C-type Lectin-like Carbohydrate Recognition of the Hemolytic Lectin CEL-III Containing Ricin-type β-Trefoil Folds*

Tomomitsu Hatakeyama1,2, Hideaki Unno3, Yoshiaki Kouzuma4, Tatsuya Uchida5, Seiichiro Eto4, Haruki Hidemura6, Norihisa Kato7, Masami Yonekura8, and Masami Kusunoki9

From the 1Department of Applied Chemistry, Faculty of Engineering, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki 852-8521, the 2Department of Resource Biological Science, College of Agriculture, Ibaraki University, 3-21-1, Chuo, Ami-machi, Inashiki-gun, Ibaraki 300-0393, and the 3Research Center for Structural and Functional Proteomics, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

CELE-III is a Ca2+-dependent hemolytic lectin, isolated from the marine invertebrate Cucumaria echinata. The three-dimensional structure of CEL-III/GalNAc and CEL-III/methyl α-galactoside complexes was solved by x-ray crystallographic analysis. In these complexes, five carbohydrate molecules were found to be bound to two carbohydrate-binding domains (domains 1 and 2) located in the N-terminal 2/3 portion of the polypeptide and that contained β-trefoil folds similar to ricin B-chain. The 3-OH and 4-OH of bound carbohydrate molecules were coordinated with Ca2+ located at the subdomains 1α, 1γ, 2α, and 2γ, simultaneously forming hydrogen bond networks with nearby amino acid side chains, which is similar to carbohydrate binding in C-type lectins. The binding of carbohydrates was further stabilized by aromatic amino acid residues, such as tyrosine and tryptophan, through a stacking interaction with the hydrophobic face of carbohydrates. The importance of amino acid residues in the carbohydrate-binding sites was confirmed by the mutational analyses. The orientation of bound GalNAc and methyl α-galactoside was similar to the galactose moiety of lactose bound to the carbohydrate-binding site of the ricin B-chain, although the ricin B-chain does not require Ca2+ ions for carbohydrate binding. The binding of the carbohydrates induced local structural changes in carbohydrate-binding sites in subdomains 2α and 2β. Binding of GalNAc also induced a slight change in the main chain structure of domain 3, which could be related to the conformational change upon binding of specific carbohydrates to induce oligomerization of the protein.

CELE-III is a hemolytic lectin isolated from the sea cucumber Cucumaria echinata (1, 2). This lectin binds to carbohydrates containing Gal/GalNAc at nonreducing ends in the presence of Ca2+. CEL-III exhibits the highest affinity for GalNAc, followed by lactose, lactulose, and methyl β-galactoside (Me-β-Gal)2 among the carbohydrates tested (3). After binding to cell surface carbohydrate chains, CEL-III oligomerizes to form membrane pores, thereby leading to colloid osmotic rupture of the cell membrane (4). In addition to hemolytic activity, this lectin exhibits a strong cytoxicity for some cultured cell lines, which is also caused by formation of oligomers in the cell membrane (5, 6). Such a cell membrane-damaging action was also known for several bacterial pore-forming toxins (7), such as α-hemolysin from Staphylococcus aureus (8, 9), aelolysin from Aeromonas hydrophila (10, 11), and the anthrax toxin from Bacillus anthracis (12). These toxins exert their pore-forming action through conformational changes that lead to oligomerization in the target cell membranes after binding to specific cell surface receptors. Recently, some hemolytic lectins have also been reported (13–16). The importance of oligomerization in the target cell membranes is also suggested for these lectins.

From the cDNA nucleotide sequence, it was inferred that CEL-III is composed of three domains as follows: two N-terminal carbohydrate-binding domains (domains 1 and 2) and a C-terminal domain (domain 3) (17). The carbohydrate-binding domains have a relatively low but definite similarity with the carbohydrate-binding domains of ricin-type (R-type) lectins, such as the B-chains of the toxic plant lectins ricin (18) and abrin (19). On the other hand, although ricin-like plant lectins contain toxic subunits, namely the A-chains that inactivate the 60 S ribosomal subunits in eukaryotic cells, CEL-III has domain 3 that contains hydrophobic segments and may be involved in the oligomerization in the target cell membrane. In fact, we have observed that domain 3 fragments, once produced by limited digestion with trypsin, spontaneously associate to form oligomers in solution. The resulting domain 3 oligomer exhibited a marked increase in β-sheet content, as measured by circular dichroism spectra (20). This is consistent with an increase in β-sheet structure of the entire CEL-III oligomer (4). Therefore, this work was conducted under the Cooperative Research Program of Institute for Protein Research, Osaka University, and supported by Grants-in-Aid for Scientific Research, Research 14560073 (to T. H.) and 18780248 (to Y. K.) from the Japan Society for the Promotion of Science. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2Z48 and 2Z49) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 To whom correspondence should be addressed: Dept. of Applied Chemistry, Faculty of Engineering, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki 852-8521, Japan. Tel.: 81-95-819-2686; Fax: 81-95-819-2684; E-mail: theta@nagasaki-u.ac.jp.

2 The abbreviations used are: Me-β-Gal, methyl β-D-galactopyranoside; Me-α-Gal, methyl α-D-galactopyranoside; TBS, Tris-buffered saline; BIS-Tris, bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane.
fore, it seems reasonable to conclude that oligomerization of CEL-III is mainly mediated by interactions between its domain 3 after binding to cell surface carbohydrates.

We have solved the crystal structure of CEL-III (21), which confirmed the domain structure as suggested from the amino acid sequence; domains 1 and 2 adopt a β-trefoil structure, each consisting of three subdomains or motifs. They showed apparent similarity with the B-chains of ricin and abrin despite their relatively low sequence identity. The basic folds of the carbohydrate-binding domains of CEL-III and ricin B-chain are similar, but there is a significant difference that CEL-III contains Ca\(^{2+}\) ions at their carbohydrate-binding sites, whereas the ricin B-chain does not. Although Ca\(^{2+}\) is known to be essential for interaction with carbohydrates in some lectins, such as C-type lectins (22–26), β-trefoil lectins generally bind carbohydrates without Ca\(^{2+}\) ions. To elucidate the carbohydrate-binding mechanism and the role of bound Ca\(^{2+}\) ions, we assessed the x-ray crystal structure of CEL-III/GalNAc and CEL-III/methyl α-galactoside (Me-α-Gal) complexes. Results revealed that CEL-III recognizes specific carbohydrates in a very similar manner as C-type lectins despite the structural similarity with R-type lectins. It was also observed that the binding of the carbohydrates induced local structural changes in carbohydrate-binding sites and domain 3 that could be related to the conformational change upon binding to cell surface carbohydrates, which leads to oligomerization of the protein in cell membranes (4).

**EXPERIMENTAL PROCEDURES**

**Materials**—Oligonucleotides used in this study were purchased from Exagen (Tokyo, Japan). Ex Taq™ DNA polymerase and the DNA ligation kit were obtained from Takara (Otsu, Japan) and used as recommended by the supplier. Restriction endonucleases and DNA-modifying enzymes were purchased from MBI Fermentas (Burlington, Canada). The plasmid vectors used in this work were as follows: pGEM-T Easy vector from Promega and pET-32a expression vector from Novagen. All other chemicals were of analytical grade for biochemical use.

**Purification and Crystallization of CEL-III/Carbohydrate Complexes**—CEL-III was purified from C. echinata according to a method reported previously (2). The proteins extracted from a homogenate of C. echinata were applied to a lactose-Cellulofine column equilibrated with 0.15 M NaCl, 10 mM Tris-HCl, pH 7.5 (TBS), containing 10 mM CaCl\(_2\). Adsorbed lectins (CEL-I, CEL-III, and CEL-IV) were eluted with TBS containing 20 mM EDTA. The lectins were then separated using a GalNAc-Cellulofine column, utilizing the differences in carbohydrate-binding specificities. After elution of CEL-III with TBS containing 0.1 M lactose, CEL-I and CEL-IV were eluted with TBS containing 20 mM EDTA. CEL-III was finally purified by gel filtration through Sephadex G-75 in TBS. Crystallization of CEL-III/carbohydrate complexes was done under similar conditions as those for native CEL-III crystals (21) in the presence of 10 mM carbohydrates. Briefly, the protein solution (7 mg/ml, 2–4 μl) in TBS containing 10 mM CaCl\(_2\) was mixed with the same amount of reservoir solution (12% (w/v) polyethylene glycol 8000, 100 mM BisTris, NaOH, pH 6.5, and 200 mM magnesium acetate) and subjected to sitting drop vapor diffusion at 20 °C. Diffraction images from the crystals were collected using synchrotron radiation on beamline BL44XU at SPring-8 (Hyogo, Japan) at 100 K using an imaging plate detector, DIP6040 (MAC Science, Japan). Diffraction images were indexed and integrated using the program Mosflm (27) and processed using the CCP4 programs (28) Scala and Truncate. Data collection statistics are summarized in Table 1.

**Structure Determination and Refinement**—The crystal structures of CEL-III/carbohydrate complexes were solved by molecular replacement method using native CEL-III (Protein Data Bank code 1VCL) (21) as a search model. Molecular replacement was performed using the program Phaser (29). The model was refined using the program Refmac (30) from the CCP4 suite (28). Manual fitting of the model was carried out by the program Xfit (31) and Coot (32). The quality of the final models for native and complexed CEL-III was assessed by Ramachandran plots and analysis of the model geometry with the program Procheck (33). The refinement statistics are summarized in Table 1. Figures for protein models were drawn by PyMOL (34).

**Expression of the Carbohydrate Recognition Domain 1 of CEL-III in E. coli Cells**—To obtain a cDNA fragment encoding domain 1 (residues 1–156) of CEL-III, PCR was performed using CEL-III cDNA (17) (DDB) Data Base accession number AB109017) as a template with forward primer 5’–CATGGA–3’ and reverse primer 5’–CGCCGAGTCGTAGAAGT–3’.
Hemolytic Lectin CEL-III/Carbohydrate Complex

GACAAGTTTTGTGCACGAAATCCA-3′ and reverse primer 5′-GGATCCCTAACCCTAGAACGCTCTGGCC-3′. The PCR products were ligated into the pGEM-T Easy vector. After confirmation of the DNA sequence, the DNA fragment was excised by digestion with NcoI and BamHI, and ligated into the expression vector pET-32a, previously digested with the same enzymes. The resulting plasmid pET-CRD1 was introduced into Escherichia coli BL21(DE3) CodonPlus RIL strain (Stratagene), and recombinant protein was induced with 1 mM isopropylthiogalactoside according to the supplier’s instruction. After induction, the culture was incubated for an additional 5 h at 37 °C. Cells were harvested and lysed by BugBuster protein extraction reagent (Novagen) and sonication. The lysate was centrifuged at 10,000 × g for 15 min, and the supernatant was loaded onto HisBind resin columns (Novagen). After washing of the column, recombinant protein was eluted by 20 mM Tris-HCl, pH 7.9, containing 1 M imidazole and 500 mM NaCl according to the supplier’s manual. Protein solution was dialyzed against 10 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl, and 10 mM CaCl2.

After washing the column, recombinant protein was eluted by 20 mM Tris-HCl, pH 7.9, containing 1 M imidazole and 500 mM NaCl according to the supplier’s manual. Protein solution was dialyzed against 10 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl, and 10 mM CaCl2. The recombinant domain 1 thus obtained was investigated by procedures identical to those described for the wild-type domain 1.

**RESULTS**

Crystalization and Structure Determination of CEL-III/Carbohydrate Complexes—Crystals were prepared under similar conditions as for native CEL-III (21). We have tried to crystallize CEL-III complexed with several carbohydrates, such as GalNAc, lactose, lactulose, Me-β-Gal, Me-α-Gal, and galactose (3). Among these, CEL-III/GalNAc and CEL-III/Me-α-Gal complexes gave good diffraction patterns, although Me-α-Gal showed lower affinity than Me-β-Gal (3), and they were analyzed. Statistics of data collection and refinement parameters are summarized in Table 1. The space group of the CEL-III/Me-α-Gal crystal was P21 with a unit cell axis of a = 53.3 Å, b = 65.5 Å, c = 127.0 Å, β = 97.1°, and two CEL-III molecules were contained in the asymmetric unit, which is similar to native CEL-III crystals. On the other hand, CEL-III/GalNAc crystals belonged to P1 with a unit cell axis of a = 52.6 Å, b = 65.2 Å, c = 66.7 Å, α = 85.3°, β = 73.6°, γ = 89.9°, which contained two CEL-III molecules in the asymmetric unit. The structures of these CEL-III/carbohydrate complexes were solved by a molecular replacement method using native CEL-III as a search model. Five Ca2+ and two Mg2+ ions were assigned to the same positions as those for native protein, considering their electron density, coordination numbers, bond distances, and ligand atoms (21). As shown in Fig. 1, the electron density map exhibited a distinct electron density corresponding to bound GalNAc and Me-α-Gal molecules at each Ca2+ ion in the five subdomains of the carbohydrate-binding domains (domains 1 and 2). The five carbohydrate molecule models were fitted to this electron density map, and refinement of the model was further performed. Finally, refinement of the model of CEL-III/Me-α-Gal and CEL-III/GalNAc complexes was performed to achieve R = 19.1% and Rfree = 23.9%, and R = 18.5%, and Rfree = 21.8%, respectively, as shown in Table 1. The resulting

**TABLE 2**

Nomenclature of domain 1 mutants and oligonucleotide primers used for mutagenesis

| Mutant | Oligonucleotides† |
|--------|-------------------|
| D23A   | 5′-CAAGCAATGCGTTGCCATAGTTGGTAATCA-3′ |
| Y36A   | 5′-AACATACGGCGCGCTGTGTTAATCA-3′ |
| D43A   | 5′-GGCGTATCTGCCCAACAGATTATCATTTGC-3′ |
| Q44A   | 5′-AACATAGCGACGGCGGATTGTGATGGGTTA-3′ |
| Q45A   | 5′-GTTCCTGCAAGGTTATGACATTGTGGGTTA-3′ |

† The primers with the reverse-complementary sequence of each primer were also used for mutagenesis.

**FIGURE 1.** Structure of carbohydrate-binding site (1α) with bound GalNAc and Me-α-Gal. 2Fobs − Fcalc electron density maps are contoured at 1.2σ in GalNAc (A) and Me-α-Gal (B) complexes. The Ca2+ ions are shown as magenta spheres.
Overall and carbohydrate-binding domain structure of the CEL-III/GalNAc complex is illustrated in Fig. 2.

**Overall Structure and Carbohydrate-binding Sites of the CEL-III/Carbohydrate Complex**—In CEL-III/carbohydrate complexes, GalNAc and Me-α-Gal molecules were bound to the five subdomains, \(1_\alpha, 1_\gamma, 2_\alpha, 2_\beta, \) and \(2_\gamma\), in domains 1 and 2, whereas subdomain \(1_\beta\) did not bind carbohydrate molecules (Fig. 2 for the CEL-III/GalNAc complex). This is consistent with the fact that these subdomains, with the exception of \(1_\beta\), have a very similar structure with \(Ca^{2+}\) ions. In fact, only subdomain \(1_\beta\) lacks amino acid residues directly involved in coordination with \(Ca^{2+}\) in other subdomains (the residues at \(a, b,\) and \(f\) positions in Fig. 3). At each carbohydrate-binding site, 3-OH and 4-OH of the bound carbohydrates are coordinated with a \(Ca^{2+}\) ion and simultaneously form hydrogen bonds with side chains of nearby amino acid residues. For example, in subdomain \(1_\gamma, 4\)-OH forms two hydrogen bonds with the carboxyl side chain of Asp-23 (position \(a\)) and the amide NH of Gly-26 (position \(c\)), and 3-OH of the carbohydrates forms a hydrogen bond with Asp-39 (position \(e\)) (Fig. 4, A and B). Asp-39 also forms a hydrogen bond with a water molecule that is fixed by a coordinate bond with \(Ca^{2+}\) and hydrogen bonds with Asp-43 (position \(f\)). Such a carbohydrate recognition mode was essentially common in all the carbohydrate-binding sites, except for the residues at position \(e\), which includes Asp, Gln, and Glu residues (Fig. 3).

In the case of the CEL-III/GalNAc complex, the binding of both anomers of GalNAc to CEL-III was observed from their electron density map (Table 3), depending on the local environment of each binding site. For example, GalNAc bound in subdomain \(1_\alpha\) was mostly α-anomer, being stabilized by a hydrogen bond with the carboxyl side chain of Asn-410 of the neighboring CEL-III molecule in the crystal. Although GalNAc shows relatively high affinity for CEL-III among simple carbohydrates, there is no specific bonding between acetamido group of GalNAc and protein atoms except for van der Waals contact. The average hydrogen bond distances for the bound carbohydrates and their anomeric state are listed in Table 3.
Besides the GalNAc molecules bound in domains 1 and 2, another GalNAc molecule was found in the pocket formed between domains 1 and 3 (Fig. 4C). Both CEL-III molecules in the asymmetric unit of the crystal contained GalNAc at this position. However, binding of this GalNAc was mediated by several hydrogen bonds with surrounding water molecules, which were fixed on the surface of the protein. Therefore, it seems reasonable to assume that this is nonspecific binding of GalNAc. In fact, no carbohydrate molecule was observed in this pocket of the CEL-III/Me-α-Gal complex.

Fig. 5A shows the superimposition of the carbohydrate-binding site residues (subdomain 2γ) of CEL-III and the ricin B-chain. As seen in this figure, orientation of GalNAc bound to CEL-III was very similar to that of galactose moiety of lactose bound to the ricin B-chain when both proteins were superimposed based on the corresponding residues (Fig. 3). Stacking interaction of the hydrophobic face of the carbohydrates with aromatic side chains (Trp-269 in CEL-III and Tyr-248 in the ricin B-chain; position d) was also common in both proteins. However, recognition of 3-OH and 4-OH of GalNAc was essentially done by coordinate bonds with Ca²⁺ ions as well as hydrogen bond networks at the binding site in CEL-III, whereas recognition of 3-0H and 4-OH was achieved only by hydrogen bonds in ricin B-chain. Despite involvement of Ca²⁺ ions,
amino acid residues involved in recognition of carbohydrates are mostly at the same positions in CEL-III and ricin B-chain. Among these, the residues at positions a, b, c, and f form coordinate bonds with Ca\(^{2+}\) in CEL-III. The residues at positions a and c also form hydrogen bonds with 4-OH of the carbohydrates, whereas the residue at position e forms a hydrogen bond with 3-OH of GalNAc as well as a water molecule, which is fixed by Ca\(^{2+}\) and the aspartate residue at position f (Fig. 6). There are some variations in the conformation of the residues at position e (Asp, Glu, and Gln) that are not directly involved in binding of Ca\(^{2+}\). Such a Ca\(^{2+}\)-mediated recognition of carbohydrates by CEL-III resembles that of C-type lectins, as was representatively shown by CEL-I from C. echinata (Fig. 5B) (25). In the CEL-I/GalNAc complex, 3-OH and 4-OH of GalNAc were coordinated with Ca\(^{2+}\) ions at the binding sites in addition to the hydrogen bond networks with nearby amino acid residues (Gln-101, Asp-103, Glu-109, Asn-123, and Asp-124). CEL-I also stabilized binding of GalNAc by stacking with the side chain of Trp-105. These results indicate that CEL-III recognizes specific carbohydrates in a similar manner as C-type lectins, although it has typical \(\beta\)-trefoil folds common in R-type lectins.

**Mutational Analyses of the Carbohydrate-binding Sites**—To examine the roles of the amino acid residues and Ca\(^{2+}\) ions in the carbohydrate-binding sites of CEL-III, several mutants of domain 1 (residues 1–156), in which amino acid residues in subdomain 1 had been replaced by Ala residues, were expressed in E. coli cells, and their carbohydrate binding activity was examined by affinity chromatography on GalNAc-Cellulofine columns. As shown in Fig. 7, the wild-type recombinant domain 1 exhibited affinity for the GalNAc-Cellulofine column, clearly indicating that domain 1 alone has carbohydrate-binding ability. On the

![Diagram of carbohydrate-binding sites](image1)

**FIGURE 5.** Comparison of carbohydrate-binding sites of CEL-III, ricin B-chain, and the C-type lectin CEL-I. A, stereo view of the superposition of the carbohydrate-binding sites (subdomain 2) of CEL-III with GalNAc (red) and the ricin B-chain with lactose (blue). B, carbohydrate-binding site of CEL-I complexed with GalNAc. Coordinate and hydrogen bonds are indicated by dotted lines.

![Diagram of carbohydrate-binding sites](image2)

**FIGURE 6.** Stereo view of the superposition of the five carbohydrate-binding sites of CEL-III/GalNAc complex. Binding site residues and bound GalNAc molecules were superposed by fitting the C\(\alpha\) atoms of the residues at positions a, b, c, and f, which were coordinated with Ca\(^{2+}\). Coordination and hydrogen bonds for the binding site in subdomain 1\(\alpha\) are indicated by dotted lines.
considering the above results, subdomain 1 as evidenced by the crystal structure (Fig. 2). However, con-
domains 1 these recombinant proteins of domain 1 contained both sub-
drivate-binding sites of CEL-III (Fig. 3). It should be noted that position is not conserved, not even among the five carbohy-
trate-binding ability. In fact, this residue of the Q

**FIGURE 7.** Affinity chromatography of CEL-III domain 1 and its mutants on GalNAc-Cellulofine column. Wild-type (WT) and five mutants (D23A, Y36A, D43A, Q44A, and Q45A) of domain 1 were subjected to affinity chromatography on a GalNAc-Cellulofine column (0.8 × 10 cm) equilibrated with TBS containing 10 mM CaCl₂. Vertical arrows indicate the point of elution with 100 mM lactose in the buffer.

other hand, mutants D23A and D43A, in which Asp-23 and Asp-43 coordinated with Ca²⁺ (Fig. 4A) had been replaced by Ala residues, showed almost no affinity for the column, confirming the importance of these residues. In addition to coordinate bonds with Ca²⁺, Asp-23 and Asp-43 also formed hydrogen bonds with 4-OH of GalNAc and a water molecule, which in turn formed a hydrogen bond with Asp-39. The Y36A mutant showed a weak affinity for the GalNAc-Cellulofine column, suggesting that the stacking interaction between Tyr-36 and GalNAc plays a supporting role in stabilizing the carbohydrate binding. The Q44A mutant showed no binding activity for the GalNAc-Cellulofine column. Although Gln-44 has no direct interaction with Ca²⁺ or carbohydrates, this residue is located at the first position of the QXW motif (Fig. 3), which is known to be a conserved motif in the β-trefoil domains (36) and contributes to stabilize the 3₁₀ helix by hydrogen bonding within the carbohydrate-binding motifs (21). Therefore, it seems likely that mutation of this residue caused destabilization of the basic structure of β-trefoil fold. In contrast to Q44A, Q45A retained a fairly high carbohydrate binding activity (Fig. 7), suggesting that the middle residue of the QXW motif contributes little to structural stabilization and carbohydrate-binding ability. In fact, this position is not conserved, not even among the five carbohydrate-binding sites of CEL-III (Fig. 3). It should be noted that these recombinant proteins of domain 1 contained both subdomains 1α and 1γ, with the latter also having binding ability as evidenced by the crystal structure (Fig. 2). However, considering the above results, subdomain 1γ alone does not seem to provide enough affinity to bind to the affinity column.

**CONFORMATIONAL CHANGES INDUCED BY BINDING OF CARBOHYDRATE**—Comparison between the structures of CEL-III and its carbohydrate complexes revealed partial conformational changes in subdomains 2α and 2β that were induced by binding of carbohydrates. As shown in Fig. 8A (subdomain 2α), binding of GalNAc induced a large movement of the side chain of Glu-184 to form a hydrogen bond with 3-OH of GalNAc. Simultaneously, Tyr-181 and Asn-186 were moved aside to accommodate Glu-184, accompanied by a movement of the main chain. On the other hand, there was a large movement of Tyr-222 in the binding site of subdomain 2β induced by binding of GalNAc (Fig. 8B), leading to a stacking interaction with the hydrophobic face of GalNAc, which would contribute to stabilization of the GalNAc binding. These changes were observed in the two CEL-III/GalNAc molecules in an asymmetric unit of the crystals, especially in domain 3. As shown in Fig. 10, in addition to the changes around the carbohydrate-binding site in subdomain 2α (residues 165–190), relatively large deviations in the C-α atoms in CEL-III/carbohydrate complexes were observed at residues 300–325 and 350–380, corresponding to the flanking regions of the two α-helices (H8 and H9). As seen in Fig. 9B, the beginning of the α-helix H8 was located near the carbohydrate-binding site in subdomain 2α, and therefore it seems possible that movement of the residues in the carbohydrate-binding site induced by the binding of GalNAc may lead to slight conformational changes in domain 3.

**DISCUSSION**

It was assumed from sequence similarity with lectins that have β-trefoil domains, such as ricin B-chain (18, 37), that the N-terminal ⅔ of CEL-III contained two carbohydrate-binding domains (domains 1 and 2). The β-trefoil domains contain a characteristic sequence motif (QXW)₃ and are thought to function as carbohydrate-binding modules not only in lectins but also in various enzymes, toxins, and cell surface receptors (36). The carbohydrate-binding ability of domains 1 and 2 was confirmed in the present x-ray crystal
structural analysis. However, there was a remarkable difference between CEL-III and ricin B-chain because CEL-III bound carbohydrate molecules at five of six subdomains in domains 1 and 2, whereas ricin B-chain only has carbohydrate-binding sites in its subdomains 1α and 2γ. As revealed previously by the crystal structure of native CEL-III (21), the main chain structures of the subdomains in CEL-III are considerably similar with the exception of 1β, reflecting their high amino acid sequence similarity (17). In particular, the amino acid residues known to be important for carbohydrate binding in other β-trefoil lectins, such as aspartate residues (Asp-23, Asp-121, Asp-168, Asp-209, and Asp-256) and aromatic residues (Tyr-36, Tyr-134, Tyr-181, Tyr-222, and Trp-269), are highly conserved in these five subdomains. These findings strongly suggest that these five subdomains have carbohydrate-binding ability, which was proven in this study. The presence of multiple binding sites in a single protein molecule should be advantageous for substantial binding affinity by cooperative interactions with carbohydrates. For example, mannos-binding lectin has relatively limited interactions with mannos-containing carbohydrate chains at each carbohydrate-binding site, but it produces sufficient affinity for mannos oligosaccharides by forming oligomers with multiple carbohydrate-binding sites (38). It has been suggested recently that the lipid rafts, which contain various glycolipids, function as receptors for several bacterial toxins (39, 40). Because lactosyl ceramide is the most effective receptor for CEL-III in the human erythrocyte membrane (41), it seems likely that the presence of multiple carbohydrate-binding sites in a protein molecule is advantageous for recognizing glycolipid clusters in the lipid rafts on target cell membranes.

We have previously presented two putative carbohydrate-binding models for CEL-III; the first in which Ca²⁺ ions were directly involved in carbohydrate binding, and the second where Ca²⁺ ions were replaced by bound carbohydrate. This study revealed that the first model was correct, and Ca²⁺ ions were essential for the recognition of carbohydrates through coordinate bonds with OH groups of carbohydrates. Despite differences in the involvement of Ca²⁺, the orientation of the carbohydrate bound to CEL-III was very similar to that for ricin B-chain as shown in Fig. 5A. The OH groups of bound carbohydrate also formed hydrogen bond networks with nearby amino acid side chains in the carbohydrate-binding sites of CEL-III. This also resembles carbohydrate
recognition by the C-type carbohydrate recognition domains, which are known to recognize carbohydrates through coordinate bonds with Ca\(^{2+}\) as well as hydrogen bond networks (Fig. 5B). To our knowledge, CEL-III is the first example of a β-trefoil lectin that uses Ca\(^{2+}\) ions to recognize the OH groups of specific carbohydrates by forming coordinate bonds. Like C-type lectins, CEL-III mainly recognizes 3- and 4-OH groups of the carbohydrates. This is consistent with the observation that fucose (6-deoxy-D-galactose) bound to CEL-III with affinity comparable with galactose (3); the hydroxyl group at the 6-position of galactose seems to have little contribution for interaction with CEL-III. For recognition of 3- and 4-OH, stacking of aromatic residues appears to play an important role to make correct orientation of the carbohydrates for the hydrogen bond network.

CEL-III oligomerizes in aqueous solution in high salt concentrations and at high pH conditions upon binding of specific carbohydrates, yielding oligomers similar to those formed in CEL-III-treated erythrocyte membranes (4). This suggests that binding of carbohydrates triggers conformational changes in CEL-III, leading to oligomerization. Oligomerization in aqueous solution was effectively induced by disaccharides with a β-galactosidic linkage, such as lactose and lactulose, whereas an α-galactoside melibiose was much less effective (4). This suggests that a β-galactoside-containing disaccharide structure is important for induction of a conformational change resulting in oligomerization of CEL-III. As seen in the crystal structures of CEL-III complexes with GalNAc and Me-α-Gal, partial conformational changes were observed in the carbohydrate-binding sites in subdomains 2α and 2β, which were also accompanied by main chain shifts. However, deviations of main chain C-α atoms between CEL-III and its carbohydrate complexes were more pronounced in domain 3, especially around the α-helices H8 and H9 flanking regions (Figs. 9B and 10). This suggests that binding of specific carbohydrates could be related to partial conformational changes leading to oligomerization of CEL-III. We have unsuccessfully tried to prepare crystals of CEL-III complexed with lactose or lactulose, which induce more efficient oligomerization in solution. This was probably because the lactose and lactulose complexes had larger structural changes compared with GalNAc and Me-α-Gal complexes, which made crystallography of the complexes difficult. Further examination of the effects of β-galactoside binding on the structure of CEL-III would provide further insights into the hemolytic mechanism of CEL-III.

REFERENCES

1. Hatakeyama, T., Kohzaki, H., Nagatomo, H., and Yamasaki, N. (1994) J. Biochem. (Tokyo) 116, 209–214
2. Hatakeyama, T., Nagatomo, H., and Yamasaki, N. (1995) J. Biol. Chem. 270, 3560–3564
3. Hatakeyama, T., Miyamoto, Y., Nagatomo, H., Sallay, I., and Yamasaki, N. (1997) J. Biochem. (Tokyo) 121, 63–67
4. Hatakeyama, T., Furukawa, M., Nagatomo, H., Yamasaki, N., and Mori, T. (1996) J. Biol. Chem. 271, 16915–16920
5. Oda, T., Tsuru, M., Hatakeyama, T., Nagatomo, H., Muramatsu, T., and Yamasaki, N. (1997) J. Biochem. (Tokyo) 121, 560–567
6. Oda, T., Shinmura, N., Nishioka, Y., Komatsu, N., Hatakeyama, T., and Muramatsu, T. (1999) J. Biochem. (Tokyo) 125, 713–720
7. Parker, M. W., and Feil, S. C. (2005) Prog. Biophys. Mol. Biol. 88, 91–142
8. Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J. E. (1996) Science 274, 1859–1866
9. Olson, R., Nariya, H., Yokota, K., Kamio, Y., and Gouaux, E. (1999) Nat. Struct. Biol. 6, 134–140
10. Parker, M. W., Buckley, J. T., Postma, J. P., Tucker, A. D., Leonard, K., Pattus, F., and Tsernoglou, D. (1994) Nature 367, 292–295
11. Rossjohn, J., Feil, S. C., McKinstry, W. J., Tsernoglou, D., van der Goot, G. F., Buckley, J. T., and Parker, M. W. (1998) J. Struct. Biol. 121, 92–100
12. Leplà, S. (2006) in The Comprehensive Sourcebook of Bacterial Protein Toxins (Alouf, J., and Popoff, M., eds) 3rd Ed., pp. 323–347, Academic Press, New York
13. Konska, G., Guillot, J., Dusser, M., Damez, M., and Botton, B. (1994) J. Biol. Chem. (Tokyo) 116, 519–523
14. Armstrong, P. B., Swarnkar, S., Srimal, S., Misquith, S., Hahn, E. A., Aimes, R. T., and Quigley, J. P. (1996) J. Biol. Chem. 271, 14717–14721
15. Satoh, F., Nakagawa, H., Yamada, H., Nagasaka, K., Nagasaka, T., Araki, Y., Tomihara, Y., Nozaki, M., Sakuraba, H., Ohshima, T., Hatakeyama, T., and Aoyagi, H. (2002) *J. Nat. Toxin* **11**, 297–304

16. Tateno, H., and Goldstein, I. J. (2003) *J. Biol. Chem.* **278**, 40455–40463

17. Nakano, M., Tabata, S., Sugihara, K., Kouzuma, Y., Kimura, M., and Yamazaki, N. (1999) *Biochim. Biophys. Acta* **1435**, 167–176

18. Funatsu, G., Kimura, M., and Funatsu, M. (1979) *Agric. Biol. Chem.* **42**, 2221–2224

19. Kimura, M., Sumizawa, T., and Funatsu, G. (1993) *Biosci. Biotechnol. Biochem.* **57**, 166–169

20. Kouzuma, Y., Suzuki, Y., Nakano, M., Matsuyama, K., Tojo, S., Kimura, M., Yamazaki, T., Aoyagi, H., and Hatakeyama, T. (2003) *J. Biochem. (Tokyo)* **134**, 395–402

21. Uchida, T., Yamasaki, T., Eto, S., Sugawara, H., Kurisu, G., Nakagawa, A., Kusunoki, M., and Hatakeyama, T. (2004) *J. Biochem.* **134**, 395–402

22. Poget, S. F., Freund, S. M., Howard, M. J., and Bycroft, M. (2001) *Biochemistry* **40**, 10966–10972

23. Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992) *Nature* **360**, 127–134

24. Shrive, A. K., Tharia, H. A., Strong, P., Kishore, U., Burns, I., Rizkallah, P. J., Reid, K. B., and Greenhough, T. J. (2003) *J. Mol. Biol.* **331**, 509–523

25. Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. A., Alvarez, R., Blixt, O., Taylor, M. E., Weis, W. I., and Drickamer, K. (2004) *Nat. Struct. Mol. Biol.* **11**, 591–598

27. Powell, H. R. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 1690–1695

28. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763

29. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 458–464

30. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255

31. McRee, D. E. (1999) *J. Struct. Biol.* **125**, 156–165

32. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132

33. Laskowski, R., MacArthur, M., Moss, D., and Thornton, J. (1993) *J. Appl. Crystallogr.* **26**, 283–291

34. Delano, W. L. (2002) The PyMOL Molecular Graphics Systems, DeLano Scientific, San Carlos, CA

35. Deng, W. P., and Nickoloff, J. A. (1992) *Anal. Biochem.* **200**, 81–88

36. Hazes, B. (1996) *Protein Sci.* **5**, 1490–1501

37. Araki, T., and Funatsu, G. (1987) *Biocim. Biophys. Acta* **911**, 191–200

38. Ng, K. K., Kolatkar, A. R., Park-Snyder, S., Feinberg, H., Clark, D. A., Drickamer, K., and Weis, W. I. (2002) *J. Biol. Chem.* **277**, 16088–16095

39. Fivaz, M., Abrami, L., and van der Goot, F. G. (1999) *Trends Cell Biol.* **9**, 212–213

40. Abrami, L., Liu, S., Cosson, P., Leplla, S. H., and van der Goot, F. G. (2003) *J. Cell Biol.* **160**, 321–328

41. Hatakeyama, T., Sato, T., Taira, E., Kuwahara, H., Niidome, T., and Aoyagi, H. (1999) *J. Biochem. (Tokyo)* **125**, 277–284