The non-typeable *Haemophilus influenzae* major adhesin Hia is a dual function lectin that binds to human-specific respiratory tract sialic acid glycan receptors.

John M Atack[1]*, Christopher J Day[1], Jessica Poole[1], Kenneth L. Brockman[2#], Jamie R L Timms[1], Linda E. Winter[3], Thomas Haselhorst[1], Lauren O Bakaletz[2], Stephen J Barenkamp[3†], and Michael P Jennings[1]*

[1]Institute for Glycomics, Griffith University, Gold Coast, Queensland, 4222, Australia

[2]Center for Microbial Pathogenesis, The Research Institute at Nationwide Children's Hospital and The Ohio State University College of Medicine, Columbus, OH 43205, USA

[3]Department of Pediatrics, Saint Louis University School of Medicine, and the Pediatric Research Institute, Cardinal Glennon Children's Medical Center, Saint Louis, MO 63104, USA

[#] current address: Department of Microbiology & Immunology, Medical College of Wisconsin, Milwaukee, Wi, USA

Running title: Hia binds human-specific glycans during colonisation

* to whom correspondence should be addressed – John M Atack or Michael P Jennings; Institute for Glycomics, Griffith University, Gold Coast, Queensland 4215, Australia, Tel: (+61) 07 56780580 (JMA) or (+61) 07 55527050 (MPJ); Fax: (+61) 07 55527050; E-mail: j.atack@griffith.edu.au or m.jennings@griffith.edu.au

[†] Deceased

This manuscript is dedicated to the memory of Dr. Stephen J. Barenkamp M.D. who passed away 17th March 2019.

Abstract

NTHi is a human-adapted pathogen that colonises the human respiratory tract. Strains of NTHi express multiple adhesins, however there is a unique, mutually exclusive relationship between the major adhesins Hia and HMW1/2. Approximately 25% of NTHi strains express Hia, a phase-variable autotransporter protein, and which has a critical role in colonisation of the host nasopharynx. The remaining 75% of strains express HMW1/2. Previous work has shown that the HMW1 and HMW2 proteins mediate binding to 2,3- and 2,6-linked sialic acid glycans found in the human respiratory tract. Here we show that that the high affinity binding domain of Hia, binding domain 1 (BD1) is responsible for binding to α2,6-sialyllactosamine glycans. BD1 is highly specific for glycans that incorporate the form of sialic acid expressed by humans, *N*-acetylneuraminic acid (Neu5Ac). We further show that Hia has lower affinity binding activity for 2,3-linked sialic acid and that this binding activity is mediated via a distinct domain. Thus, Hia with its dual binding activities functionally mimics the combined activities of the HMW1 and 2 adhesins. In addition, we show that Hia has a role in biofilm formation by strains of NTHi that express the adhesin. Knowledge of the binding affinity of a major NTHi adhesin, and putative vaccine candidate, will direct and inform development of future vaccines and therapeutic strategies for this important pathogen.

Importance

Host-adapted bacterial pathogens like NTHi have evolved specific mechanisms to colonize their restricted host niche. Relatively few of the adhesins expressed by NTHi have been characterized as regards their binding affinity at the molecular level. In this work we show that the major NTHi adhesin, Hia, preferentially binds to Neu5Ac-α2,6-sialyllactosamine, the form of sialic acid expressed in humans. The receptors targeted by Hia in the human airway mirror those targeted by influenza A virus and indicates the broad importance of sialic acid glycans as receptors for airway pathogens.

Introduction

Non-typeable *Haemophilus influenzae* (NTHi) is a human-adapted pathogen, responsible for multiple acute and chronic infections of the respiratory tract, including otitis media (OM) (1) community acquired pneumonia (2), and chronic obstructive pulmonary disease (COPD) exacerbations (3). Each year there are 31 million new cases of the most severe form of OM, chronic suppurative OM, are diagnosed (4), 60% of whom suffer an associated hearing loss. Globally, there are over 700 million cases of acute OM every year (4); in the USA alone, each year there are ~25 million episodes of acute OM, >13 million antibiotic prescriptions, and public health costs estimated at $3- $5 billion (5, 6). According to WHO estimates, approximately 65 million people have moderate to severe COPD. Over 3 million people died of COPD in 2005, which corresponded to 5% of all deaths globally (7). Invasive disease caused by NTHi has increased significantly in recent years, in part due to vaccines against *Haemophilus influenzae* type b, and *Streptococcus pneumoniae* (8). At present, there is no effective vaccine against NTHi.

NTHi is commonly carried the human nasopharynx asyptomatically. Many bacterial pathogens express outer-surface proteins that target specific host molecules to allow them to adhere to and persist in specific niches in the host. Examples of bacterial adhesins recognising particular host proteins include the type IV pilus of *Neisseria gonorrhoeae*, which recognises host integrins (9); the type IV pilus of NTHi, which recognizes ICAM1 (10); the curli pili of S*almonella enterica*, which binds host TLR2 receptors; and the FimH protein of uropathogenic *Escherichia coli*, which binds to mannosylated glycoproteins (11). Many bacteria also express virulence factors that belong to the auto-transporter protein family. These proteins have a diverse array of functions including adhesion to host surfaces (12). Auto-transporter proteins are characterised by a large barrel-like C-terminal domain which inserts into the outer-membrane, forming a pore through which the N-terminal effector

78 portion passes to reach the extracellular environment (13, 14). NTHi express many
79 autotransporter proteins (15) that fulfil a variety of roles in NTHi pathobiology. One of these
80 autotransporters, Hia, is an adhesin that is expressed by approximately 25% of NTHi strains
81 (16). The remaining ~75% of NTHi strains express the HMW1/2 proteins (17), which have
82 previously been demonstrated to be involved in adhesion of NTHi to human cells (18). It is
83 unclear why strains encode genes for Hia or HMW but never both. The HMW1 protein binds
84 to host cell glycans as cellular receptors, specifically α2,3-sialyllactosamine (2-3 SLN) (19).
85 We recently demonstrated that HMW2, which is ~65% identical to HMW1, binds the related
86 glycan α2,6-sialyllactosamine (2-6 SLN), with high specificity for 2-6 SLN containing *N*-
87 acetylneuraminic acid (Neu5Ac), the form of sialic acid expressed by humans (20).
88 Intriguingly, 2-3 SLN is found mainly in the lower human respiratory tract, whereas 2-6 SLN
89 is found throughout the entire respiratory tract, but predominates in the upper airway (21). It
90 has previously been demonstrated that Hia is required for adherence to Chang epithelial cells
91 (22), and we have demonstrated that Hia is required for colonisation of the host nasopharynx
92 (23). However, the cellular receptor for Hia is currently unknown. We hypothesized that Hia
93 may also recognize host-specific glycans found in the human respiratory tract. In the current
94 study we present an investigation to identify and characterized the Hia cellular receptor.
95
96 **Results**
97 *Hia is a lectin that recognizes Neu5Ac-α2,6-lactosamine (2-6 SLN-Ac) with high affinity.*
98 In order to determine whether Hia had glycan binding activity, we cloned and over-expressed
99 Hia from NTHi strain R2866, in *E. coli* BL21. Heterologous over-expression of Hia in *E. coli*
100 was used previously to investigate Hia binding activity (22). Hia over-expression was
101 confirmed by Western blot and whole cell ELISA (Supplementary Figure 1). The glycan
102 binding ability of *E. coli* strain BL21 cells expressing Hia (BL21-Hia) was compared to wild

type BL21 cells, using glycan array analysis. The background binding of BL21 only was subtracted from BL21-Hia in order to deduce the glycans bound in an Hia-dependent manner. A subset of the identified glycans were characterised for their binding affinity to BL21-Hia using surface plasmon resonance (SPR; Table 1). These studies demonstrated that Hia bound to a number of sialylated glycans, with the greatest affinity for Neu5Ac-α2,6-lactosamine (2-6 SLN-Ac), with a disassociation constant (K_D) of 185 nM. A comparison of the binding affinity of Hia to matched glycan pairs containing either a terminal *N*-acetylneuraminic acid (Neu5Ac; the only form expressed in humans) or *N*-glycolylneuraminic acid (Neu5Gc; which is expressed in most mammals), showed that Hia preferentially binds to structures containing a terminal Neu5Ac (Table 1), with a ~7-fold preference for 2,6-SLN-Ac over Neu5Gc-α2,6-lactosamine (2,6 SLN-Gc) (185 nM vs 1.39 μM; Table 1). Whilst some binding to 2,3 SLN-Ac (2.03 μM; Table 1) was observed, this occurred with approximately 11-fold lower affinity than for 2-6 SLN-Ac (185 nM).

Modelling shows key interactions between BD1 residues D618 and A620 and the Neu5Ac moiety of 2-6 SLN-Ac

The Hia protein has previously been shown to contain high- and low-affinity host cell binding domains (BD) termed BD1 and BD2 respectively (22, 24). BD1 and BD2 are proposed to bind a common, but unknown cellular receptor (24). Hia BD1 consists of amino acids 541-714 inclusive (22), with residues in the BD1 shown to be essential for binding to Chang epithelial cells when Hia is expressed in *E. coli* (22). To determine the molecular basis of the interactions between Hia BD1 and 2-6 SLN-Ac, we carried out molecular docking studies using the previously published Hia BD1 structure (22). All docking structures of 2-6 SLN-Ac with HiaBD1 indicated interaction of the ligand at the interface of chain A and chain C of HiaBD1. Figure 1 shows a bound structure of 2-6 SLN-Ac that represents a sialic acid-

specific binding mode with the negatively charged carboxylate group of the Neu5Ac residue engaging in strong electrostatic interaction with R674. The glycerol side chain of the sialic acid moiety of 2-6 SLN-Ac plays an important role as it engages in hydrogen bonds with D618 and A620. Importantly, the high flexibility of the α(2-6)-linkage of 2-6 SLN-Ac allows the coordination of the lactosamine disaccharide moiety. In addition, our docking studies indicate that residue R674 is involved in coordinating 2-6 SLN-Ac in all potential 25 docked conformations.

Hia BD1 is the site of high affinity interactions with the cellular receptor 2-6 SLN-Ac.

Using purified Hia BD1 (aa 514-714 inclusive) (22) we investigated BD1 binding specificity using SPR. Table 1 shows that Hia BD1 binds with high affinity and specificity to 2-6 SLN-Ac, with a K_D of 64.9 nM ± 5.6. This value is in a similar range to the affinity we observe with full-length Hia (185 nM ± 59.9). Hia BD1 interacts with 2-6 SLN-Gc with ~1000 lower affinity (61.28 μM ± 9.1; see Table 1) than with 2-6 SLN-Ac. In order to determine the specific region of BD1 responsible for the interaction with 2-6 SLN-Ac, we constructed a peptide library of BD1 aa 541-714, consisting of peptides of 15 amino acids in length, overlapping consecutive peptides by 10 aa each (for example, peptide one consisted of residues 541-555; peptide two of residues 546-560, etc; Table 2). We used these peptides to block the interaction between BL21-Hia and 2-6 SLN-Ac using an SPR competition assay. Using this methodology, we show that a peptide comprised of twenty amino acid residues containing both D618 and A620 (p16+17; residues 616-635) blocks 100% of the interaction between BL21-Hia and 2-6 SLN-Ac (Table 2). Peptide 16 and Peptide 17 individually result in blocking of 95% and 85% of interactions, respectively, between BL21-Hia and 2-6 SLN-Ac (Table 2). Peptides flanking the region of 16+17 (peptide 15 = aa 611-625; peptide 18 = 626-640; Table 2) only block ~50% of interactions, with no other peptide 15mer of BD1

153 blocking interactions between BL21-Hia and 2-6 SLN-Ac (data not shown). Residues D618
154 and A620 were previously shown to be key for binding to host cells, as when these residues
155 were mutated (D618K and A620R), binding was lost (22). Our blocking studies provide
156 strong evidence that additional residues, and likely secondary structure around these residues
157 that can only form in the 20mer p16+17, mediate direct interaction between 2-6 SLN-Ac and
158 Hia, leading to high-affinity binding.
159 In order to confirm these findings, we generated recombinant Hia with the single mutations
160 D618K and A620R, and a double mutant of Hia lacking both of these residues
161 (D618K/A620R double). SPR analysis was used to compare the binding of this panel of Hia
162 mutants with wild type Hia and BD1, using the same subset of glycans (see Table 3). These
163 findings demonstrated that the A620R Hia mutant and the D618K/A620R Hia double mutant
164 (all located in BD1) completely lose the ability to bind 2-6 SLN-Ac, while still maintaining
165 binding to 2-3 SLN-Ac. Collectively, these data demonstrated that the binding site of 2-3
166 SLN-Ac is not BD1, and confirmed the role of BD1 in binding specificity to 2-6 SLN-Ac.
167
168 *Hia is involved in interactions between NTHi and epithelial cells*
169 In order to demonstrate a biological role for Hia in attachment of NTHi to host epithelium,
170 we performed adherence assays using Chang epithelial cells. Prior to carrying out these
171 adherence assays, we confirmed 2-6 SLN was localized on the surface of these cells using
172 Dylight 649 Conjugated SNA, a lectin specific for 2-6 SLN (Figure 2A). Following treatment
173 with sialidase to remove sialylated glycans, 2-6 -SLN was no longer detected on the cell
174 surface by SNA (Figure 2A). Using NTHi strain R2866 that expressed Hia (wild type R2866;
175 R2866 *hia*+), and an isogenic mutant lacking Hia (R2866 *hia::tet*), we showed that the ability
176 of NTHi to adhere to Chang cells decreased when NTHi lacked Hia. Wild type R2866 is

177 unable to bind Chang cells treated with sialidase, which removes sialylated glycan structures
178 (Figure 2B).
179
180 *Residues D618 and A620 are critical to the interaction of Hia with Chang cells*
181 In order to determine the contribution of the key 2-6 SLN-Ac interacting residues (D618,
182 A620), and residue R674, indicated as important from our modelling studies, we carried out
183 adherence assays using a Chang epithelial cell model (23) to determine relative adherence of
184 *E. coli* BL21 strains expressing wild type (wt) Hia and our panel of Hia point mutants.
185 Adherence of *E. coli* BL21 cells to Chang cells was significantly greater when cells
186 expressed wt Hia (14.44% adherence, Figure 3) compared to control cells that did not express
187 Hia (empty BL21; 1.26% adherence; P = 0.0007). BL21 that express the Hia D618K/A620R
188 double mutant exhibited an approximately 4.5-fold decrease in relative adherence (3.21%
189 adherence; P = 0.002) compared to BL21 that express wt Hia. BL21 that expressed Hia
190 R674A, showed an approximate 2-fold decrease in relative adherence compared to cells that
191 express wt Hia (8.4% adherence), but this was not statistically significant compared to cells
192 expressing wt Hia (P = 0.06). These data indicated that the interaction between Hia and 2-6
193 SLN-Ac is critical to bacterial interactions with epithelial cells, and demonstrate the key
194 contribution of residues D618 and A620 of Hia BD1 in mediating this interaction.
195
196 *Expression of Hia in NTHi results in larger more robust biofilms by strains where the hia*
197 *gene is present*
198 The role of Hia in biofilm formation by two NTHi strains (R2866 and strain 11; both
199 encoding the *hia* gene) was tested using our well defined static biofilm model for NTHi (25).
200 Biofilms were formed for 24 hours at 37°C. Both strain R2866 and strain 11 formed much
201 larger biofilms when Hia was expressed (*hia*+) compared to when it was absent, as assessed

by confocal microscopy (Figure 4). NTHi that expressed Hia (*hia*+) formed biofilms with significantly more biomass (P = <0.0001 strain R2866 Figure 4A; P = <0.01 strain 11; Figure 4B) and were significantly thicker (P = <0.0001 strain R2866 Figure 4A; P = <0.05 strain 11; Figure 4B) compared to those formed by strains that did not express Hia (*hia::tet*). Descriptively, biofilms formed by NTHi that expressed Hia were significantly denser, and had a lawn-like architecture, compared to those formed by the respective isogenic mutant strain that did not express Hia. Biofilms of the two *hia::tet* isogenic mutant strains were significantly rougher (e.g. had a greater difference in overall surface height and topography) with dense tower-like regions of bacteria surrounded by open water channels (indicated by black in the representative top-down images) (Figure 4). The differences in bacterial distribution within the biofilm are apparent from the area occupied by layer (AOL) graphs (Figure 4). The area occupied by layer is a calculation of the amount (or percentage) of bacterial biomass that is present within each 1-μm optical section of the biofilm taken from the base of the biofilm to the top. These data are plotted wherein the layer closest to the glass surface is at the bottom of the y axis and the top of the biofilm (farthest from the surface) is at the top of the y axis. The relative shift of the blue lines to the right and upward, compared to the red lines (Figure 4), indicated that biofilms formed by NTHi that express Hia are substantially taller and more-dense than those formed by NTHi that do not express Hia. These results indicated that the Hia adhesin was a critical determinant of biofilm structure and organization in these strains, possibly due to increased inter-bacterial associations.

Discussion

In this work, we have demonstrated that the NTHi adhesin Hia is a lectin, with high specificity for host-specific glycans. Hia mediates high-affinity binding to 2-6 SLN-Ac. Molecular modelling studies using the crystal structure of Hia BD1 (22) in complex with 2-6

SLN-Ac showed that Hia residues D618 and A620, and to some extent R674, were critical to this interaction. We experimentally confirmed our modelling using a diverse and comprehensive array of complementary *in vitro* studies. Using a combination of *E. coli* expressing Hia, purified HiaBD1, and a peptide library derived from BD1, we determined that residues D618 and A620 of Hia are required for the high affinity interaction between Hia and 2-6 SLN-Ac. Interestingly, our SPR data also confirmed that Hia recognises 2-3 SLN-Ac, but this interaction was approximately 10-fold lower than that for 2-6 SLN-Ac. However, using the HiaBD1 protein, we confirmed that the interaction of Hia with 2-3 SLN-Ac is not mediated by BD1, which is consistent with previous findings which proposed that Hia contains two binding domains (22). Therefore, contrary to previous work which stated that BD1 and BD2 bind the same ligand (24), we have shown that the two binding domains of Hia interact with distinct ligands; BD1 with 2-6 SLN-Ac and BD2 with 2-3 SLN-Ac. Moreover, in both cases, the preference is for the form of sialic acid expressed by humans (Neu5Ac). This offers an insight into the evolution of NTHi as a human-specific pathogen: although Neu5Gc and Neu5Ac (the precursor to Neu5Gc) can be expressed by most mammals, humans only make Neu5Ac linked glycans, due to a mutation in the CMAH gene responsible for the conversion of Neu5Ac to Neu5Gc (26). As it appears that Hia preferentially binds Neu5Ac linked glycans over Neu5Gc linked glycans, this finding strongly suggests that Hia has evolved to preferentially bind glycans most likely to be present in its human host. Preference for the Neu5Ac form of sialic acid has also been observed in the utilisation of Neu5Ac for macromolecular biosynthesis of bacterial cell surface glycans in NTHi (27), which further supported the central role of sialic acid in the adaptation of NTHi to its human host.

NTHi strains that do not possess the gene encoding *hia* instead encode genes for and express the adhesins HMW1 and HMW2 (18). Previous work has demonstrated that NTHi strains

252 either encode genes for Hia or HMW1/2, but never both, with approximately 75% strains
253 expressing HMW1/HMW2, and the remaining 25% expressing Hia. We have recently
254 demonstrated that HMW2 preferentially binds 2-6 SLN-Ac (20), whereas HMW1 has a
255 preference for 2-3 SLN structures (19). However, HMW1 showed no preference for either
256 Neu5Ac or Neu5Gc containing structures, and had a much lower affinity than HMW2 for 2-6
257 SLN-Ac (20). Therefore two distinct NTHi adhesins, Hia and HMW1/2, that show a discrete
258 lineage distribution in the NTHi population, have evolved to bind the same subset of glycans:
259 Hia binds 2-6 SLN-Ac preferentially over 2-3 SLN-Ac; HMW2 specifically binds 2-6 SLN-
260 Ac; HMW1 binds a broader range of 2-3 and 2-6 linked glycan structures compared to
261 HMW2, but with lower overall affinity (Figure 5). Production of both 2-6 SLN-Ac and 2-3
262 SLN-Ac linked glycan structures is found throughout the human airway, but these receptors
263 are not evenly distributed throughout the upper and lower airway; 2-6 SLN-Ac is found in
264 both the upper and lower respiratory tract, while 2'-3' SLN-Ac linked glycans are found
265 predominantly in the lower respiratory tract (28-30). Thus, NTHi strains that express Hia or
266 HMW1/2 are able to adhere to the entire human respiratory tract. It would be interesting to
267 study the distribution of NTHi strains expressing only HMW1 or HMW2 – based on their
268 differing binding affinities, it may be that HMW2-only expressing strains are more prevalent
269 in upper respiratory tract infections, whereas HMW1-only expressing strains are more
270 disposed to infecting the lower respiratory tract. It is also intriguing to note that the binding
271 specificity of Hia (and HMW1/2) mirrors perfectly that of human influenza A viruses (31-
272 33), indicating that specificity for human specific glycans has evolved in both viral and
273 bacterial human-adapted airway pathogens. A common strategy to block these interactions
274 may therefore serve as a general therapy for both these types of infection. It is well known
275 that infection with the influenza virus predisposes individuals to colonization and infection
276 by *Streptococcus pneumoniae*; although there are likely multiple aspects behind the increased

277 severity of pneumococcal disease following influenza virus infection (34), it is thought that
278 desialylation of the host epithelia by the viral neuraminidase allows for more efficient
279 colonization by the pneumococcus (35). This desialylation in turn increases the susceptibility
280 of these patients to pneumococcal pneumonia following influenza virus infection. The
281 lethality of the 1918 'Spanish flu' outbreak was mainly due to secondary infections by
282 bacterial pathogens, including both *S. pneumoniae* and *H. influenzae*, with up to 95% of the
283 deaths from this pandemic attributable to secondary bacterial infections (36, 37). Thus, the
284 sharing of common receptors indicates the possibility of direct interaction between influenza
285 virus and NTHi occurs during co-infection, and may provide a fruitful area of investigation in
286 the study of the dynamics of this polymicrobial interaction.
287 As well as playing a key role in host colonisation through recognition of human specific
288 glycan structures, we have demonstrated that Hia also plays a key role in biofilm formation
289 by NTHi. Biofilm formation by NTHi has been shown to increase the resistance of bacteria to
290 antibiotics (38), and killing by neutrophils (39) when compared to planktonic counterparts.
291 Increased resistance of bacteria within biofilms to antibiotics has been demonstrated for a
292 number of major human pathogens, such as *Pseudomonas aeruginosa* (40) and
293 *Staphylococcus aureus* (41). Biofilm formation also plays a key role in NTHi disease
294 pathologies, such as middle ear infections (42) and exacerbations of cystic fibrosis (43).
295 Using two diverse strains of NTHi, we showed that Hia is a critical determinant of biofilm
296 development and structure and that the potential to block Hia function through knowledge of
297 its specific binding affinities could play a key role in targeting biofilm formation during
298 disease caused by NTHi.
299 To summarize, we have provided an in-depth characterization of the binding affinity
300 of the NTHi adhesin Hia, by determining the major human cellular receptors it has evolved to
301 bind and by demonstrating the molecular basis of these interactions. We also demonstrate

302 that Hia has a role in biofilm formation by NTHi, and therefore likely contributes to
303 antibiotic resistance and chronicity by this mechanism. Knowledge of the factors required by
304 NTHi to colonise and cause disease will be key to developing both vaccines and treatments
305 against this organism. Our demonstration that the major NTHi adhesins HMW1 and HMW2
306 bind the same host glycans as Hia (20), and that these adhesins are expressed by nearly 100%
307 of NTHi strains is a key step towards the development of a rationally designed vaccine
308 against NTHi, and to the production of novel treatments against this pathogen.
309
310 **Materials and methods**
311 *Bacterial strains and growth conditions*
312 NTHi strains expressing Hia have been described previously [R2866 (44) and strain 11 (45)].
313 NTHi strains were routinely grown in Brain-Heart Infusion (BHI) broth supplemented with
314 1% hemin and 20 µg NAD^+/mL (sBHI), and grown aerobically at 37°C with 150 rpm
315 shaking. For solid media, 1.5% agar was added to sBHI broth. sBHI media were
316 supplemented with tetracycline (5 µg/mL) as required. Plates were grown at 37°C in
317 atmosphere containing 5% CO_2. *Escherichia coli* were grown using Luria-Bertani (LB)
318 media at 37°C, and supplemented with tetracycline (5 µg/mL) as required.
319
320 *Generation of a hia knockout mutant in NTHi strains R2866 and 11*
321 A region of NTHi R2866 chromosome containing the *hia* promoter and the ATG start and
322 5′ region of the gene were generated by PCR using primer pair hia-UP-F / hia-UP-R, and
323 cloned into pGEM Teasy according to manufacturer's instructions (Promega) to generate
324 plasmid vector Teasy::hiaUP. Inverse PCR was used to linearise this vector at the *hia* start
325 codon using primers hia-INV-F / hia-INV-R. A tetracycline resistance cassette, encoding
326 *tetM*, was generated from plasmid vector pGEM-TetM(B) using M13F and M13R primers.

This was cloned into the linearised Teasy::hiaUP vector so the gene was in the same orientation as the *hia* gene, and orientation confirmed using PCR and sequencing. This vector was designated Teasy::hiaUP::TetM. Following linearization with NgoMIV (New England Biolabs), DNA was transformed into NTHi strains R2866 and strain 11 using the MIV method (46). Transformants were selected on BHI media containing 5 µg tetracycline /mL, and positive colonies confirmed by sequencing and Western blotting using an anti-Hia monoclonal antibody 1F4 (47). Strains were designated as R2866 or strain 11 *hia::tet.*

Cloning and over-expression of full length Hia in E. coli

Primers HiaFULL-F and HiaFULL-R (Supplementary Table 1) were used to amplify full length wild-type *hia* (R2866_0725) including the signal sequence (residues 1-49) from genomic DNA prepared from NTHi strain R2866. PCR was carried out using KOD hot-start polymerase (EMD Millipore) according to manufacturer's instructions. Following digestion with BspHI and XhoI (NEB) and clean up, DNA was cloned into pET15b digested with NcoI and XhoI. the resulting plasmid was designated pET15b::Hia. Following confirmation of correct clones by sequencing, over-expression was carried out in *E. coli* BL21 following by inducing cells with 0.5 mM IPTG overnight at 37°C with 200 rpm shaking. Over-expression was confirmed by Western blot as previously described (23) using anti-Hia monoclonal antibody 1F4 (47). Whole cell ELISA using standard methods (48) with modifications as previously described (23) and starting with 1:10,000 dilution of primary antibody anti-Hia monoclonal antibody 1F4 confirmed the location of Hia at the cell surface.

Generation of Hia point mutants for over-expression

Inverse PCR was carried out using primer pairs designed to introduce point mutations as previously described and used here to abrogate binding of *E. coli* expressing Hia to Chang

cells (22). D618K, A620A and a 618/620 double mutant were generated using specific forward primers Hia-D618K-F, Hia-A620R-F, or Hia-618/620-double-F and common reverse primer Hia-618/620-R. A R674A mutant was generated using primer pair Hia-R674A-F and Hia-R674A-R. All inverse PCR reactions were carried out using KOD hot-start polymerase (EMD Millipore) according to manufacturer's instructions, and a plasmid mini-prep (Qiagen) of pET15b::Hia as template. All primer sequences are listed in Supplementary Table 1. Clones were sequenced using primers either side of the point mutation Hia-screen-F and Hia-screen-R using BigDye 3.1 according to manufacturer's instructions (Thermo Fisher), and sequenced at Australian Genome Analysis Facility (AGRF, Brisbane, Australia). Over-expression was carried out as described above for Hia wild-type, and cell surface localization confirmed using whole cell ELISA as above.

Over-expression and purification of Hia BD1

Primers to clone Hia binding domain 1 (BD1; amino acid residues 540-714) were designed based on those from (22). HiaBD1-F and HiaBD1-R were used to amplify BD1 from NTHi strain R2866 genomic DNA using KOD hot-start polymerase (EMD Millipore) according to manufacturer's instructions. Following digestion with NdeI and BamHI (NEB) and clean up, DNA was cloned into pET15b digested with the same enzymes. This strategy would clone the gene in frame with an N-terminal 6xHis tag for purification. The resulting plasmid was designated pET15b::HiaBD1. Over-expression was carried out in *E. coli* BL21 following by inducing cells with 0.5 mM IPTG overnight at 37C with 200 rpm shaking. Cells were pelleted, resuspended in 1x binding buffer (50 mM NaPO4, 300 mM NaCl, pH7.4), lysed using 0.1 mm glass beads and a Tissue lyser (Qiagen) for 30 mins at 50 osc^{-1} min^{-1}. Purification was carried out using TALON gravity flow resin in 1x binding buffer. Protein was eluted from the resin using step wise concentrations of imidazole in 1x binding buffer

(10-500mM imidazole), fractions analysed by SDS PAGE, and fractions containing pure BD1 pooled and concentrated using centrifugal concentrators (Millipore, 10kDa cut-off). Pure concentrated BD1 was buffer exchanged into 1x phosphate buffered saline (1x PBS) using the same centrifugal concentrators. Protein was analysed by SDS PAGE, and quantified using an extinction coefficient of 8480 M^{-1} cm^{-1} and MW of 20471.43 Da (based on the sequence of Hia BD1+6xHis tag)

Glycan array

Glycan array slides were printed using OPEpoxy (CapitalBio) activated substrates with the glycan library as previously described (49) using an ArrayIt Spotbot Extreme 3 contact printer with solid metal pins. The glycan array binding experiments were performed and analysed as previously described (50). Briefly, 1 mL of OD_{600} 0.2 *E. coli* BL21 with heterologous expression of Hia wild-type or Hia point-mutants in PBS were incubated with 15μL of 50 μM Bodipy methylester for 15 minutes, centrifuged at 900g for 3 minutes and the pellet washed 3 times with PBS to removed excess dye. The cell pellet was resuspended in 1 mL of array PBS (1x PBS containing 1 mM $CaCl_2$ and 1 mM $MgCl_2$) and 300 μL was applied to the slide in a 65 μL gene frame without a coverslip. Slides were washed three times for 2 minutes in array PBS, dried by centrifugation and scanned and analysed using the Scan Array Express software package (Perkin Elmer) and Microsoft Excel for statistical analysis. Binding of Hia was defined as both above the background of the slide (cut off of 550 fluorescence units) and 2-fold and significantly ($p<0.05$) above the background of empty vector BL21 binding to the array by Student's unpaired t-test of fluorescence of BL21 empty vector controls versus Hia expressing strains of BL21. All glycan array binding data is presented in Supplementary Table 2 and the MIRAGE compliant information is listed in Supplementary Table 3.

Surface Plasmon Resonance (SPR)

Surface plasmon resonance (SPR) experiments of the full length wild-type Hia expressed on the surface of *E. coli* BL21was performed using a GE Biacore T100 system and a Series S C1 sensor chip using a modification of methods previously described (51, 52). *E. coli* (BL21 strains expressing full length wild-type Hia, point mutants, or BL21 only) cells at $1x10^6$ bacteria/mL were immobilised to the chip surface following the C1 NHS/EDC method template with a contact time of 900 seconds at a flow rate of 5 µL/minute in 10 mM sodium acetate pH 5.5. Interaction of glycans with the bacteria was performed using five-fold serial dilutions with maximum concentration of 20 µM on first analysis and 5 µM when affinities were better defined using single cycle kinetics in 1x PBS pH 7.4 at 20 µL/minute with 60 second contact time and a final dissociation time of 10 minutes. A blank ethanolamine immobilisation was used as a control flow cell and 1x PBS pH 7.4 was used as the zero-concentration control. Regeneration of the bacterial surface was performed by flushing 10 mM Tris 1mM EDTA over the surface for 5 minutes at 30 µL/minute. Affinities (K_D) were determined using the Biacore T100 evaluation software analysis of double baseline subtracted data. All interactions were measured in triplicate and displayed plus/minus 1 standard deviation of the measured mean.

Purified Hia BD1 protein was immobilised onto a CM5 sensor chip amine capture on a Biacore T100 with a contact time of 600 seconds at a flow rate of 5 µL/minute in 10 mM sodium acetate pH 4.5. Glycans were run at the optimised concentrations outlined above. With the analysis performed as outlined above.

Peptide binding region identification was performed using a modified version of a previously described method (53), competition assays using immobilised Hia expressing cells and flowed peptides and glycan. *E. coli* BL21 expressing full-length wild-type Hia were

immobilised onto a H1 sensor chip using a ForteBio Pioneer using a contact time of 720 seconds at a flow rate of 10 µL/minute in 1x PBS at $1x10^8$ bacteria/mL Assays were set up using the NextStep injection feature as previously described (54, 55) with combinations of 2,6-SLN, Hia overlapping peptides and PBS as the negative control used to determine the Hia region that interacted with 2,6-SLN. Analysis was performed using the QDat analysis software package.

Distribution of 2,6-SLN on Chang cells

Chang cells ($1x10^4$ cells) in 100 µL total volume were seeded into Transwell inserts with a 6.5 mm diameter and 0.4 um pore size (Corning Incorporated, Corning, New York). Cell culture medium (DMEM, 10% heat-inactivated calf serum, 2 mM L-glutamine) was replaced daily until cells reached confluence, 2 to 3 days. The apical surface of the cells was rinsed twice with sterile DPBS and then incubated with 0.1 units of neuraminidase (Sigma) in 100 µL DPBS or with DPBS alone for 2 hours at 37°C. The cells were then rinsed twice with DPBS and 100 µL of 10 µg/mL Dylight 649 Conjugated *Sambucus nigra* (EY Laboratories, San Mateo, California) was added to the apical surface and incubated for 15 minutes. The cells were rinsed twice with DPBS, incubated with 3 units of Alexa Fluor™ 594 Phalloidin (Thermo Fisher Scientific, Waltham, Massachusetts) for 30 minutes and rinsed twice. The membrane of the Transwell was excised and then mounted with ProLong™ Glass Antifade Mountant with NucBlue™ Stain (Thermo Fisher Scientific, Waltham, Massachusetts). Images were captured on a LSM 700 laser scanning microscope and rendered with Zeiss Zen software (Zeiss).

Biofilm formation

Biofilms were formed by NTHI cultured within chambers of eight-well-chambered coverglass slides (Thermo Scientific, Waltham, MA) as described previously (56). Briefly, mid-log phase cultures of NTHI strains were diluted with sBHI. NTHI were inoculated at 4 × 10^4 c.f.u. in 200 µl total volume per well and slides were incubated at 37°C with 5% atmospheric CO2. Biofilms were grown for a total of 24 h, with the growth medium replaced after 16 h. To visualize, biofilms were stained with LIVE/DEAD BacLight stain (Life Technologies) and fixed overnight in fixative (1.6% paraformaldehyde, 2.5% glutaraldehyde, 4% acetic acid in 0.1 M phosphate buffer, pH 7.4). Fixative was replaced with saline before imaging with a Zeiss 510 Meta-laser scanning confocal microscope. Images were rendered with Zeiss Zen software.

Analysis of biofilm formation and architecture

Z-stack images acquired at 63x with a Zeiss 510 Meta-laser scanning confocal microscope were analyzed by COMSTAT2 to determine biomass (µm3/µm2), average thickness (µm), roughness (Ra) and percent area occupied by layers. Area occupied by layer was plotted as percent bacterial biomass coverage per 1 µm optical section from the base of the biofilm. Standard error of the mean for replicate biofilms was calculated for each individual layer with GraphPad Prism version 6.0 (GraphPad Software, San Diego, CA).

Adherence of NTHI strain R2866 to Chang cells

Chang cells (1x10^4 cells) in 100 µL total volume were seeded into Transwell inserts with a 6.5 mm diameter and 0.4 um pore size (Corning Incorporated, Corning, New York). Cell culture medium (DMEM, 10% heat-inactivated calf serum, 2 mM L-glutamine) was replaced daily until cells reached confluence, 2 to 3 days. The apical surface of the cells was rinsed twice with sterile DPBS. Strain R2866 strains were added to the apical surface of the Chang

476 cells at an MOI of 100 in 50 µL of DPBS and incubated for 30 minutes at 37°C. The cells
477 were rinsed twice with DPBS, incubated with 3 units of Alexa Fluor™ 594 Phalloidin
478 (Thermo Fisher Scientific, Waltham, Massachusetts) for 30 minutes and rinsed twice. The
479 membrane of the Transwell was then excised and mounted with ProLong™ Glass Antifade
480 Mountant with NucBlue™ Stain (Thermo Fisher Scientific, Waltham, Massachusetts).
481 Images were captured on a LSM 700 laser scanning microscope and rendered with Zeiss Zen
482 software (Zeiss).
483
484 *Adherence assays with BL21 strains*
485 *E. coli* BL21 strains expressing either wild-type Hia, D618K/A620R double mutant, R674A
486 mutant, or containing the empty pET15b expression vector were grown to mid-log (OD_{600} =
487 ~0.6) in LB broth containing ampicillin (100 µg/mL), and cfu calculated by serially diluting.
488 Approximately $5x10^5$ cfu (100 µl) each mid-log culture were added to wells of a 24 well
489 plate containing differentiated Chang cells as described previously (23). Six wells were used
490 per strain per experiment. Following addition of BL21, plates were incubated at 37°C for 2
491 hours. Following incubation, supernatant was removed, and non-adherent cells removed by
492 gentle washing with 1x PBS four times. Adherent cells were released by incubating with
493 0.05% Trypisn in 1x PBS for 10mins at room-temperature. Adherent bacteria were quantified
494 by serial dilution and plating. Percent adherence was calculated as the number of adherent cfu
495 in relation to the total input cfu per strain. Student's t-test was carried out using total number
496 of each output and comparing to wild-type cells. All data (total cfu and percent adherence) is
497 presented as Supplementary Data 1.
498
499 *Modelling interaction of HiaBD1 with 2-6 SLN*

Docking of 2-6 SLN to HiaBD1 was performed using the AutoDock Vina protocol (57) that has the highest scoring power among commercial and academic molecular docking programs (58) and is implemented in the YASARA Structure molecular modelling package (Ver. 16.46) (59). The docking experiment was set up by using the X-ray crystal structure of HiaBD1 (pdb code 1S7M (22) with a box centered D620 using a grid size of 50 Å × 50 Å 50 Å (x, y, z) covering chain C. A total number of 25 Vina docking runs were performed. The 3D topology of the 2-6 SLN glycan was generated using the carbohydrate builder available at GLYCAM-Web server (http://glycam.org) (60).

Acknowledgements

Work was funded by the Australian National Health and Medical Research Council (NHMRC) Program Grant 1071659 and Principal Research Fellowship 1138466 to MPJ, Project Grant 1099279 to JMA, an Australian Research Council Discovery Project 180100976 to JMA, a Garnett Passe and Rodney Williams Grant-in-Aid (Supplementation) to JMA, and a National Institutes of Health (NIH; USA) R01 Grant DC015688 to LOB and MPJ. We thank Patrick Azzari for a critical readthrough of, and valuable input to, the manuscript.

References

1. Haggard M. 2008. Otitis media: prospects for prevention. Vaccine 26 Suppl 7:G20-4.
2. Johnson RH. 1988. Community-acquired pneumonia: etiology, diagnosis, and treatment. Clin Ther 10:568-73.
3. Sethi S, Murphy TF. 2008. Infection in the pathogenesis and course of chronic obstructive pulmonary disease. N Engl J Med 359:2355-65.
4. Monasta L, Ronfani L, Marchetti F, Montico M, Vecchi Brumatti L, Bavcar A, Grasso D, Barbiero C, Tamburlini G. 2012. Burden of Disease Caused by Otitis Media: Systematic Review and Global Estimates. PLOS ONE 7:e36226.
5. A.A.P. 2004. Diagnosis and management of acute otitis media. Pediatrics 113:1451-65.
6. Alsarraf R, Jung CJ, Perkins J, Crowley C, Alsarraf NW, Gates GA. 1999. Measuring the indirect and direct costs of acute otitis media. Arch Otolaryngol Head Neck Surg 125:12-8.
7. Heinz E. 2018. The return of Pfeiffer's bacillus: Rising incidence of ampicillin resistance in Haemophilus influenzae. Microbial genomics 4:e000214.
8. Langereis JD, de Jonge MI. 2015. Invasive Disease Caused by Nontypeable Haemophilus influenzae. Emerg Infect Dis 21:1711-8.
9. Edwards JL, Apicella MA. 2005. I-domain-containing integrins serve as pilus receptors for *Neisseria gonorrhoeae* adherence to human epithelial cells. Cell Microbiol 7:1197-211.
10. Novotny LA, Bakaletz LO. 2016. Intercellular adhesion molecule 1 serves as a primary cognate receptor for the Type IV pilus of nontypeable Haemophilus influenzae. Cell Microbiol 18:1043-55.
11. Hanson MS, Brinton CC, Jr. 1988. Identification and characterization of E. coli type-1 pilus tip adhesion protein. Nature 332:265-8.
12. Benz I, Schmidt MA. 2011. Structures and functions of autotransporter proteins in microbial pathogens. Int J Med Microbiol 301:461-8.
13. Grijpstra J, Arenas J, Rutten L, Tommassen J. 2013. Autotransporter secretion: varying on a theme. Res Microbiol 164:562-82.
14. Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D. 2004. Type V protein secretion pathway: the autotransporter story. Microbiol Mol Biol Rev 68:692-744.
15. Spahich NA, St. Geme I, Joseph W. 2011. Structure and function of the Haemophilus influenzae autotransporters. Frontiers in Cellular and Infection Microbiology 1.
16. Barenkamp SJ, St Geme JW, 3rd. 1996. Identification of a second family of high-molecular-weight adhesion proteins expressed by non-typable Haemophilus influenzae. Mol Microbiol 19:1215-23.
17. St Geme JW, 3rd, Kumar VV, Cutter D, Barenkamp SJ. 1998. Prevalence and distribution of the hmw and hia genes and the HMW and Hia adhesins among genetically diverse strains of nontypeable *Haemophilus influenzae*. Infect Immun 66:364-8.
18. St Geme JW, Falkow S, Barenkamp SJ. 1993. High-molecular-weight proteins of nontypable Haemophilus influenzae mediate attachment to human epithelial cells. Proceedings of the National Academy of Sciences 90:2875-2879.
19. St Geme JW. 1994. The HMW1 adhesin of nontypeable Haemophilus influenzae recognizes sialylated glycoprotein receptors on cultured human epithelial cells. Infection and Immunity 62:3881-3889.

20. **Atack JM**, Day CJ, Poole J, Brockman KL, Bakaletz LO, Barenkamp SJ, Jennings MP. 2018. The HMW2 adhesin of non-typeable Haemophilus influenzae is a human-adapted lectin that mediates high-affinity binding to 2-6 linked N-acetylneuraminic acid glycans. Biochem Biophys Res Commun doi:10.1016/j.bbrc.2018.06.126.
21. Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y. 2006. Influenza virus receptors in the human airway. Nature 440:435.
22. Yeo HJ, Cotter SE, Laarmann S, Juehne T, St Geme JW, Waksman G. 2004. Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter The EMBO Journal 23:1245-1256.
23. Atack JM, Winter LE, Jurcisek JA, Bakaletz LO, Barenkamp SJ, Jennings MP. 2015. Selection and counter-selection of Hia expression reveals a key role for phase-variable expression of this adhesin in infection caused by non-typeable *Haemophilus influenzae*. J Infect Dis 212:645-53.
24. Laarmann S, Cutter D, Juehne T, Barenkamp SJ, St Geme JW. 2002. The Haemophilus influenzae Hia autotransporter harbours two adhesive pockets that reside in the passenger domain and recognize the same host cell receptor. Molecular Microbiology 46:731-743.
25. Brockman KL, Azzari PN, Branstool MT, Atack JM, Schulz BL, Jen FE-C, Jennings MP, Bakaletz LO. 2018. Epigenetic Regulation Alters Biofilm Architecture and Composition in Multiple Clinical Isolates of Nontypeable *Haemophilus influenzae*. mBio 9.
26. Varki NM, Varki A. 2007. Diversity in cell surface sialic acid presentations: implications for biology and disease. Laboratory Investigation 87:851.
27. Ng PSK, Day CJ, **Atack JM**, Hartley-Tassell LE, Winter LE, Marshanski T, Padler-Karavani V, Varki A, Barenkamp SJ, Apicella MA, Jennings MP. 2019. Nontypeable *Haemophilus influenzae* Has Evolved Preferential Use of N-Acetylneuraminic Acid as a Host Adaptation. mBio 10.
28. Baum LG, Paulson JC. 1990. Sialyloligosaccharides of the respiratory epithelium in the selection of human influenza virus receptor specificity. Acta Histochem Suppl 40:35-8.
29. Couceiro JNSS, Paulson JC, Baum LG. 1993. Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. Virus Research 29:155-165.
30. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk H-D. 2004. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proceedings of the National Academy of Sciences of the United States of America 101:4620-4624.
31. Connor RJ, Kawaoka Y, Webster RG, Paulson JC. 1994. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. Virology 205:17-23.
32. Matrosovich M, Tuzikov A, Bovin N, Gambaryan A, Klimov A, Castrucci MR, Donatelli I, Kawaoka Y. 2000. Early Alterations of the Receptor-Binding Properties of H1, H2, and H3 Avian Influenza Virus Hemagglutinins after Their Introduction into Mammals. Journal of Virology 74:8502-8512.
33. Ng PS, Böhm R, Hartley-Tassell LE, Steen JA, Wang H, Lukowski SW, Hawthorne PL, Trezise AE, Coloe PJ, Grimmond SM, Haselhorst T, von Itzstein M, Paton AW, Paton JC, Jennings MP. 2014. Ferrets exclusively synthesize Neu5Ac and express naturally humanized influenza A virus receptors. Nat Commun 5:5750.
34. Rudd JM, Ashar HK, Chow VT, Teluguakula N. 2016. Lethal Synergism between Influenza and Streptococcus pneumoniae. Journal of infectious pulmonary diseases 2:10.16966/2470-3176.114.

35. Nita-Lazar M, Banerjee A, Feng C, Amin MN, Frieman MB, Chen WH, Cross AS, Wang LX, Vasta GR. 2015. Desialylation of airway epithelial cells during influenza virus infection enhances pneumococcal adhesion via galectin binding. Mol Immunol 65:1-16.
36. Morens DM, Taubenberger JK, Fauci AS. 2008. Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness. The Journal of infectious diseases 198:962-970.
37. Rynda-Apple A, Robinson KM, Alcorn JF. 2015. Influenza and Bacterial Superinfection: Illuminating the Immunologic Mechanisms of Disease. Infection and immunity 83:3764-3770.
38. Slinger R, Chan F, Ferris W, Yeung SW, St Denis M, Gaboury I, Aaron SD. 2006. Multiple combination antibiotic susceptibility testing of nontypeable Haemophilus influenzae biofilms. Diagn Microbiol Infect Dis 56:247-53.
39. Juneau RA, Pang B, Weimer KE, Armbruster CE, Swords WE. 2011. Nontypeable Haemophilus influenzae initiates formation of neutrophil extracellular traps. Infect Immun 79:431-8.
40. Drenkard E. 2003. Antimicrobial resistance of Pseudomonas aeruginosa biofilms. Microbes Infect 5:1213-9.
41. Belbase A, Pant ND, Nepal K, Neupane B, Baidhya R, Baidya R, Lekhak B. 2017. Antibiotic resistance and biofilm production among the strains of Staphylococcus aureus isolated from pus/wound swab samples in a tertiary care hospital in Nepal. Annals of clinical microbiology and antimicrobials 16:15-15.
42. Hall-Stoodley L, Hu FZ, Gieseke A, Nistico L, Nguyen D, Hayes J, Forbes M, Greenberg DP, Dice B, Burrows A, Wackym PA, Stoodley P, Post JC, Ehrlich GD, Kerschner JE. 2006. Direct Detection of Bacterial Biofilms on the Middle-Ear Mucosa of Children With Chronic Otitis Media. JAMA 296:202-211.
43. Starner TD, Zhang N, Kim G, Apicella MA, McCray PB, Jr. 2006. Haemophilus influenzae forms biofilms on airway epithelia: implications in cystic fibrosis. Am J Respir Crit Care Med 174:213-20.
44. Nizet V, Colina KF, Almquist JR, Rubens CE, Smith AL. 1996. A virulent nonencapsulated Haemophilus influenzae. J Infect Dis 173:180-6.
45. Barenkamp SJ, Bodor FF. 1990. Development of serum bactericidal activity following nontypable Haemophilus influenzae acute otitis media. Pediatr Infect Dis J 9:333-9.
46. Herriott RM, Meyer EM, Vogt M. 1970. Defined nongrowth media for stage II development of competence in Haemophilus influenzae. J Bacteriol 101:517-24.
47. Winter LE, Barenkamp SJ. 2009. Antibodies specific for the Hia adhesion proteins of nontypeable *Haemophilus influenzae* mediate opsonophagocytic activity. Clin Vaccine Immunol 16:1040-1046.
48. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A laboratory manual, second edition. Cold Spring Harbour Laboratory Press.
49. Waespy M, Gbem TT, Elenschneider L, Jeck AP, Day CJ, Hartley-Tassell L, Bovin N, Tiralongo J, Haselhorst T, Kelm S. 2015. Carbohydrate recognition specificity of trans-sialidase lectin domain from *Trypanosoma congolense*. PLoS Negl Trop Dis 9:e0004120.
50. Day CJ, Tran EN, Semchenko EA, Tram G, Hartley-Tassell LE, Ng PSK, King RM, Ulanovsky R, McAtamney S, Apicella MA, Tiralongo J, Morona R, Korolik V, Jennings MP. 2015. Glycan:glycan interactions: High affinity biomolecular interactions that can mediate binding of pathogenic bacteria to host cells. Proceedings of the National Academy of Sciences 112:E7266-E7275.

51. Mubaiwa TD, Hartley-Tassell LE, Semchenko EA, Day CJ, Jennings MP, Seib KL. 2018. The Bexsero Neisseria meningitidis serogroup B vaccine antigen NHBA is a high-affinity chondroitin sulfate binding protein. Sci Rep 8:6512.
52. Tromp AT, Van Gent M, Abrial P, Martin A, Jansen JP, De Haas CJC, Van Kessel KPM, Bardoel BW, Kruse E, Bourdonnay E, Boettcher M, McManus MT, Day CJ, Jennings MP, Lina G, Vandenesch F, Van Strijp JAG, Jan Lebbink R, Haas PA, Henry T, Spaan AN. 2018. Human CD45 is an F-component-specific receptor for the staphylococcal toxin Panton-Valentine leukocidin. Nat Microbiol doi:10.1038/s41564-018-0159-x.
53. Poole J, Day CJ, Haselhorst T, Jen FE, Torres VJ, Edwards JL, Jennings MP. 2020. Repurposed Drugs That Block the Gonococcus-Complement Receptor 3 Interaction Can Prevent and Cure Gonococcal Infection of Primary Human Cervical Epithelial Cells. mBio 11.
54. Coll RC, Hill JR, Day CJ, Zamoshnikova A, Boucher D, Massey NL, Chitty JL, Fraser JA, Jennings MP, Robertson AAB, Schroder K. 2019. MCC950 directly targets the NLRP3 ATP-hydrolysis motif for inflammasome inhibition. Nat Chem Biol 15:556-559.
55. Hartley-Tassell LE, Awad MM, Seib KL, Scarselli M, Savino S, Tiralongo J, Lyras D, Day CJ, Jennings MP. 2019. Lectin Activity of the TcdA and TcdB Toxins of Clostridium difficile. Infect Immun 87.
56. Jurcisek JA, Dickson AC, Bruggeman ME, Bakaletz LO. 2011. *In vitro* biofilm formation in an 8-well chamber slide. Journal of Visualized Experiments doi:10.3791/2481:pii: 2481. doi: 10.3791/2481.
57. Trott O, Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455-61.
58. Wang Z, Sun H, Yao X, Li D, Xu L, Li Y, Tian S, Hou T. 2016. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. Phys Chem Chem Phys 18:12964-75.
59. Krieger E, Koraimann G, Vriend G. 2002. Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. Proteins 47:393-402.
60. Woods RJ. 2005-2020. GLYCAM Web. http://glycam.org. Accessed

Table 1 – Surface Plasmon Resonance analysis of glycan binding affinity of BL21-Hia and purified recombinant Hia-BD1.

	Hia	**HiaBD1**
2-3 SLN-Ac	2.03 μM ± 0.443	NCDI
2-3 SLN-Gc	10.1 μM ± 0.320	NCDI
2-6 SLN-Ac	185 nM ± 59.9	64.9 nM ± 5.6
2-6 SLN-Gc	1.39 μM ± 0.411	61.28 μM ± 9.1
SLeX-Ac	NCDI	19.9 μM ± 14.5
SLeX-Gc	NCDI	NCDI
LNnT	NCDI	NCDI
2-3 Ac-SLNnT	2.03 μM ± 1.49	NCDI
2-3 Gc-SLNnT	5.44 μM ± 1.11	NCDI
LNT	NCDI	NCDI
GM1-Ac	9.61 μM ± 3.85	NCDI
GM1-Gc	10.3 μM ± 5.63	NCDI

NCDI = No concentration dependent interaction. Indicates no binding between Hia expressing bacteria and the structure at a maximum concentration of 100 μM.

Table 2 - Surface Plasmon Resonance analysis of the blocking activity of peptides to interfer with the Hia : 2-6 SLN-Ac interaction.

	Sequence	Block 2-6 SLN-Ac
p15	DNLTKQN**D**D**A**YKGLT	54% ± 6.9
p16	QN**D**D**A**YKGLTNLDEK	95% ± 4.8
p17	YKGLTNLDEKGTDKQ	85% ± 8.6
p18	NLDEKGTDKQTPVVA	48% ± 1.7
p16/17 common	YKGLTNLDEK	95% ± 4.3
p16+17	QN**D**D**A**YKGLTNLDEKGTDKQ	100% ± 2.5

D618 and A620

Table 3 - Surface Plasmon Resonance analysis of glycan binding affinity of *E. coli* BL21 expressing wild-type Hia and Hia isogenic mutants.

	2-3 SLN-Ac	2-6 SLN-Ac
Hia wild-type	2.47 μM ± 0.8	110 nM ± 3.0
Hia D618K	842 nM ± 178	21 nM ± 1.0
Hia A620R	3.04 μM ± 0.7	NB
Hia D618K/A620 double	3.42 μM ± 1.2	NB

NB: No binding using a OneStep injection at 10 μM, indicates a K_D above 10 μM.

Figure 1 – molecular docking of Hia binding domain 1 (BD1) to 2-6 SLN. The previously published structure of Hia BD1 was used (Yeo et al, 2004); PDB accession number 1S7M. Docking structure of 2,6SLN-Ac into Hia (22). **A)** solid surface of 1S7M and **B)** magnified region of 1S7M shown as secondary structure and bound 2-6 SLN-Ac. Key amino acids are labelled.

Figure 2 – 2-6 SLN presence on Chang cells, and adherence of NTHi strain R2866 expressing Hia to Chang cells.

A) 2-6 Sialyl-N-acetyllactosamine expressed on the surface of Chang cells. Upper panel, top-down view of 2-6 SLN distribution on Change cells or cells pre-treated with neuraminidase. SNAi shown in white, phalloidin shown in red, nuclear DNA shown in blue. Scale bar, 25 µm. Lower panel, representative side view of an optical section through SNAi labelled Chang cells. SNAi shown in white, nuclear DNA shown in blue; **B) Chang cell R2866 adherence.** Adherence of wild type R2866 expressing Hia (*hia*+) and the R2866 *hia::tet* mutant to Chang cells. Left panels, Chang cell monolayer with phalloidin shown in red, nuclear DNA shown in blue. Middle panels, distribution of strain R2866 mutants that constitutively express GFP, shown in green. Bacteria that express Hia (*hia*+) bound markedly better to Chang cells than those that do not express Hia (*hia::tet*). Right panels, merged images that shows distribution of strain R2866 mutants across the surface of the Change cells. Scale bar, 100 µm.

Figure 3 - Percent adherence to Chang cells of *E. coli* BL21 strains expressing wild-type Hia or isogenic mutant variants. Percent adherence of each strain is calculated as adherent cfu following 2 hrs incubation / total input cfu. All raw data is presented in Supplementary

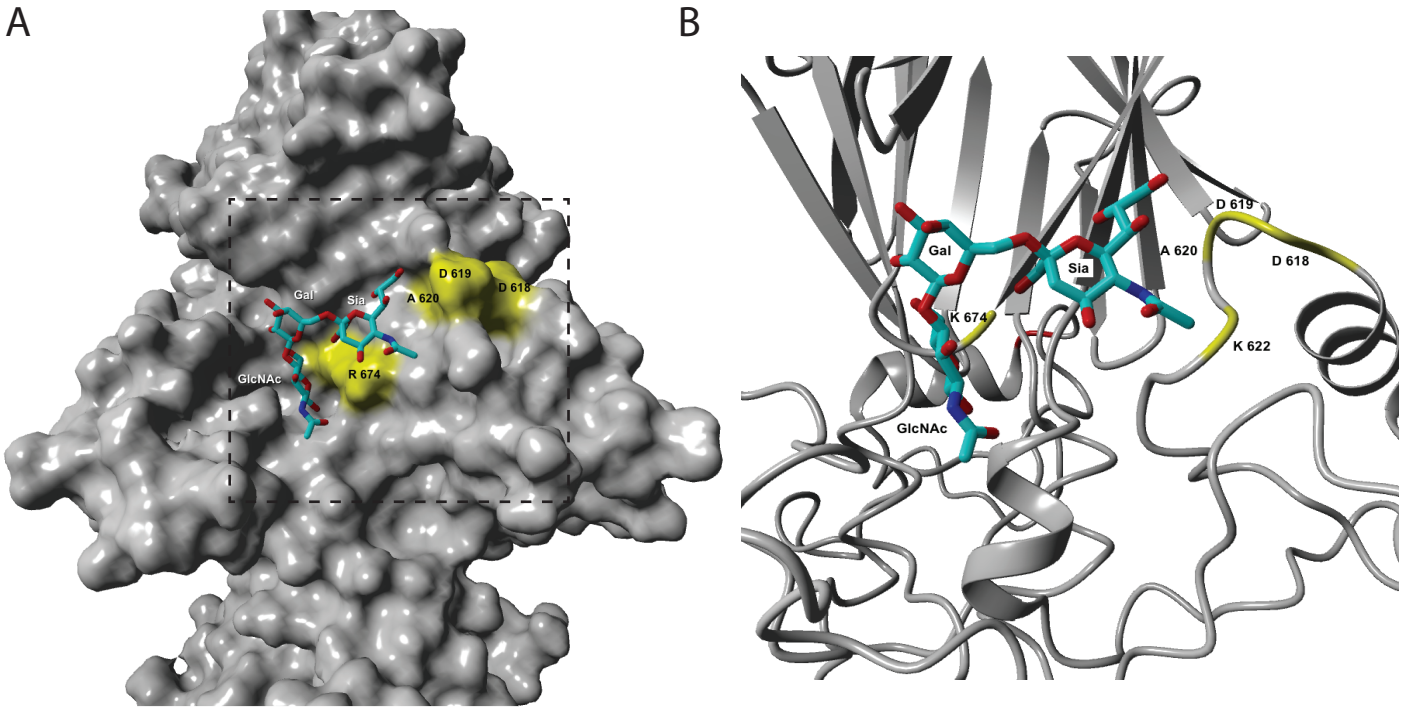
Data 1. * = P-value = <0.005. NSD = no significant difference. P-values calculated using Student's t-test

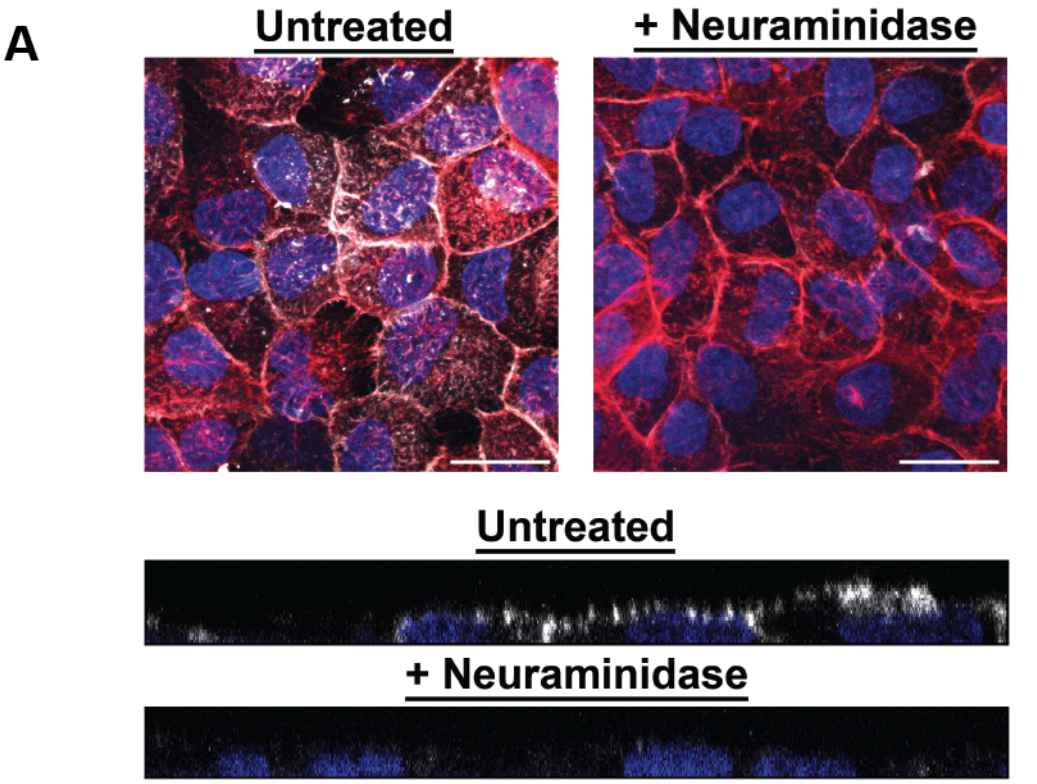
Figure 4 – biofilm formation by NTHi strains R2866 and strain 11

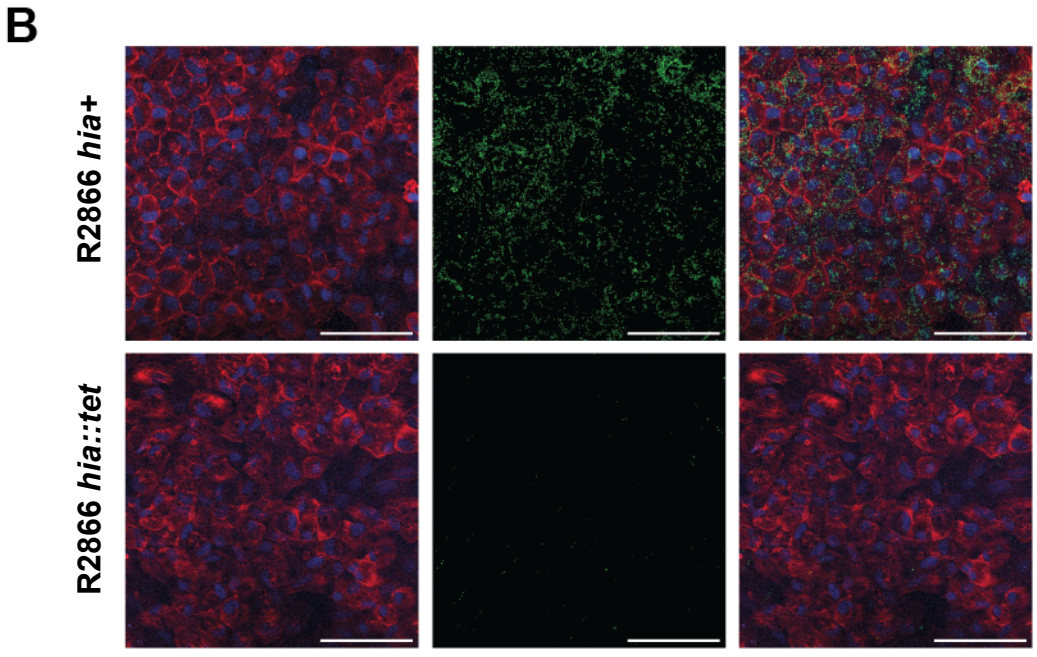
Representative orthogonal image renderings of biofilm formation by NTHI strain R2866 and strain 11 *hia*::*tet* and *hia*+ biofilms. Scale bars, 100 μm. Biomass, average thickness and roughness of *hia*::*tet* and *hia*+ biofilms grown for 24 hrs were analyzed by COMSTAT2 and values are shown as mean ± standard error of the mean. *p < 0.05, **p < 0.01, ****p < 0.0001, student's t-test. Average percent area occupied by bacteria at each individual 1 μm optical section ('layer') were determined. Dashed lines indicate standard error of the mean.

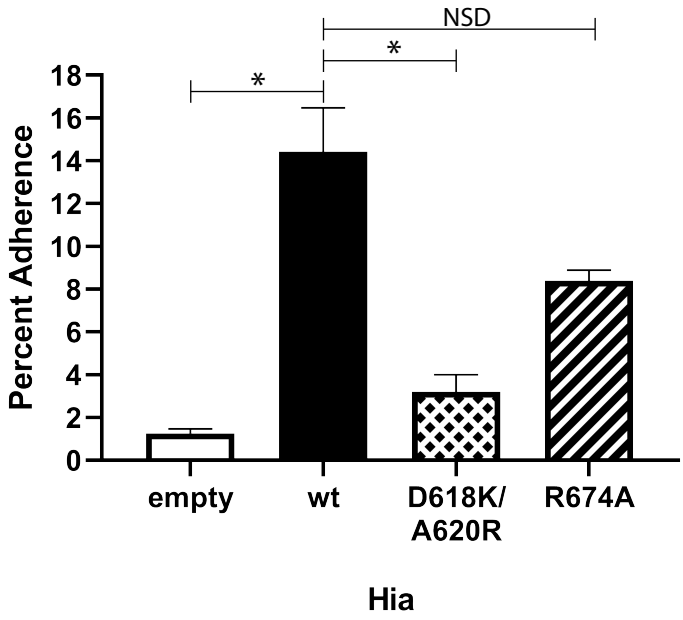
Figure 5 – illustration showing distribution of 2-6 SLN and 2-3 SLN in human airway. Hia and HMW2 both bind 2-6 SLN with high affinity, with both showing a very high preference for the human specific sialic acid Neu5Ac (2-6 SLN-Ac) over Neu5Gc (2-6 SLN-Gc). This means NTHi strains expressing either Hia or HMW2 are able to colonise the entire respiratory tract (upper in red, lower in blue). HMW1 preferentially binds 2-3 SLN, and with no affinity for Neu5Ac over Neu5Gc, which means NTHi strains only expressing HMW1 may have a preference for the lower respiratory tract (blue). Schematic diagram taken from GetDrawings.com (http://getdrawings.com/respiratory-system-with-label-drawing#respiratory-system-with-label-drawing-13.jpg) under a CC BY-NC 4.0 Licence.

Supplementary Figure 1 A) Western blot showing the over-expression of wild-type Hia in *E. coli* BL21; and **B)** Whole cell ELISA showing that Hia expressed in *E. coli* BL21 is located on the bacterial cell surface

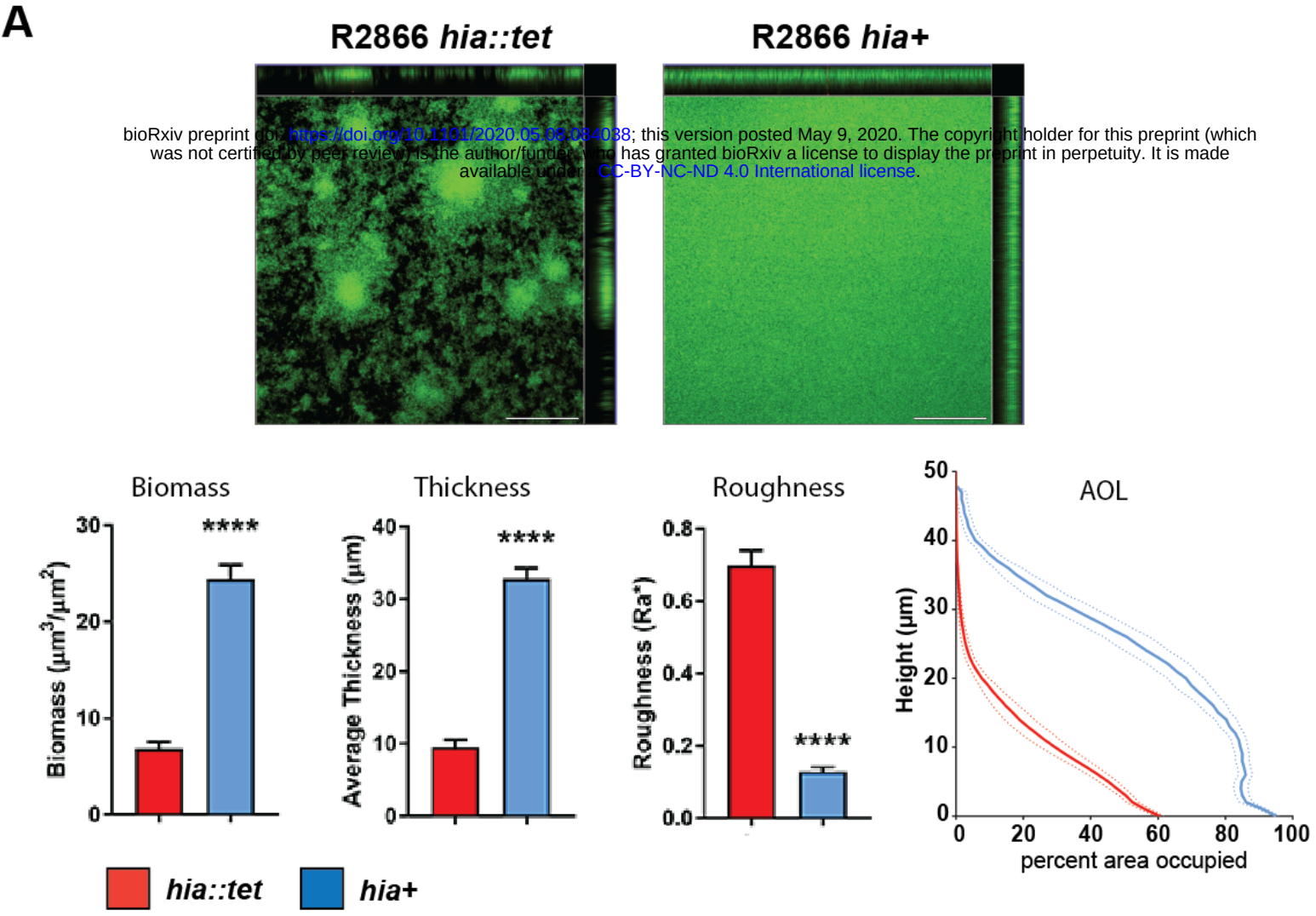








A



B