isolation width of 3 *m/z*, normalized collision energy of 40% (RF amplitude in percent used for fragment ions), activation *Q* of 0.25, and an activation of 30 ms. All spectra were recorded using automatic gain control to control the number of laser shots and the automatic spectrum filter tool.

Bactericidal Assay—NTHi strains 2019 and 2019*siaP* were grown to an A_{600} of 0.2 in sBHI broth with and without supplemental Neu5Ac. A 0.5-ml aliquot of each was centrifuged for 1 min at 10,000 rpm and 25 °C. The pellet was resuspended in 1 ml of phosphate-buffered salt solution consisting of 10 mM K_2HPO_4, 10 mM KH_2PO_4, 136 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 0.3 mM $MgCl_2 \cdot 6H_2O$, 1 mM $MgSO_4 \cdot 7H_2O$, and 0.01% bovine serum albumin, pH 7.0. The bactericidal assay, modified from that reported by Andreoni *et al.* (25) was carried out in duplicate in a 96-well plate with a 200-μl final volume. Pooled normal human serum, a 20-donor pool of serum from human volunteers who had no previous history of serious infections, was diluted to 10% in phosphate-buffered salt solution. A control containing pooled normal human serum that had been heat-inactivated for 30 min at 56 °C was included in each experiment. The resuspended bacteria were serially diluted and plated on sBHI with or without appropriate antibiotic selection to determine the initial CFU. For the assay 10 μl (1×10^6 organisms) of the bacterial stock were incubated in the diluted serum at 37 °C with shaking at 200 rpm for 30 min. The reaction mixtures were serially diluted into phosphate-buffered salt solution and plated on sBHI with or without appropriate antibiotic selection, and emerging colonies were counted the next day. The resulting CFU value corresponded to the number of bacteria that survived after 30 min. Killing was assessed by comparing the number of CFU from the 30-min serum incubation with the number of the initial CFU. Results were expressed as the $\log_{10}$ change in CFU at 30 min compared with the initial CFU. Data from six experiments were pooled, and statistical analysis was carried out using the paired *t* test and analysis of variance functions from Prism, version 4 (Graphpad Software, San Diego).

Biofilm Growth in a Continuous Flow Chamber—NTHi strains 2019, 2019*siaP*, 2019*siaT*, and 2019*siaPsiaT* were grown in a continuous-flow chamber as described previously (26). Strains were grown to mid-log phase in RPMI 1640 media (Invitrogen) supplemented with protoporphyrin IX (1 μg/ml), hypoxanthine (0.1 mg/ml), uracil (0.1 mg/ml), β-NAD (10 μg/ml), sodium pyruvate (0.8 mM), and Neu5Ac (100 μM). Cultures were diluted to an A_{600} of 0.25 to inoculate chambers. Chambers were filled with the prepared inoculum and incubated at 37 °C for 1 h to allow for adherence of bacteria to the coverslip. Supplemented RPMI media, diluted 1:10 with PBS, was pumped into the chambers with a flow rate of 180 μl/min. Biofilms were allowed to grow for 1–2 days. Chambers were photographed and then stained with LIVE/DEAD *Bac*Light Viability Stain

(Molecular Probes, Eugene, OR) following the manufacturer's protocol. The biofilm was visualized with a Nikon Digital Eclipse C1 laser scanning confocal microscope (Nikon Instruments Inc., Melville, NY) at a magnification of ×20. The EZ-C1 Gold Version 2.20 confocal software (Nikon) was used to produce a three-dimensional orthogonal image of the biofilm.

RESULTS

Cloning and Mutagenesis of siaP—The HI0146 (*siaP*) gene from NTHi strain 2019 was cloned and shown to encode a sialic acid-binding protein of 329 amino acids. This gene is adjacent to the *siaT* sialic acid transporter (*siaQM*) identified previously in both NTHi 2019 and 7502 (12). The sialic acid transporter and binding protein belong to the TRAP-type family of transporters. Analysis of the SiaP protein indicated that it has a 23-amino acid leader sequence that is cleaved upon secretion into the periplasmic space. The functional *siaP* of NTHi 2019 was disrupted by insertion of a kanamycin resistance gene. The resulting mutant, NTHi 2019*siaP*, was constructed to ensure the transcription of *siaT* and subsequent genes. The *siaP* mutant was complemented by the insertion of *siaP* into a pseudogene in the NTHi 2019 genome, which restored the *siaP* mutant to a wild-type phenotype.

Alignment of primary sequences indicates that a number of DctP family proteins are involved in the specific transport of different sialic acid isoforms for diverse species of bacteria including *Vibrio cholerae*, *Pasteurella multocida*, and *Fusobacterium nucleatum*. The alignment reveals a number of residues that are highly conserved among other extracytoplasmic solute-binding receptor (ESR) proteins believed to bind sialic acid (supplemental Fig. S1). These data, together with the crystal structure described below, was used to identify specific residues that likely contribute to Neu5Ac binding. *siaP* genes with defined point mutations were introduced into 2019*siaP* using the complementation system described above to further study these residues.

Characterization of the Neu5Ac-binding Site of SiaP—SiaP from NTHi binds to Neu5Ac with high affinity and is crucial for pathogenicity (12, 14). High selectivity is achieved by the arrangement of residues in the binding pocket. We have determined the structure of Neu5Ac-bound SiaP from NTHi strain 2019 to 1.4 Å resolution. The structure was refined to an *R*-factor of 0.16 and free *R*-factor of 0.2 (Table 2). There is one molecule in the asymmetric unit, and electron density clearly shows the presence of Neu5Ac at the binding site (Fig. 1). The structure reveals a two-domain protein with sialic acid bound between the two domains. Before we determined this structure, two structures of SiaP from the unencapsulated serotype d laboratory strain Rd (both unbound and with the Neu5Ac analog, Neu5Ac2en) were determined (15). The two SiaP proteins have a molecular mass of about 37 kDa, and the sequences are highly conserved. SiaP from the NTHi strain differs by two amino acids from that of the nonpathogenic strain. However, the crystallization conditions of the proteins are different as are the methods of expression and purification. We expressed SiaP with an N-terminal thioredoxin tag and a C-terminal histidine tag. A majority of the expressed protein lacked the thioredoxin tag, likely removed by the host cell prior to purification. Subsequent anion exchange chromatography resulted in the removal of the small amount of thioredoxin-tagged protein. This was confirmed by N-terminal sequencing. Our structure of SiaP in complex with Neu5Ac reveals different interactions that stabilize the sugar.

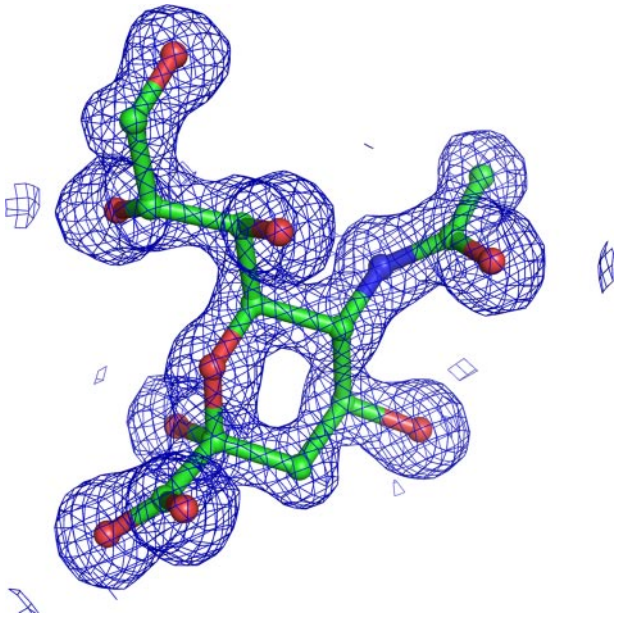

FIGURE 1. **Electron density map of the Neu5Ac-binding site of SiaP.** The figure shows the ($2F_o - F_c$) electron density (at 1.2 $\sum$) and the modeled Neu5Ac in the binding site.

Overall, our structure is very similar to the Neu5Ac2en-bound structure determined previously. The refined model contains 310 residues and 481 water molecules. Both ligand-bound structures reveal that the highly charged carboxylate at the C-1 position of Neu5Ac initiates binding through direct interactions with positively charged residues Arg-127 and Arg-147 (Fig. 2). The orientation is probably governed by a conserved interaction between the glycol group hydroxyls and the carboxyl of Glu-67. An interaction between O-10 of Neu5Ac and Asn-10 of SiaP is also seen from both ligand-bound structures. However, our structure reveals an interaction between the O-2 of Neu5Ac and Asn-187 and between O-2 and Arg-127, both of which are absent in the structure with Neu5Ac2en. Asp-49 interacts with O-7 directly in the structure bound to Neu5Ac, whereas in the structure of the analog this interaction happens through a water molecule. The presence of the O-2 hydroxyl group moves the sugar toward the Asp-49, while preserving the other interactions. Additional stabilizing interactions are mediated by water molecules in the binding pocket. Fig. 2 shows a number of differences in SiaP interactions to Neu5Ac2en and Neu5Ac.

To investigate the basis of SiaP specificity for Neu5Ac, mutations were introduced throughout the binding site. Arg-127 and Asn-187, which interact with the C-1 carboxylate and O-2 of Neu5Ac, were mutated to R127A or R127K and N187Q. Arg-147, which interacts with the C-1 carboxylate of Neu5Ac, was mutated to R147A or R147K. Glu-67, which binds the hydroxyls of the glycol group, was mutated to an alanine. An N10A mutation was introduced to abolish the interaction between Asn-10 and O-10 of Neu5Ac. Asp-49 interacts directly

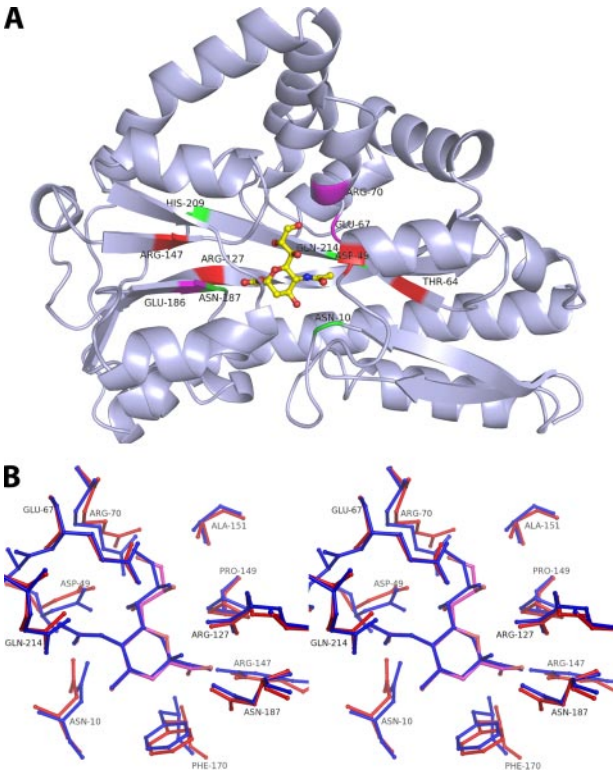

FIGURE 2. **Structure of SiaP and the binding site.** *A,* ribbon diagram showing SiaP with Neu5Ac bound. The residues marked in *red* show no complementation; the ones in *maroon* show partial complementation, and those in *green* show complete complementation. Neu5Ac is shown as a *ball-and-stick* model. *B,* stereo diagram showing Neu5Ac (in *blue*) and Neu5Ac2en (in *red*) and the interacting residues.

with O-7 of Neu5Ac and was mutated to alanine. Gln-214, mutated to Q214A, interacts with the ligand through water molecules. Some binding site residues make interactions with another residue that stabilizes Neu5Ac or water binding. Glu-186, mutated to E186Q, interacts with Arg-127. His-209, located in proximity to Arg-127 and Glu-186, was mutated to H209A. Arg-70, which interacts with Asp-49, was mutated to alanine. A large pocket near Thr-64 in the binding site is occupied by water molecules. Mutations T64K and T64R were introduced to displace some of the water molecules in the pocket.

Whole cell ELISA was performed to quantify the effect of the point mutations on sialylation of the LOS. Monoclonal antibody 3F11 recognizes the acceptor for sialic acid, and this acceptor is inaccessible to the antibody when the LOS is sialylated. The point mutations were categorized into three classes based on the reactivity of 3F11 as follows: complete complementation, partial complementation, and no complementation (Table 3). Point mutations that resulted in complete complementation did not affect LOS sialylation and thus are able to transport sialic acid. We hypothesize that mutations to Arg-127, Arg-147, or Glu-67 will have a significant impact on the enthalpy of binding. The results of these studies reveal that mutations N10A, N187Q, H209A, and Q214A show complete complementation, suggesting that these mutations have no effect on function. However, the mutations T64K, E67A, R70A, R127K, and E186Q show a partial complementation effect. The mutations D49A, T64R, R127A, R147A, and R147K completely

TABLE 3
ELISA analysis of point mutations

NTHi 2019 strains complemented with a mutated *siaP* gene were used in whole cell ELISA to quantify sialylation of the LOS using the monoclonal antibody 3F11. Data are reported as the A_{405} in both neuraminidase-treated and untreated wells. The ratio neuraminidase-treated over untreated wells is indicative of the relative sialylation.

Strain	No neuraminidase, mean ± S.D.	Neuraminidase, mean ± S.D.	Ratio
2019wt	0.013 (0.004)	1.039 (0.005)	80
2019*siaP*	0.788 (0.012)	0.738 (0.017)	0.94
2019*siaP* complemented	0.016 (0.006)	0.782 (0.017)	49
Complete complementation			
2019*siaP* N010A	0.016 (0.003)	1.089 (0.031)	68
2019*siaP* N187Q	0.045 (0.003)	1.171 (0.012)	26
2019*siaP* H209A	0.014 (0.003)	0.823 (0.084)	59
2019*siaP* Q214A	0.006 (0.003)	0.885 (0.036)	148
Partial complementation			
2019*siaP* T064K	0.272 (0.005)	1.270 (0.247)	4.7
2019*siaP* E067A	0.307 (0.036)	1.050 (0.013)	3.4
2019*siaP* R070A	0.266 (0.028)	0.536 (0.037)	2
2019*siaP* R127K	0.422 (0.023)	0.690 (0.030)	1.6
2019*siaP* E186Q	0.340 (0.017)	1.263 (0.086)	3.7
No complementation			
2019*siaP* D049A	0.604 (0.017)	0.733 (0.023)	1.2
2019*siaP* T064R	0.673 (0.074)	0.792 (0.070)	1.2
2019*siaP* R127A	0.630 (0.019)	0.576 (0.034)	0.91
2019*siaP* R147A	0.567 (0.031)	0.483 (0.019)	0.85
2019*siaP* R147K	0.592 (0.012)	0.609 (0.001)	1.0

abolish incorporation of Neu5Ac into LOS. Although the assay measures incorporation, it does not reveal details of the mechanism involved. It is possible that mutations have altered binding affinity (thermodynamics) or kinetics of transport or a combination of both.

Determination of the SiaP Binding Affinity to Neu5Ac—ITC was used to determine the affinity of SiaP for Neu5Ac (supplemental Fig. S2). The data indicate that binding is enthalpically driven with a globally fit K_d of 28 ± 1 nM ($\Delta H = -15.3 \pm 0.2$ kcal/mol; $-T\Delta S = 5.0$ kcal/mol). The *n* value as calculated from the individual data sets indicated that SiaP bound to Neu5Ac with 1:1 stoichiometry. ITC was also used to investigate the binding of other sugars by SiaP, including *N*-glycolylneuraminic acid, a sialic acid found in many animal species except humans, and ketodeoxyoctanoate (KDO), a charged eight-carbon keto-sugar. SiaP bound to *N*-glycolylneuraminic acid similar to Neu5Ac with comparable thermodynamics and a K_d of about 98 ± 6 nM ($\Delta H = -13.5 \pm 0.5$ kcal/mol; $-T\Delta S = 3.9$ kcal/mol). In contrast, SiaP had no measurable interaction with KDO. Binding of Neu5Ac to the SiaP mutant R127A and Neu5Ac was not detectable by ITC. The binding affinity (K_d) of SiaP to Neu5Ac2en was reported to be about 200-fold weaker than to Neu5Ac (15).

MALDI-LIT-MS of O-Deacylated LOS—The sialylation of LOS in wild type, 2019*siaP*, and selected point mutation strains was investigated further. LOS samples were *O*-deacylated with anhydrous hydrazine (*O*-LOS) and analyzed by MALDI-LIT-MS. Based on previous studies we expected to observe a complex mixture of glycoforms in the wild-type NTHi 2019, the major glycoform consisting of a lactose moiety (Galβ1,4-Glc-) linked in a β1,4-linkage to Hep^{I} of the conserved core structure of *H. influenzae*, Hep$^{III}\alpha$1,2-Hep$^{II}\alpha$1,3-Hep$^{I}\alpha$1,5(P)KDO-lipid A. Fig. 3, *A* and *B*, shows wild-type NTHi 2019 expressing a mixture of glycoforms with the lactose-containing "B"-glycoforms with 2 or 3 phosphoethanolamine moieties (B_2 and B_3,

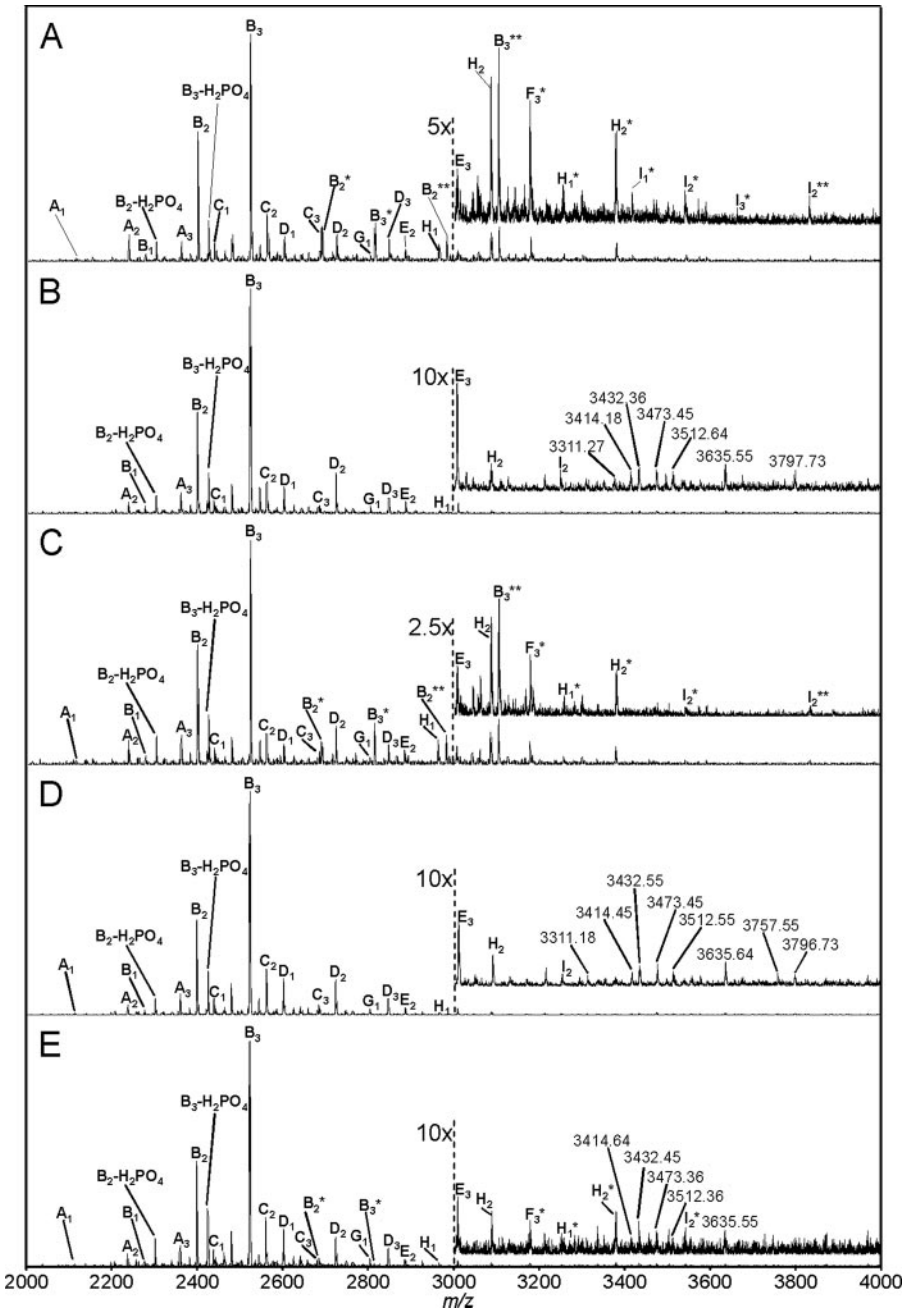

FIGURE 3. **Negative ion MALDI-LIT mass spectra of *O*-deacylated LOS from *H. influenzae*.** *A,* 2019 wild type; *B,* 2019*siaP*; *C,* 2019*siaP* complemented with exogenous copy of *siaP* gene; *D,* 2019*siaP* R127A; *E,* 2019*siaP* R127K grown in sBHI broth supplemented with 100 μM Neu5Ac. See supplemental Table S2 for molecular weights and proposed compositions. *Capital letters* designate glycoforms; *asterisks* associated with *letters* indicate addition and number of sialic acid residues; *subscripts* indicate the number of PEA moieties. Molecular mass is given for novel unidentified LOS glycoforms, but whose masses are consistent with addition of additional hexose sugars (see supplemental Table S2 for proposed compositions).

respectively) dominating the spectrum. Other peaks correlate to LOS glycoforms that differ from the B-glycoform by the addition of as many as four hexoses and a single *N*-acetylhexosamine (for identities and *m/z*, see Table 4 and supplemental Table S2). In addition to these larger glycoforms, there are a number of smaller glycoforms that differ from the B-glycoform by the loss of a single hexose. A number of low abundance LOS molecular ions are also observed whose mass is consistent with sialylation, even though the media contained no supplemental Neu5Ac. These latter sialylated glycoforms are likely because of trace amounts of sialic acid present in the media used for growth. Upon the addition of supplemental Neu5Ac to the media, these sialylated glycoforms become more abundant, and additional higher mass sialylated LOS glycoforms become apparent (Fig. 3*B* and supplemental Table S2). The glycoforms present in the *O*-LOS of the NTHi 2019*siaP* mutant are very similar to that of the wild type, with a very similar distribution of glycoforms (Fig. 3*C* and supplemental Table S2). However, the *O*-LOS from the mutant lacks sialic acid-containing glycoforms, even though the bacteria were grown on media supplemented with Neu5Ac, confirming a role for SiaP in the uptake or processing of sialic acid. When the *siaP* gene was complemented in the mutant strain, sialic acid was once again incorporated into the LOS (Fig. 3*D* and supplemental Table S2). Reintroduction of *siaP* appears to have little effect on the expression of the various sialylated LOS glycoforms when compared with the wild type (compare Fig. 3, *B* and *C*), suggesting that the addition of the functional *siaP* has not disrupted any genes involved in the expression of specific glycoforms. The *O*-LOS from NTHi 2019 expressing the *siaP* point mutation R127A showed similar glycoforms to the wild type, with the exception that no sialic-containing species were formed (Fig. 3*D* and supplemental Table S2). A second point mutant R127K gave *O*-LOS expressing low levels of several sialylated glycoforms, but unlike the wild-type strain, only singly sialylated glycoforms were observed (Fig. 3*E* and supplemental Table S2). The phenotypes of both point mutants suggest a critical role for Arg-127 in the function of SiaP. In addition to the glycoforms described in Table 4, there are number of unidentified LOS glycoforms that appear in the 2019*siaP* mutant and the R127A mutant, which are not expressed in the wild-type or complemented mutant strains (Fig. 3, *B* and *D*, *m/z* of ions as indicated). These glycoforms appear to be sodium adducts containing additional hexose moieties based on the differences in their observed masses relative to the previously characterized glycoforms E and I. However, further studies are required to confirm to exact

TABLE 4
O-LOS glycoforms

Glycoform[a]	HexNAc[b]	Hex[b]	Calculated [M - H]$^{-}$[c]
Core = Hep-(Hep$_2$)-KDO(P)-*O*-deacyl LA			
A$_1$ (A$_2$; A$_3$)	0	1	**2112.72** (2235.73; 2358.74)
B$_1$ (B$_2$; B$_3$)	0	2	**2274.78** (2397.89; 2520.79)
C$_1$ (C$_2$; C$_3$)	0	3	**2436.83** (2559.84; 2682.84)
D$_1$ (D$_2$; D$_3$)	0	4	**2598.88** (2721.89; 2844.90)
E$_2$ (E$_3$)	0	5	**2883.94** (3006.95)
G$_1$	1	4	**2801.96**
H$_1$ (H$_2$)	1	5	**2964.01** (3087.02)
Sialylated glycoforms[d]			
B$_2$* (B$_3$*)	0	2	**2688.88** (2811.89)
B$_2$** (B$_3$**)			**2979.97** (3102.98)
F$_3$*	1	3	**3177.02**
H$_1$* (H$_2$*)	1	5	**3255.11** (3378.12)
I$_1$* (I$_2$*; I$_3$*)	1	6	**3417.16** (3540.14; 3663.18)
I$_2$**			**3831.26**

[a] Subscript number indicates the number of PEA moieties.
[b] This is the proposed number of Hex and HexNAc sugars in Core-substituted branch structure.
[c] Glycoform molecular ions are monoisotopic and contain the conserved Core, plus additional branch Hex. (1–6), HexNAc (0–1) and sialic acid sugars (NeuAc; 0–2) and PEA moieties (1–3).
[d] Asterisks indicate the number of sialic acid residues.

composition of these low abundance, high mass glycoforms. Interestingly, these glycoforms are also present to a lesser extent in the partially functional R127K point mutant, which contains a reduced level of sialylated structures compared with the wild type, suggesting that the pathway for the formation of these high mass hexose and/or *N*-acetylhexosamine-extended glycoforms are competing with the sialylation pathway for the same acceptor(s) (Fig. 3*E*). The MALDI-MS studies also validate the ELISA studies as the basis of categorizing the complemented site-directed mutants into the different groupings.

Bactericidal Assay—A correlation has been previously demonstrated between LOS sialylation and protection of *H. influenzae* from complement-mediated killing by normal human serum (6–8). It appears that the NTHi 2019*siaP* mutant is incapable of incorporating sialic acid into its LOS, suggesting that the mutant will have increased susceptibility to killing by normal human serum. To address this hypothesis we carried out bactericidal assays on wild-type NTHi 2019 and the NTHi 2019*siaP* mutant. When grown in the absence of Neu5Ac, both the wild-type and *siaP* mutant were susceptible to killing by normal human serum (Fig. 4). Upon addition of supplemental Neu5Ac to the media, the wild-type strain became resistant to serum killing, whereas the *siaP* mutant remained susceptible (Fig. 4). As expected, strains incubated with heat-inactivated serum were not killed, demonstrating the complement dependence of the reaction (data not shown). Studies with the complemented *siaP* mutant indicated that it was resistant to killing in this assay (data not shown).

Biofilm Formation—Previous work with the 2019*siaT* mutant showed that inactivation of sialic acid transport has a detrimental effect on the viability of cells within a biofilm (12). A continuous-flow cell assay was used to determine whether the biofilm formation phenotype of 2019*siaP* and 2019*siaPsiaT* was similar to 2019*siaT*. Biofilm chambers were inoculated and examined at 24 and 48 h using the LIVE/DEAD *Bac*Light stain and confocal microscopy. Chambers examined after 24 h revealed little difference between wild-type and mutant strains (data not shown). By 48 h post-inoculation, there was a significant difference between the wild-type strain and the sialic acid transporter mutants (Fig. 5). Although NTHi 2019 produced a biofilm with a majority of viable cells present (Fig. 5*A*), biofilms produced by the three mutants, NTHi 2019*siaP*, NTHi 2019*siaT*, and NTHi 2019*siaPsiaT*, predominantly consisted of dead cells (Fig. 5, *B–D*, respectively). Clearly, sialic acid is not required for the formation of a biofilm; however, the transport of sialic acid is absolutely required for long term viability of bacteria within the biofilm. These results are consistent with previous observations of 2019*siaT* and confirm the necessity for sialic acid for the survival of NTHi within the biofilm (12).

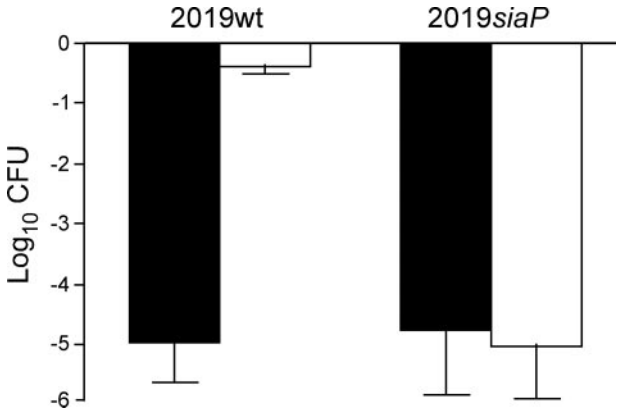

FIGURE 4. **Resistance to serum killing of wild-type NTHi 2019, NTHi 2019*siaP*, and complemented NTHi 2019*siaP*.** Bacteria were grown on sBHI media in the absence (*closed bars*) or presence (*open bars*) of supplemental Neu5Ac (100 μg/ml). Bacteria were exposed to 10% normal human serum for 30 min at 37 °C. The data were the result of duplicate reactions from the assay repeated six times. The scale in all panels is at $\log_{10}$ intervals.

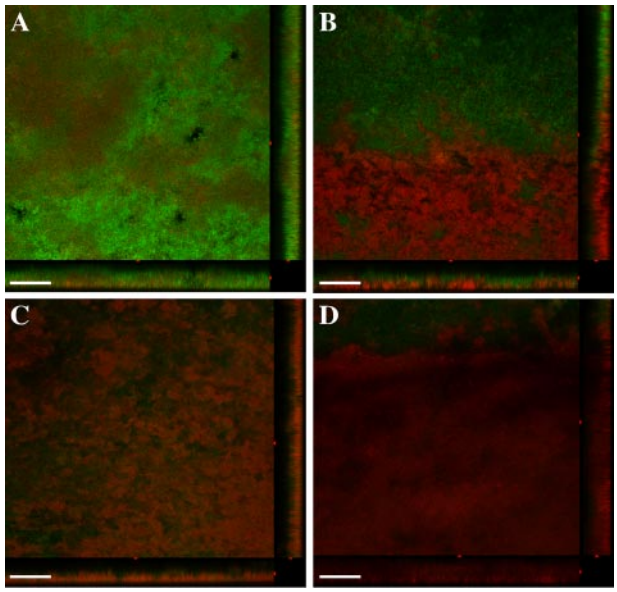

FIGURE 5. **Biofilm formation of sialic acid transporter mutants.** Strains were grown for 2 days in supplemented RPMI diluted 1:10 with PBS, stained with the LIVE/DEAD *Bac*Light viability stain and examined by confocal microscopy. Three-dimensional orthogonal images of 2019 (*A*), 2019*siaP* (*B*), 2019*siaT* (*C*), and 2019*siaPsiaT* (*D*) were rendered using Nikon EZ-C1 confocal software. The three mutant strains (*B–D*) had a significantly increased proportion of dead cells compared with the wild type (*A*). *White bar* represents 100 μm.

DISCUSSION

In this study we characterized the Neu5Ac-binding site of SiaP, the periplasmic solute-binding protein for the sialic acid TRAP transporter of NTHi strain 2019. Our previous studies demonstrated that deletion of *siaT* in NTHi 2019 creates a mutant incapable of transporting sialic acid and thus unable to sialylate its LOS (12). Studies by Severi *et al.* (14) showed that the phenotype of the *siaP* mutant is very similar to that of a *siaT* mutant. In this study we described the crystal structure of SiaP in complex with Neu5Ac and have investigated the Neu5Ac-binding site of SiaP by site-directed mutagenesis. We confirmed the functionality of each modified *siaP* by complementing a *siaP* mutant and examining the extent of surface sialylation using ELISA. The results were supported by mass spectrometry, which was also able to identify two novel sets of high molecular weight LOS glycoform populations in the *siaP* mutants, one containing sialic acid and the other not. Our studies further demonstrated that reducing Neu5Ac uptake results in increased serum sensitivity and impairs the ability to form a viable biofilm. This would likely impact the ability of the bacterium to evade the host innate immune response.

Our Structure Reveals the Basis for SiaP Specificity to Neu5Ac—Although Muller *et al.* (15) previously reported the structure of SiaP bound to Neu5Ac2en, a comparison of the two structures reveals interesting differences in the mode of binding. These differences account for the observed difference in the binding affinity between the real ligand and the analog. The difference in the structures of the ligands is the addition of the O-2 hydroxyl group to C-2 of Neu5Ac that results in a pucker of the sugar ring. There are both global effects as well as local effects that are observed between the two structures. Our structure with Neu5Ac bound is more closed than the reported structure with Neu5Ac2en. There is an additional rotation of 3.7° between the two domains toward each other. This suggests that the increase in affinity not only comes from more numerous and stronger interactions with the ligand but also because of interactions between the two domains in the closed conformation. Mutagenesis studies tested both these hypotheses and found that changes in both amino acids that directly interact with the ligand as well as those on the interface affect complementation.

One major contribution to the direct interaction with Neu5Ac comes from Arg-127 whose δ-guanido interacts with both the C-1 carboxyl and the C-2 hydroxyl of the sugar (forming a salt bridge). Not only is there an extra interaction, the interaction between the C-1-carboxyl oxygen and the ϵ-nitrogen is also shorter by 0.3 Å. Although the R127A mutation showed no complementation, R127K does not abolish binding. This suggests that the primary role of Arg-127 is electrostatic stabilization of the charge on the sugar. Glu-186 interacts with Arg-127 and appears to optimize its position because E186Q showed partial complementation. Asn-187 makes direct interactions with Neu5Ac at the C-1 carboxyl and the C-2 hydroxyl, however, the mutation N187Q did not affect incorporation. It is possible that the same interactions are preserved by this conservative mutation.

The guanidino group of Arg-147 seems to be key to the binding interaction, both in terms of its charge interaction with the C-1-carboxyl group as well as in the orientation of the ligand. Both R147A and R147K mutations showed no complementation. Based on the structures of the two sugar-bound forms, one would have predicted an abolition of binding for the E67A mutation. Surprisingly, the E67A mutation does show partial complementation. Previous studies indicate that SiaP is specific to nine carbon sugars and does not bind to eight (KDO) or six carbon sugars (15). These suggest that the role of Glu-67 is to induce specificity, and the lack of Glu-67 would probably result in reduced specificity rather than abolition of binding. Asn-10 makes two interactions with the acetyl oxygen in both ligand-bound structures, from the side chain and a water-mediated interaction by the main chain amide. However, Asn-10 is not well conserved, and the N10A mutation showed complete complementation. This mutation would abolish the direct interaction with the acetyl oxygen; however, the main chain amide would still be present. His-209 is located in proximity to Arg-127 and Glu-186 near the binding site and is conserved among ESR proteins that interact with TRAP transporters. Although His-209 does not interact with the ligand, it is somewhat surprising that both H209A and H209Q mutations showed complete complementation. Gln-214 is a conserved residue located in the binding site from β-10; however, mutation of this residue does not have an effect on SiaP function. Thr-64 is a completely buried residue and is not directly involved in an interaction with the sugar. Mutation of a small buried residue to a much larger residue close to the binding site will have the effect of introducing a number of conformational changes, resulting in affected binding. Likewise, the T64K mutation had partial complementation; however, T64R showed no complementation.

Neither Asp-49 nor Arg-70 make direct interaction with Neu5Ac at O-7. They are involved, however, in interface stabilization on ligand binding. Both of these residues interact with residues in the other domain in the ligand-bound form. The result of our D49A mutation taken together with only partial complementation of R70A suggests that these two residues are crucial to stabilize the closed form of the protein.

The phenotype of the R127A and R127K mutants was further examined by mass spectrometry. As expected, the R127A mutant failed to sialylate LOS, similar to the phenotype of the *siaP* mutant. Interestingly, the LOS glycoform profile of the R127K mutant (see Fig. 3*E*) showed a population of two competing sets of higher mass glycoforms as follows: those that are consistent with the wild-type strain and terminate with sialic acid (Fig. 3*A*), and others with the same masses as the glycoforms present in the *siaP* mutant (see Fig. 3*B*) that lack sialic acid and are extended by a hexose sugar. Many of these higher mass LOS glycoforms have not been reported previously. These glycoforms appear to represent two alternative biosynthetic pathways. The first involves capping of the oligosaccharide branch with sialic acid under normal conditions where intracellular sialic acid is present, and the second leads to extended higher mass glycoforms formed in the absence of intracellular sialic acid. In cases where there are low intracellular sialic acid

TABLE 5

Comparison of SiaP to DctP family proteins

SiaP sequence was aligned with selected DctP family ESR proteins using ClustalW.

Gene name	Organism	Identity	Similarity	Putative solute
		%	%	
COG1638	*H. somnus* 2336	71	85	Neu5Ac
PM1709	*P. multocida* PM70	71	85	Neu5Ac
FN1472	*F. nucleatum* ATCC 25586	65	80	Neu5Ac
VVA1202	*Vibrio vulnificus* YJ016	47	63	Neu5Ac
VC1779	*V. cholerae* El Tor N16961	46	62	Neu5Ac
PBPRA2281	*Photobacterium profundum* SS9	46	62	Neu5Ac
SMb20295	*Sinorhizobium meliloti* 1021	35	53	Neu5Ac
RSP0910	*Rhodobacter sphaeroides* 2 4 1	26	42	C_4-dicarboxylates
YiaO	*E. coli* K12-MG1655	26	48	Pentose

levels, a mixture of the two pathways is manifest as seen with the R127K mutant.

The affinity of SiaP for Neu5Ac was determined using ITC. We calculated a K_d of 28 nM and a stoichiometry of binding of 1:1. Severi *et al.* (14) reported that the SiaP from *H. influenzae* RM118 had a stoichiometry of 1:1 and a K_d of 120 nM, similar to our values. The low K_d value indicates a high affinity interaction between SiaP and Neu5Ac. The failure of SiaP to bind KDO, an eight-carbon keto sugar, is indicative of the high specificity of this binding protein. Although our binding affinities are in the same range reported by Severi *et al.* (14), we used ITC to measure the data. ITC allows for the analysis of complete thermodynamic parameters of the binding reaction. The interaction of SiaP with Neu5Ac and Neu5Gc are both enthalpically driven with substantial unfavorable contributions from entropy. This is not surprising because of the fact that the binding of SiaP to Neu5Ac is stabilized by several hydrogen bonds with favorable enthalpic contributions and has a large number of buried water molecules, which results in a significant unfavorable entropic contribution to the binding.

Different types of sialic acid transporters from *E. coli* and *Haemophilus ducreyi* have been described (27, 28). Instead of the TRAP transporter found in NTHi, *E. coli* uses a major facilitator type transporter encoded by the *nanT* gene (27). The major facilitator family of transporters does not utilize ESR proteins, which may impact the relative binding affinities of the NanT and SiaPT transporters. The use of a high affinity sialic acid TRAP transporter either indicates that the level of free sialic acid is limited within the host sites colonized by NTHi, provides NTHi with an advantage over competing organisms, or both. *H. ducreyi* does not possess a TRAP transporter homologous to SiaPT but instead uses an ATP-binding cassette transporter for the import of sialic acid (28). Like the TRAP transporter of NTHi, ATP-binding cassette transporters use ESR proteins; however, the energy for transport is provided by the hydrolysis of an ATP molecule. The presence of an ESR associated with the *H. ducreyi* transporter (SatABCD) indicates that this transporter likely has a higher binding affinity than NanT and may be similar to SiaPT.

We have previously demonstrated that *siaT*, encoding the membrane-spanning component of a TRAP transporter, is involved in sialic acid import in NTHi strains 2019 and 7502 (12). This TRAP system is composed of a membrane transporter and a binding protein with homologs in many bacterial species. The primary sequences of SiaP and other members of the DctP family of solute-binding proteins were aligned (Table 5 and supplemental Fig. S1). The comparison suggests three major groups based on sequence similarity and the putative substrate: sialic acid, C_4-dicarboxylates, and pentose sugars. Homologs that putatively bind sialic acid have highly conserved regions. Each homolog was encoded by a gene found in the vicinity of a *nanA* homolog. The percent similarity of SiaP to other members of this group ranged from 53 to 85% (Table 5). Most variability between these proteins was observed in the signal peptide at the N terminus and ~80 residues at the C terminus. A number of highly conserved, positively charged residues were demonstrated to be involved in solute binding. Sequence analysis indicates that one of the more divergent SiaP homologs in this group is from *Vibrio cholerae.* Our studies indicate that this homolog also binds Neu5Ac with an affinity of about 300 nM.[5]

H. influenzae is a member of the family Pasteurellaceae, which is a diverse family of proteobacteria. This family includes pathogens of a number of vertebrates, including humans. *Haemophilus somnus* and *P. multocida,* both members of this family, are very closely related to *H. influenzae*. They are both veterinary pathogens, but *P. multocida* can be a human pathogen when associated with animal bite wounds. Both *H. somnus* and *P. multocida* incorporate sialic acid into their LOS, live in a very similar biological niche, and share a similar sialobiology with *H. influenzae*. All are mucosal pathogens infecting the mucin/sialic acid-rich environment of the respiratory and genital tracts. All incorporate sialic acid into their LOS, can degrade sialic acid to hexosamines, and yet are unable to synthesize sialic acid. To acquire sialic acid, they have adopted almost identical sialic acid uptake systems using a TRAP transport system to move sialic acid from their environment into the cytoplasm. Both have a sialic acid-binding protein with high homology to SiaP (Table 5). *P. multocida* has several neuraminidases that can cleave sialic acid from α2–3 and α2–6 linkages (29). The mechanism by which *H. influenzae* acquires sialic acid from the respiratory mucosal environment has not yet been resolved. *F. nucleatum* subsp. *vincentii* is another species shown to possess a TRAP transport system with a binding protein that shares high homology with SiaP (30, 31). Like the Pasteurellaceae described above, this bacterium produces a neuraminidase and incorporates sialic acid into its LPS, yet it appears to lack a homolog of Neu5Ac synthase (30, 31). It also lives in the sialic acid-rich environment of the human oral cavity.

[5] C. Cho, J. Mowers, N. Coussens, A. Zaleski, M. A. Apicella, and S. Ramaswamy, unpublished data.

Taken together, these data strongly suggest that high affinity, acidic sugar-binding proteins have evolved to scavenge a nutrient that is important for the survival of these species in their respective environments. Inhibition of this binding protein may serve as a potential target for antimicrobial activity. The characterization of the binding interactions presented in this study will be beneficial to future drug design.

Acknowledgments—Use of the IMCA-CAT beamline 17-ID (or 17-BM) at the Advanced Photon Source was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with the Center for Advanced Radiation Sources at the University of Chicago.

REFERENCES

1. Rao, V. K., Krasan, G. P., Hendrixson, D. R., Dawid, S., and St. Geme, J. W., III (1999) *FEMS Microbiol. Rev.* **23,** 99–129
2. Mandrell, R. E., McLaughlin, R., Aba Kwaik, Y., Lesse, A., Yamasaki, R., Gibson, B., Spinola, S. M., and Apicella, M. A. (1992) *Infect. Immun.* **60,** 1322–1328
3. Masoud, H., Moxon, E. R., Martin, A., Krajcarski, D., and Richards, J. C. (1997) *Biochemistry* **36,** 2091–2103
4. Phillips, N. J., Apicella, M. A., Griffiss, J. M., and Gibson, B. W. (1992) *Biochemistry* **31,** 4515–4526
5. Schweda, E. K., Hegedus, O. E., Borrelli, S., Lindberg, A. A., Weiser, J. N., Maskell, D. J., and Moxon, E. R. (1993) *Carbohydr. Res.* **246,** 319–330
6. Hood, D. W., Makepeace, K., Deadman, M. E., Rest, R. F., Thibault, P., Martin, A., Richards, J. C., and Moxon, E. R. (1999) *Mol. Microbiol.* **33,** 679–692
7. Hood, D. W., Cox, A. D., Gilbert, M., Makepeace, K., Walsh, S., Deadman, M. E., Cody, A., Martin, A., Mansson, M., Schweda, E. K., Brisson, J. R., Richards, J. C., Moxon, E. R., and Wakarchuk, W. W. (2001) *Mol. Microbiol.* **39,** 341–350
8. Figueira, M. A., Ram, S., Goldstein, R., Hood, D. W., Moxon, E. R., and Pelton, S. I. (2007) *Infect. Immun.* **75,** 325–333
9. Mansson, M., Bauer, S. H., Hood, D. W., Richards, J. C., Moxon, E. R., and Schweda, E. K. (2001) *Eur. J. Biochem.* **268,** 2148–2159
10. Jones, P. A., Samuels, N. M., Phillips, N. J., Munson, R. S., Jr., Bozue, J. A., Arseneau, J. A., Nichols, W. A., Zaleski, A., Gibson, B. W., and Apicella, M. A. (2002) *J. Biol. Chem.* **277,** 14598–14611
11. Greiner, L. L., Watanabe, H., Phillips, N. J., Shao, J., Morgan, A., Zaleski, A., Gibson, B. W., and Apicella, M. A. (2004) *Infect. Immun.* **72,** 4249–4260
12. Allen, S., Zaleski, A., Johnston, J. W., Gibson, B. W., and Apicella, M. A. (2005) *Infect. Immun.* **73,** 5291–5300
13. Vimr, E., Lichtensteiger, C., and Steenbergen, S. (2000) *Mol. Microbiol.* **36,** 1113–1123
14. Severi, E., Randle, G., Kivlin, P., Whitfield, K., Young, R., Moxon, R., Kelly, D., Hood, D., and Thomas, G. H. (2005) *Mol. Microbiol.* **58,** 1173–1185
15. Muller, A., Severi, E., Mulligan, C., Watts, A. G., Kelly, D. J., Wilson, K. S., Wilkinson, A. J., and Thomas, G. H. (2006) *J. Biol. Chem.* **281,** 22212–22222
16. Alexeyev, M. F. (1995) *BioTechniques* **18,** 52, 54, and 56
17. Johnston, J. W., Zaleski, A., Allen, S., Mootz, J. M., Armbruster, D., Gibson, B. W., Apicella, M. A., and Munson, R. S. (2007) *Mol. Microbiol.* **66,** 26–39
18. Pflugrath, J. W. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 1718–1725
19. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
21. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
22. Vriend, G. (1990) *J. Mol. Graphics* **8,** 52–56
23. Dam, J., and Schuck, P. (2005) *Biophys. J.* **89,** 651–666
24. Gibson, B. W., Engstrom, J. J., John, C. M., Hines, W., and Falick, A. M. (1997) *J. Am. Soc. Mass Spectrom.* **8,** 645–658
25. Andreoni, J., Kayhty, H., and Densen, P. (1993) *J. Infect. Dis.* **168,** 227–231
26. Davies, D. G., Parsek, M. R., Pearson, J. P., Iglewski, B. H., Costerton, J. W., and Greenberg, E. P. (1998) *Science* **280,** 295–298
27. Martinez, J., Steenbergen, S., and Vimr, E. (1995) *J. Bacteriol.* **177,** 6005–6010
28. Post, D. M., Mungur, R., Gibson, B. W., and Munson, R. S., Jr. (2005) *Infect. Immun.* **73,** 6727–6735
29. Mizan, S., Henk, A., Stallings, A., Maier, M., and Lee, M. D. (2000) *J. Bacteriol.* **182,** 6874–6883
30. Kapatral, V., Anderson, I., Ivanova, N., Reznik, G., Los, T., Lykidis, A., Bhattacharyya, A., Bartman, A., Gardner, W., Grechkin, G., Zhu, L., Vasieva, O., Chu, L., Kogan, Y., Chaga, O., Goltsman, E., Bernal, A., Larsen, N., D'Souza, M., Walunas, T., Pusch, G., Haselkorn, R., Fonstein, M., Kyrpides, N., and Overbeek, R. (2002) *J. Bacteriol.* **184,** 2005–2018
31. Kapatral, V., Ivanova, N., Anderson, I., Reznik, G., Bhattacharyya, A., Gardner, W. L., Mikhailova, N., Lapidus, A., Larsen, N., D'Souza, M., Walunas, T., Haselkorn, R., Overbeek, R., and Kyrpides, N. (2003) *Genome Res.* **13,** 1180–1189
32. Campagnari, A. A., Gupta, M. R., Dudas, K. C., Murphy, T. F., and Apicella, M. A. (1987) *Infect. Immun.* **55,** 882–887