Structure of the Major Oligosaccharide from the Lipooligosaccharide of *Haemophilus ducreyi* Strain 35000 and Evidence for Additional Glycoforms†

William Melaugh,‡ Nancy J. Phillips,‡ Anthony A. Campagnari,§ Michael V. Tullius,‡ and Bradford W. Gibson*,‡

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446, and Infectious Disease Institute, State University of New York at Buffalo, Buffalo, New York 14215

Received May 20, 1994; Revised Manuscript Received August 12, 1994®

ABSTRACT: *Haemophilus ducreyi* is a sexually transmitted pathogen that colonizes the genital epithelium in humans, causing genital ulcers or chancroid. Its surface lipooligosaccharides (LOSs) have been shown to play a role in ulcer formation and may also be important in cell adhesion and invasion of host tissue. Earlier we presented a preliminary structure of the major LOS from strain 35000 that suggested the presence of terminal lactosamine [Melaugh, W., Phillips, N. J., Campagnari, A. A., Karalus, R., & Gibson, B. W. (1992) *J. Biol. Chem. 267*, 13434–13439]. We have now confirmed this structure and assigned the anomeric linkages by 2D NMR studies. In addition to this major structure, analysis by electrospray ionization mass spectrometry of both *O*-deacylated LOSs and the oligosaccharides released after treatment with mild acid indicates the presence of several other LOS glycoforms. These glycoforms constitute a series of both truncated and elongated analogs of the major oligosaccharide determined by NMR. One of these glycoforms exists as a smaller oligosaccharide corresponding to the major structure minus terminal galactose. Three other glycoforms appear as larger molecular weight species formed by the addition of phosphoethanolamine, *N*-acetylhexosamine, and *N*-acetylhexosamine plus hexose. Two sialylated glycoforms were also identified and subsequently confirmed by treatment with neuraminidase, but these glycoforms were not found in the released oligosaccharide pool due to the acid lability of sialic acid. This study clearly indicates that the LOSs from *H. ducreyi* strain 35000 exist as a heterogeneous population whose structures differ primarily in their phosphorylation states and terminal sugars and whose terminal glycan structures can resemble those of human antigens.

Haemophilus ducreyi, a Gram-negative human mucosal pathogen, is one of the principal causes of genital ulcer disease or chancroid (Johnson et al., 1988; Albritton, 1989; Morse, 1989). Chancroid is endemic to large parts of Asia, Africa, and India (Kibukamusoke, 1965; Weisner et al., 1983) and has recently made a comeback in the United States (Becker et al., 1987; Morse, 1989). Chancroid has also been identified as an independent risk factor in the transmission of AIDS, which makes it a potentially life-threatening sexually transmitted disease (Pepin et al., 1989). The prevalence of many antibiotic-resistant *H. ducreyi* strains is making treatment increasingly difficult, particularly in patients who also have HIV infection (Jessamine & Ronald, 1990). Clearly, other approaches are needed that will lead to the development of new therapeutic strategies to combat the spread of this disease.

Studies in rabbits and mice using both live and heat-killed organisms have shown that lipopolysaccharide (LPS[1]) is an important factor in establishing infection (Tuffrey et al., 1990; Campagnari et al., 1991; Purcell et al., 1991). Like the LPSs from *Neisseria* and other *Haemophilus* species (Jennings et al., 1983; Gibson et al., 1989; DiFabio et al., 1990; John et al., 1991; Phillips et al., 1992), the LPSs from *H. ducreyi* appear to lack the repeating *O*-antigenic side chains characteristic of LPSs from most Gram-negative enteric bacteria and are often referred to more correctly as lipooligosaccharides (LOSs). Monoclonal antibodies developed against gonococcal LOS, which recently have been found to contain structures that are similar or identical to the terminal tetrasaccharide of paragloboside (John et al., 1991; Yamasaki et al., 1991), also bind strongly to the LOSs of a large majority of *H. ducreyi* strains (Campagnari et al., 1990). These paragloboside structures are also partially sialylated (Gibson et al., 1993; Yamasaki et al., 1993). Given that paragloboside is a major component of human antigens, it has been suggested that the expression of these same or similar epitopes on bacterial surfaces is a form of host mimicry, perhaps enabling the organisms to attach to host cells and/or evade the immune system (Mandrell et al., 1988; John et al., 1991).

Despite substantial immunochemical data, very little structural data have been presented that might support the

† This work was supported by grants from the University of California University AIDS Research Program (R91SF243 and R93SF038, B.W.G.) and the National Institutes of Health (AI 30006, A.A.C.). We also acknowledge support for the Mass Spectrometry Facility through grants from the National Center for Research Resources (RR 01614; A. Burlingame, Director). The scientific contributions of W.M. and N.J.P. were equivalent and their relative order in authorship is arbitrary.

* Author to whom correspondence and reprint requests should be addressed: School of Pharmacy 926-S, 513 Parnassus Ave., University of California, San Francisco, CA 94143-0446. Telephone: (415) 476-5320. FAX: (415) 476-0688.

‡ University of California at San Francisco.

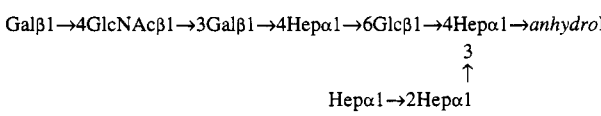
§ State University of New York at Buffalo.

® Abstract published in *Advance ACS Abstracts,* October 1, 1994.

[1] Abbreviations: 1D, one-dimensional; 2D, two-dimensional; COSY, 2D *J*-correlated spectroscopy; DQF-COSY, double-quantum-filtered COSY; ESI-MS, electrospray ionization mass spectrometry; Gal, galactose; GalNAc, *N*-acetylgalactosamine; Glc, glucose; GlcNAc, *N*-acetylglucosamine; Hep, L-*glycero*-D-*manno*-heptose or D-*glycero*-D-*manno*-heptose; Hex, hexose; HexNAc, *N*-acetylhexosamine; HOHAHA, 2D homonuclear Hartmann–Hahn spectroscopy; HPLC, high-performance liquid chromatography; Kdo, 3-deoxy-D-*manno*-octulosonic acid; LOS, lipooligosaccharide; LPS, lipopolysaccharide; LSIMS, liquid secondary ion mass spectrometry; $(M - nH)^{n-}$, deprotonated molecular ion; Neu5Ac, 5-*N*-acetylneuraminic acid or sialic acid; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; PEA, phosphoethanolamine; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis.

0006-2960/94/0433-13070$04.50/0

role of specific saccharides in these important biological processes. Previously, we had reported a preliminary structure for the major LOS from *H. ducreyi* strain 35000 that contained several unusual features, including the likely presence of terminal lactosamine (Melaugh et al., 1992). In this current report, we present the complete structure of this major oligosaccharide as determined by NMR studies, which now confirms the presence of lactosamine on the nonreducing terminus and other aspects of the preliminary structure:



Furthermore, chemical and mass spectrometric data are presented that verify the presence of sialic acid in the LOS from strain 35000, as well as additional glycoforms of lower abundances.

EXPERIMENTAL PROCEDURES

Isolation, Purification, and SDS–PAGE Analysis of LOSs. LOS from *H. ducreyi* strain 35000 was isolated using a modified phenol/water extraction procedure (Westphal & Jahn, 1965). Organisms were dried in acetone and suspended in a mixture of water/phenol, which was kept at 65 °C for 10–15 min. The mixture was cooled to 10 °C and centrifuged at 80000*g* for 45 min, and the top aqueous layer was saved. The bottom phenol layer was heated to 65 °C, hot water was added, and the mixture was shaken vigorously at 65 °C for 15 min. The mixture was again centrifuged, and the aqueous layers were combined and dialyzed against running water for 72 h. The LOS sample was then centrifuged for 8 h at 80000*g*. The pellet was resuspended in water, centrifuged at 105000*g* for 3 h, and lyophilized. The LOS was analyzed by SDS–PAGE on a 14% polyacrylamide gel, as we have previously described (Campagnari et al., 1990). The gels were silver-stained according to the method of Tsai and Frasch (1982). The apparent molecular weights of the major LOS bands were estimated by comparison with the relative migration of the LOS from *Neisseria gonorrhoeae* strain PID_2 (data not shown).

Preparation of Chemically Modified LOS. To prepare water soluble LOS that was more amenable to mass spectrometric analysis, ≈1 mg of LOS was *O*-deacylated under mild hydrazine conditions as previously described (Melaugh et al., 1992). To distinguish the presence of HexNAc (203 Da) from the isobaric mass of phosphate plus phosphoethanolamine (80 + 123 Da, respectively), a portion of the *O*-deacylated LOS was treated further with 48% aqueous HF to remove phosphate esters. Briefly, 100 μg of *O*-deacylated LOS was dissolved in 10 μL of 48% aqueous HF for 16 h at 4 °C. The resulting *O*-deacylated, dephosphorylated LOS was used as is after excess HF was removed in a vacuum desiccator.

Isolation and Purification of the Oligosaccharide Fraction. LOS from *H. ducreyi* strain 35000 (≈20 mg) was taken up in a solution of 1% acetic acid (2 mg/mL) and heated at 100 °C for 2 h. The hydrolysate was centrifuged at 5000*g* for 20 min at 4 °C, and the supernatant was removed. The pellet was washed twice with 1–5 mL of H_2O followed by centrifugation at 5000*g* for 20 min at 4 °C. The supernatant and washings were pooled and lyophilized. The resulting oligosaccharide fraction was dissolved in 300 μL of 50 mM pyridinium acetate (pH 5.2), centrifuge-filtered, and loaded onto two Bio-Gel P-4 columns (1.6 × ≈80 cm, <400 mesh, 30 °C) that were connected in series. Samples were eluted in the pyridinium acetate buffer at a flow rate of 10–12 mL/h. Fractions were collected every 10 min and dried on a Speed-Vac concentrator. Chromatograms were recorded and stored with a Shimadzu C-R3A Chromatopac integrator.

For the determination of oligosaccharide heterogeneity, an alternate gel filtration high-performance liquid chromatography (HPLC) system was employed consisting of a Waters 6000A HPLC pump and two 600 × 7.5 mm Bio-Sil TSK-125 columns (Bio-Rad) connected in series. Samples were eluted in 50 mM pyridinium acetate (pH 5.2) at a flow rate of 1 mL min^{-1}. Detection was as described above. Fractions were collected in 0.5–1-mL volumes.

Mass Spectrometric Analysis. Fractions from the gel filtration column were first analyzed by liquid secondary ion mass spectrometry (LSIMS) as described in detail elsewhere (Phillips et al., 1990). *O*-Deacylated LOS and the unfractionated oligosaccharide pool were also analyzed by electrospray ionization mass spectrometry (ESI-MS) on a VG/Fisons Platform quadrupole mass spectrometer in the negative-ion and/or positive-ion modes (Gibson et al., 1993). The total oligosaccharide pool obtained after mild acid hydrolysis of LOS was first dissolved in water to make a 10 μg/μL solution. Approximately 1 μL of this solution was then mixed with 4 μL of the electrospray running solvent and injected via a Rheodyne injector into a running solvent of H_2O/acetonitrile (1/1, v/v) containing 1% acetic acid at a flow rate of 10 μL/min. *O*-Deacylated LOS was analyzed in a similar fashion. Mass calibration was carried out with an external carbohydrate reference standard using the supplied VG/Fisons software.

Nuclear Magnetic Resonance Spectroscopy. All 1D and 2D ^1H NMR spectra were recorded on a GE GN-500 spectrometer at 25 °C. The major oligosaccharide (M_r 1677, approximately 0.5 mg) was lyophilized several times from 99.96% D_2O (Aldrich) and then dissolved in 0.3 mL of 99.996% D_2O (Cambridge Isotope Laboratories). A trace of acetone was added to the sample as an internal reference (δ 2.225). The phase-sensitive double-quantum-filtered COSY (DQF-COSY) spectrum (Rance et al., 1983) was acquired with time-proportional phase incrementation of the first pulse (Redfield & Kuntz, 1975; Marion & Wüthrich, 1983). The data matrix consisted of 800 × 1024 data points, acquired with 16 scans per t_1 value and a spectral width of ±2000 Hz. The matrix was apodized with a 45-deg-shifted sine-bell window function in both dimensions and zero-filled to produce a 2000 × 2000 real data matrix. The 2D homonuclear Hartmann–Hahn (HOHAHA) spectra were obtained using the MLEV-17 sequence (Bax & Davis, 1985) and quadrature detection (States et al., 1982). A spectral width of ±2000 Hz was used with 90- and 120-ms relay mixing times. Data matrices of 560 × 2048 and 512 × 1024 data points were acquired for the 90- and 120-ms spectra, respectively. Both spectra were recorded with 16 scans per t_1 increment, and the matrices were zero-filled and apodized with Gaussian window functions in both dimensions to give 2000 × 2000 real points. The 2D nuclear Overhauser effect (NOESY) spectrum was acquired in the phase-sensitive mode (States et al., 1982), using a 350-ms mixing time and 3.0-s interscan delay. The spectral width was ±2000 Hz, and the data matrix contained 512 × 1024 data points, with 16 scans per t_1 value. Zero-filling and Gaussian apodization functions were used in both dimensions to give a final 2000 × 2000 matrix. All data processing was done on a VAX computer with a UNIX operating system, using programs developed and modified at the UCSF NMR

facility as previously described (Basus et al., 1988), including some recent improvements (M. Day and D. Kneller, unpublished data).

Sialic Acid Linkage Determination. LOS was dissolved in 0.1% (v/v) Triton X-100 to give a 10 mg/mL solution. A 200-μg (20-μL) aliquot of LOS was then added to 70 μL of 50 mM NaOAc buffer (pH 5.5), followed by the addition of 10 milliunits of Newcastle disease virus sialidase dissolved in 10 μL of 50 mM NaOAc at pH 5.5. This final reaction mixture was incubated at 37 °C for 10 h, and small aliquots of the reaction mixture (5 μL) were removed at various time points (t = 4, 6, 7, and 10 h) for the determination of the released sialic acid (Neu5Ac), as described elsewhere (Mandrell et al., 1992). After 7 h, a larger 10-μL aliquot was removed, 100 milliunits of *Clostridium perfringens* sialidase dissolved in 10 μL of 50 mM NaOAc (pH 5.5) was added, and the reaction mixture was again incubated at 37 °C.

RESULTS

Our preliminary structure for the major oligosaccharide from *H. ducreyi* strain 35000 LOS was based on tandem mass spectrometric and methylation analyses, which yielded several possible structures (Melaugh et al., 1992):

Gal1→4GlcNAc1→(3Gal or 6Glc)1→(2/4)Hep1→(3Gal or 6Glc)1→(3/4)Hep1→*anhydro*Kdo
(3/4)
↑
Hep1→(2/4)Hep1

These proposed structures contain anhydro-Kdo on their reducing termini, which is formed by β-elimination of a phosphate group from the 4-position of intact Kdo during the acetic acid hydrolysis step (Auzanneau et al., 1991). The uncertainties in the sequence involve the placement of the dilinked hexoses and heptoses and the arrangement of the two oligosaccharide branches.

On the basis of its strong binding to monoclonal antibody 3F11, which recognizes the terminal region of the lacto-*N*-neotetraose epitope of paragloboside, Galβ1→4GlcNAcβ1→3Galβ1→4Glcβ1→ceramide (Mandrell et al., 1988; Campagnari et al., 1990), we originally suggested that the terminal trisaccharide of *H. ducreyi* 35000 LOS was probably Galβ1→4GlcNAcβ1→3Gal. This tentative assignment dictated that the remaining monosubstituted hexose, 1,6-Glc, be linked to the 1,3,4-Hep. From GC/MS analyses of alditol acetates and partially methylated alditol acetates, it was established that 1-Hep, 1,2-Hep, and 1,3,4-Hep were L-*glycero*-D-*manno*-heptoses and 1,4-Hep was a D-*glycero*-D-*manno*-heptose (Melaugh et al., 1992). Three L-*glycero*-D-*manno*-heptoses were previously identified in the Hepα1→2Hepα1→3(4)Hepα1→5Kdo core of the LOS from a related bacterium, *Haemophilus influenzae* nontypable strain 2019 (Phillips et al., 1992), suggesting that this core structure may be conserved in *Haemophilus* LOS. This argued for assigning the 1,4-linked D-*glycero*-D-*manno*-heptose as the unusual heptose in the longer branch. From these data, a most likely preliminary oligosaccharide structure was proposed, although no further structural data were available to confirm or complete the assignment (Melaugh et al., 1992):

Galβ1→4GlcNAcβ1→3Gal1→4Hep1→6Glc1→4Hepα1→*anhydro*Kdo
3
↑
Hep1α→2Hepα1

NMR Analysis of the Major Oligosaccharide from Strain 35000. To define the precise linkages and anomeric configurations of the major oligosaccharide from *H. ducreyi* 35000, a series of NMR experiments was carried out. The 1D ^{1}H NMR spectrum of the major oligosaccharide (M_r 1677) showed multiple anomeric resonances, consistent with an oligosaccharide sample containing several diastereomeric forms of anhydro-Kdo at the reducing terminus (Figure 1). As was previously observed in the hexasaccharide released from the LOS of *H. influenzae* strain 2019 (Phillips et al., 1992), this reducing terminal microheterogeneity causes the proton resonances from the saccharides closest to the reducing terminus to appear as multiple signals, whereas the more distant residues are not noticeably affected. Integration of the anomeric resonances permitted a total of eight anomeric protons to be accounted for as follows: I (δ 5.103, 5.083, 5.037; J sm), II (δ 5.786, 5.764, 5.752; J sm), III (δ 5.115; J sm), IV ($\delta \approx$4.50; $J \approx$ 8 Hz), V (δ 4.925; J sm), VI (δ 4.497; J = 8.0 Hz), VII (δ 4.723; J = 8.0 Hz), and VIII (δ 4.481; J = 8.0 Hz). The four high-field-shifted anomeric signals with large coupling constants (J = 8.0 Hz) suggested β-linked residues, most likely including the three hexoses (1-Gal, 1,3-Gal, and 1,6-Glc) and one 1,4-GlcNAc present in the oligosaccharide. The remaining four low-field-shifted anomeric resonances were tentatively assigned to the four Hep residues, all of which were expected to have the *manno* configuration, including three L-*glycero*-D-*manno*-heptoses and one D-*glycero*-D-*manno*-heptose.

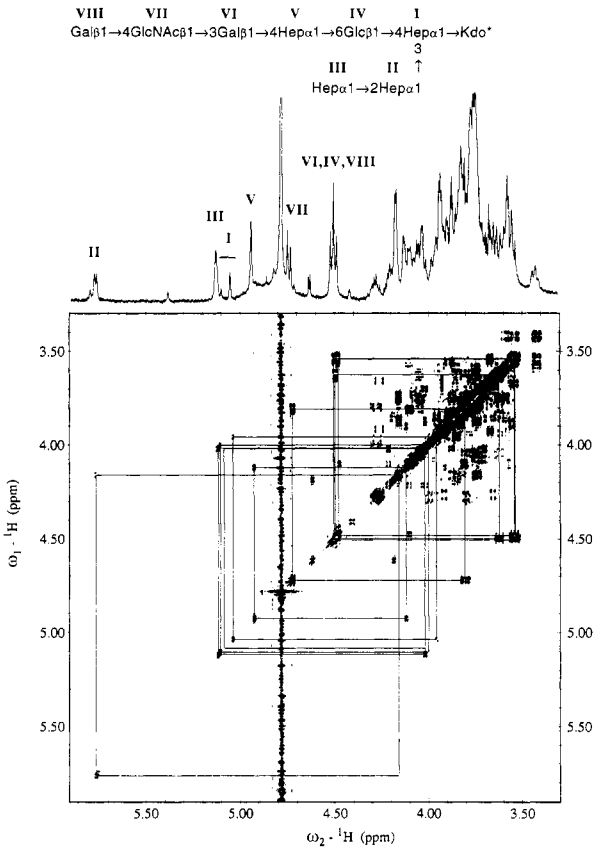


FIGURE 1: Section of the 500-MHz phase-sensitive DQF-COSY spectrum of the major oligosaccharide from *H. ducreyi* strain 35000 (M_r 1677) and the corresponding region of the 1D ^{1}H NMR spectrum. Anomeric protons are labeled as shown in Table 1, and H-1/H-2 cross-peaks are connected with lines on the 2D spectrum. The peak at δ 4.78 is HOD, and the peaks at δ 4.617, 4.475, and 4.409 are from anhydro-Kdo.

To assign most of the ring protons in sugar residues I–VIII, DQF-COSY (Figure 1) and 2D HOHAHA (Figure 2A) spectra were obtained and analyzed in combination. Where possible, apparent vicinal coupling constants were measured

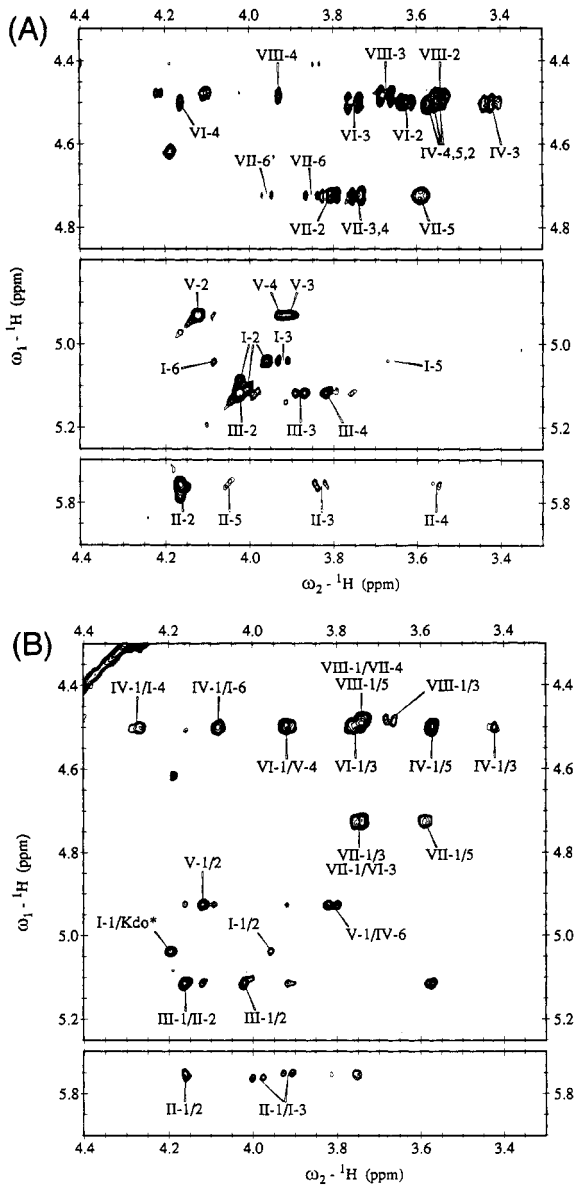


FIGURE 2: (A) Selected regions of the 2D HOHAHA spectra of the major oligosaccharide from *H. ducreyi* strain 35000 acquired with 90-ms (top) and 120-ms (middle and bottom) mixing times. (B) Corresponding regions of the 2D NOESY spectrum run with a 350-ms mixing time. An NOE cross-peak between III-1 and II-1 is present in the full spectrum. The assignment of VIII-5 is tentative. Additional unassigned NOESY cross-peaks to the anomeric protons of III (δ 4.120, 3.916, and 3.570) and II (δ 3.750) are present in the spectrum.

from DQF-COSY cross-peaks and used to establish monosaccharide identities. Working from the low-field to the high-field anomeric proton resonances, the monosaccharide spin systems were analyzed as will be outlined here.

The most low-field-shifted anomeric proton resonance, II, consisted of three broad singlets that were all coupled to H-2 protons at $\delta \approx 4.16$. For this residue, H-3, H-4, and H-5 were assigned from cross-peaks to H-1 in the HOHAHA spectrum (Figure 2A). The small $J_{1,2}$ and $J_{2,3}$ coupling constants measured from DQF-COSY cross-peaks (Table 1) were consistent with the *manno* configuration. Protons H-1–H-3 of residue III could be assigned from DQF-COSY connectivities, with the assignment of H-4 coming from the H-1/H-4 HOHAHA cross-peak. Again, the $J_{1,2}$ and $J_{2,3}$ values measured for this residue were appropriate for a *manno*-heptose. The anomeric proton of residue I consisted of three broad singlets, one of which was partially overlapping with H-1 of III. This pattern of three resolved anomeric resonances was seen for the 1,3,4-linked heptose from the *H. influenzae* 2019 hexasaccharide (Phillips et al., 1992). As was the case with that residue, this spin system was difficult to follow. The HOHAHA spectrum suggested that the H-2 and H-3 protons of I were strongly coupled in each case, with only the most intense anomeric signal (δ 5.037) showing a barely resolved H-1/H-3 cross-peak. However, this information was sufficient to follow coupling connectivities to H-4 and H-5 in the DQF-COSY spectrum, as was done in the assignment of the *H. influenzae* 2019 hexasaccharide. The remaining low-field-shifted anomeric proton resonance, V, was associated with the fourth *manno*-heptose spin system. Coupling connectivity from H-1 to H-3 of V could be followed in the DQF-COSY spectrum, and the HOHAHA spectrum suggested that H-4 of this spin system was down-field-shifted and strongly coupled to the H-3 resonance.

Of the high-field-shifted anomeric resonances, VII was the farthest down field. Spin system VII was mapped from H-1 to H-2 and from H-4 to H-6/6′ using the HOHAHA cross-peaks to follow DQF-COSY connectivities from both ends of the spin system. While the chemical shift of H-3 could not be established from the DQF-COSY data, the intense cross-peak at $\delta \approx 3.74$ in the HOHAHA spectrum suggested that H-3 was strongly coupled to H-4. This and the measured J values suggested a monosaccharide with the *gluco* configuration, which in this case could be either the 1,4-GlcNAc or 1,6-Glc residue. Since its anomeric proton was not affected by reducing terminal microheterogeneity, this residue presumably was not proximal to the inner core area. At 500 MHz, the remaining three anomeric protons, IV, VI, and VIII, overlapped to give what looks like a large broad triplet, evidently arising from two sharp doublets and one complex set of closely occurring doublets. Spin system VI, whose H-1 proton was the low-field-shifted sharp doublet, could only be mapped from H-1 to H-4 in the DQF-COSY and HOHAHA spectra, suggesting a sugar with the *galacto* configuration. This was supported by the small $J_{3,4}$ value measured. The same was true for spin system VIII, whose anomeric proton was the high-field-shifted sharp doublet. Thus, these spin systems represent the two galactose residues present in the oligosaccharide. The final spin system, IV, was mapped from H-1 to H-4 and from H-5 to H-6/6′ in the DQF-COSY spectrum after the H-6 and H-6′ resonances were clearly located from cross-peaks to H-2 in the HOHAHA spectrum. In this case, the H-6 and H-6′ protons gave very weak HOHAHA cross-peaks to H-1. Nonetheless, combined data suggested the *gluco* configuration for this residue.

The oligosaccharide sequence was established by analysis of interresidue NOEs observed in the 2D NOESY spectrum (Figure 2B). Starting with the core region, the anomeric resonances of the three core heptoses (residues I, II, and III) all showed intraresidue NOEs to their respective H-2 protons, consistent with α-linked heptoses in the *manno* configuration. The anomeric proton of heptose III showed strong interresidue NOEs to H-1 and H-2 of residue II, indicating that Hep III was the terminal heptose linked to the 2-position of Hep II. An interresidue NOE between two anomeric protons is characteristic of 1,2-linkages in the α-configuration (Romanowska et al., 1988) and was also observed for the Hepα1→2Hepα1→ disaccharide branch in the *H. influenzae* 2019 hexasaccharide (Phillips et al., 1992). The anomeric proton of Hep II also showed an interresidue NOE to H-3 of Hep I, establishing that Hep II was linked to the 3-position of Hep I. Thus, the linkages of this *H. ducreyi* Hep_3 core

Table 1: Partial Proton NMR Assignment of the Major Oligosaccharide (M_r 1677) from *H. ducreyi* Strain 35000 (D_2O, 25 °C)[a]

VIII VII VI V IV I
Galβ1----→4GlcNAcβ1----→3Galβ1-----→4Hepα1-------→6Glcβ1----------→4Hepα1→Kdo*
3
↑
Hepα1---------→2Hepα1------
III II

proton	VIII	VII	VI	V	IV	III	II	I
H-1	4.481	4.723	4.497	4.925	≈4.50[b]	5.115	5.786	5.103
							5.764	5.083
							5.752	5.037
H-2	3.542	3.806	3.623	4.120	≈3.54[b]	4.020	4.163	4.002
							4.158	4.019
							4.164	3.958
H-3	3.672	≈3.74	3.749	3.900	3.420	3.876	3.830	3.990
								4.000
								3.917
H-4	3.928	3.743	4.161	3.922	3.576	3.818	3.555	4.275
								4.288
								4.264
H-5	≈3.74[c]	3.587			3.559		4.051	3.797
								3.791
								3.656
H-6		3.851			3.809			4.081[c]
H-6′		3.957			4.102			
NAc		2.041						
$J_{1,2}$	8.0[d]	8.0[d]	8.0[d]	<3	≈8	<3	<3	<3
$J_{2,3}$	10	10	10	4	≈9	4	4	
$J_{3,4}$	4		4		≈9	8	10	10
$J_{4,5}$		10						10

[a] Chemical shifts are reported in ppm, and *J* values are apparent coupling constants (Hz) obtained from DQF-COSY cross-peaks, unless otherwise noted. Kdo* indicates anhydro-Kdo. [b] Average of three overlapping signals arising from reducing terminal microheterogeneity. [c] Tentative assignments made from NOESY data. [d] *J* values measured from the 1D spectrum.

appear to be identical to those of the Hep_3 core of *H. influenzae* 2019 LOS.

The anomeric proton of β-galactose residue VIII showed an intraresidue H-1/H-3 NOE cross-peak and an interresidue cross-peak to H-4 of VII, which may be overlapping with the expected intraresidue H-1/H-5 cross-peak. The absence of additional interresidue NOE cross-peaks to VIII indicated that VIII was the nonreducing terminal β-Gal linked to the 4-position of VII, which established VII as the 1,4-GlcNAc moiety. Residue VII showed an intraresidue H-1/H-5 NOE cross-peak in addition to an interresidue cross-peak to H-3 of β-Gal VI, which is presumably overlapping with an intraresidue H-1/H-3 cross-peak. Thus, the linkage of β-GlcNAc residue VII to β-Gal residue VI established the *H. ducreyi* 35000 terminal trisaccharide as Galβ1→4GlcNAcβ1→3Galβ1→.

H-1 of VI showed an intraresidue NOE cross-peak to H-3 and an interresidue cross-peak to H-4 of V, which linked β-Gal VI to the 4-position of Hep V. The expected H-1/H-5 intraresidue cross-peak for VI probably overlaps with another signal and could not be located with certainty. Consistent with the other α-linked *manno*-heptoses, Hep V showed an intraresidue H-1/H-2 NOE cross-peak. An interresidue cross-peak between H-1 of V and H-6 of IV established the linkage of Hep V to the 6-position of β-Glc IV. Residue IV showed intraresidue H-1/H-3 and H-1/H-5 cross-peaks and interresidue cross-peaks to H-4 and a resonance tentatively assigned to H-6 of Hep I. These same interresidue cross-peaks were observed in the NOESY spectrum of the *H. influenzae* 2019 hexasaccharide, again supporting the similarities of the two *Haemophilus* species core structures. Thus, the NOESY data support a complete carbohydrate structure for the *H. ducreyi* 35000 oligosaccharide as shown in Table 1, where the core heptoses are L-*glycero*-D-*manno*-heptoses and the 1,4-linked heptose is a D-*glycero*-D-*manno*-heptose.

Analysis of Oligosaccharide Heterogeneity. SDS–PAGE and mass spectrometric analyses of *O*-deacylated LOS from

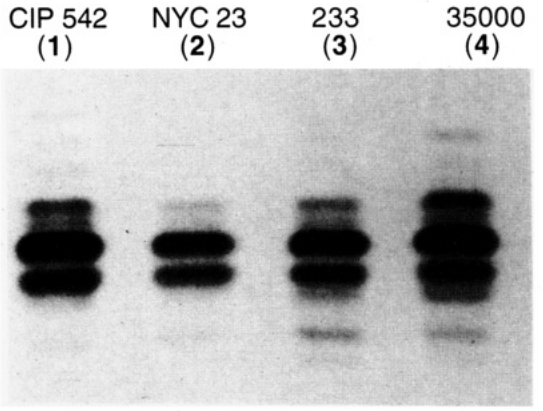


FIGURE 3: SDS–PAGE of the LOS from *H. ducreyi* strains CIP 542 (lane 1), NYC 23 (lane 2), 233 (lane 3), and 35000 (lane 4). The two bands present in NYC 23 correspond to LOSs with M_r's of ≈5100 and 4500. These two bands also appear as the major two bands in strain 35000, the smaller of which at M_r ≈4500 binds the monoclonal antibody 3F11.

strain 35000 reveal the presence of several other LOSs besides the major species, as determined by the current NMR studies and previously published LSIMS data (Melaugh et al., 1992). For example, SDS–PAGE analysis clearly shows the presence of three major LOS glycoforms and several minor forms of lower molecular weight (see Figure 3). The major LOS band at 4.5 kDa previously had been shown to react with the monoclonal antibody 3F11 (Campagnari et al., 1991) and likely corresponds to the LOS containing the terminal lactosamine group. The band just above with an estimated mass of 5.1 kDa is sensitive to neuraminidase and is likely the sialylated analog of this major component.

Electrospray ionization mass spectrometry of the total *O*-deacylated LOSs initially indicated the presence of five separate glycoforms (Gibson et al., 1993). On the basis of their molecular weights, these additional LOS species were

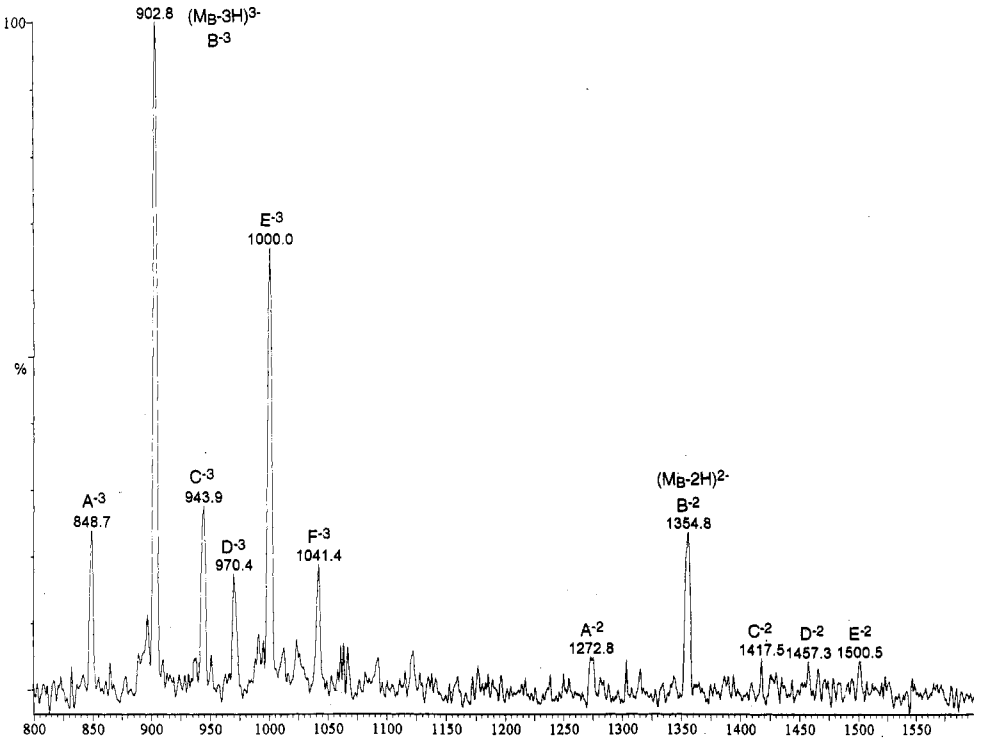


FIGURE 4: Negative-ion ESI-MS of *O*-deacylated LOS from *H. ducreyi* strain 35000. The LOS forms and charge states are indicated as follows: A^{-3} is $(M_A - 3H)^{3-}$, B^{-3} is $(M_B - 3H)^{3-}$, etc.

Table 2: ESI-MS Analyses of *O*-Deacylated LOSs from *H. ducreyi* Strain 35000

LOS	molecular weight (M_r) obsd[a]	calcd	relative ion abundance	proposed composition
A	2549.1	2549.3	24	$Hex_2HexNAcHep_4Kdo(P)$-lipid A[b]
B	2711.4	2711.5	100	$Hex_3HexNAcHep_4Kdo(P)$-lipid A
C	2834.7	2834.5	28	$Hex_3HexNAcHep_4PEAKdo(P)$-lipid A
D	2914.2	2914.7	17	$Hex_3HexNAc_2Hep_4Kdo(P)$-lipid A
E	3003.0	3002.7	67	$Neu5AcHex_3HexNAcHep_4Kdo(P)$-lipid A
F	3127.2	3125.7	20	$Neu5AcHex_3HexNAcHep_4PEAKdo(P)$-lipid A

[a] All molecular weights are reported as their average mass values. [b] After *O*-deacylation, the lipid A moiety is converted into diphosphoryldiacyl-lipid A containing two *N*-linked β-hydroxymyristic acid chains with an average M_r of 953.0089.

interpreted as representing both extended and truncated analogs of the major lactosamine-containing oligosaccharide now confirmed by NMR. Further ESI-MS analysis of this LOS preparation has now revealed a sixth LOS component (LOS-F) at m/z 1041.4 (M_r 3127.2), as can be seen from Figure 4 and the composition assignments listed in Table 2.

In contrast to the SDS–PAGE and ESI-MS data of LOS and *O*-deacylated LOS, the original LSIMS and tandem mass spectrometric data of the individual oligosaccharide fractions clearly defined only one major oligosaccharide structure at M_r 1677 corresponding to the LOS-B glycoform (Melaugh et al., 1992). However, frequently LSIMS is not sensitive enough to detect weak ions at higher masses. Indeed, reexamination of later eluting fractions provided only indirect evidence for an additional oligosaccharide at M_r 1514, corresponding to the loss of one hexose unit (162 Da). However, the assignment of even this second oligosaccharide is in some doubt, due to the fact that LSIMS spectra of even pure oligosaccharides produce small fragment ions at lower mass formed by cleavage at glycosidic oxygens (Y-type ions) (Domon & Costello, 1988). These ions have nominal masses that are identical to the molecular weights of truncated species. Furthermore, unless these truncated oligosaccharides are chromatographically separated or judged to be too large to originate from gas phase fragmentation of the higher mass oligosaccharide species, it is very difficult to assign their origin as fragment ions or chemically distinct oligosaccharides.

Since it was not possible to unambiguously verify the additional glycoforms by LSIMS, a more thorough investigation of the oligosaccharide fraction was carried out by ESI-MS. ESI-MS is generally considered to be a much more sensitive analytical method that, under the conditions used here, does not produce fragment ions. Analysis of the total oligosaccharide fraction from *H. ducreyi* strain 35000 (see Figure 5) yielded a number of peaks that were consistent with the LOS heterogeneity assessed from the ESI-MS and SDS–PAGE gel data (see Figures 3 and 4 and Table 2). Furthermore, the precise masses obtained are mostly in agreement with the proposed compositions suggested from the *O*-deacylated LOS data (see Table 2). For example, the base peak at m/z 838.2 can be assigned as the doubly charged peak for the previously identified major oligosaccharide terminating in lactosamine with M_r 1677. The peak at m/z 757.1 contains one less hexose, and the higher mass peak at m/z 939.7 contains an additional *N*-acetylhexosamine (or phosphate and PEA). The peak at m/z 898.3 has been tentatively assigned as the addition of PEA to the major oligosaccharide (M_r 1677 + 123), but its observed mass is somewhat lower than expected (see Table 3). Lastly, another extended analog not observed in the ESI-MS spectrum of the *O*-deacylated LOS is seen at m/z 1021.0. This latter peak is consistent with the addition of *N*-acetylhexosamine (or phosphate and PEA) and hexose to the major oligosaccharide (M_r 1677 + 203 + 162). The remaining peaks in this spectrum correspond to various molecular ion adducts of these oligosaccharides, such as the peak at m/z 856.5, which arises from the addition of HCl to the doubly charged deprotonated

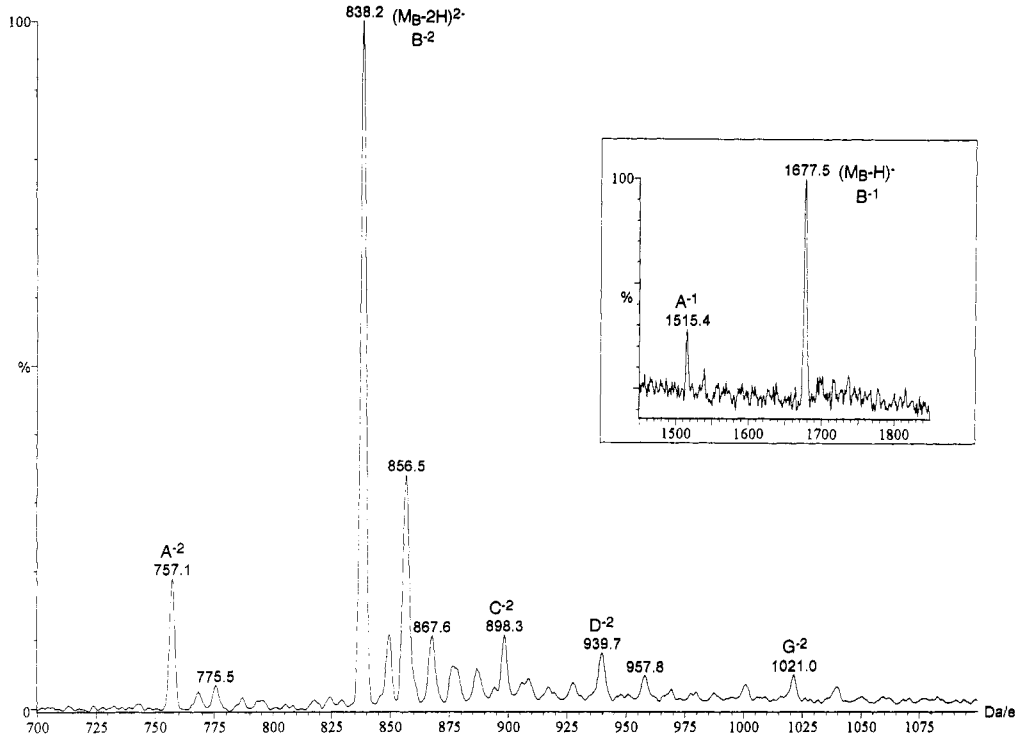


FIGURE 5: Negative-ion ESI-MS of the pooled, size-exclusion oligosaccharide fractions of *H. ducreyi* strain 35000. The oligosaccharide forms and charge states are indicated as follows: A^{-2} is $(M_A - 2H)^{2-}$, B^{-2} is $(M_B - 2H)^{2-}$, etc. Note that oligosaccharide A is derived from LOS A, etc. The inset shows the singly charged m/z region.

Table 3: ESI-MS Analyses of Oligosaccharides from LOSs of *H. ducreyi* Strain 35000

LOS	molecular weight (M_r) obsd[a]	calcd	relative ion abundance	proposed composition
A	1516.2	1516.3	19	$Hex_2HexNAcHep_4$-anhydro-Kdo
B	1678.4	1678.5	100	$Hex_3HexNAcHep_4$-anhydro-Kdo
C	1798.6	1801.5	11	$Hex_3HexNAcHep_4$PEA-anhydro-Kdo
D	1881.4	1881.7	9	$Hex_3HexNAc_2Hep_4$-anhydro-Kdo
G	2044.0	2043.8	5	$Hex_4HexNAc_2Hep_4$-anhydro-Kdo

[a] All molecular weights are reported as their average mass values. Additional peaks were seen that corresponded to the addition of sodium, potassium, and HCl, i.e., $(M - 2H + Na)^-$, $(M - 2H + K)^-$, $(M - H + HCl)^-$.

molecular ion, i.e., $(M - 2H + HCl)^{2-}$. As expected, the sialylated oligosaccharides are absent from this spectrum since they are converted to their asialo analogs during acetic acid hydrolysis.

Analysis of O-Deacylated and HF-Treated O-Deacylated LOSs. Since there was still some uncertainty regarding the identification of HexNAc (203 Da) versus an alternative assignment of phosphate and phosphoethanolamine (80 + 123 = 203 Da) for the LOS glycoforms D and G, the *O*-deacylated LOS preparation was further treated with aqueous HF to remove phosphate and PEA groups. A comparison of the before and after spectra (data not shown) of these two LOS preparations clearly shows that the molecular weights of all LOS components, except for LOS-C, shifted by the loss of 240 Da, the mass difference expected for removal of the three phosphates that are known to reside on Kdo and lipid A, i.e., diphosphoryl-lipid A and 4-phospho-Kdo (Melaugh et al., 1992). LOS-C, however, was no longer present as a distinct peak, as one would have predicted once its single PEA group was removed, making it identical to the LOS-B glycoform. Therefore, the mass spectral data of the *O*-deacylated, dephosphorylated LOSs support the presence of an additional HexNAc on the LOS-D and -G glycoforms, as well as the assignment of a PEA moiety to LOS-C as listed in Table 3.

Sialic Acid Linkage Analysis. Lastly, to determine how the sialic acid is linked to the terminal lactosamine acceptor, Galβ1→4GlcNAc, the LOS was treated with sialidase. First, Newcastle disease virus sialidase was used, an enzyme specific for Neu5Acα2→3 linkages, followed by treatment with *Clostridium perfringens* sialidase, a enzyme with broader specificities, i.e., Neu5Acα2→3, 6, or 8. The amount of sialic acid released by the NDV sialidase digestion was 6 pmol of Neu5Ac/μg of LOS, as measured by high-pH anion-exchange chromatography of the enzyme digests. The amount released by adding extra *C. perfringens* sialidase was 8.7 pmol of Neu5Ac/μg of LOS.

DISCUSSION

In this study, we have completed the structure determination of the major oligosaccharide from the LOS of *H. ducreyi* strain 35000 and have confirmed several features proposed in our preliminary characterization (see Figure 6). The core Hep_3Kdo region of this oligosaccharide is identical to that in *H. influenzae* nontypable strain 2019 (Phillips et al., 1992) and consists of three α-linked L-*glycero*-D-*manno*-heptoses and reducing terminal Kdo, which is phosphorylated in the intact LOS structure. Two strains of *H. influenzae* type b have also been found to produce LOS with this common core structure, although additional branching from the Hep_3 core was observed in these strains (Phillips et al., 1993; Schweda et al., 1993). Furthermore, the major *H. influenzae* oligosaccharide structures contain one or two PEAs on the Hep_3 core, whereas the major *H. ducreyi* 35000 oligosaccharide is only partially substituted with PEA.

The 1,4-linked D-*glycero*-D-*manno*-heptose present in the *H. ducreyi* 35000 LOS was confirmed by 2D NMR studies to exist on the oligosaccharide branch, separated from the core heptoses by a 1,6-linked β-glucose residue. The oligo-

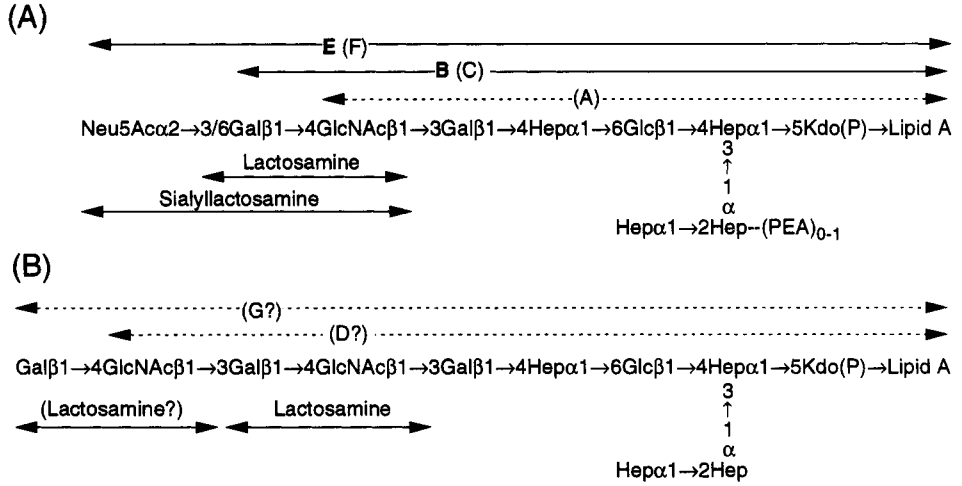


FIGURE 6: Proposed total structures of *H. ducreyi* 35000 LOS glycoforms. LOS structures listed in A are based on the lactosamine and sialyllactosamine glycans (LOS-A, -B, -C, -E, and -F). LOS structures in part B are more speculative (LOS-D and -G) and may consist of the addition of a second lactosamine disaccharide as a biosynthetic alternative to sialic acid.

saccharide branch terminates in a lactosamine disaccharide, which most likely is the recognition site for monoclonal antibody 3F11, which recognizes the terminal region of paragloboside (Mandrell et al., 1988). Indeed, the branch heptose breaks up what otherwise could have been a lacto-*N*-neotetraose tetrasaccharide, a structure seen in both *Neisseria meningitidis* (Jennings et al., 1983; Michon et al., 1990; Gamian et al., 1992) and *Neisseria gonorrhoeae* LOS (John et al., 1991; Yamasaki et al., 1991). Enzymatic susceptibility of the LOS to the Newcastle disease virus sialidase shows that this lactosamine is largely sialylated via an $\alpha 2 \rightarrow 3$ linkage to lactosamine, although a small amount may also be present in an $\alpha 2 \rightarrow 6$ or $\alpha 2 \rightarrow 8$ linkage due to the incomplete release of sialic acid. Complete release of sialic acid was accomplished with the nonspecific sialidase from *C. perfringens*. It should be noted that sialic acid linked $\alpha 2 \rightarrow 3$ to lactosamine has been recently reported in an LOS from *N. meningitidis* (Yamasaki et al., 1993).

We have also investigated the source of the heterogeneity in the terminal glycan regions of these LOSs and have concluded that, besides the major oligosaccharide as determined from NMR and mass spectrometric data (LOS-B), several other glycoforms are present, which for the most part can be explained as biosynthetic intermediates of this major LOS structure (see Figure 6). In addition to these intermediates, three additional glycoforms apparently arise through an alternative biosynthetic pathway leading to the addition of sialic acid (sialyllactosamine) or HexNAc and Hex (polylactosamine?). It should be noted that a similar divergent biosynthetic pathway has been reported for *N. gonorrhoeae* LOS (Yamasaki et al., 1991).

The significance of such a diverse set of LOS glycoforms in one strain is not clear, but evidence to date suggests that *H. ducreyi* does not undergo significant phase variation of surface antigens, including LOS, and therefore may require different LOS structures to perform specific and distinct biological functions. Recent structure studies on the oligosaccharide regions of lipooligosaccharides from pathogenic *Neisseria* and *Haemophilus* provide convincing evidence that the LOSs are mimicking human glycoconjugates (Mandrell et al., 1988; Gibson et al., 1993; Phillips et al., 1993; Yamasaki et al., 1993). These structural similarities to host antigens may represent the key to understanding the roles LOSs play in the pathology of the diseases these bacteria cause. For example, LOS species that terminate in lactosamine may be necessary for the adhesion and invasion of host cells. In contrast, LOS terminating in sialyllactosamine might be better explained as necessary for providing an evasion mechanism against host immune and defense responses. In support of this latter interpretation, recent studies of pathogenic *Neisseria* have suggested that sialylation of LOS confers serum resistance to strains that are otherwise serum-sensitive (Parsons et al., 1989; Frangipane & Rest, 1993), as well as reducing opsonized phagocytosis and/or killing by human neutrophils (Estabrook et al., 1992; Kim et al., 1992; Rest & Frangipane, 1992). The development of mutant strains that express only a single or small subset of this LOS repertoire should allow us the means to better unravel the functions of these different LOSs. Toward this latter goal, we are currently engaged in the construction and evaluation of *H. ducreyi* mutants that are defective in LOS biosynthesis (Campagnari et al., 1994).

ADDED IN PROOF

After the submission of this paper, two structures were published by Schweda et al. (1994) that are of the oligosaccharides obtained from the LOS of two additional strains of *H. ducreyi*, both of which are similar to one or more of the structures identified in this work.

REFERENCES

Albritton, W. L. (1989) *Microbiol Rev. 53*, 377–89.
Auzanneau, F.-I., Charon, D., & Szabó, L. (1991) *J. Chem. Soc., Perkin Trans. 1*, 509–517.
Basus, V. J., Billeter, M., Love, R. A., Stroud, R. M., & Kuntz, I. D. (1988) *Biochemistry 27*, 2763–2771.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Becker, T. M., Dewitt, W., & Van Dusen, G. (1987) *South. Med. J. 80*, 182–184.
Campagnari, A. A., Spinola, S. M., Lesse, A. J., Abu Kwaik, Y., Mandrell, R. E., & Apicella, M. A. (1990) *Microb. Pathog. 8*, 353–362.
Campagnari, A. A., Wild, L. M., Griffiths, G. E., Karalus, R. J., Wirth, M. A., & Spinola, S. M. (1991) *Infect. Immun. 59*, 2601–8.
Campagnari, A. A., Karalus, R., Apicella, M. A., Melaugh, W., Lesse, A. J., & Gibson, B. W. (1994) *Infect. Immun. 62*, 2379–2386.
DiFabio, J. L., Michon, F., Brisson, J.-R., & Jennings, H. J. (1990) *Can. J. Chem. 68*, 1029–1034.
Domon, B., & Costello, C. E. (1988) *Glycoconjugate J. 5*, 397–409.

Estabrook, M. M., Christopher, N. C., Griffiss, J. M., Baker, C. J., & Mandrell, R. E. (1992) *J. Infect. Dis. 166*, 1079–1088.
Frangipane, J. V., & Rest, R. F. (1993) *Infect. Immun. 61*, 1657–1666.
Gamian, A., Beurret, M., Michon, F., Brisson, J.-R., & Jennings, H. J. (1992) *J. Biol. Chem. 267*, 922–925.
Gibson, B. W., Webb, J. W., Yamasaki, R., Fisher, S. J., Burlingame, A. L., Mandrell, R. E., Schneider, H., & Griffiss, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 17–21.
Gibson, B. W., W., M., Phillips, N. J., Apicella, M. A., Campagnari, A. A., & Griffiss, J. M. (1993) *J. Bacteriol. 175*, 2702–2712.
Jennings, H. J., Johnson, K. G., & Kenne, L. (1983) *Carbohydr. Res. 121*, 233–241.
Jessamine, P. G., & Ronald, A. R. (1990) *Med. Clin. North Am. 74*, 1417–1431.
John, C. M., Griffiss, J. M., Apicella, M. A., Mandrell, R. E., & Gibson, B. W. (1991) *J. Biol. Chem. 266*, 19303–19311.
Johnson, A. P., Abeck, D., & Davies, H. A. (1988) *J. Infect. 17*, 99–106.
Kibukamusoke, J. W. (1965) *Trans. R. Soc. Trop. Med. Hyg. 59*, 642–648.
Kim, J., Zhou, D., Mandrell, R. E., & Griffiss, J. M. (1992) *Infect. Immun. 60*, 4439–4442.
Mandrell, R. E., Griffiss, J. M., & Macher, B. A. (1988) *J. Exp. Med. 168*, 107–126.
Mandrell, R. E., McLaughlin, R., Abu Kwaik, Y., Lesse, A., Yamasaki, R., Gibson, B., Spinola, S. M., & Apicella, M. A. (1992) *Infect. Immun. 60*, 1322–1328.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967.
Melaugh, W., Phillips, N. J., Campagnari, A. A., Karalus, R., & Gibson, B. W. (1992) *J. Biol. Chem. 267*, 13434–13439.
Michon, F., Beurret, M., Gamian, A., Brisson, J.-R., & Jennings, H. J. (1990) *J. Biol. Chem. 265*, 7243–7247.
Morse, S. A. (1989) *Clin. Microbiol. Rev. 2*, 137–57.
Parsons, N. J., Andrade, J. R. C., Patel, P. V., Cole, J. A., & Smith, H. (1989) *Microb. Pathog. 7*, 63–72.
Pepin, J., Plummer, F. A., Brunham, R. C., Piot, P., Cameron, D. W., & Ronald, A. R. (1989) *AIDS 3*, 3–9.
Phillips, N. J., John, C. M., Reinders, L. G., Gibson, B. W., Apicella, M. A., & Griffiss, J. M. (1990) *Biomed. Environ. Mass Spectrom. 19*, 731–745.
Phillips, N. J., Apicella, M. A., Griffiss, J. M., & Gibson, B. W. (1992) *Biochemistry 31*, 4515–4526.
Phillips, N. J., Apicella, M. A., Griffiss, J. M., & Gibson, B. W. (1993) *Biochemistry 32*, 2003–2012.
Purcell, B. K., Richardson, J. A., Radolf, J. D., & Hansen, E. J. (1991) *J. Infect. Dis. 164*, 359–67.
Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.
Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson. 19*, 250.
Rest, R. F., & Frangipane, J. V. (1992) *Infect. Immun. 60*, 989–997.
Romanowska, E., Gamian, A., Lugowski, C., Romanowska, A., Dabrowski, J., Hauck, M., Opferkuch, H. J., & von der Lieth, C.-W. (1988) *Biochemistry 27*, 4153–4161.
Schweda, E. K. H., Hegedus, O. E., Borrelli, S., Lindberg, A. A., Weiser, J. N., Maskell, D. J., & Moxon, E. R. (1993) *Carbohydr. Res. 246*, 319–330.
Schweda, E. K., Sundstrom, A. C., Eriksson, L. M., Jonasson, J. A., & Lindberg, A. A. (1994) *J. Biol. Chem. 269*, 12040–12048.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Tsai, C. M., & Frasch, C. E. (1982) *Anal. Biochem. 119*, 115–119.
Tuffrey, M., Alexander, F., Ballard, R. C., & Taylor, R. D. (1990) *J. Exp. Pathol. (Oxford) 71*, 233–44.
Weisner, P., Brown, S., Kraus, S., & Perine, P. (1983) in *Perspectives in Sexually Transmitted Diseases* (Holmes, K., & Mardh, P., Eds.) pp 219–234, McGraw-Hill, New York.
Westphal, O., & Jahn, K. (1965) in *Methods in Carbohydrate Chemistry* (Whistler, R., Ed.) p 83, Academic Press, New York.
Yamasaki, R., Bacon, B. E., Nasholds, W., Schneider, H., & Griffiss, J. M. (1991) *Biochemistry 30*, 10566–10575.
Yamasaki, R., Griffiss, J. M., Quinn, K. P., & Mandrell, R. E. (1993) *J. Bacteriol. 175*, 4565–4568.