The Fine Structure of Caenorhabditis elegans N-Glycans

Received for publication, August 6, 2002, and in revised form, October 1, 2002
Published, JBC Papers in Press, October 1, 2002, DOI 10.1074/jbc.M208020200

John F. Cipollo‡, Catherine E. Costello§, and Carlos B. Hirschberg¶

From the ‡Department of Molecular and Cell Biology, Boston University, Goldman School of Dental Medicine and §Mass Spectrometry Resource, Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118-2526

We report the fine structure of a nearly contiguous series of N-glycans from the soil nematode Caenorhabditis elegans. Five major classes are revealed including high mannose, mammalian-type complex, hybrid, fucopaussimannosidic (five mannose residues or fewer substituted with fucose), and phosphocholine oligosaccharides. The high mannose, complex, and hybrid N-glycan series show a high degree of conservation with the mammalian biosynthetic pathways. The fuco-paussimannosidic glycans contain a novel terminal fucose substitution of mannose. The phosphocholine oligosaccharides are high mannose type and are multiply substituted with phosphocholine. Although phosphocholine oligosaccharides are known immunomodulators in human hematode and trematode infections, C. elegans is unique as a non-parasitic nematode containing phosphocholine N-glycans. Therefore, studies in C. elegans should aid in the elucidation of the biosynthetic pathway(s) of this class of biomedically relevant compounds. Results presented here show that C. elegans has a functional orthologue for nearly every known enzyme found to be deficient in congenital disorders of glycosylation types I and II. This nematode is well characterized genetically and developmentally. Therefore, elucidation of its N-glycome, as shown in this report, may place it among the useful systems used to investigate human disorders of glycoconjugate synthesis such as the congenital disorders of glycosylation syndromes.

Much of the current understanding of glycosylation processes in eukaryotes has been derived from yeast, mammalian cell lines, and organ tissues. Although these systems have been useful, they are not suited for the analysis of these processes in multicellular organisms as they relate to genetic, developmental, or environmentally interactive processes.

Caenorhabditis elegans is a developmentally well characterized model organism. Its genome has been sequenced, and the organism can be genetically manipulated. Several gene ablation services are currently available, and these should increase the rate of genetic and biochemical progress. The wealth of genetic information has allowed the cloning of some glycosyltransferases through reverse genetics (1, 2). However, the lack of C. elegans N-glycan structural information has made substrate specificity and enzyme activity difficult to identify. Recently, the major O-linked glycan structures were reported (3). The observation of a novel β1,6Gal substitution facilitated the identification of the elusive activity of the previously cloned enzyme (4). Therefore, the elucidation of the C. elegans N-glycan structures is warranted. Once this system is in place, the effect that aberrant glycan biosynthesis has at the multicellular level can be more clearly understood. Such an effect has been shown by the study of the sqv mutants where metabolic error in chondroitin biosynthesis leads to malformation of the nematode vulva (5, 6).

There is a growing number of human genetic disorders wherein an error in biosynthesis or metabolism leads to dysregulation of glycoconjugate synthesis and developmental impairment. This group of diseases has been classified under the general heading congenital disorders of glycosylation (CDG)1 (7, 8). Yeast has facilitated the discovery of several causative gene lesions associated with these human disorders. However, because yeasts are simple eukaryotes, information provided by yeast systems is only inferential when extrapolated to the multicellular organism. C. elegans may aid in such investigations because the effects of errors in glycosylation manifest themselves throughout cell lineages and resultant tissue formation.

Here we present an overview of the N-glycans produced by C. elegans. From mixed stage worms we have characterized a nearly contiguous series of N-glycans that suggest the major N-glycosylation pathways in this nematode. Together with the sequenced genome and mapped cell fates of the organism, this information will facilitate understanding the role of glycosylation in development.

EXPERIMENTAL PROCEDURES

Materials

All chemicals were ACS grade or better. Phosphocholine calcium salt was from Fluka (Fliuka Chemie GmbH, Industriestrasse 25, CH-9471 Buchs SG, Switzerland). TEPC-15 antibody and anti-mouse horseradish peroxidase conjugate were from Sigma, and the chemiluminescence detection kit was from PerkinElmer Life Sciences. The PNGase F used in this study was the kind gift of Dr. Thomas Plummer.

Western and Carbohydrate Blotting

SDS-PAGE and blotting to polyvinylidene difluoride membranes was performed as described previously (9). Wheat germ agglutinin was purchased from EY (EY Laboratories, Inc., San Mateo, CA) and

† To whom correspondence should be addressed: Dept. of Molecular and Cell Biology, Boston University Goldman School of Dental Medicine, 715 Albany St., Evans 437, Boston, MA 02118. Tel.: 617-414-1040; Fax: 617-414-1041; E-mail: chirschb@bu.edu.

§ The on-line version of this article (available at http://www.jbc.org) contains Figs. 1 and 2.

¶ The abbreviations used are: CDG, congenital disorders of glycosylation; MALDI-TOF MS matrix-assisted laser desorption ionization/time of flight mass spectrometry; PNGase, peptide N-glycosidase; GP, glycopeptides; HF, hydrofluoric acid; dH2O, distilled water; GC/MS, gas chromatography/mass spectrometry; Pc, phosphorylcholine; ER, endoplasmic reticulum; HPAEC, high pH anion exchange; PAD, pulse amperometric detection; GNT, N-acetylgalactosamine transferase.

This paper is available on line at http://www.jbc.org

49143

1 The abbreviations used are: CDG, congenital disorders of glycosylation; MALDI-TOF MS matrix-assisted laser desorption ionization/time of flight mass spectrometry; PNGase, peptide N-glycosidase; GP, glycopeptides; HF, hydrofluoric acid; dH2O, distilled water; GC/MS, gas chromatography/mass spectrometry; Pc, phosphorylcholine; ER, endoplasmic reticulum; HPAEC, high pH anion exchange; PAD, pulse amperometric detection; GNT, N-acetylgalactosamine transferase.

□ This work was supported by National Institute of Health Grant GM 30365 (to C. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

☐ The on-line version of this article (available at http://www.jbc.org) contains Figs. 1 and 2.
C. elegans N-Glycans

TEPC-15 from Sigma. Protein loading was normalized to protein concentration.

Oligosaccharide Purification

Glycan Extraction—The procedure was monitored by Bearden’s protein assay (10) and phenol-sulfuric assay for neutral heoxo (11). C. elegans N2 worms (100 g wet weight) were suspended in 10 mm phosphate buffer, pH 7.0, which included 1 mm phenylmethylsulfonyl fluoride, 1 mm EDTA, and 0.3% sodium azide. Worms were disrupted by bead beating with 425–600-μm glass beads in a BioSpec Bead Beater® (6 times for 1 min and were submerged in ice for 5 min between repetitions. Worm homogenate was lyophilized, followed by delipidation by 4000-fold extraction from 10:10:1 chloroform/methanol/water (CMW) and removal of free glycans by extensive solvent extraction from 50% methanol, which was monitored for neutral hexose by phenol-sulfuric acid. A LiF salt precipitate was detectable in the methanol extract by Bearden’s protein assay, indicating that the detected carbohydrate was not co-valently linked to protein. Protein was solubilized in 125 mM Tris chloride, pH 6.8, 1% SDS, 5% β-mercaptoethanol buffer (TSB), by boiling for 10 min followed by agitation on a rotatory shaker for 3 h and then centrifugation at 10,000 × g for 10 min. The procedure was performed three times, and the soluble contaminants were combined, dialyzed, and digested with 1-1-tosylamido-2-phenyl ethyl chloromethyl ketone-treated trypsin (Trypsin) for 4 h at 37 °C in 50 mM ammonium bicarbonate, pH 8.5. The reaction was stopped by boiling two times for 10 min. An adaptation of the solvent precipitation method for the isolation of oligosaccharides of Verostek et al. (12) was employed. The majority of peptides and glycopeptides (GP) were precipitated in 80% acetone. To recover the remaining peptides in the 80% acetone, the solvent was rotary evaporated to dryness, the residue resuspended in dH2O, and peptides further purified over C-18 Sep-Pak® by step elution using 2 column volumes each of 20, 40, 60, and 100% isopropl alcohol subse- quent to an aqueous wash of dH2O to first remove residual contami- nants from the C-18-bound PG. Solvent and C-18-purified GP were combined and glycans released with PNGase F activity. The reaction was adjusted to pH 5.5 with acetic acid, and oligosaccharide was recovered by precipitation in 50% methanol at –20 °C for 4 h followed by centrifugation. The 50% methanol extract was rotary evaporated to dryness and reconstituted in dH2O, and residual peptide/glycopeptides were removed using the same Sep-Pak® C-18 procedure described above. This step the dH2O eluate contained glycan, and residual GP was recovered in the isopropl alcohol elution. The recovered GP was rotary evaporated, resuspended in 50 μl pH 5.5 ammonium acetate buffer, and digested with PNGase A at 2.5 milliunits/ml (Roche Diag- nostics). The released oligosaccharides and GP were collected using the same Sep-Pak® C-18 procedure employed for the PNGase F PNGase A released N-glycans with α,3Fuc linked to the reducing end GlcNAc residue, whose linkages are not accessible to PNGase F. These oligosaccharides were trimethylsialylated and analyzed.

For the permethylation of HF-treated glycans, the reaction was ended after the 1st h and the chloroform/dH2O wash procedure was performed. The sample was evaporated by speed evaporation and again permethylated for an additional hour followed by reaction termination and processing as above.

Monosaccharide Analysis

Monosaccharide composition of glycopeptide/peptide and released glycans was determined by GC/MS on a Finnigan TRACE GC 2000 series coupled to the GCQ plus electron ionization ion trap mass spectrometer. Monosaccharides generated by acid hydrolysis of glycans were trimethylsilylated and analyzed.

Mass Spectrometry

MALDI-TOF MS was performed on permethylated glycans using either a Finnigan Vision 2000 or Bruker Refex IV instrument fitted with a nitrogen laser (337 nm, 3-ns pulse width) operated in positive ion linear and/or reflectron mode(s). To acquire branch information on select glycan species, jack bean α-mannosidase digestions were per- formed using a method published previously (15) followed by permeth- ylation and analysis by MALDI/TOF MS. Targets were spotted in a layered fashion allowing each layer to dry before adding the next. First, 1 μl of 6-aza-2-thiotyramine (15 mg/ml in 70% methanol) was spotted. Next, an equal volume of 20 μm NaOAc was layered over the top, and finally 15–30 pmol of the permethylated oligosaccharide mixture in 70% methanol was layered. The TOF MS mass scales were calibrated exter- nally with a mixture of malto-oligosaccharides in the size range of Glc3-10 purchased from Sigma. Permethylated authentic αMan3GlcNAcβ (Sigma) was also used as external standard. Average masses were reported for spectra collected in linear mode, and monoisotopic values for the 1H peak were reported for those collected in reflectron mode. Mass accuracy was ±0.1%.

1H NMR Spectroscopy

Polysaccharide samples were exchanged from 99.9% 2H2O (Cam- bridge Isotope Laboratories, Andover, MA) three times followed by three additional exchanges from 99.96% 2H2O (Cambridge Isotopes) three times. The samples were then reconstituted in 99.96% 2H2O and lyophilized for 2 days and stored over P2O5 in vacuo for several days. Samples were reconstituted in 99.996% 1H2O (Cambridge Isotopes) with 5 mm acetone added as internal standard. All spectra were col- lected on a Bruker Avance 500 MHz spectrometer. Spectral width for one-dimensional experiments was 3008 Hz collected with 8 K points digitization. Total scans were 1024. The two-dimensional DQF-COSY experiments were performed with spectral widths of 1608 Hz in both t1 and t2 dimensions. Digitization was 512 and 4096 points in t1 and t2, respectively. Line broadening of 3–5 Hz was used in both dimen- sions, and a sine bell squared adiposation function was applied in t2. One-dimensional spectra were collected at 300 and 318 K and internally referenced to acetone at 2.225 and 2.217 ppm, respectively. DQF-COSY spectra were collected at 300 K and internally referenced to acetone at 2.225 ppm. For resonance intensity integrations, FIDs were collected for ~5 times t1 relaxation time.

RESULTS

Preparation of C. elegans N-Glycans for Structural Analysis

By using wheat germ agglutinin for detection, we found that C. elegans N-glycans are releasable with PNGase F. Only a trace amount of N-glycan was detectable by 1H NMR analysis of PNGase A released material performed subsequent to PNGase F treatment, suggesting that C. elegans mixed stage N-glycans contain little α1,3-linked core Fuc (data not shown).

TEPC-15 IgA detects phosphorylcholine (Pc). A TEPC-15 blot of the glycoprotein of the nematode showed a loss of antibody binding subsequent to PNGase F treatment suggesting that a portion of C. elegans N-glycans contain the Pc functional group.

The overall goal of this work was to provide an overview of N-glycosylation in C. elegans. PNGase F-released glycans were separated by size exclusion Bio-Gel P-4, and detection was by the phenol-sulfuric assay. The glycans were first pooled widely (Fig. 1, pools a–d) and analyzed by MALDI-TOF MS to deter- mine major and minor component glycan compositions (Table 1). Narrow pooling (Fig. 1, pools A–E) for fine structural deter-


Little was known about N-linked oligosaccharide composition in *C. elegans* prior to this study. Seven μg of oligosaccharide from the included volume of the Bio-Gel P-4 eluate was analyzed by GC/MS. Fuc, Man, and GlcNAc were identified on the basis of their GC retention times and electron ionization mass spectra. It is noteworthy that no sialic acid was seen, a result consistent with the previous report for *C. elegans* (16).

**Structural Characterization**

**Characterization of *C. elegans* N-Glycans**—A major concern in this work was the detectability of both abundant and rare glycans. As our glycan source was from mixed stage worms, it was expected that glycans derived from the lower mass larval stages would be under-represented. To address this issue properly, an aliquot of Bio-Gel P-4 separated glycans was pooled widely to include all fractions of the included volume. The glycans were then derivatized by permethylation.

We have employed an HF hydrolysis and permethylation strategy in order to aid in providing an overview of glycoconjugates in this study. Permethylation tends to impart similar chemical properties over a wide range of oligosaccharides, which translates into similar detector response with added sensitivity in MALDI-TOF MS analysis. HF treatment has been used in combination with permethylation to aid in characterizing oligosaccharides in complex mixtures, where some components contain Pc (17, 18). As *C. elegans* protein extracts bind TEPC-15, we anticipated that some glycoconjugates would bind TEPC-15, we anticipated that some glycoconjugates would

Table I summarizes the MALDI-TOF MS spectra of permethylated and HF-treated and permethylated glycans and their compositions assigned on the basis of observed *m/z* values (Fig. 1, pools a–d). The compositions indicate the presence of five classes of N-glycan compounds, which include high mannose, mammalian-type hybrid and complex, fuco-pausimannosidic, and those containing Pc. Permethylation alone produced ions representative of all classes. Complex and hybrid glycans were more numerous than high mannose and fuco-pausimannosidic. HF treatment prior to permethylation produced a change in relative abundance of detected glycan types, where high mannose glycans were most abundant followed by fuco-pausimannosidic, and complex glycans were absent. The appearance of an ion at *m/z* 1768.4 was in agreement with Hex₆GlcNAc₂Pi₂, which is consistent with loss of two choline groups and substitution with methyl groups during the permethylation procedure, a modification that has been noted (19).

At this point in our investigation the following three lines of evidence pointed toward the presence of Pc oligosaccharides: a Western blot result that was positive for TEPC-15 reactive glycans; a change in glycan class detection upon HF treatment, presumably resulting from a change in surface activity due to release of phosphate esters; and appearance of the *m/z* 1768.4 ion described above. Further investigation of the Pc as well as the other classes of N-glycans is described below.

**Characterization of High Abundance N-Glycans**—For a more detailed structural analysis of the major glycans present, the Bio-Gel P-4 eluate was pooled more narrowly as shown in Fig. 1. Permethylation ± HF treatment was performed, and the glycans were analyzed by MALDI-TOF MS to estimate pool complexity and relative abundance of glycan species. The nar-
rowly pooled glycans (Fig. 1, pools A–E) contained the major N-glycans from mixed stage *C. elegans* (see Table I and Supplemental Material Fig. 2).

One-dimensional and two-dimensional 1H DQF-COSY NMR spectroscopy of each pool revealed the glycosidic linkage and sequence of the pools A–E and confirmed the monosaccharide components determined previously by GC/MS. Anomeric proton chemical shifts and their relative proton intensities, as integrated from one-dimensional spectra, are shown in Table II. The two-dimensional 1H DQF-COSY experiment allows C1-H/C2-H correlations to be well resolved which, in combination with the library of these chemical shifts and associated coupling constant data, allows the linkage arrangement of oligosaccharides in complex mixtures to be assigned accurately (20–31). The DQF-COSY C1-H/C2-H region of pools A–E is shown in Fig. 8.

The C1-H/C2-H correlation of glycan constituents are monosaccharide- and linkage-dependent. The results in Fig. 8 confirm the identities and linkage arrangement of the monosaccharide constituents of the pool. Apportionment of the anomeric resonance intensity of the one-dimensional NMR spectra gave each pool the constituent isomers and relative abundance (see Table II).

HPAEC was used to isolate each pool component, and the compositions of individual glycan species were verified by linear mode MALDI-TOF MS analysis (HPAEC + MS) (Figs. 2–5 and Fig. 7), which agreed well with the sequence and abundance of each isomer as determined by NMR analysis (see Tables II and III and Scheme II). In Scheme I four representative oligosaccharide structures, containing all residue types detected, are presented for the classes of N-glycans found in the high resolution study of rowly pooled *C. elegans* glycans. Anomeric chemical shifts, linkage type, and the residue numbering scheme are provided in Scheme I. The NMR results will be discussed here along with HPAEC and MALDI-TOF MS results as these data are highly complementary. The isomers that were determined in each pool are shown in Scheme II.

### Detailed Analysis of pool A Oligosaccharides—The majority (96%) of the pool A components were high mannose glycans. 70% were Man9GlcNAc2 and were identical to the unprocessed form found in the ER prior to mannosidase trimming. Another 26% were Man8GlcNAc2 and differed from the previous glycoform only by the absence of the middle arm 1,2-Man residue after, presumably, being trimmed by ER 1,2-mannosidase. Three trace components were detected by HPAEC + MS in-
including two isomers of Fuc2Man5GlcNAc2, eluting at separate positions by HPAEC, and one Man$_n$GlcNAc$_2$ isomer. The trace components were not in sufficient abundance to determine their fine structure. The pool isomers are shown in Scheme II. Key NMR proton resonances used in isomer assignments are underlined in Table II. Mass spectrometry, HPAEC + MS, exoglycosidase data, and NMR analyses that resulted isomer assignments are discussed in detail below.

HPAEC separation of pool A glycans produced four well resolved peaks, which were centered at 4.6 (2%), 7.4 (2%), 15.1 (35%), and 19.7 (61%) min. MALDI-TOF MS analysis of the permethylated pool A glycans produced ions at $m/z$ 1931.2, 1930.3, 2192.4, and 2396.0, respectively (see Fig. 2). These are consistent with the sodium-adducted ions of two isomers of Fuc$_2$Man$_n$GlcNAc$_2$, Man$_n$GlcNAc$_2$, and Man$_n$GlcNAc$_2$, respectively. Reflectron mode MALDI-TOF MS analysis of the permethylated Bio-Gel P4 glycans did not detect the low abundance monoisotopic ions for the Fuc$_2$Man$_n$GlcNAc$_2$ isomers, but a previously undetected low abundance ion consistent with 2-O-mannosidase activity here. Reflectron mode MALDI-TOF MS analysis of the permethylated pool A glycans produced ions at $m/z$ 1931.2, 1930.3, 2192.4, and 2396.0, respectively (see Fig. 2). These are consistent with the sodium-adducted ions of two isomers of Fuc$_2$Man$_n$GlcNAc$_2$, Man$_n$GlcNAc$_2$, and Man$_n$GlcNAc$_2$, respectively. Reflectron mode MALDI-TOF MS analysis of the permethylated Bio-Gel P4 glycans did not detect the low abundance monoisotopic ions for the Fuc$_2$Man$_n$GlcNAc$_2$ isomers, but a previously undetected low abundance ion consistent with 2-O-mannosidase activity here.

The one-dimensional $^1$H NMR spectra of pool A contained 10.47 mol of anomeric proton resonance intensity distributed among 5.41 (0.70 mol), 5.34 (0.96 mol), 5.31 (0.96 mol), 5.19 (0.43 mol), 5.14 (0.96 mol), 5.12 (0.02 mol), 5.09 (0.29 mol), 5.05 (2.64 mol), 4.87 (0.98 mol), 4.76 (1.00 mol, measured at 318 K), 4.70 (1.00 mol), and 4.60 ppm (0.55 mol). The above chemical shift values, along with the C1-H/C2-H correlations discussed below, are well documented identifiers for residues found within high mannose glycans (23, 27). Resonance intensity values, along with the C1-H/C2-H correlations discussed below, are well documented identifiers for residues found within high mannose glycans.

Subtracting the resonance intensity in the two partial structures from pool resonance intensities leaves 0.70 mol at 5.40 ppm, indicating that 70% of pool A glycans escaped this activity here. Reflectron mode MALDI-TOF MS analysis of the permethylated pool A glycans produced ions at $m/z$ 1931.2, 1930.3, 2192.4, and 2396.0, respectively (see Fig. 2). These are consistent with the sodium-adducted ions of two isomers of Fuc$_2$Man$_n$GlcNAc$_2$, Man$_n$GlcNAc$_2$, and Man$_n$GlcNAc$_2$, respectively. Reflectron mode MALDI-TOF MS analysis of the permethylated Bio-Gel P4 glycans did not detect the low abundance monoisotopic ions for the Fuc$_2$Man$_n$GlcNAc$_2$ isomers, but a previously undetected low abundance ion consistent with 2-O-mannosidase activity here.

The one-dimensional $^1$H NMR spectra of pool A contained 10.47 mol of anomeric proton resonance intensity distributed among 5.41 (0.70 mol), 5.34 (0.96 mol), 5.31 (0.96 mol), 5.19 (0.43 mol), 5.14 (0.96 mol), 5.12 (0.02 mol), 5.09 (0.29 mol), 5.05 (2.64 mol), 4.87 (0.98 mol), 4.76 (1.00 mol, measured at 318 K), 4.70 (1.00 mol), and 4.60 ppm (0.55 mol). The above chemical shift values, along with the C1-H/C2-H correlations discussed below, are well documented identifiers for residues found within high mannose glycans (23, 27). Resonance intensity values, along with the C1-H/C2-H correlations discussed below, are well documented identifiers for residues found within high mannose glycans.
structure defined above is assigned the resonance intensity from residues 7 and 10 to be added yielding archetypal Man₉GlcNAc₂ as isomer A₁ and is 70% of pool isomers, which closely agrees with the HPAEC + MS analysis of the pool that gave a 19.7-min peak with a relative abundance of ~61% and m/z consistent with the Man₉GlcNAc₂ composition (see Fig. 2). This leaves 0.26 mol of intensity at 5.09 ppm and 0.02 mol of resonance each at 5.05, 5.12, and 5.09 ppm. To the remaining 0.26 mol of the Man₇GlcNAc₂ partial structure is added 0.26 mol of 5.09 ppm yielding Man₈GlcNAc₂ isomer A₂ as 26% of pool glycans, in good agreement with mass (m/z 2192.4) and relative abundance (34%) detected for the HPAEC + MS peak eluting at 15.1 min. This oligosaccharide is identical to ER-processed Man₈GlcNAc₂. On balance 0.02 mol each of 5.05, 5.12, and 4.87 ppm and 0.03 mol of 5.09 ppm resonance remains to be assigned.

As stated above, the HPAEC + MS analysis of pool A showed the presence of trace amounts of Fuc₂Man₅GlcNAc₂ and Man₃GlcNAc₄ ranging from ~1 to 2% of total pool isomers. In addition to the above remaining resonance intensity, a trace amount of NAc resonance was seen at 2.09 ppm close to those of core GlcNAc residues 1 and 2, consistent with β1,2GlcNAc₁₃ (see Table II). A slight foot at the downfield edge of the resonance at 4.87 ppm was seen, which may be from terminal α₁,6-linked residue 4 and α₁,6-Fuc (33). As discussed below in pools B and C, anomeric resonance was seen centered at 5.33 for α₁,2Fuc ppm, and thus a trace amount of this Fuc linkage may be present in this pool. A trace amount of Fuc methyl proton resonance was seen at 1.20 ppm. To the Manβ₁,4GlcNAcβ₁,4GlcNAcα₁β partial structure the remaining resonance is assigned. However, the remaining resonance does not allow linkage assignment with completeness and certainty. Trace resonance at 5.12 ppm hints that the Man₃GlcNAc₄ likely contains residue 1₃ and is assigned here as A₃, which is supported by the trace amount of m/z 1661.4 ion seen in the reflectron MALDI-TOF MS analysis of the pool A isomers (see Table I and Schemes I and II). The low intensity DQF-COSY 4.63 C₁-H/3.85 C₂-H resonance shows the presence of trace α₁,6Fuc, which causes a spatially dependent downfield shift of GlcNAc residue 2 (34). The Fuc residue was likely to be present in the Fuc₂Man₅GlcNAc₂ and is assigned here as A₄, which is consistent with the compositions detected by MS (m/z 1931.2 and 1930.3) of pool HPAEC peaks seen at 4.6 and 7.4 min, respectively. As seen in Table III, the glycan isomer compositions and relative abundance derived by ¹H NMR closely match those detected by HPAEC + MS.

**Detailed Analysis of Pool B Oligosaccharides**—Pool B contained high mannose, hybrid, and fuc-panmannosidic gly-
cans. The high mannose isomer, Man$_7$GlcNAc$_2$, was most abundant. This glycan was apparently trimmed by mannosidases with activities similar to mammalian ER and Golgi α1,2-mannosidases. The hybrid isomer was Fuc$_1$Man$_5$GlcNAc$_2$ and contained core-bisecting 1,4GlcNAc and core-linked α1,6Fuc. The fuco-pausimannosidose isomers retained the ER Man$_9$GlcNAc$_2$ central or lower arm, which was terminally substituted with Fuc. The pool isomers are shown in Scheme II. Key NMR proton resonances used in isomer assignments are underlined in Table II. Mass spectrometry data, HPAEC data, and NMR analysis that resulted in isomer assignments are discussed in detail below.

Reflectron MALDI-TOF MS of the Bio-Gel P4 pool B glycans was performed (Table I). Abundant molecular ions were seen at m/z 1987.3, 1753.3, and 1590.4, which were consistent with sodium-adducted ions of Man$_7$GlcNAc$_2$, Fuc$_1$Man$_5$GlcNAc$_2$, and Fuc$_1$Man$_5$GlcNAc$_3$. Low abundance ions were seen at m/z 1549.4, 1497.3, and 1007.0, which are consistent with Fuc$_1$Man$_5$GlcNAc$_2$, Man$_7$GlcNAc$_2$, and Man$_7$GlcNAc$_3$, respectively. With HF treatment of pool glycans prior to permethylation, ions were seen at m/z 1989.4, 1754.8, 1580.4, 1550.4, 1376.0, and 1006.6, which are consistent with the sodium-adducted ions of Man$_7$GlcNAc$_2$, Fuc$_1$Man$_5$GlcNAc$_2$, Man$_7$GlcNAc$_2$, Fuc$_1$Man$_5$GlcNAc$_2$, Man$_7$GlcNAc$_2$, and Man$_7$GlcNAc$_3$, respectively.

HPAEC + MS analysis of pool B isomers produced four resolved peaks (see Fig. 3), which eluted at 3.5, 4.3, 11.4, and 12.3 min. The first peak eluted at 3.5 min in the void volume, resolved peaks (see Fig. 3), which eluted at 3.5, 4.3, 11.4, and 12.3 min. The first peak eluted at 3.5 min in the void volume, contained a low abundance ion at m/z 1510.1, which is consistent with Fuc$_1$Man$_5$GlcNAc$_2$, indicating that two isoforms with this composition exist in the pool B. Therefore, by HPAEC + MS pool B contains Fuc$_1$Man$_5$GlcNAc$_2$ (13%), Man$_7$GlcNAc$_2$ (56%), and two isomers of Fuc$_1$Man$_5$GlcNAc$_2$ (23 and 8%) as the four major components.

The one-dimensional $^1$H NMR spectra of pool B contained resonance intensity at 5.41 (0.15 mol, residue 10), 5.34 (0.65 mol, residue 5), 5.33 (0.27 mol, residue 15, see below), 5.31 (0.12 mol, residue 8), 5.19 (0.44 mol, residue 14), 5.14 (0.53 mol, residue 6), 5.09 (0.88 mol, residue 7, and 5), 5.05 (1.33 mol, residue 8, 9, and 11), 4.92 (0.32 mol, terminal residue 4), 4.89 (0.20 mol, residue 12), 4.87 (0.68 mol, residue 4,), 4.70 (0.56 mol, residue 1b), 4.60 (1.00 mol, residue 2), and 4.46 ppm (0.20 mol, residue 14). At 318 K residue 3 was present at 1.00 mol of integrated resonance intensity at 4.76 ppm but was obscured at 300 K by the HDO peak (see Table II). This gives a total integrated anomer intensity of 8.33 mol. Residues 1–3 are present in all structures and will not be considered in the allocations below. In the DQF-COSY spectra of the pool, a strong cross-peak was seen at 5.34 C1-H/4.10C2-H and was from 2-O-substituted residue 5. A unique cross-peak was seen at 5.33 C1H/3.86 C2-H ppm, and the C2-H/C3-H cross-peak of the scalar coupled system was seen at 3.86/4.00 ppm. The measured coupling constants for the system were $J_{1,2}$, 3.5 Hz and $J_{2,3}$, 9.9 Hz, which are consistent with α-1,Fuc (35). The presence of 2-O-substituted α1,6-linked upper arm residue 6 assignment is confirmed by the presence of a strong DQF-COSY cross-peak at 5.14 C1H/4.02C2-H ppm. The cross-peak at 5.05 C1-H/4.07 C2-H ppm confirms the presence of terminal α1,2Man residues 8–10. A low cut close to base line in the DQF-COSY spectrum of the pool revealed a cross-peak at 4.89 C1-H/3.78 C2-H ppm, coinciding with the 0.20 mol resonance at 4.89 ppm in the one-dimensional spectrum (boxed in Fig. 8). This is diagnostic of core-linked α1,6-Fuc (33). The cross-peak at 4.87 C1-H/4.13 C2-H ppm confirms the presence of 3, 3-O-di-O-substituted Man residue 4 (25). The presence of core bisecting GlcNAc residue 14 is confirmed by DQF-COSY cross-peak 4.45 C1-H/3.65 C2-H ppm (36).

Residues 7 and 8 are present at only 0.15 and 0.12 mol in 2-O-substituted forms, as indicated by their relative resonance intensities at 5.41 and 5.31 ppm, respectively. Therefore, the major pool isomer, Man$_7$GlcNAc$_2$, detected in the Bio-Gel P4 pool by MS and subsequently by pool HPAEC + MS analysis (~56% peak area at 11.4 min) must contain 2-O-substituted residues 6 (at 5.14 ppm substituted by residue 9 at 5.05 ppm) and 5 (at 5.34 ppm substituted by terminal residue 8 at 5.05 ppm) in order to reach the requisite size. The only Fuc$_1$Man$_5$GlcNAc$_2$ that can contain 6 or 9 must be devoid of residue 5 in order to contain the required number of Man residues. The absence of 5 causes the C2-H/C3-H resonance of 3 to shift to 4.09/3.66 ppm (37). The moderate $J_{2,3}$ (~3.5) and strong $J_{2,4}$ (~10 Hz) coupling constants of mannose allow detection of low abundance resonance in the C2-H/C3-H region of the two-dimensional NMR spectrum. Because no resonance

![Fig. 3. HPAEC + MS analysis of pool B glycans.](image-url)
was seen at 4.09 C2-H/3.66 C3-H ppm, all pool glycans contain residue 5. Thus, due to size constraints the Fuc1Man5GlcNAc2 components in the pool cannot contain 2-O-substituted 6 or its substituent residue 9. Therefore, all of the 0.53 mol of resonance at 5.14 ppm from 2-O-substituted 6 exists in the Man5GlcNAc2 pool B isomer. The 56% pool abundance of the 11.4 min HPAEC + MS detected Man5GlcNAc2 peak supports the size and 53% relative abundance of this NMR assignment. Hence, the major pool isomer is Manα1,2Manα1,3(Manα1,2Manα1,6(Manα1,3Manα1,6)Manβ1,4GlcNAcβ1,4GlcNAcβ assigned here as isomer B1 (see Scheme II and Table II). This assignment consumes 0.53 mol of intensity at 5.34, 5.14, 5.09, and 4.87 ppm for residues 5–7 and 4 and 1.06 mol at 5.05 ppm for residues 8 and 9. On balance this leaves the following intensities to assign: 0.12 mol at 5.34; 0.27 mol at 5.33 ppm; 0.12 mol at 5.31 ppm; 0.35 mol at 5.09 ppm; 0.27 mol at 5.05 ppm; 0.32 mol at 4.92 ppm; 0.20 mol at 4.89 ppm; 0.15 mol at 4.87 ppm; and 0.20 mol at 4.46 ppm, respectively. The HPAEC + MS detected Fuc1Man5GlcNAc2 peak represented ~13% of components. NMR analysis revealed 0.20 mol each of α1,6Fuc residue 12 and β1,4GlcNAc residue 14. Assignment of these residues to the archetypal core mannosidase yields Manα1,3(Manα1,6(β1,4GlcNAcβ1,4Manβ1,4GlcNAcβ1,4FucO1,6)βGlcNAcαβ assigned here as isomer B2. This assignment consumes 0.20 mol of resonance intensity for terminal α1,3Man 5 (5.09 ppm), terminal α1,6Man 4 (4.92 ppm), β1,4GlcNAc 14 (4.46 ppm), and core α1,6Fuc 12 (4.89 ppm). On balance this leaves 0.12 mol at 5.34, 5.31, and 4.92 ppm, 0.15 mol at 5.41, 5.09, and 4.87 ppm, and 0.27 mol at 5.33 ppm, and 0.05 ppm with 27% of components left to assign.

The two remaining components have the composition Fuc1Man5GlcNAc2 as determined by HPAEC + MS. Both contain α1X-Fuc present at 5.33 ppm. Guerardel et al. (3) recently reported a terminal α1,2Fuc substitution in C. elegans mucin-type glycans, and its 5.325 C1-H/3.815 C2-H ppm chemical shift is virtually identical to that observed for the Fuc studied here. We tentatively assigned the Fuc here as terminal α1,2Fuc, an assignment that is supported in the following paragraphs. The presence of 0.15 mol of resonance at 5.41 for 2-O-substituted 7 and its substituting terminal residue 10 (0.15 mol of the 0.27 mol at 5.05 ppm) requires that one of these isomers contain this Manα1,2Manα1,3-disaccharide moiety. The presence of 0.12 mol each of 2-O-substituted 5 at 5.34 ppm, 2-O-substituted α1,2-linked 8 at 5.31 ppm, and terminal α1,2-linked Man 11 (the balance of 0.27-0.15 mol of 10 at 5.05 ppm) indicates that ~0.12 mol of lower arm Manα1,2Manα1,2-Manα1,3- is present. Assigning the 0.12 mol of the 5 plus 7 Manα1,2Manα1,3-disaccharide moiety to the core Manα1,GlcNAc yields the Manα1,GlcNAc partial structure Manα1,2Manα1,3(Manα1,6)Manβ1,4GlcNAcβ1,4GlcNAcβ assigned here as Ba. This uses 0.12 mol of the remaining resonances at 5.05, 5.34, 5.31, and 4.92 ppm leaving 0.15 mol each at 5.41, 5.09, 5.05, and 4.87 ppm and 0.27 mol at 5.33 ppm. Assigning all but the α1,2Fuc resonance at 5.33 ppm to the core Manβ1,4GlcNAcβ1,4GlcNAcαβ gives Manα1,2Manα1,3-Manα1,6(Manα1,3Manβ1,4GlcNAcβ1,4GlcNAcαβ assigned here as Manα1,GlcNAc partial structure Bβ.

We cannot say with absolute certainty where the α1,2Fuc is in Ba and Bβ from the NMR data presented here. However, both contain complete arms of the ER Manα1,GlcNAc with intact α1,2Man residues, and this observation led us to hypothesize the existence of a capping substitution of terminal α1,2Man by α1,2Fuc.

To investigate further the position of Fuc substitution, we performed jack bean α-mannosidase digestion of the Bio-Gel P-4 pool B glycans, and the products were permethylated and analyzed by linear mode MALDI-TOF MS. Sodiated molecular ions were seen at m/z 765.0, 1550.3, and 1754.2 (data not shown). The m/z 765.0 ion corresponds to Manα1,GlcNAc from complete digestion of pool high mannose glycans. The detection of ions at m/z 1550.3 and 1754.4 is consistent with Fuc1Man α1,GlcNAc2 and Fuc1Manα1,GlcNAc that are α1,2Fuc substituted at α1,2Man as hypothesized. The assignment of 0.15 mol of α1,2Fuc as a substituent of residue 10 of partial structure Bβ gives isomer B3 and is 15% of the pool. The assignment of 0.12 mol of α1,2Fuc as substituting residue 11 gives isomer B4 and is 12% of pool glycans. These assignments complete the anomeric proton allocations of the pool. B4 has a terminal α1,6Man, which is accessible to jack bean α-mannosidase. Isomer B3 has terminal α1,3Man, which is slowly accessible to the enzyme. This situation would produce species that, upon permethylation, would exhibit [M + Na] at both m/z 1550.3 and 1754.4, just as observed. We should note that the predicted products of pool glycans with β1,4GlcNAc and α1,6Fuc core substituents were absent from the MALDI-TOF MS spectra of the digestion product, but their presence in the pool was detected by composition and linkage signatures in the pool B MALDI-TOF MS and NMR spectra, respectively. The pool B isomer distribution and relative abundance assigned by NMR are in close agreement with the relative abundance and compositions detected by HPAEC + MS (see Tables II and IV and Scheme II).

**Detailed Analysis of Pool C Oligosaccharides**—Pool C contained high mannose, hybrid, and fuco-pausimannosidic oligosaccharides. The high mannose glycans were Manα1,GlcNAc2 and were probably the result of mammalian-type ER and Golgi mannosidase trimming. The hybrid glycans were Manα2,4-GlcNAc3 and contained either lower arm α1,2GlcNAc or core-bisecting α1,4GlcNAc substitutions. The fuco-pausimannosidic oligosaccharides were Fuc1Manα2,4-GlcNAc2, which contained either the novel terminal α1,2Fuc or core α1,6Fuc additions. In the case of α1,2Fuc substitution, the isomers retained the ER Manα1,GlcNAc2 central or lower arm, which was terminally substituted with Fuc. The α1,6-fucosylated isomer was incompletely trimmed and retained a lower arm α1,2Man residue, which suggested that a C. elegans α1,6-fucosyltransferase acts in the absence of lower arm β1,2GlcNAc, unlike that of mammalian systems (38). The pool isomers are shown in Scheme II. Key NMR proton resonances used in isomer assignments are underlined in Table II. Mass spectrometry, HPAEC + MS, exoglycosidase data, and NMR analyses that resulted in isomer assignments are discussed in detail below.

MALDI reflectron TOF MS analysis of the permethylated Bio-Gel P4 pool C glycans contained monoisotopic ions at m/z 1783.4, 1579.3, 1549.0, 1538.3, 1416.3, and 1171.4 (see Table I and Supplemental Material Fig. 2), which are consistent with the following respective sodiated glycan compositions: Manα1,GlcNAc2, Manα1,GlcNAc2 Fuc1Manα1,GlcNAc2, Manα1,GlcNAc1, Manα1,GlcNAcβ, and Manα1,GlcNAc (see Table I). HF hydrolysis of pool C glycans prior to permethylation and MALDI reflectron TOF MS analysis gave a nearly identical unidentified monoisotopic ion series except that Manα1,GlcNAc2 was not seen. Judging from the Bio-Gel P-4 elution position and HPAEC + MS analysis (Figs. 1 and 5), it was concluded that the Manα1,GlcNAc2 was produced during the permethylation procedure by hydrolysis of other pool glycans.

HPAEC + MS analysis of pooled and permethylated glycans (Fig. 5) yielded 9 glycans wholly consistent with Bio-Gel P-4 MS detected compositions with the addition of minor glycan species appearing at m/z 1212. 3 and 1754.2. The pool C HPAEC + MS elution position, composition, and relative pool abundance as seen by electrochemical detection and m/z were
as follows: 4.6 min, Fuc1Man1GlcNAc2 (6%, m/z 1549.9); 5.3 min, Man1GlcNAc2 (6%, m/z 1212.3); 5.7 min, Man1GlcNAc3 (4%, m/z 1416.7); 6.5 min, Man1GlcNAc3 (13%, m/z 1416.6); 8.6 min, Man2GlcNAc3 (4%, m/z 1579.4); 9.1 min, Man2GlcNAc3 (2%, m/z 1622.5); 10.3, Man2GlcNAc2 (21%, m/z 1783.8), 11.8 min, Man1GlcNAc2 (40%, m/z 1783.7); and 13.3 min, Fuc1Man1GlcNAc2 (5%, m/z 1754.21). Note that there are two species of Man2GlcNAc2 (eluting at 5.70 and 6.52 min), and two species of Man1GlcNAc2 (eluting at 10.29 and 11.83 min). The linkages were derived by NMR as described below.

Total integrated anemic proton intensity of pool C was 7.46 mol, which was distributed in peaks centered at 5.41 (0.04 mol, residue 7), 5.33–5.34 (0.48 mol, residues 5 and 15), 5.19 (0.41 mol, residue 1a), 5.14 (0.26 mol, residue 6), 5.12 (0.18 mol, residue 5), 5.09 (1.09 mol, residues 7 and 5), 5.05 (0.69 mol, residues 8, 9, and 11), 4.92 (0.68 mol, residues 4 and 6), 4.89 (0.66 mol, residue 12), 4.87 (0.75 mol, residue 4), 4.76 (1.0 mol, residue 3, integrated at 318 K), 4.70 (0.60 mol, residue 1b), 4.60/4.63 (1.00 mol, residue 2), 4.55 (0.18 mol, residue 13), and 4.46 ppm (0.04 mol, residue 14). The corresponding DQF-COSY C1-H/C2-H cross-peaks for the one-dimensional NMR resonances of the above pool were seen at the following positions: 5.34 C1-H/4.10 C2-H ppm for 2-O-substituted residue 5; 5.19 C1-H/3.83 C2-H ppm for residue 1a; 5.14 C1-H/4.02 C2-H for 2-O-substituted a1,6-linked residue 6; 5.12 C1-H/4.06 C2-H ppm for 2-O-substituted (a1,2GlcNAc 13 substituted) a1,3-linked residue 5 (39); 5.09 C1-H/4.06 C2-H ppm for terminal a1,3-linked residue 5; 5.05 C1-H/4.06 C2-H ppm for terminal a1,2-linked residues 8 and 9; 4.92 C1-H/3.98 C2-H ppm for terminal a1,6-linked residues 4 and 6; 4.87 C1-H/4.13 C2-H ppm for 3, (6) di-O-substituted residue 4; 4.70 C1-H/3.63 C2-H ppm for residue 1b; resonance of residue 2 was split between 4.63 C1-H/3.81 C2-H and 4.60 C1-H/3.78 C2-H ppm due to the presence in the pool of compounds containing Fuc residue 12, as in pool B; 4.54 C1-H/3.70 C2-H for β1,2-linked GlcNAc residue 13 (39); and 4.46 C1-H/3.70 C2-H 3.65 ppm for bisecting β1,4GlcNAc (34). The anemic protons of residues 1–3 are present in all glycans, and their proton allocations will not be discussed further. The α1,6Fuc residue 12 was only 0.06 mol and its DQF-COSY cross-peak was not seen due to low abundance.

It should be noted that no resonance intensity was seen at 4.09 C2H/3.66 C3-H ppm indicating that all pool components contain residue 5 (see above). Man1GlcNAc2 accounts for ~60% of pool components by HPAEC + MS analysis and is distributed between two species at ~20 and 40% of total pool components (see Fig. 5). 2-O-Substituted residue abundance is low with only 0.05 mol as 7 and none as 8, which would be observed at 5.41 and 5.31 ppm, respectively. Therefore, any Man1-, GlcNAc2 partial structure must contain residues 3, 4, 5, and 7 (see Scheme I), giving a core of the following linkage sequence Man1,6(Mana1,3)Man1,6(Mana1,3)Man1,4GlcNAc1,4GlcNAc/b. The difference between the two Man1GlcNAc2 iso- mers is thus, in the addition of one of the terminal α1,2-linked residues 8 or 9, which substitute residues 5 and 6, respectively.

At 5.14 ppm 0.26 mol of resonance is integrated. Assigning 26% of the pool glycans as isomer C1 consumes all the resonance of residue 6 and 0.26 mol each of residue 4 at 4.87 ppm, terminal α1,2-linked Man 9 resonance at 5.05 ppm, and 0.52 mol at 5.09 ppm for residues 5 and 7. After this isomer assignment, the remaining resonance intensities left to assign are as follows: at 5.41 ppm, 0.05 mol (residue 7); 5.34 ppm, 0.48 mol (residue 5); 5.12 ppm, 0.18 mol (residue 5); 5.09 ppm, 0.57 mol (residues 7 and 5); 5.05 ppm, 0.43 mol (residues 8, 9, and 11); 4.92 ppm, 0.68 mol (residues 4 and 6); 4.89 ppm, 0.06 mol (residue 12); 4.87 ppm, 0.49 mol (residue 4); 4.55 ppm, 0.18 mol (residue 13); and 4.46 ppm, 0.04 mol (residue 14). This leaves 74% of the pool components left to assign.

The 0.06 mol of resonance at 4.89 ppm is from core-linked α1,6-Fuc residue 12 (see Table II). Exactly 0.08 mol of resonance is present at 5.33 ppm from a1,2Fuc residue 15. The presence of these two Fuc residues is supported by the HPAEC + MS analysis of pool C that detected Fuc1Man1GlcNAc2 (6% integrated area, Fig. 4B) and Fuc1Man1GlcNAc2 (5% of integrated area, Fig. 4J). This leads to the partial structural assignment Fuc1,X: Man1a1,3(Mana1,6)Man1β1,4GlcNAc1,4GlcNAc/b giving partial structure Ca as 10% of the pool. The NMR analysis (see Table II) revealed 0.18 mol of β1,2GlcNAc (4.55 ppm, residue 13) and 0.04 mol of core bisecting β1,4GlcNAc (4.46 ppm, residue 14). These assignments allow the following partial structures to be assigned: GlcNAcβ1,2Man1a1,3Man1β1,4GlcNAcβ1,4GlcNAc/b as partial structure Cβ, and Man1a1,3(Mana1,6)GlcNAcβ1,4Man1β1,4GlcNAcβ1,4GlcNAc/b as partial structure Cγ as 18 and 4% of pool components, respectively. The above assignments consume all resonance intensity at 5.12 ppm for 5 when substituted by 13 (39), 4.89 (α1,6-Fuc residue 12), 4.55 (β1,2GlcNAc residue 13), and 4.46 ppm (β1,4GlcNAc residue 14) and 0.05 mol of intensity at 5.34 ppm (α1,2Fuc residue 15). The resonance intensity arising from the remaining partial structure’s residues 4 and 5 in Ca, Cβ, and Cγ will be addressed below as they can be affected by their substitution state. The remaining reso-
nance intensity left to assign are from 2-O-substituted Man residues 7 (5.41 ppm, 0.05 mol) and 5 (5.34 ppm, 0.43 mol), terminally linked Man 5 (5.09 ppm, 0.57 mol), terminally α1,2-linked Man residues 8 and 9 (5.05 ppm, 0.43 mol), terminally α1,6-linked residues 4 and 6 (4.92 ppm, 0.68 mol), and 3, (±)-di-O-substituted residue 4 (4.87 ppm, 0.49 mol).

The assignment of Fuc- and GlcNAc-substituted partial structures Ca (10%), Cβ (18%), and Cy (4%) above can be subtracted from the 74% of isomers left to assign, revealing the remaining 42% as the high mannose isomers Manα,GlcNAc2 plus Manα,GlcNAc2. By FAD detection Manα,GlcNAc2 is ~4% of the pool leaving 96% as Manα,GlcNAc2. As described above this Manα,GlcNAc2 isomer differs from the C1 Manα,GlcNAc2 isomer only in the absence of residue 9, and the presence of residue 8, which 2-O-substitutes residue 5 and leads to the assignment of isomer C2 (see Scheme II). This assignment leaves 0.05, 0.05, 0.19, 0.05, 0.30, and 0.11 mol left to assign, respectively, at 5.41, 5.34, 5.09, 5.05, 4.92, and 4.87 ppm.

Assignment of the 4% Cy as C3 consumes all of this partial structure and 0.04 mol of terminal residues α1,3Man 5 (5.09 ppm) and α1,6Man 4 (4.92 ppm), leaving 0.05, 0.06, 0.15, 0.05, 0.26, and 0.11 mol of resonance at 5.41, 5.34, 5.09, 5.05, 4.92, and 4.87 ppm left to assign. HPAEC + MS analysis detected Manα,GlcNAc2 (13%), Manβ,GlcNAc2 (6%), and Manγ,GlcNAc2 (2%). To the 0.18 mol of Cβ 0.12 mol of 4.92 ppm resonance (residue 4) is assigned giving C4 (see Scheme II). Of the remaining 0.06 mol of Cβ, 0.04 is assigned without further residue addition yielding isomer C5 (see Scheme II). To the remaining 0.02 mol of Cβ, 0.02 mol of terminal α1,3Man 7 is assigned giving isomer C6, which also necessitates that resonance of residue 4 be present at 4.87 ppm due to its α-2-o-substitution (see Scheme I). Thus, 0.02 mol is consumed at 4.87 ppm. Assignment of C4, C5, and C6 completes the allocation of C1,3Man 5GlcNAc2, Fuc 1Man4GlcNAc2 and Fuc 1Man5GlcNAc2 products were produced by the incomplete digestion of Fucα,Mannα,GlcNAc2 pool isomer C9 (4% of pool). C9 is identical to B3 and contains α1,3Man, which is slowly accessible to the α-mannosidase and results in the appearance of both Fucα,Mannα,GlcNAc2 and Fucα,Mannα,GlcNAc2 ions. Although product ions from glycans with α1,6Fuc and β1,4GlcNAc core substituents were not detected, their compositions and unique substituent signatures were seen by MALDI-TOF MS and NMR analysis, respectively. These data strongly support the presence of terminally linked α1,2Fuc as well as high mannose, and lower arm GlcNAc-substituted glycans derived here by NMR above. For every NMR-derived structure there was a corresponding pool isomer of the same predicted mass and nearly identical pool abundance as detected by HPAEC + MS (see Table III).

Detailed Analysis of Pool D Oligosaccharides—Pool D contained high mannose, hybrid, Pc-, and fuco-pausimannosidic glycans. The most abundant isomer was Manα,GlcNAc2. It was fully trimmed at the central, upper, and lower arm α1,2Man residues, consistent with activities identical to mammalian ER and Golgi α1,2-mannosidases. The hybrid glycan was Manα2-GlcNAc2 and contained core-bisecting β1,4GlcNAc. The Pc glycan was identical to the Manα,GlcNAc2 isomer of the pool but was substituted with three Pc groups. Fuco-pausimannosidic glycans included Fucα,Mannα,GlcNAc2 and Fucα,Mannα,GlcNAc2. The former was identical to the Manα,GlcNAc2 pool isomer but was substituted with core α1,6Fuc. The later retained the core mannotriose (Manα1,3(Manα1,6)Manβ1,4-) and contained core-linked α1,6Fuc. The pool isomers are shown in Scheme II. Key NMR proton resonances used in isomer assignments are underlined in Table II. Mass spectrometry, HPAEC + MS, and NMR analysis that resulted in isomer assignments are discussed in detail below.

MALDI-TOF MS analysis of the Bio-Gel P4 pool D permethylated glycans revealed ions at m/z 1579.2, 1497.2, 1375.3, 1345.3, and 1171.4, which are consistent with the sodium-added ions of Manα,GlcNAc2, Manβ (see Table I), Manβ,GlcNAc2, Fucα,Mannα,GlcNAc2, and Manβ,GlcNAc2. MALDI-TOF MS analysis performed on pool glycans with HF hydrolysis prior to permethylation produced ions consistent with the compositions above with the addition of low abundance ions at m/z 1066.5 and 1417.0, which are consistent with Manα1,GlcNAc2 and Manβ,GlcNAc2.

HPAEC analysis of pool glycoconjugates produced four peaks (Fig. 5A). MALDI-TOF MS analysis of the HPAEC pooled and permethylated glycans produced sodium-added ions at m/z 1755.8 (4.6 min), 1345.0 (10.3 min), and 1580.2 (18.1 min) consistent with Fucα,Mannα,GlcNAc2, Fucα,Mannβ,GlcNAc2, Manβ,GlcNAc2, respectively. MALDI-TOF MS of the 22.7 min peak in native form produced ions at m/z 1755.4, 1345.4, 1278.9, and 867.0, which are consistent with Pcα,Mannα,GlcNAc2 [Na+], Pi2,Pcα,Mannα,GlcNAc2 [Na+], Manα,GlcNAc2 [2Na-H+], and Pcα,Mannα,GlcNAc2 [H2+] (see Table I). These results strongly suggested the presence of a Pcα,Mannα,GlcNAc2 pool glycan.

In the one-dimensional NMR spectrum in addition to 1 mol of core residues 1–3, resonance intensity for the following anomeric chemical shifts were observed: 5.10 ppm, 1.74 mol (residues 5 and 7), 4.92 ppm, 1.00 mol (residues 4 and 6), 4.89 ppm, 0.35 mol (residue 12), 4.87 ppm, 0.74 mol (residue 4), and 4.45 ppm, 0.03 mol (residue 14), and at 3.23 ppm 2.0 mol of Pc methyl proton intensity was detected (see Table II). The indicated identity of each residue was confirmed by their respective DQF-COSY cross-peaks as follows: 5.19CH3-H/3.87CH2-H ppm, residue 1a; 5.10 C1-H/4.07 C2-H ppm, residues 5 and 7; 4.91 C1-H/3.99 C2-H ppm, residues 4 and 6; 4.89 C1-H/3.80 C2-H ppm, residue 12; 4.87 C1-H/4.14 C2-H ppm, residue 4; 4.78 ppm.
C1-H/4.26 C2-H ppm, residue 3; 4.70 C1-H/3.69 C2-H ppm, residue 1β; 4.67 C1-H/3.78 C2-H ppm, residue 2 in presence of 12 (40); 4.60 C1-H/3.78C 2-H ppm, residue 2 in absence of 12. As described for pools B and C, no resonance was detected at 4.09 C2-H/3.66 C3-H for residue 3 indicating that all pool isomers have residue 5, which substitutes 3. Because there is no detectable resonance intensity at 5.34 or 5.12 ppm for α, 1,3-linked lower arm substitution of residue 5 by residue 8 or 13, 5 is unsubstituted in all pool isomers. These observations allow the assignment of the partial structure Man3β,1Man3β,1,4GlcNAc αFucα1,6GlcNAcβ, which is present in 35% of pool isomers and is designated here as partial structure Do. This assignment leaves 1.39, 1.00, 0.74, and 0.03 mol of resonance at 5.10, 4.92, 4.87, and 4.45 ppm left to assign, respectively.

The 0.03 mol of intensity at 4.45 ppm allows the assignment of partial structure Manα3,1Manβ1,4GlcNAcαFucα1,6GlcNAcβ in 3% of pool isomers and is designated as partial structure Df. The partial assignment consumes the 4.45 ppm resonance and 0.03 mol of resonance at 5.10 ppm. This leaves 1.36, 1.00, and 0.74 mol of resonance at 5.10, 4.92, and 4.87 ppm left to assign, respectively.

In addition to a terminal α, 1,3-linked lower arm residue 5 as described above, the upper arm 6-O-linked α, 1,6Man 6 residue is also unsubstituted in all pool isomers as no resonance is detected at 5.14 or 4.54 ppm for residues 9 or upper arm α, 1,2-linked Man, respectively. Also, no resonance at 5.41 ppm for residue 10 was detected indicating that central arm residue 7 is also unsubstituted. These observations provide that the maximum high mannose structure that may exist in the pool is the Manα3,1GlcNAcα isomer Manα1,6(Manα1,3)Manα1,6(Manα1,3)Manβ1,4GlcNAcβ,1,4GlcNAcαβ and is designated as partial structure Dy. All non-mannose resonances are accounted for by Do and Df and total to 38% of pool isomers. By difference Dy is representative of 62% of the pool. The assignment of Dy consumes 0.62 mol of intensity at 4.92 and 4.87 ppm for terminal α, 1,6Man 6 and 3,6-di-O-substituted α, 1,6-linked Man 4. At 5.10 ppm 1.24 mol of intensity is consumed with the assignments of α, 1,3-linked residues 5 and 7. On balance this leaves 0.12 mol at 5.10 and 4.87 ppm and 0.38 mol at 4.92 ppm left to assign. PAD detection and HPAEC + MS analysis of the pool gave 38% Fuc-substituted isomers (13% FucαMan3β,1GlcNAcβ and 25% FucαMan3β,1GlcNAcβ) and 62% high mannose (55% Man6GlcNAcα and 7% PeMan3β,1GlcNAcβ), and HF-treated and permethylated pool glycan indicated a trace amount of Man3β,1GlcNAcβ. These compositions and relative abundance agree closely with the above distribution indicated by assignment of partial structures Da (35%, containing Fuc), Db (3%, containing bisecting GlcNAc), and Dy (62%, containing a Manα3GlcNAcβ core without Fuc or GlcNAc). With the pool’s HPAEC distribution of Man3β,1GlcNAcβ and PeMan3β,1GlcNAcβ and the relative abundance of Pe methyl proton intensity at 3.23 ppm (1.83 mol, see Table II), 55% of the pool is assigned as D1 and is structurally identical to Dy. On balance, the remaining 7% of Dy is assigned 3 Pe moieties yielding D2. As there are 9 methyl protons on each Pe moiety, the 1.83 mol integrated closely matches this 7% assignment (0.07 × 9 protons × 3 Pc = 1.89 mol). To confirm the presence of the Pc oligosaccharide, we performed one-dimensional and twodimensional DQF-COSY NMR experiments on 5 mM phospho-choline chloride calcium salt solution. DQF-COSY cross-peaks for methylene protons were seen at 4.16/3.57 and 3.73/3.63 ppm (Fig. 6A). Examination of the DQF-COSY spectrum of pool D revealed Pc methyl protons at 4.19/3.60 and 3.74/3.64 ppm, which are slightly deshielded compared with free Pc (compare Fig. 6, A and B). Along with the MALDI-TOF MS results of the pool above, these results are conclusive for the presence of PeMan3β,1GlcNAcβ isomer D2.

Approximately 13% pool glycans were FucαMan3β,1GlcNAcβ by HPAEC + MS analysis. To Da is added 0.12 mol each of 4.92, 4.87, and 5.10 ppm resonance for terminally linked α, 1,6Man 6 and 3,6-di-O-substituted α, 1,6-linked Man 4 and terminal α, 1,3-linked Man 7, respectively, yielding isomer D3 (see Scheme III). On balance this leaves 0.26 mol of resonance left to assign at 4.96 ppm.

The HPAEC + MS analysis of the pool showed that 25% of the pool was FucαMan6GlcNAcα by PAD detection. The structure of this compound is given by the assignment of 0.23 mol of the remaining resonance at 4.92 ppm from residue 4 to partial structure Do. This yields glycan D4 (see Scheme III). The HPAEC + MS analysis of pool isomers did not provide obvious evidence of any structure bearing bisecting α, 1,4GlcNAc. However, the HF-hydrolyzed and -permethylated glycan revealed low abundance ions at m/z 1417.0 and 1006.5 (see Table I), which are consistent with Manα1,3(Manα1,6)GlcNAcβ,1,4-Manβ1,4GlcNAcβ1,4GlcNAcαβ. The remaining 0.03 mol of 4.92 ppm of resonance is assigned to the remaining 0.03 mol of Do yielding isomer D5 (see Scheme II) completing the anomeric proton resonance assignment of pool D. Glycan compositions and relative pool abundance derived here by NMR were in close agreement with those found by HPAEC + MS (Table III).
Detailed Analysis of Pool E Oligosaccharides—The major pool isomer (96%) was Man₃GlcNAc₂ and contained the core mannotriose. A minor isomer was Fuc₁Man₃GlcNAc₂, which was identical to the major pool glycan with the addition of core-linked α₁,6Fuc. The pool isomers are shown in Scheme II. Key NMR proton resonances used in isomer assignments are underlined in Table II. Mass spectrometry, HPAEC + MS, and NMR analysis that resulted in isomer assignments are discussed in detail below.

The Bio-Gel P-4 permethylated glycans of pool E were examined by MS after treatment with HF, it is tempting to conclude that the 4.52/4.12 ppm resonance may be due to an undefined phosphorylated functional group. However, we cannot present definitive evidence for such a substitution.

DISCUSSION

We have determined that *C. elegans* contains a nearly contiguous series of N-glycans. Five classes were observed including high mannose, complex, hybrid, fuco-paucimannosidic, and Pc glycans. The high mannose, complex, and hybrid glycans show a high degree of conservation with those of mammals, whereas the last two are unique to *C. elegans*. This series provides insight into the nature, conservation, and uniqueness of N-glycosylation in this model organism.

The mannan Man₃₋₅GlcNAc₂ series suggest that *C. elegans* possesses mannosidase activities that are identical to mammalian ER α₁,2-mannosidase and Golgi mannosidases I and II (41). High field NMR analysis revealed the presence of hybrid glycans with GNT I- and GNT III-type additions, the enzymes that add β₁,2GlcNAc to the lower arm and bisecting β₁,4GlcNAc to the β₁,4Man, respectively. Three *C. elegans* homologues of GNT I have been cloned and their encoded activities demonstrated (1). Although chitobiose core-linked α₁,6Fucose was detected, the *C. elegans* enzyme appears to have looser substrate specificity than that of mammalian systems, where the lower arm GlcNAc (residue 13, Scheme I) is required for fucosyltransferase activity (38). Here we have found a series of core-fucosylated N-chains lacking the lower arm GlcNAc substitution, bisecting GlcNAc, or no GlcNAc at all. However, we cannot rule out the existence in *C. elegans* of a highly active Golgi β-N-acetylgalactosaminidase acting on nascent N-glycans in vivo as has been reported in a lepidopteran insect cell system (42).

A novel high mannose branch Fuc substitution was seen, which caps α₁,2-linked Man residues. This addition is novel, as we were unable to find any example in nature where fucose terminally substitutes mannose and, thus, strongly suggests the existence of a novel fucosyltransferase. It is reasonable to predict that the novel fucosylation seen in *C. elegans* mixed stage N-glycans is developmentally regulated, and we are presently investigating this possibility. The trematode *Schistosoma mansoni* differentially expresses fucose-containing moieties on its N-glycans throughout development, where core α₁,3Fuc is seen in eggs and miracidia; Lewis X glycans are seen in cercaria, and mainly core α₁,6Fuc is seen in adult worms (43). The differential expression is developmentally regulated (44).
Novel distal core and outer chain fucosylation patterns on *Hemonchus contortus* N-glycans have been reported, and their expressions are, apparently, developmentally regulated as well (18, 45). By using two-dimensional chromatographic mapping, Natsuka *et al.* (46) have reported that a minor portion of *C. elegans* mixed stage N-glycans released from mixed stage worms by hydrazinolysis contain core-linked α1,3Fuc. Although, according to our results, mixed stage *C. elegans* worms contain α1,6- and novel α1,2Fuc substitutions, we did not detect core α1,3-fucosylated N-glycans (Fig. 8).

Pe substituted Man₅GlcNAc₂ was detected in *C. elegans*. These N-glycans have been reported in *Acanthocheilonema viteae, Onchocerca volvulus*, and *Trichinella spiralis* (17, 18, 47, 48). The Pe-containing oligosaccharides are able to modulate B-lymphocyte activation through activation of protein tyrosine kinase and mitogen-activating protein kinase signal transduction pathways resulting in desensitization to downstream activation (49). It is a conserved modification across parasitic nematode and trematode species, and the biosynthetic components are now being pursued as novel drug targets (53).

**Fig. 7.** HPAEC + MS analysis of pool E glycans. The HPAEC trace is shown in A. MALDI-TOF MS analysis of permethylated glycans pooled at 4.7 min are shown in B.

GNT II adds β₁,2GlcNAc to the upper arm-ranked α₁,6Man and GNT IV adds β₁,4GlcNAc to the lower arm α₁,3Man, respectively, in preparation for tri- and tetra-antennary N-glycans (36, 38). Although complex glycans were not present as the major glycans, here MALDI-TOF MS analysis of size exclusion chromatography pooled glycans (pools a–d, Fig. 1 and Table I) strongly suggests that *C. elegans* produces complex N-glycans with up to six HexNAc residues (Hex₅HexNAc₅) and hybrid and/or complex chains with up to five HexNAc residues (Hex₅HexNAc₅). These oligosaccharides suggest that GNT II and IV type activities are present. Also, LacdiNAC- (43, 44) and chito-oligomer branches (17) have been reported in nematodes glycans. Therefore, it is possible that *C. elegans* produces these oligosaccharides.

Recently *C. elegans* gly-2 was found to possess GNT V activity *in vitro* (6-O-substitutes residue 4, see Scheme I). The gly-2 gene restored *Phaseolus vulgaris* (PHA) lectin-binding affinity in *lec-4* cells. The PHA lectins can detect the complex glycan determinant Galβ₁,4GlcNAcβ₁,2Man α₁,6-. It is possible that *C. elegans* produces the galactosylated branch, as some low abundance glycans, such as the Hex₅GlcNAc₅ glycan seen here by MALDI-TOF MS, may have contained it, but the Gal abundance was too low to detect by GC/MS.

Although novel carbohydrate modifications were seen, it is shown here that much of the *N*-glycan biosynthesis of *C. elegans* is consistent with that seen in mammalian systems. Addition of GlcNAc residues by activities similar or identical to GNT I, II, and III and likely other mammalian-type HexNAc transferases were seen. The N-glycan series defined here shows that there is a functional *C. elegans* orthologue for nearly every CDG type I and II studied to date. Currently, there is a growing constellation of human congenital disorders of glycosylation in which CDG type I and CDG type II classify cases involving N-glycan dolichol-linked precursor synthesis and transfer and subsequent processing, respectively. Presently, these include CDG types Ia–g and IIa–c for each of which the biochemical lesion has been identified (8, 51–54). The affected proteins are as follows: Ia, phosphomannomutase-2; Ib, phosphomannomutase-1; Ic, dolichyl-P-Glc:Man₉GlcNAc₂-PP-dolichylglycosyltransferase; Id, dolichyl-P-Man:Man₉GlcNAc₂-PP-dolichylmannosyltransferase; Ie, Dol-P-Man synthase (catalytic subunit); If, MPDU1 (Man-P-Dol-dependent utilization); Ig, dolichyl-P-Man:7GlcNAc₂-PP-dolichylmannosyltransferase Ia, UDP-GlcNAc:α-6-d-mannoside-β-1,2-N-acetylglucosaminyltransferase II; Ib, ER glucosidase I; Ic, GDP-fucose transporter; and Id, UDP-galactose:N-acetylgalcosamine-β-1,4-galactosyltransferase I. In fact, when expressed in CDG type Ic fibroblasts (LAD-II), a *C. elegans* GDP-fucose transporter reestablished expression of fucosylated glycoconjugates with high efficiency, demonstrating the existence of some level of functional compatibility between *C. elegans* and human systems (55).

In this communication we have reported the most complete characterization of *C. elegans* N-glycans to date. We are currently characterizing stage-specific glycosylation patterns in the worm. The present series has revealed a high degree of conservation with mammalian systems as well as novel substitutions including Pc and Man-linked Fuc. This nematode is well characterized genetically and developmentally, and its genome has been completely sequenced. The work presented here has characterized the N-glycans of *C. elegans*, which, along with its capacity as a developmental, genetic, and
FIG. 8. C1-H/C2-H section of DQF-COSY 1H NMR spectra of pools A–E. A–E are C1-H/C2-H sections of pools A–E, respectively. The boxed region in B was taken from a lower slice in the two-dimensional spectrum. For the identity of the numbered C1-H/C2-H resonances see Scheme I and Table III.
genomic model system, has revealed its suitability as a model system for the study of glycosylation processes such as those involved in CDG and Pto oligosaccharide biosynthetic pathways.

**Acknowledgments**—We thank the staff of the National Institutes of Health Mass Spectrometry Resource, (supported by National Institute of Health Grant P41 RR10888 and S10 RR15942), and we particularly thank Shiu-Yung Chan for the expertise in GC/MS and helpful suggestions. We thank Dr. Phillips Robbins for helpful suggestions in preparing this manuscript. We also thank the members of the NMR Core Facility for Structural NMR Spectroscopy at the Boston University School of Medicine for their resources.

**Note Added in Proof**—In a recent review article and at the Annual Conference of The Society for Glycobiology held November 9–12, 2002, others reported some similar structures: 1) Altman, F., Fabini, G., Ahorn, H., and Wilson, I. B. (2001) Biochimie 83, 703–712; 2) Geiser, H. A., Hanneman, A. J., and Reinhold, V. (2002) Glycobiology 12, 650 (Abstr. 27); and 3) Della, A., Haslam, S., Hitchen, P., Morris, H. R., Panico, M., and Smith, M. S. (2002) Glycobiology 12, 648 (Abstr. 23).

**REFERENCES**

1. Chen, S., Zhou, S., Sarkar, M., Spence, A. M., and Schacht, H. (1999) J. Biol. Chem. 274, 288–297
2. Hagen, F. K., and Nehre, K. (1996) J. Biol. Chem. 271, 8268–8277
3. Guerardel, Y., Balanzino, L., Maes, E., Leroy, Y., Coddeville, B., Oris, R., and Strecker, G. (2001) Biochimie 83, 167–182
4. Warren, C. E., Kiyona, A., Partridge, A. E., and Dennis, J. W. (2002) Glycobiology 12, 643–649
5. Bulk, D. A., Wei, G., Toyoda, H., Kinoshita-Toyoda, A., Waldrup, W. R., Esko, J. D., Robbins, P. W., and Selleck, S. B. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 10838–10843
6. Berninsone, P., Hwang, H. Y., Zemtseva, I., Harvitz, H. R., and Hirscheberg, C. B. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 3738–3743
7. Aebi, M., Helenius, A., Schenk, B., Barone, R., Fiumara, A., Berger, E. G., Hennet, T., Imbach, T., Stutz, A., Bjursell, C., Ullner, A., Wahlstrom, J. G., Briones, P., Cardo, E., Clayton, P., Winchester, B., Cormier-Dalre, V., de Lonlay, P., Cuer, M., Dupre, T., Seta, N., de Koning, T., Dorland, L., de Leos, P., Kupers, L., and et al. (1999) Glycoconjug. J. 16, 669–678
8. Freeze, H. H. (2001) Glycobiology 11, R129–R143
9. Bollag, M. D., Edelstein, S. J. (1990) Protein Methods, pp. 95–139, 181–208, Wiley-Liss, Inc., New York
10. Bearden, J. C., Jr. (1973) Biochim. Biophys. Acta 297, 9–59
11. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956) Anal. Biochem. 8, 350–356
12. Verostek, M. F., Lubowski, C., and Trimbble, R. B. (2000) Anal. Biochem. 278, 111–122
13. Townsend, R. R., and Hardy, M. R. (1991) Glycobiology 1, 139–147
14. Ciucanu, I., and Kerek, F. (1984) Carbohydr. Res. 131, 209–217
15. Berman, E., and Allerhand, A. (1981) J. Biol. Chem. 256, 6657–6662
16. Bacz, A., Kahane, J., and Zuckermandel, B. M. (1990) Exp. Parasitol. 71, 483–488
17. Haslam, S. M., Houston, K. M., Harnett, W., Reason, A. J., Morris, H. R., and Dell, A. (1999) J. Biol. Chem. 274, 7359–7369
18. Haslam, S. M., Coles, G. C., Munn, E. A., Smith, T. S., Smith, H. F., Morris, H. R., and Dell, A. (1996) J. Biol. Chem. 271, 30561–30570
19. Gerdt, S., Denny, R. D., Bergonie, G., Schnabel, R., and Geyer, H. (1999) Eur. J. Biochem. 266, 952–963
20. Verostek, M. F., and Trimbble, R. B. (1995) Glycobiology 5, 671–681
21. Verostek, M. F., Atkinson, P. H., and Trimbble, R. B. (1993) J. Biol. Chem. 268, 12104–12115