Structural Basis for the Substrate Specificity of Endo-β-*N*-acetylglucosaminidase F_3[†,‡]

Christopher A. Waddling, Thomas H. Plummer, Jr., Anthony L. Tarentino, and Patrick Van Roey*

Division of Molecular Medicine, Wadsworth Center, New York State Department of Health, Albany, New York 12201

Received January 31, 2000; Revised Manuscript Received April 26, 2000

ABSTRACT: Endo-β-*N*-acetylglucosaminidase F_3 cleaves the $\beta(1-4)$ link between the core GlcNAc's of asparagine-linked oligosaccharides, with specificity for biantennary and triantennary complex glycans. The crystal structures of Endo F_3 and the complex with its reaction product, the biantennary octasaccharide, Gal-$\beta(1-4)$-GlcNAc-$\beta(1-2)$-Man-$\alpha(1-3)$[Gal-$\beta(1-4)$-GlcNAc-$\beta(1-2)$-Man-$\alpha(1-6)$]-Man-$\beta(1-4)$-GlcNAc, have been determined to 1.8 and 2.1 Å resolution, respectively. Comparison of the structure of Endo F_3 with that of Endo F_1, which is specific for high-mannose oligosaccharides, reveals highly distinct folds and amino acid compositions at the oligosaccharide recognition sites. Binding of the oligosaccharide to the protein does not affect the protein conformation. The conformation of the oligosaccharide is similar to that seen for other biantennary oligosaccharides, with the exception of two links: the Gal-$\beta(1-4)$-GlcNAc link of the $\alpha(1-3)$ branch and the GlcNAc-$\beta(1-2)$-Man link of the $\alpha(1-6)$ branch. Especially the latter link is highly distorted and energetically unfavorable. Only the reducing-end GlcNAc and two Man's of the trimannose core are in direct contact with the protein. This is in contrast with biochemical data for Endo F_1 that shows that activity depends on the presence and identity of sugar residues beyond the trimannose core. The substrate specificity of Endo F_3 is based on steric exclusion of incompatible oligosaccharides rather than on protein−carbohydrate interactions that are unique to complexes with biantennary or triantennary complex glycans.

Flavobacterium meningosepticum secretes three endo-β-*N*-acetylglucosaminidases, Endo[1] F_1, F_2, and F_3 (*1*, *2*). The three enzymes release N-linked oligosaccharides from glycoproteins and glycopeptides by cleaving the $\beta(1-4)$ glycosidic bond between the two GlcNAc's of the *N*,*N*′-diacetylchitobiose core but have specificities for distinct oligosaccharide structures. Endo F_1 cleaves only high-mannose and hybrid oligosaccharides. Endo F_2 preferentially processes biantennary complex oligosaccharides, and Endo F_3 is specific for bi- and triantennary complex oligosaccharides. Endo F_1 is closely related to Endo H, which is commonly used as a tool for the release of oligosaccharides from glycoproteins (*3*). Endo H and Endo F_1 have nearly identical substrate specificity and 32% sequence identity. In contrast, Endo F_2 and Endo F_3 have about 20% sequence identity with each other but only about 15% with Endo H or Endo F_1. It is unclear why *F. meningosepticum* and a few other bacteria produce and secrete such highly specialized glycohydrolases because the N-linked oligosaccharide substrates are not normally found in their environment. Regardless, further exploitation of the substrate specificity of these enzymes will lead to the development of improved tools for the analysis of N-linked oligosaccharides and for the specific deglycosylation of glycoproteins for structural studies (*4*, *5*).

The endoglycosidases of *F. meningosepticum* belong to family 18 of the glycohydrolases (*6*), all of which have $(\beta/\alpha)_8$-barrel folds and function with retention of configuration (*7*). The structure of Endo F_1 (*8*) and that of Endo H (*9*) show a β-hairpin in the loop of β-strand/loop/α-helix unit 2 at one end of a shallow substrate binding cleft. Chitinases (*10*, *11*) of family 18 also have a β-hairpin in loop 2, and the importance of this area for substrate recognition has been demonstrated by the structure of the complex of the chitinase hevamine with *N*,*N*′,*N*″-triacetylchitotriose (*12*), which shows that the trisaccharide binds in specific binding pockets in this area. The active site of the family 18 glycohydrolases contains two conserved acidic residues at the end of β-strand 4, Asp126 and Glu128 in Endo F_3. The glutamic acid has been identified as the proton donor, and the aspartic acid has been assigned a secondary role, stabilizing the intermediate in a substrate-assisted hydrolysis mechanism in which the carbonyl group of the C2-acetamido of the leaving GlcNAc acts as the nucleophile (*13*, *14*).

Here, we describe the crystal structure of Endo F_3, both in free form and in complex with its product, the complex biantennary octasaccharide shown in Figure 1. Differences between the structures of Endo F_3 and Endo F_1 are consistent with the differences in substrate specificity. The structure of the complex shows that direct contacts with the oligosaccharide are limited to the core structure and that specificity is defined through steric exclusion.

[†] Research supported by Grants GM34071 (T.H.P. and A.L.T.) and GM50431 (P.V.R.) from the National Institute of General Medical Sciences, National Institutes of Health.

[‡]The atomic coordinates and experimental structure factors have been deposited with the Protein Data Bank, RCSB (PDB entry codes 1EOK and 1EOM).

[1] Abbreviations: Endo, endo-β-*N*-acetylglucosaminidase; PEG, poly(ethylene glycol); rmsd, root-mean-square deviation.

10.1021/bi0001731 CCC: $19.00
Published on Web 06/14/2000

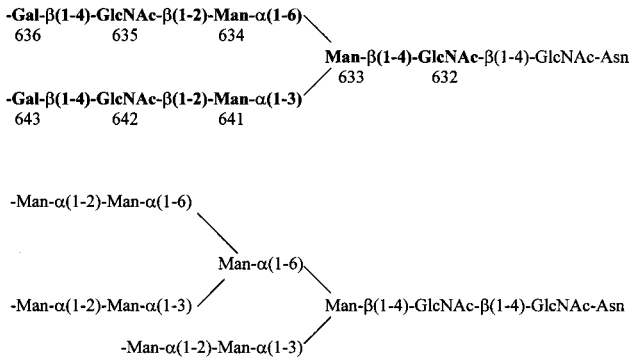

FIGURE 1: Schematic diagrams of a biantennary complex oligosaccharide (top) and a high-mannose oligosaccharide (bottom), the substrates of Endo F_3 and Endo F_1, respectively. The product biantennary octasaccharide used in the crystallographic studies is shown in bold. The numbers indicate the numbers used to identify the sugar groups in the text and figures and correspond to the numbering in the PDB file.

MATERIALS AND METHODS

Recombinant Endo F_3 was expressed in *Escherichia coli* and purified as previously described (*15*). Crystals of Endo F_3 were grown by hanging-drop vapor-diffusion methods. Initial crystallization conditions were identified from sparse-matrix screening methods (*16*), using the Crystal Screen kits I and II (Hampton Research), followed by typical optimization. The crystals did not grow very reproducibly, and macroseeding techniques were employed to obtain large quantities of crystals suitable for X-ray crystallography. The protein solution was at approximately 5 mg/mL in 10 mM Tris·HCl, pH 7.0. The crystallization buffer consisted of 2.2 M $(NH_4)_2SO_4$ and 4% 2-methyl-2,4-pentanediol. Needle-shaped crystals, with dimensions up to 1.0 × 0.2 × 0.1 mm grow to full size in about 10 days. The crystals belong to space group $P2_12_12_1$ with cell parameters $a = 69.35$ Å, $b = 118.33$ Å, and $c = 40.18$ Å with one molecule in the asymmetric unit and a solvent content of about 53%.

The product octasaccharide was prepared by Endo F_3 processing of a dansylated glycopeptide obtained from porcine fibrinogen, as previously described (*17*). Crystals of the complex were grown under conditions similar to those of free Endo F_3. The protein and octasaccharide were mixed at the time of crystallization by combining 2 μL of 3 mM octasaccharide solution with 2 μL of 8 mg/mL protein solution and 3 μL of the crystallization buffer.

The crystals did not tolerate addition of cryoprotectants directly to the drop or transfer to cryoprotectant-containing solutions. Therefore, the crystals were first transferred to 100 μL of crystallization buffer. Between 300 and 500 μL of a 50% PEG 400 solution, which is not completely miscible with the high-salt crystallization buffer, was then layered on top of this solution. After equilibration, the crystal was pulled through the low-salt/high-PEG phase with a cryoloop and flash cooled immediately to 100 K in the cold nitrogen gas stream. Data were measured on a Rigaku R-Axis IV detector with a Rigaku RU-200 rotating anode generator, MSC-Yale mirrors, and an Oxford Cryosystems low-temperature device and processed with DENZO and SCALEPACK (*18*). Attempts at determining the structure by molecular replacement techniques, using various models based on the structures of Endo H and Endo F_1, were unsuccessful. Therefore, the structure was determined by multiple isomorphous replacement methods, program PHASES (*19*), using MeHgAc, KI_3, and UO_2Ac_2 as derivatives. The final figure of merit for the phasing to 2.5 Å resolution is 0.573. Phases were improved by solvent flattening using the program DM, as implemented in the CCP4 package (*20*). The structures were modeled using O version 6.1.2 (*21*) and refined using X-PLOR version 3.851 (*22*). Water molecules were included on the basis of the waterpick routine of X-PLOR. Water molecules that had refined thermal parameters greater than 60.0 Å^2 were considered noise and removed. The final model consists of residues 8−289. The average thermal parameters of the sugar groups range from 18.6 Å^2 for the α(1−6)-linked mannose and 20.8 Å^2 for the reducing-end GlcNAc to 56.9 Å^2 for the Gal of the α(1−3) branch. Figure 2 shows the final $2(F_o - F_c)$ electron density map for the oligosaccharide. The density is clear for all sugar groups except for the terminal Gal's. No peptide conformation falls outside the allowed regions of the Ramachandran plot, as defined in PROCHECK (*23*), with only one residue in the generously allowed region and 89.2% in the most favored region. Data collection, phasing, and final refinement statistics are listed in Table 1.

RESULTS

Endo F_3 Structure and Comparison with Endo F_1. The structure of Endo F_3, shown in Figure 3, is a typical $(\beta/\alpha)_8$-barrel, a cyclic 8-fold repeat of β-strand/loop/α-helix units in which the β-strands form a central eight-stranded β-barrel. All β-strands are parallel, with their carboxy termini at the top surface of the molecule. The active sites of $(\beta/\alpha)_8$-barrel enzymes are located above the barrel on the top surface. The helices run to the bottom of the molecule and are located on the outside of the molecule. When viewed along the β-barrel axis, Endo F_3 is circular in shape with a diameter of about 42 Å. The β-barrel is elliptical, with a major axis of about 18 Å and a minor axis of about 13 Å. In Endo F_1 and Endo H, the overall shape of the molecule is more elliptical, with the short axis running between β/loop/α units 2 and 6. The β-barrel is also elliptical with its major axis at an angle of about 35° with the shortest dimension of the molecule.

In all $(\beta/\alpha)_8$-barrel structures, the loops connecting the β-strands to the α-helices have the greatest variation and define the substrate binding site and active site of the enzyme, while the β-barrel, α-helices, and the short turns at the carboxy-terminal end of the helices provide the structural scaffold. Consistently, the structures of Endo F_3 and Endo F_1, compared in Figure 4, are very similar in the β-barrel area but differ significantly in the loops and in some of the α-helices.

The most important differences between Endo F_3 and Endo F_1 are observed in loops 2, 3, and 4 and in the area of α-helices 5 and 6. The loop of unit 2 of Endo F_3 (Figure 4a,c) includes two small 1.5 turn α-helices. The first one, residues 42−47, lies above the barrel and is directed radially away from the center of the molecule. This loop geometry differs greatly from the β-hairpin geometry seen in Endo F_1 (Figure 4b,d). This short loop 2 helix of Endo F_3 is complemented by a more structured loop 3, which also contains a short α-helix, residues 90−93, in Endo F_3 but not in Endo F_1. These two small helices flank a deep gap above α-helix 3 (residues 104−114), exposing the central

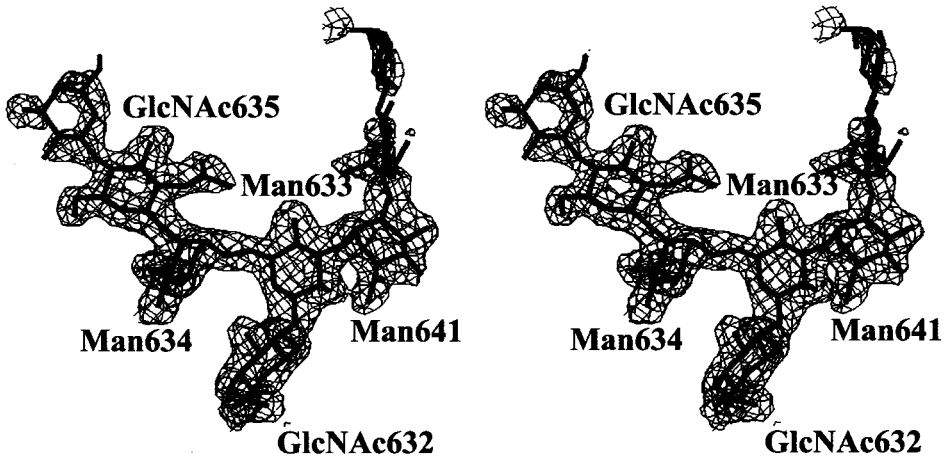

FIGURE 2: Stereo diagram of the electron density of the oligosaccharide in the final ($2F_o - F_c$) map, contoured at the 1.0σ level. This figure is drawn with the program SETOR (*29*).

Table 1: Data Collection, Phasing, and Refinement Statistics

	Endo F_3	complex	MeHgAc	KI_3	UO_2Ac_2
data collection					
resolution (Å)	20.−1.8	2.1	2.6	2.4	2.8
R_{mer}	0.077	0.095	0.076	0.066	0.070
$\langle I/\sigma(I)\rangle$, last shell	3.5	5.6	6.0	6.7	5.3
completeness (%)	98.2	98.2	89.5	98.7	98.3
redundancy	4.7	5.5	3.5	2.4	2.3
phasing					
concn (mM)			5	2	15
time			24 h	48 h	10 days
sites			1	1	3
phasing power			1.05	1.09	1.51
R_{Cullis}			0.652	0.645	0.608
R_{Kraut}			0.180	0.062	0.131
refinement					
R_{cryst} (R_{free})	0.174 (0.208)	0.166 (0.215)			
atoms ($\langle B\rangle$, Å^2)					
protein	2176 (27.2)	2185 (17.3)			
water	287 (27.8)	304 (30.4)			
sulfate	5 (23.3)	30 (48.6)			
ligand	0	98 (38.1)			
rmsd					
bond lengths (Å)	0.007	0.013			
angles (deg)	1.3	1.68			
dihedrals (deg)	23.9	23.3			

section of α-helix 3 to the binding cleft. In Endo F_1 helix 3, which is the most conserved in geometry and location of all helices in the molecule, is completely covered by the loop 2 β-hairpin, which runs tangential to the barrel. The difference in the transition from β-strand 2 to a β-hairpin in Endo F_1 and to an α-helix in Endo F_3 results from highly specific structural and conformational features. The last residue of β-strand 2 is a conserved aromatic residue, Phe45 in Endo F_1 and Phe39 in Endo F_3, that is important for the substrate recognition. This residue is followed by a *cis*-peptide link to a Ser in Endo F_1 and also in Endo H. In Endo F_3, residue 40 is a Gly, a flexible residue that allows the backbone to make a sharp turn away from the barrel.

The amino acid composition of this area also differs significantly, as shown in Figure 4c,d. In Endo F_3, residues from the small helices in loops 2 and 3 and from helix 3 form a cluster of aromatic and hydrophobic residues. In Endo F_1, and in Endo H, the area of the β-hairpin of loop 2 consists mostly of polar neutral residues, mainly asparagines. Carbohydrate binding sites typically consist of a majority of polar neutral residues that provide hydrogen-bonding contacts and some aromatic residues that can stack against the hydrophobic face of the sugar ring. In this context, the Endo F_1 recognition site is typical, but the cluster of aromatic residues in Endo F_3, combined with the limited number of polar neutral residues, is unusual.

Large differences are also seen in the conformations of the loops of units 4, 5, and 6 at the opposite end of the substrate binding cleft. Endo F_1 and Endo H differ from most other $(\beta/\alpha)_8$-barrel proteins in that the α-helices of units 5 and 6 are absent and the loops of these units are very short. We have previously proposed that the lack of bulk in this area of Endo F_1 and Endo H was required to allow an intact glycoprotein substrate to bind to the enzyme (*9*). In contrast, units 5 and 6 of Endo F_3 include α-helices, 2.5 and 3.5 turns long, respectively, and the loop of unit 6 is 8 residues longer than the corresponding loop in Endo F_1. Loop 4 of Endo F_3, residues 127−159, is 12 residues longer than the corresponding loop in Endo F_1 and lies on top of the molecule, extending over the area of units 5 and 6. The addition of the unit 5 and 6 α-helices and the longer loops 4 and 6 leads to a more bulky molecule in this area, resulting in the more circular shape as well as in a higher rim surrounding the barrel (Figure 4a,b), making the active site of Endo F_3 less accessible for large glycoprotein substrates. This is consistent with our interpretation of the Endo H and Endo F_1 structures, because Endo F_3 processes glycopeptides as well as Endo H and Endo F_1 but is not efficient at cleaving biantennary oligosaccharides from intact glycoproteins.

The N-terminal ends of loops 3 and 4 of Endo F_1 adopt conformations that align somewhat with the loop 2 β-hairpin, and the three loops form β-sheet-like hydrogen-bonding interactions but that are limited to no more than two residues for each loop. In the absence of this β-hairpin in Endo F_3, the first four or five residues of loops 3 and 4 are located more toward the area of loops 5 and 6. The most important result of this difference in conformation is that the relative arrangement of the active site residues, Asp126 and Glu128 in Endo F_3 and Asp130 and Glu132 in Endo F_1, is not identical. While the Asp residues, which are the last residues of β-strand 4, are in identical locations and conformations,

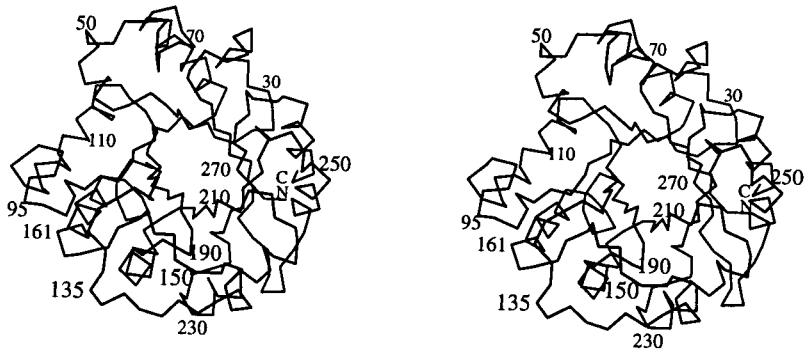

FIGURE 3: Stereo diagram of the α-carbon tracing of Endo F_3. This figure and Figures 4−6 were made with the program MOLSCRIPT, version 2.1.2 (*30*).

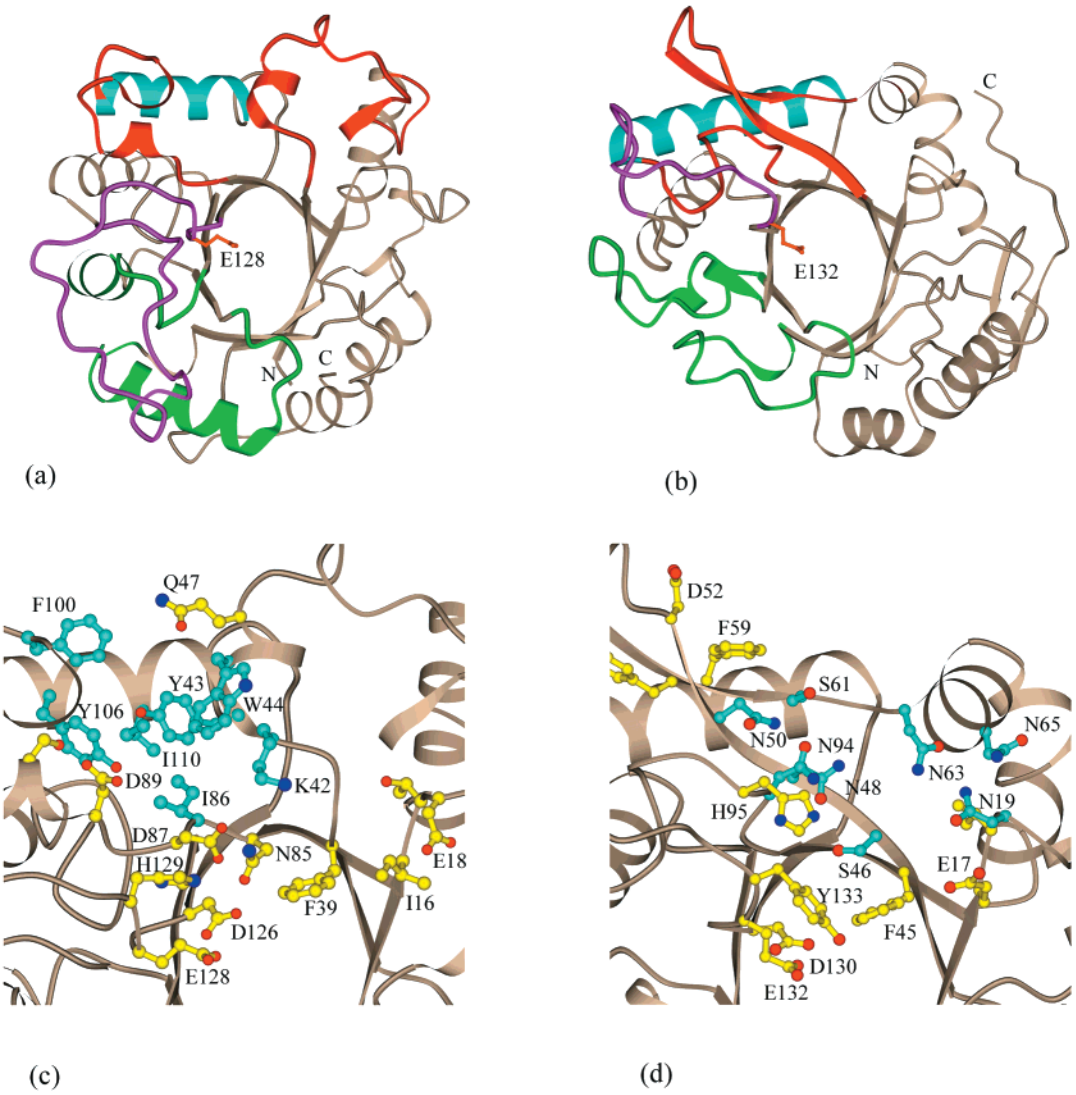

FIGURE 4: Ribbon diagrams of (a) Endo F_3 and (b) Endo F_1 with the catalytic acid residue shown. Highlighted areas are as follows: loops 2 and 3, red; helix 3, light blue; loop 4, magenta; and, loops and α-helices 5 and 6, green. Loop 2/loop 3 area of (c) Endo F_3 and (d) Endo F_1, showing the distinct fold and amino acid composition of the end of the substrate binding cleft. The clusters of hydrophobic and aromatic residues in Endo F_3 and polar neutral residues in Endo F_1 are highlighted as light blue residue colors.

the relatively small changes in the main chain conformation following the Asp residues result in a 3.65 Å shift in the location of the α-carbon of the Glu residues (Figure 5). However, the Glu side chains are sufficiently large and flexible to allow the carboxylate groups to almost overlap. As a result, there appears a small but not insignificant difference in the shape of the substrate binding cleft in the area of the active site. However, the relative locations of the carboxylate groups of the two acidic residues involved in the mechanism of action are conserved, and Asp126 and Glu128 remain close enough to be in hydrogen-bonding contact. Some variation in the conformations of the N-terminal ends of loops 3 and 4 has been observed in different crystal forms of Endo H mutants (*24*). This suggests that these loops may be flexible enough in Endo F_1 and/or Endo F_3 to allow some adjustment upon substrate binding.

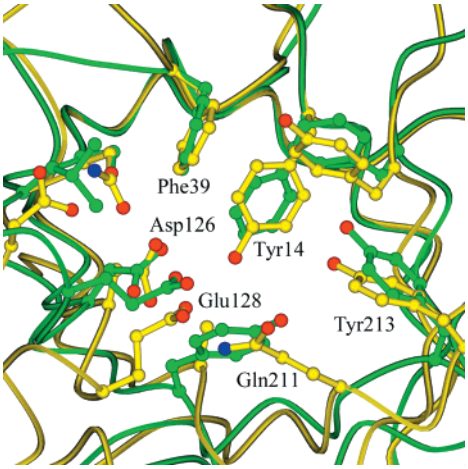

FIGURE 5: Overlap of the upper levels of the β-barrels of Endo F_3 (yellow) and Endo F_1 (green). Residue numbers indicated correspond to the residues of Endo F_3. Note the position of the Endo F_3 Gln211 side chain overlapping with the hydroxyl group of Tyr171 of Endo F_1 and the relatively distant locations of the α-carbon atoms of the catalytic residues Glu128 (Endo F_3) and Glu132 (Endo F_1) but the close proximity of their carboxylate groups.

The β-barrel is somewhat narrower in Endo F_3 than in Endo F_1 (*8*) (Figure 4a,b) or Endo H (*9*). The root-mean-square deviation (rmsd) for the fit of the main chain atoms of the β-strand residues of Endo F_3 and Endo F_1 is 0.77 Å, with the largest deviations occurring in strands 6 and 7, in part related to a kink in strand 6 of Endo F_1 (and Endo H) that is absent in Endo F_3. The binding cleft in the area of the active site is defined by residues from the upper layer of the barrel and by residues contributed by several of the surrounding loops. As shown in Figure 5, these residues are highly conserved with only one exception: Tyr171 of Endo F_1 is replaced by Ala186 in Endo F_3. However, a difference in the conformation of loop 6 allows Gln211 of Endo F_3 to position its amide oxygen atom in the same location of the hydroxyl of Endo F_1 Tyr171, preserving the hydrogen-bonding contact site for substrate binding. This conserved conformation and maintenance of polar groups is necessary for substrate binding and surprisingly sensitive to mutations, as evidenced by the structures of mutants of the acidic residues of Endo H (*24*). The primary catalytic residue of Endo H, Glu132, does not interact with any of the barrel residues except for Asp130. However, the Glu132Asp mutation distorts the upper level of the barrel similarly, as mutations of Asp130 and the conformational changes are detrimental to substrate binding.

Endo F_3−Octasaccharide Complex. The oligosaccharide is located in the binding cleft as expected (Figure 6a), with its reducing-end GlcNAc (632) above the central area of the β-barrel, adjacent to the two active site acidic residues, and the trimannose core close to the β-strands and loops of units 1, 2, and 3. The oligosaccharide binds tightly to the protein, as evidenced by the low thermal parameters of the reducing-end GlcNAc and the trimannose core, which are similar to those of the surrounding protein residues, and consistent with the fact that only low concentrations of oligosaccharide are required in cocrystallization experiments. The conformation of the protein does not change upon binding of the oligosaccharide, with an overall rmsd of 0.24 Å for the main chain atoms of the complexed and uncomplexed Endo F_3. Side chain conformations are also not affected, and the only apparent result of the binding of the oligosaccharide in the cleft is the displacement of eight water molecules, seven of which are replaced by hydroxyls from the oligosaccharide.

The carbohydrate makes direct or indirect, water-bridged contacts with residues from loops 1, 2, 3, 4, 7, and 8 of the protein (Figure 6c and Table 2). However, the direct contacts are limited to the reducing-end GlcNAc (632) and two Man's of the trimannose core, namely, the central Man (633) and the α(1−6)-linked Man (634). These three sugar groups are set in well-defined binding pockets and are bound through an extensive network of hydrogen bonds and hydrophobic interactions. All hydroxyl groups, except O6 of GlcNAc 632 and O4 of Man 633, are in hydrogen-bonding contact with the protein. The side chain of Phe39 sits below the three sugar groups, not only stacking against the hydrophobic face of the α(1−6)-linked Man (634) but also making a close contact with O3 of the GlcNAc (632). GlcNAc 632 forms hydrogen bonds with residues from loops 4 (Asp126 and Glu128), 7 (Gln211 and Tyr213), and 8 (Tyr272). The O1-hydroxyl is observed in the β-configuration, which is the minor configuration (approximately 35%) in the equilibrium between the spontaneously interchangeable α- and β-configurations of a free O1-hydroxyl of GlcNAc in solution but the same as the configuration of the glycosidic link of the substrate. O1 interacts with the carboxylate group of the catalytic residue Glu128 and also with the amide oxygen of Gln211. Asp126 forms a hydrogen bond with the nitrogen of the C2-acetamido group and is in the correct position to stabilize a positive charge on the nitrogen. This conformation confirms the relative roles of the two acidic residues in the active site: Glu128 as the proton donor and Asp126 as stabilizing the intermediate in the substrate-assisted mechanism (*13*). The central Man (633) interacts through hydrogen bonds and van der Waals contacts to residues from loops 1, 7, and 8. O2 is in hydrogen-bonding contact with the carboxylate and hydroxyl oxygen atoms of Glu245 and Tyr272, and O3 interacts with Arg20. The α(1−6)-linked Man (634) is sandwiched between Phe39, Asp87, and His129. In contrast, the α(1−3)-linked Man (641) only makes very weak contacts, consisting of stacking with the side chains of Arg20 and Glu245 at distances of about 3.8 Å. Beyond the trimannose core, both branches extend away from the protein, and neither the GlcNAc (635 and 642) nor the Gal (636 and 643) of either branch interacts directly with the protein. Overall, binding to the protein reduces the solvent-accessible surface of the oligosaccharide to 37% of that of the free oligosaccharide. The accessibilities of the individual sugar groups to a probe of 1.4 Å radius (Table 2) range from almost 0 for the α(1−6)-Man (634) to over 88% for the Gal (636) on the α(1−6) branch.

Oligosaccharide Conformation. The conformation of the oligosaccharide (Figure 6b) is similar to those reported for other protein−carbohydrate crystal structures (*25*), except for the GlcNAc-β(1−2)-Man link of the α(1−6)-branch (635−634) and the Gal-β(1−4)-GlcNAc link of the α(1−3)-branch (643−642). All other dihedral angles of the oligosaccharide links, listed in Table 3, are within about 10° of the averages of the first or second most commonly observed conformations for that link. The conformation of the uncommon Gal-GlcNAc link differs by about 20° in φ and about 50° in ψ from the average for other glycan

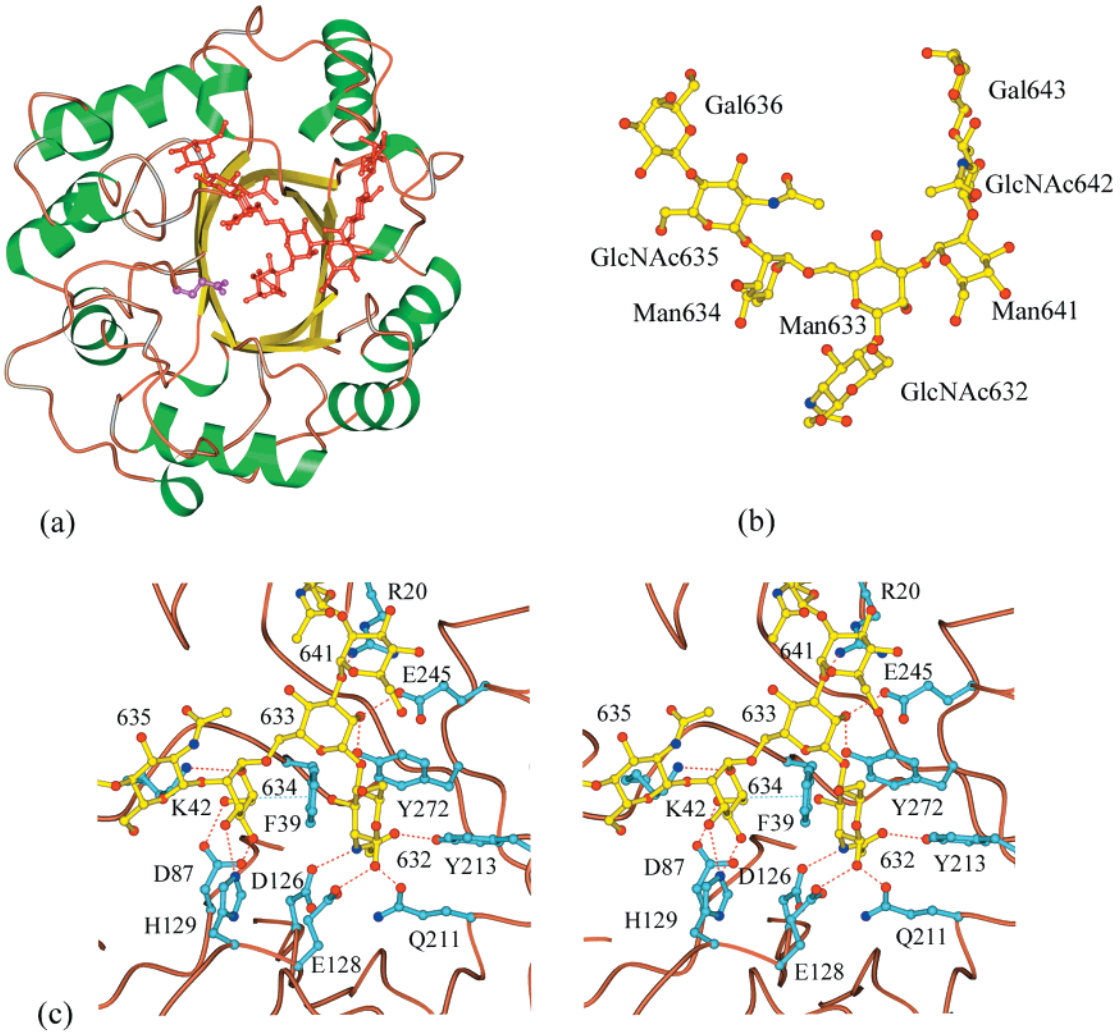

FIGURE 6: (a) Structure of the complex of Endo F_3 with the biantennary oligosaccharide. The carbohydrate is shown in red and the Glu128 side chain in magenta. (b) Three-dimensional structure of the biantennary oligosaccharide. (c) Stereo diagram showing a detailed view of the hydrogen-bonding (red lines) and hydrophobic (blue lines) interactions of the oligosaccharide with Endo F_3.

structures ($\varphi = -70.4° \pm 9.1°$, $\psi = 129.5° \pm 7.1°$). However, the conformation is within the low-energy well for this link in the corresponding disaccharide (*26*) and is less than 1 kcal/mol above the global minimum. In addition, the electron density for the Gal is relatively poor as the result of disorder. Therefore, the geometry of this link may be poorly determined.

The observed conformation of the GlcNAc-β(1−2)-Man link of the α(1−6) branch ($\varphi = -148.5°$, $\psi = -144.2°$) is more interesting. The electron density for both sugars groups, Man 634 and GlcNAc 635, is very well defined, and the thermal parameters of both groups are low. However, the conformation is at least 10 kcal/mol above the global minimum for the link in a disaccharide ($\varphi = -90°$, $\psi = -78°$) (*26*) and quite far removed from any one of the five low-energy mimima for the disaccharide or any one of the five distinct conformations observed in crystal structures (*25*). Among the previously observed conformations, only two have been seen more than once: $\varphi = -80.2°$, $\psi = -97.2°$, observed 36 times, and $\varphi = 58.3°$, $\psi = -87.2°$, observed 8 times, corresponding to the global minimum and the next lowest energy minima of the disaccharide, respectively. Of the three conformations that have been reported once, only the conformation of the link in a biantennary complex oligosaccharide bound to a human Fc fragment ($\varphi = -145.8°$, $\psi = -106.5°$) (*27*) somewhat resembles the conformation seen in this structure. However, the conformation of the link when bound to the Fc fragment is not well established because, as reported by the author, the electron density in that area of the map is very poor and the model was built on the basis of stereochemical criteria only.

The most common conformation of the GlcNAc-β(1−2)-Man link is not accessible for the Man 634 and GlcNAc 635 sugars in the complex with Endo F_3 due to a steric conflict with the protein, which would place the *N*-acetyl group too close to the side chains of Lys42 and Met58. Of the remaining previously observed conformations, two place either the *N*-acetyl group or the O6-hydroxyl too close to a sulfate ion that is located in a positively charged area formed by His129 and two lysines from a symmetry-related molecule. Likewise, most of the calculated low-energy conformations lead to some weak steric conflicts for either GlcNAc 635 or Gal 636. Most importantly, none of the previously observed conformations nor any of the calculated low-energy conformations results in energetically favorable interactions. However, in the conformation observed in this structure, Gal 636 is in hydrogen-bonding contact with a symmetry-related molecule, both directly and through a sulfate group. One must

Table 2: Average Thermal Parameters, Percentage of Solvent Accessibility of the Sugar Group in the Complex vs the Free Oligosaccharide, and Intermolecular Contacts of the Sugar Residues in the Endo F_3−Octasaccharide Complex

residue	⟨*B*⟩ (Å^2)	access (%)	atom	contact 1[a]	distance (Å)	contact 2[a]	distance (Å)
GlcNAc	20.79	8.2	O1	Glu128-Oϵ1	2.58		
632				Gln211-Oϵ1	2.45		
				Glu128-Oϵ2	3.02		
				Wat573-OH2	3.04		
			N2	Asp126-Oδ2	3.02		
				Glu128-Oϵ2	2.96		
			O3	Phe39-Cϵ2	3.14		
				Phe39-Cζ	3.16		
				Tyr272-Oη	3.18		
			O4	Tyr272-Oη	3.04		
			O7	Tyr213-Oη	2.59		
Man	24.26	13.8	O2	Glu245-Oϵ2	2.52		
633			O2	Tyr272-Oη	2.75		
			O3	Arg20-Nζ2	3.16		
			O4	Wat491-OH2	2.66	Asp18-Oδ2	2.43
Man	18.60	0.3	O2	Lys42-Nζ	3.12		
634			O3	His129-Nϵ2	2.97		
			O3	Wat482-OH2	2.94		
			O3	Wat506-OH2	2.85		
			O4	His129-Nϵ2	2.80		
			O4	Asp87-Oδ2	2.49		
			C6	Phe39-Cζ	3.49		
			O6	Lys42-Nζ	2.81		
			O6	Asp87-Oδ1	2.66		
			O6	Wat411-OH2	2.67	Asn85-Nδ2	3.05
						Leu41-O	2.68
						Gly40-N	2.76
GlcNAc	34.10	16.9	N2	Wat551-OH2	2.98		
635			O5	Wat482-OH2	2.79		
			O6	Trp44-Cη2	3.31		
				Wat444-OH2	3.21	Asp89-Oδ2	2.88
				Wat482-OH2	2.72		
Gal	43.40	88.7	O2	(SO_4)705-O3*	2.79	Lys25-Nζ*	2.54
636						Lys63-Nζ*	3.01
			O4	Lys27-Nζ*	2.99		
Man	39.69	47.0	O3	Wat546-OH2	2.38		
641			O6	Glu245-Oϵ1	3.38		
			O6	Tyr243-Oη	3.53		
GlcNAc	51.46	56.8	N2	Thr175-O*	2.89		
642			O5	Wat546-OH2	3.14		
Gal	56.96	54.1	O2	Asp18-O	2.99		
643			O4	Wat542-OH2	2.33		
			O4	Asn179-N*	3.16		
			O6	Ala177-O*	2.61		
			O6	Wat544-OH2	2.61	Lys181-N*	2.71
						Gly174-O*	2.81
						Ala177-O*	2.75

[a] An asterisk indicates a symmetry-related molecule.

Table 3: Dihedral Angles of the Oligosaccharide Links: φ = O5−C1−O(*x*)′−C(*x*)′, ψ = C1−O(*x*)′−C(*x*)′−C(*x* − 1)′, and ω = O6′−C6′−C5′−C4

link	φ (deg)	ψ (deg)	ω (deg)
Man-β(1−4)-GlcNAc 633−632	−74.9	116.3	
Man-α(1−6)-Man 634−633	78.7	155.3	201.4
Man-α(1−3)-Man 641−633	56.0	−122.1	
GlcNAc-β(1−2)-Man 635−634	−148.5	−144.1	
Gal-β(1−4)-GlcNAc 636−635	−79.5	124.4	
GlcNAc-β(1−2)-Man 642−641	−58.9	−94.0	
Gal-β(1−4)-GlcNAc 643−642	−89.7	79.7	

conclude that this energetically favorable interaction, together with the tight binding of the Man, is sufficient to compensate for the high energy of the GlcNAc-β(1−2)-Man link conformation.

DISCUSSION

The structure of the complex of Endo F_3 with the biantennary oligosaccharide shows that only residues from the Man-α(1−3)[Man-α(1−6)]-Man-β(1−4)-GlcNAc core, shared by all N-linked oligosaccharides, make direct contact with the protein. Therefore, the substrate specificity of Endo F_3 is not determined by protein−carbohydrate interactions that are unique to the binding of bi- or triantennary complex oligosaccharides but rather by steric exclusion of alternate oligosaccharide structures. Specifically, the structure shows that biantennary complex oligosaccharides, as well as triantennary complex oligosaccharides with the third branch attached β(1−4) to the α(1−3)-linked Man, can be bound by the enzyme. However, a branch linked β(1−4) to the α-(1−6)-linked Man is not possible because O4 of that Man is directed toward the protein. This is consistent with the fact that Endo F_3 does not process tetrantennary oligosaccharides. Likewise, typical high-mannose oligosaccharides are not suitable substrates for Endo F_3 because both O3 and O6 of the α(1−6)-linked Man are in hydrogen-bonding contact with the protein and not available for substitution.

The absence of direct contact beyond the trimannose core is surprising in view of other biochemical and structural data. Endo F_1 and Endo H are known to require at least one Man beyond the trimannose core, the Man linked α(1−3) to the α(1−6)-Man of the core, for substrate binding and processing (*28*). While no biochemical evidence for a similar requirement exists for Endo F_3, it was anticipated that an enzyme specific for complex oligosaccharides would interact with the GlcNAc's beyond the trimannose core. This absence of specificity-defining contacts, together with the highly strained conformation of the oligosaccharide, leads us to question if bi- and triantennary complex oligosaccharides are the intended substrates for Endo F_3. Two other features of the structure, the absence of interaction with the cluster of aromatic residues formed by the small helices in the loops of units 2 and 3 and the availability of O4 on the central Man, suggest that Endo F_3 is capable of processing other oligosaccharides. These might include oligosaccharides that carry a substituent on O4 of the central Man of the core, such as a bisecting GlcNAc, as well as alternative substituents on the α(1−6)-linked Man. In addition, although Endo F_3 processes bi- and triantennary oligosaccharides from glycopeptides very efficiently, absence of these types of oligosaccharides in the normal environment of its source further supports the possibility that other oligosaccharides are the intended target substrates of the enzyme. Identification of such alternative substrates could lead to an understanding of why *F. meningosepticum* produces and secretes this enzyme.

REFERENCES

1. Tarentino, A. L., Quinones, G., Schrader, W. P., Changchien, L. M., and Plummer, T. H., Jr. (1992) *J. Biol. Chem. 267*, 3868−3872.
2. Tarentino, A. L., Quinones, G., Changchien, L. M., and Plummer, T. H., Jr. (1993) *J. Biol. Chem. 268*, 9702−9708.
3. O'Neill, R. A. (1996) *J. Chromatogr. A 720*, 201−215.
4. Tarentino, A. L., and Plummer, T. H., Jr. (1994) *Methods Enzymol. 230*, 44−57.
5. Grueninger-Leitch, F., D'Arcy, A., D'Arcy, B., and Chene, C. (1996) *Protein Sci. 5*, 2617−2622.
6. Henrissat, B. (1991) *Biochem. J. 280*, 309−316.
7. Davies, G., and Henrissat, B. (1995) *Structure 3*, 853−859.
8. Van Roey, P., Rao, V., Plummer, T. H., Jr., and Tarentino, A. L. (1994) *Biochemistry 33*, 13989−13996.
9. Rao, V., Guan, C., and Van Roey, P. (1995) *Structure 3*, 449−457.
10. Terwisscha van Scheltinga, A. C., Hennig, M., and Dijkstra, B. W. (1996) *J. Mol. Biol. 262*, 243−257.
11. Aronson, N. N., Blanchard, C. J., and Madura, J. D. (1997) *J. Chem. Inf. Comput. Sci. 37*, 999−1005.
12. Terwisscha van Scheltinga, A. C., Kalk, K. H., Beintema, J. J., and Dijkstra, B. W. (1994) *Structure 2*, 1181−1189.
13. Tews, I., Terwisscha van Scheltinga, A. C., Perrakis, A., Wilson, K. S., and Dijkstra, B. W. (1997) *J. Am. Chem. Soc. 119*, 7954−7959.
14. Brameld, K. A., Shrader, W. D., Imperiali, B., and Goddard, W. A. I. (1998) *J. Mol. Biol. 280*, 919−923.
15. Tarentino, A. L., Quinones, G., and Plummer, T. H., Jr. (1995) *Glycobiology 5*, 599−601.
16. Jancarik, J., and Kim, S. H. (1991) *J. Appl. Crystallogr. 24*, 409−411.
17. Plummer, T. H., Jr., Phelan, A. W., and Tarentino, A. L. (1996) *Anal. Biochem. 235*, 98−101.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
19. Furey, W., and Swaminathan, S. (1997) *Methods Enzymol. 277*, 590−620.
20. CCP4 (1994) *Acta Crystallogr. D50*, 760−763.
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
22. Brunger, A. T. (1988) *J. Mol. Biol. 203*, 803−816.
23. Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 282−291.
24. Rao, V., Cui, T., Guan, C., and Van Roey, P. (1999) *Protein Sci. 8*, 2338−2346.
25. Petrescu, A. J., Petrescu, S. M., Dwek, R. A., and Wormald, M. R. (1999) *Glycobiology 9*, 343−352.
26. Imberty, A., Delage, M.-M., Bourne, Y., Cambillau, C., and Perez, S. (1991) *Glycoconjugate J. 8*, 456−483.
27. Deisenhofer, J. (1981) *Biochemistry 20*, 2361−2370.
28. Trimble, R. B., and Tarentino, A. L. (1991) *J. Biol. Chem. 266*, 1646−1651.
29. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI0001731