J. Biochem. **135**, 101–107 (2004)
DOI: 10.1093/jb/mvh012

Characterization of Recombinant CEL-I, a GalNAc-Specific C-Type Lectin, Expressed in *Escherichia coli* Using an Artificial Synthetic Gene

Tomomitsu Hatakeyama*[,1], Kouhei Shiba[1], Noriaki Matsuo[1], Tokiko Fujimoto[1], Tatsuya Oda[2], Hajime Sugawara[3] and Haruhiko Aoyagi[1]

[1]Department of Applied Chemistry, Faculty of Engineering, and [2]Division of Biochemistry, Faculty of Fisheries, Nagasaki University, 1-14 Bunkyo-machi, Nagasaki 852-8521; and [3]Metabolic Function Research Group, RIKEN Plant Science Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045

Received September 6, 2003; accepted November 20, 2003

CEL-I is a C-type lectin isolated from the Holothuroidea *Cucumaria echinata*. This lectin shows very high *N*-acetylgalactosamine-binding specificity. We constructed an artificial gene encoding recombinant CEL-I (rCEL-I) using a combination of synthetic oligonucleotides, and expressed it in *Escherichia coli* cells. Since the recombinant protein was obtained as inclusion bodies, the latter were solubilized using urea and 2-mercaptoethanol, and the protein was refolded during the purification and dialysis steps. The purified rCEL-I showed comparable hemagglutinating activity to that of native CEL-I at relatively high Ca^{2+}-concentrations, whereas it was weaker at lower Ca^{2+}-concentrations due to decreased Ca^{2+}-binding affinity. rCEL-I exhibited similar carbohydrate-binding specificity to native CEL-I, including strong GalNAc-binding specificity, as examined by hemagglutination inhibition assay. Comparison of the far UV-CD spectra of recombinant and native CEL-I revealed that the two proteins undergo a similar conformational change upon binding of Ca^{2+}. Single crystals of rCEL-I were also obtained under the same conditions as those used for the native protein, suggesting that they have similar tertiary structures. Although native CEL-I exhibited strong cytotoxicity toward cultured cells, rCEL-I showed low cytotoxicity. These results indicate that rCEL-I has a tertiary structure and carbohydrate-binding specificity similar to those of native CEL-I. Howeger, there is a subtle difference in the properties between the two proteins probably due to the additional methionine residue at the N-terminus of rCEL-I.

Key words: artificial gene, crystal, C-type lectin, *Cucumaria echinata*, cytotoxicity.

Abbreviations: CD, circular dichroism; EDTA, ethylenediamine tetraacetate; FBS, fetal bovine serum; GalNAc, *N*-acetylgalactosamine; MBP, mannose (mannan)-binding protein; α-MEM, α-minimal essential medium; RHL, rat hepatic lectin; TBS, Tris-buffered saline.

Lectins play important roles in various tissues and body fluids as carbohydrate recognition molecules (*1*, *2*). Among animal lectins, the C-type lectin family (*3*) includes various proteins that exhibit carbohydrate-binding activity depending on Ca^{2+} ions. The C-type lectins in vertebrates are categorized into seven groups according to their functional domain structure (*4*). In contrast, invertebrate C-type lectins are mostly single-domain proteins, which form a multimeric structure. Some of them are thought to be involved in host defense systems through binding to carbohydrate chains on the surface of foreign pathogens. We have isolated four Ca^{2+}-dependent galactose/*N*-acetylgalactosamine (GalNAc)-specific lectins (CEL-I, II, III, and IV) from the marine invertebrate *Cucumaria echinata* (Holothuroidea) (*5*). Among them, CEL-I and CEL-IV belong to the C-type lectin family, while CEL-III is a novel Ca^{2+}-dependent lectin that exhibits strong hemolytic activity and cytotoxicity (*6–8*). The smallest lectin, CEL-I, is composed of two identical subunits of 16 kDa linked by a single disulfide bond. This lectin exhibits very high specificity for *N*-acetylgalactosamine; the binding affinity of CEL-I for GalNAc is estimated to be approximately 1,000-fold higher than that for galactose (*5*, *9*). So far, there is only limited knowledge concerning the carbohydrate-recognition mechanism of diverse C-type lectins (*10–15*). Therefore, investigation for the GalNAc-specificity of CEL-I by means of mutational analyses should provide valuable information concerning the structural basis of the carbohydrate-recognition mechanism of C-type lectins. The carbohydrate-recognition mechanism of C-type lectins has been extensively investigated in the case of mannose (mannan)-binding protein (MBP) (*16–18*), by means of various mutants and X-ray crystallographic analyses (*19*). Among invertebrate C-type lectins, a tunicate (*Polyandrocarpa misakiensis*) lectin (TC14) (*20*) is the only one whose carbohydrate-binding mode has been elucidated by X-ray crystallography (*21*). This lectin binds to galactose essentially in the same manner as in the case of MBP, whereas

*To whom correspondence should be addressed. Fax: +81-95-819-2684, E-mail: thata@net.nagasaki-u.ac.jp

the orientation of the bound galactose is considerably different from that in the case of MBP.

Here we report the expression and characterization of recombinant CEL-I (rCEL-I) in *Escherichia coli* cells using an artificial synthetic gene. Although the protein was initially obtained as inclusion bodies, it was successfully refolded after solubilization with a denaturant. The resulting lectin exhibited Ca^{2+}-dependent carbohydrate-binding ability with specificity similar to that of native CEL-I. However, some differences in its properties were also observed compared to those of the native protein.

MATERIALS AND METHODS

Construction and Expression of Recombinant CEL-I in E. coli Cells—A nucleotide sequence for rCEL-I was designed based on the amino acid sequence previously determined by means of protein chemical techniques (*9*), taking into account the codon usage for *E. coli*. The 5′-terminal and 3′-terminal halves of rCEL-I DNA were prepared by PCR using eight synthetic oligonucleotides (Espec Oligo Service). The two double-stranded fragments were ligated at the *Bgl*I site to form the entire coding region, and amplified by PCR using oligonucleotides FS1 and RS2. The amplified fragment was cloned into *E. coli* JM109 using the pGEM-T vector (Promega). The nucleotide sequence of the rCEL-I gene was confirmed by sequencing with a Hitachi DNA sequencer SQ5500E. The inserted rCEL-I gene was digested with *Nde*I and *Bam*HI, and then ligated with the pET-3a vector (Novagen) previously digested with the same enzymes. Since the recombinant protein was expressed in *E. coli* BL21(DE3)pLysS (Novagen) as inclusion bodies, they were collected by centrifugation from the cells after sonication in TBS, and immediately solubilized with 20 mM Tris-HCl buffer, pH 8.5, containing 8 M urea and 1% 2-mercaptoethanol. The solubilized protein was directly applied to a DEAE-Cellulofine column (3 × 3 cm) in 20 mM Tris-HCl, pH 8.5, containing 4 M urea, and eluted with a NaCl gradient, 0 to 0.5 M. The rCEL-I fractions were pooled and dialyzed against TBS for three days. The active protein exhibiting carbohydrate-binding activity was purified by affinity chromatography on a GalNAc-Cellulofine column (1.5 × 3 cm) (*6*). The rCEL-I dimer was finally purified by gel filtration on a Sephadex G-50 column (2.5 × 46 cm) in TBS.

Purification of Native CEL-I—Native CEL-I was purified from the protein extract of *C. echinata* by column chromatography on lactosyl-Cellulofine, GalNAc-Cellulofine, and Sephadex G-75, essentially as reported earlier (*6*).

N-Terminal Amino Acid Sequence Analysis—The N-terminal amino acid sequence of the expressed protein was determined with a protein sequencer PPSQ-21 (Shimadzu).

Determination of Protein Concentrations—Protein concentrations were determined from the molar absorption coefficients at 280 nm calculated from the amino acid compositions of the proteins.

Circular Dichroism (CD) Spectroscopy—Far-UV CD spectra of the proteins were obtained with a JASCO J-720 spectropolarimeter. The spectra were measured in a quartz cell of 1-mm path length at 20°C. The protein concentration was 0.12 mg/ml.

Hemagglutination Assay—Serial two-fold dilutions of a sample (50 μl) were mixed with the same volume of a 5% (v/v) suspension of rabbit erythrocytes in round-bottomed microtiter plate wells. Incubation was performed in TBS containing 10 mM $CaCl_2$. The extent of agglutination was determined visually after incubation for 1 h at room temperature. The hemagglutinating activity was expressed as a titer, *i.e.*, the reciprocal of the highest dilution giving detectable agglutination. The hemagglutination inhibition assay was performed by incubating 50 μl aliquots of the protein solutions (titer 2) containing various concentrations of carbohydrates with the same volume of a 5% (v/v) suspension of rabbit erythrocytes in TBS containing 10 mM $CaCl_2$.

Crystallization and X-Ray Diffraction of rCEL-I—Crystallization of rCEL-I was performed by the hanging drop vapor diffusion method under similar conditions to as previously reported (*22*). The protein solution (2 μl, 5 mg/ml) in TBS containing 10 mM $CaCl_2$ was mixed with an equal volume of the reservoir solution [62% (v/v) 2-methyl-2,4-pentanediol, 0.1 M Tris-HCl, pH 8.0], followed by equilibration in hanging drops against 0.2 ml of the reservoir solution at 20°C. Single crystals suitable for data collection were obtained in two months. Data collection for X-ray diffraction of rCEL-I was performed on the beamline BL45XU of the SPring-8 (Hyogo), using a Rigaku RAXIS-V imaging-plate area detector. Image data were processed with the program suite *HKL2000* (*23*).

Cell Culture—HeLa (human epithelial carcinoma) cells were cultured in α-minimal essential medium (α-MEM) supplemented with 10% fetal bovine serum (FBS), 10 μg each of adenosine, guanosine, cytidine, and thymidine per ml of medium, penicillin (100 μg/ml), and streptomycin (100 μg/ml), as described (*8*).

Cytotoxicity Assay—The cytotoxicity of CEL-I was evaluated by MTT assay (*24*), which detects mitochondrial succinate dehydrogenase present in living cells. In brief, 2×10^4 cells/well in a 96-well plate in α-MEM containing 35 μM bovine serum albumin (BSA) were incubated with varying concentrations of recombinant or native CEL-I for 24 h at 37°C, and then assayed as described (*8*).

RESULTS

Construction, Expression, and Purification of Recombinant CEL-I—Since the amino acid sequence of CEL-I was determined by the protein chemical techniques using the native protein purified from *C. echinata* (*9*), the gene for rCEL-I was designed as shown in Fig. 1A, based on the codon usage of *E. coli*. The rCEL-I gene was constructed as two fragments; the 5′- and 3′-halves were synthesized by PCR using forward and reverse long oligonucleotides with overlaps of 18 bases (FL1/RL1 and FL2/RL2, respectively) (Fig. 1B). They were further amplified using short oligonucleotide pairs (FS1/RS1 and FS2/RS2, respectively). The resulting double-stranded DNA fragments were ligated at the *Bgl*I site, and the entire gene was finally amplified by PCR using oligonucleotides FS1 and RS2. The nucleotide sequence of the rCEL-I gene was

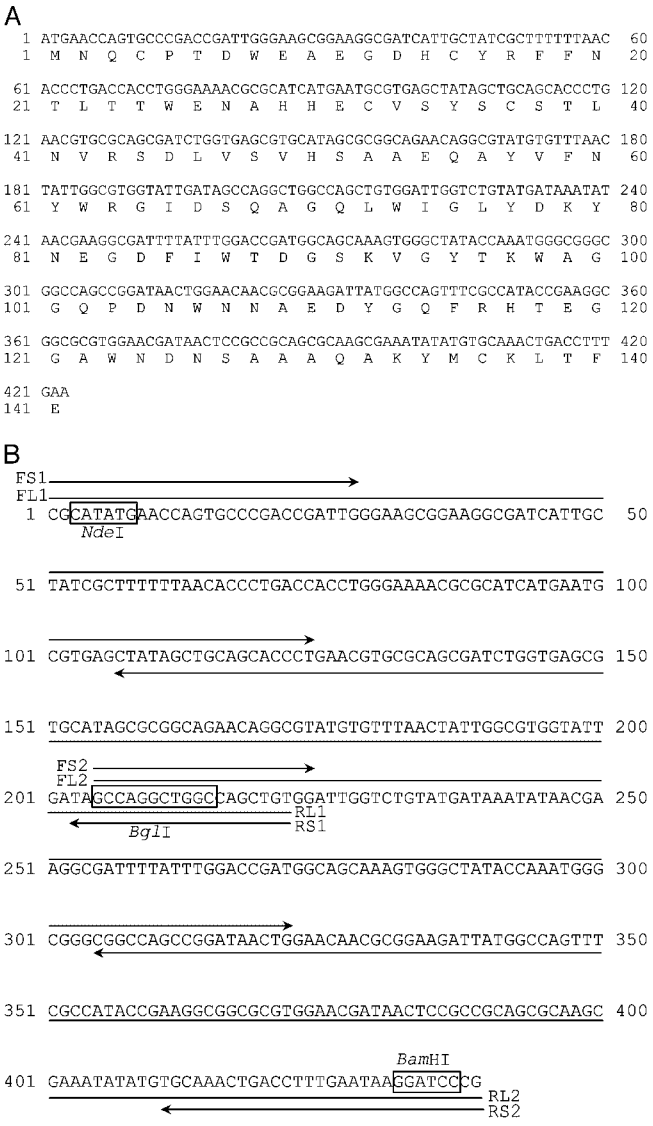

Fig. 1. **The nucleotide sequence of rCEL-I (A) and the synthetic strategy for the entire gene (B).** (A) A nucleotide sequence of recombinant CEL-I (rCEL-I) was designed based on the amino acid sequence previously reported (lower line) (*9*), taking into account the codon usage in *E. coli*. An initiator methionine residue was added to the N-terminus. (B) An artificial gene for rCEL-I was prepared using eight synthetic oligonucleotides, which are denoted by horizontal arrows with their names, as described under "MATERIALS AND METHODS." Restriction sites are enclosed in boxes. After separately preparing 5′-terminal and 3′-terminal double-stranded fragments, they were ligated at the *Bgl*I site. The obtained gene was inserted into the pET-3a vector at the *Nde*I and *Bam*HI sites.

confirmed after cloning using the pGEM-T vector in *E. coli* JM109, and expressed using the pET-3a vector in *E. coli* BL21(DE3)pLysS, as described under "MATERIALS AND METHODS." After induction with isopropyl β-D-thiogalactoside, the recombinant protein was obtained exclusively in inclusion bodies. Therefore, the latter were solubilized in 20 mM Tris-HCl buffer, pH 8.5, containing 8 M urea and 1% 2-mercaptoethanol, and then applied to a DEAE-Cellulofine column pre-equilibrated with 20 mM Tris-HCl buffer, pH 8.5, containing 4 M urea (Fig. 2A). Proteins were eluted with a linear gradient of NaCl, 0 to

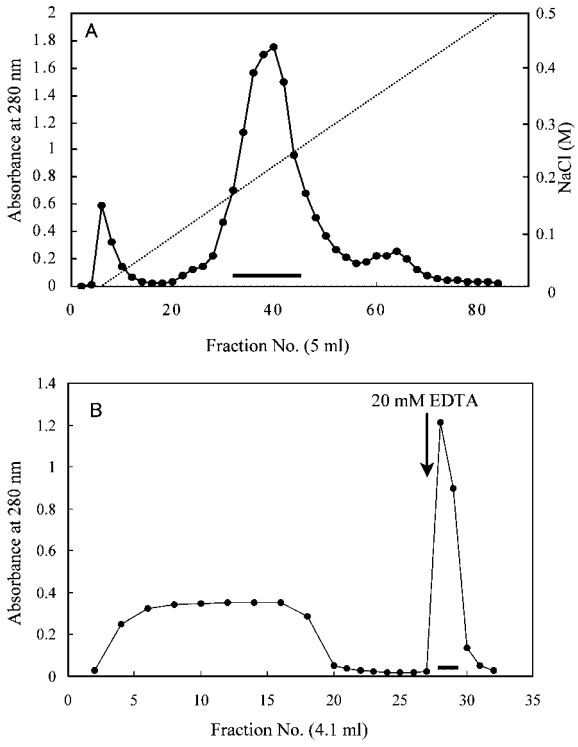

Fig. 2. **Purification of rCEL-I expressed in E. coli cells.** rCEL-I was expressed as inclusion bodies, and after solubilization with 20 mM Tris-HCl buffer, pH 8.5, containing 8 M urea and 1% 2-mercaptoethanol, it was separated on a DEAE-Cellulofine column (A). The fractions indicated by the horizontal bar, which contained rCEL-I, were pooled and dialyzed against TBS for three days to complete refolding of the recombinant protein. rCEL-I having carbohydrate-binding activity was purified by affinity chromatography on a GalNAc-Cellulofine column (B). Elution of the adsorbed protein was performed with TBS containing 20 mM EDTA.

0.5 M. The fractions containing rCEL-I were pooled and then dialyzed against TBS for three days to complete refolding and formation of disulfide bonds. Active rCEL-I with carbohydrate-binding ability was purified by affinity chromatography on a GalNAc-Cellulofine column (Fig. 2B). The adsorbed protein was eluted with 20 mM EDTA. On N-terminal sequence analysis of rCEL-I, an initiator methionine residue was found at the N-terminus. The yield of rCEL-I was 20 mg per liter of culture. Approximately 30% of the protein purified on the DEAE-Cellulofine column was adsorbed to the GalNAc-Cellulofine column. As shown in Fig. 3, the adsorbed fraction gave a band corresponding to the rCEL-I dimer linked by an interchain disulfide bond (between Cys37), which appeared at the same position as the native protein (*9*), while the flow-through fraction gave only a monomer band slightly lower than that of the adsorbed rCEL-I monomer. There appeared to be an additional faint band at around 30 kDa on an SDS-gel in the absence of 2-mercaptoethanol (Fig. 3A). This might also be a dimer of rCEL-I with a different conformation in the SDS-gel, since only a single monomer band was detected in the presence of 2-mercaptoethanol (Fig. 3B). Approximately 30% of the adsorbed protein was estimated to be the dimer, as determined on densitometry of the SDS-gel. The rCEL-I dimer was finally purified by gel filtration on

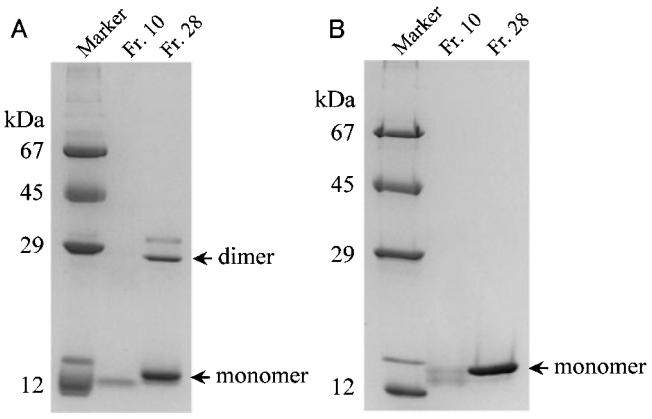

Fig. 3. **SDS-PAGE of the proteins separated on the GalNAc-Cellulofine column.** The flow-through (Fr. 10) and adsorbed (Fr. 28) fractions were examined by SDS-PAGE in the absence (A) and presence (B) of 2-mercaptoethanol.

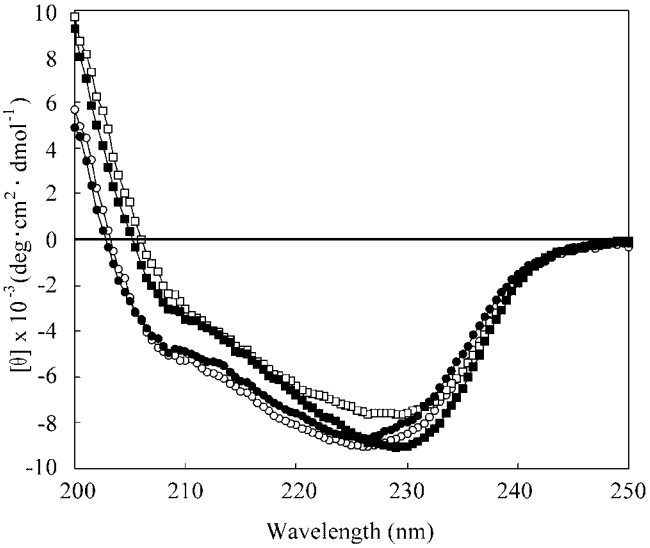

Fig. 4. **Comparison of the far UV-CD spectra of rCEL-I and native CEL-I.** Measurement was performed with 0.12 mg/ml protein in TBS at 20°C. The spectra of rCEL-I (solid circles, solid squares) and native CEL-I (open circles, open squares) were recorded in the presence (solid squares, open squares) and absence (solid circles, open circles) of 10 mM $CaCl_2$ in TBS. The values ([θ]) are expressed as the mean residue molar ellipticity.

a column of Sephadex G-50 (data not shown) and then used for the following experiments.

Hemagglutinating Activity of rCEL-I—The hemagglutinating activity of rCEL-I was determined using rabbit erythrocytes. rCEL-I induced hemagglutination at 4.2 μg/ml in the presence of 10 mM $CaCl_2$, while no agglutination was observed in the absence of $CaCl_2$, indicating that rCEL-I has Ca^{2+}-dependent carbohydrate-binding activity. The minimum concentration of rCEL-I for hemagglutination was twice that of native CEL-I (2.1 μg/ml in this experiment). The effects of several simple carbohydrates on hemagglutination by these proteins were compared, as shown in Table 1. Hemagglutination by rCEL-I was strongly inhibited by GalNAc, while other galactose-related carbohydrates caused much lower inhibition. Carbohydrates containing β-galactosides (methyl β-galactoside, *p*-nitrophenyl β-galactoside, lactose, and lactulose) seemed to show slightly higher inhibition than α-galactosides (methyl α-galactoside, melibiose, and raffinose), which is the same tendency as for native CEL-I. However, the inhibition was still more than 250-fold lower than that by GalNAc. It should be noted that galactosamine caused no inhibition even at 25 mM, which is in sharp contrast with GalNAc, and suggests that the positive charge of the amino group of galactosamine might strongly interfere with the interaction in the carbohydrate-binding sites of these proteins. Although there were some discrepancies in the order of inhibition between rCEL-I and native CEL-I, these results indicate that rCEL-I has similar carbohydrate-binding specificity to that of native CEL-I.

Interaction of Recombinant and Native CEL-I with Ca^{2+}—Figure 4 shows the far UV-CD spectra of rCEL-I and native CEL-I. Both lectins exhibited similar spectra with a negative peak at 226 nm in the absence of Ca^{2+}, suggesting that they have similar secondary structures.

Table 1. **Inhibition of hemagglutinating activity of rCEL-I toward rabbit erythrocytes by simple carbohydrates.** All the carbohydrates used were of the D-configuration.

Carbohydrate	Minimum concentration of carbohydrate inhibiting hemagglutination (mM)	
	rCEL-I	Native CEL-I[a]
GalNAc	0.0061	0.0061
Galactose	6.25	6.3
Glucose	>25	>25
Mannose	>25	>25
Fucose	6.25	6.3
Galactosamine	>25	>25
Methyl α-galactoside	12.5	12.5
Methyl β-galactoside	3.13	3.1
p-Nitrophenyl β-galactoside	1.56	6.3
Lactose (Galβ1-4Glc)	3.13	1.6
Lactulose (Galβ1-4Fru)	3.13	12.5
Melibiose (Galα1-6Glc)	12.5	12.5
Raffinose (Galα1-6Glcα1-2βFru)	12.5	25

[a]The values for native CEL-I reported previously (*5*).

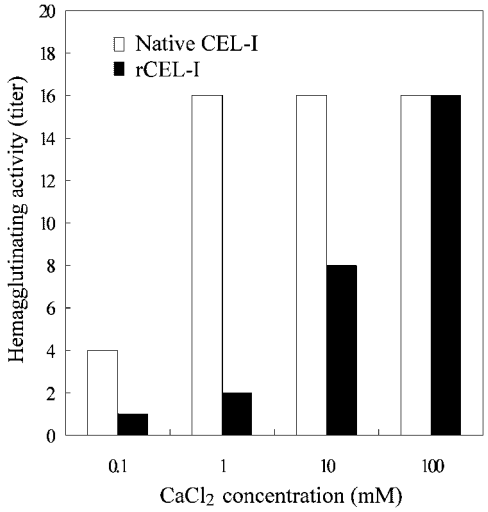

Fig. 5. **Ca^{2+}-dependence of hemagglutination by rCEL-I and native CEL-I.** The hemagglutinating activity of rCEL-I and native CEL-I was compared in the presence of the indicated concentrations of $CaCl_2$. The activity was expressed as a titer, *i.e.* the reciprocal of the highest dilution giving detectable agglutination. No activity was detected for either protein at $CaCl_2$ concentrations lower than 0.01 mM.

The shape of the spectra obviously changed upon the addition of 10 mM Ca^{2+}; the mean residue molar ellipticities in the 200–220 nm region shifted to positive values, indicating the binding of Ca^{2+} induced a definite conformational change. However, the spectra of rCEL-I and native CEL-I also exhibited some differences in intensity around 220–240 nm after binding of Ca^{2+}, suggesting that there is a minor difference in the secondary structures of the Ca^{2+}-bound forms of these proteins. The effects of Ca^{2+} on these proteins were also examined by hemagglutination assay. As shown in Fig. 5, the hemagglutinating activities of rCEL-I and native CEL-I were compared with various concentrations of Ca^{2+}. While native CEL-I exhibited maximum hemagglutinating activity at Ca^{2+} concentrations above 1 mM, the activity of rCEL-I decreased to one-eighth with a decrease in the Ca^{2+} concentration from 100 mM to 1 mM. This indicates that rCEL-I has lower Ca^{2+}-binding ability, compared to the native protein.

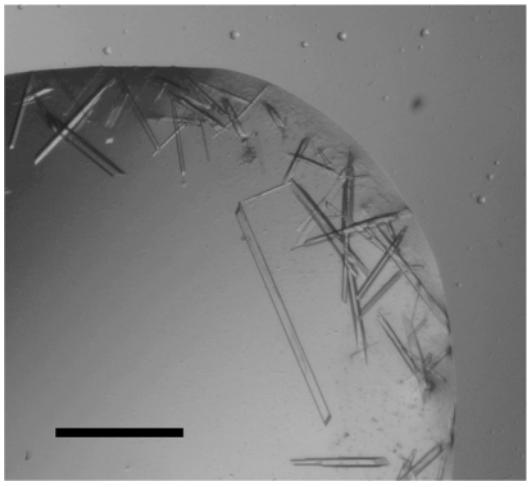

Fig. 6. **Crystals of rCEL-I grown by the hanging drop vapor diffusion method.** A rCEL-I solution (2 μl, 5 mg/ml) in TBS containing 10 mM $CaCl_2$ was mixed with an equal volume of the reservoir solution, followed by equilibration against 0.2 ml of the reservoir solution at 20°C. The bar represents 0.5 mm.

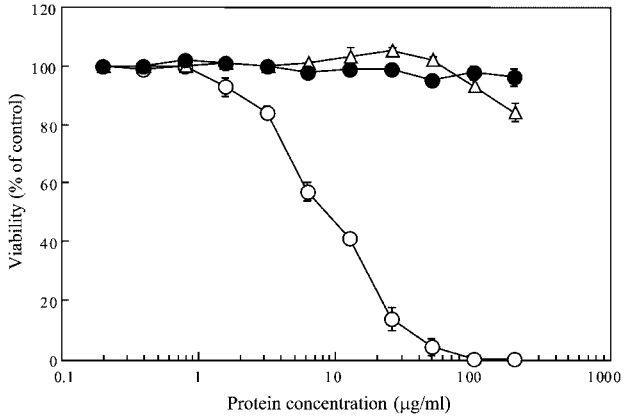

Fig. 7. **Dose-response curves for the cytotoxicity of rCEL-I and native CEL-I toward HeLa cells.** Cells in 96-well plates (2×10^4 cells/well) were incubated with various concentrations of native (open circles) and recombinant (open triangles) CEL-I in α-MEM containing 35 μM bovine serum albumin for 24 h at 37°C. Incubation with native CEL-I was also performed in the presence of 0.1 M GalNAc (solid circles). Cell viability was assessed by MTT assay (*24*). Each point is the mean of duplicate determinations.

Crystallization of rCEL-I and Measurement of Its X-Ray Diffraction—Needle-like crystals of rCEL-I were formed in a week with the hanging drop vapor diffusion method. After two months, single crystals of up to $0.02 \times 0.04 \times 0.9$ mm in size were obtained (Fig. 6). X-ray diffraction data were corrected at SPring-8 beamline BL45XU. The crystals belong to space group *C*2 with unit cell dimensions of $a = 90.77$, $b = 70.20$, $c = 63.87$ Å, and $\beta = 122.60°$. Assuming one molecule of CEL-I, which comprises two identical subunits per asymmetric unit, the Matthews constant V_m value (*25*) is 2.67 $Å^3$ Da^{-1}, corresponding to a solvent content of 53.94%. The space group of rCEL-I crystals is the same as that of native CEL-I. However, the unit cell parameters were slightly different from those for native CEL-I ($a = 92.38$, $b = 69.94$, $c = 76.69$ Å, and $\beta = 136.46°$) (*22*). However, the volumes of the unit cells in native CEL-I and rCEL-I are almost the same (3.41×10^6 and 3.43×10^6 $Å^3$, respectively).

Cytotoxicity of CEL-I—As shown in Fig. 7, native CEL-I exhibited strong cytotoxicity toward HeLa cells. The concentration giving 50% viability was approximately 8 μg/ml (0.25 μM), and most of the cells died with above 50 μg/ml (1.6 μM). In the presence of 100 mM GalNAc, this cytotoxic effect of native CEL-I was completely abolished, indicating that it is mediated by the binding of CEL-I to the cell surface carbohydrate chains containing GalNAc or related carbohydrates. On the other hand, the cytotoxicity of rCEL-I was much lower than that of the native protein; only weak cytotoxicity was observed with above 100 μg/ml of the protein. This might be partly due to the decreased carbohydrate-binding activity of rCEL-I re-

sulting from its lower Ca^{2+} affinity. It seems also possible that a slight change in the carbohydrate-binding site of the protein might affect its specificity for complex carbohydrate chains on the cell surface.

DISCUSSION

Since it was rather difficult to obtain fresh CEL-I-expressing tissues from *C. echinata* for preparation of its cDNA, we successfully attempted to construct an artificial CEL-I gene from single-stranded synthetic oligonucleotides and expressed it in *E. coli* cells. The protein was expressed exclusively as inclusion bodies. Therefore, the latter were solubilized with a denaturant, and the protein was refolded during the purification and dialysis steps. The refolding efficiency was moderately high (30%), and most of the proteins that failed to bind to the affinity column gave lower bands on SDS-PAGE (Fig. 3). Such shorter polypeptides might have resulted from unsuccessful synthesis or partial digestion by proteases in *E. coli* cells. During the purification and dialysis steps, a single interchain disulfide bond (Cys37) linking two identical subunits was formed, as judged on SDS-PAGE (Fig. 3), although it was not complete. The carbohydrate-binding activity of rCEL-I was estimated to be about half that of the native protein based on the hemagglutination assay results. However, this was found to be mostly due to a decrease in the Ca^{2+}-binding ability of rCEL-I (Fig. 5). When the Ca^{2+} concentration was increased to ten-fold (100 mM) the standard concentration we were using (10 mM), rCEL-I exhibited similar hemagglutinating activity to the native protein. The reason for this lower Ca^{2+}-binding ability of rCEL-I is uncertain, but it seems possible that the additional methionine residue at the N-terminus might have affected the Ca^{2+}-binding site and thereby reduced its affinity. Although its Ca^{2+}-binding ability was lower than that of the native protein, the hemagglutination inhibition assay demonstrated the similar carbohydrate-binding specificities of recombinant and native CEL-I, including strong GalNAc-specificity. Therefore, it seems reasonable to conclude that rCEL-I has essentially the same carbohydrate-binding site structure as the native protein. Furthermore, comparison of the far-UV CD spectra of rCEL-I and native CEL-I demonstrated similar secondary structures in the absence of Ca^{2+}. Upon the addition of 10 mM Ca^{2+}, the spectra of both proteins changed, indicating a conformational change on binding of Ca^{2+}. Although there was a subtle difference between the CD spectra of their Ca^{2+}-bound forms, these results suggest that rCEL-I and native CEL-I have similar three-dimensional structures. On the other hand, we also examined the effect of 100 mM Ca^{2+} on both proteins. However, little change was observed in their CD spectra (data not shown). This indicates that the increase in hemagglutinating activity at higher Ca^{2+} concentrations may not parallel the change in the secondary structure of rCEL-I.

rCEL-I was crystallized under the same conditions as native CEL-I (*22*). The space group of rCEL-I crystals (*C*2) was the same as that of the native protein, and the unit cell volumes were very similar. These results also seem to support the similarity of the structures of the two proteins. Differences in the unit cell dimensions between the two proteins might be due to a slight change in the conformation caused by the additional methionine residue at the N-terminus of rCEL-I. Since the crystal structure of native CEL-I was recently solved by means of the molecular replacement method using human lithostathine, which belongs to the C-type lectin superfamily, as a search model (H. Sugawara *et al*., unpublished), the crystal structure of rCEL-I should be solved in the near future by using the data for native CEL-I.

Some plant lectins are known to exhibit cytotoxicity toward cultured cells (*26–28*). Such cytotoxicity is assumed to be mediated by intracellular signals that are triggered by the binding of lectins to cell surface carbohydrates. As demonstrated in this study, native CEL-I was found to exhibit strong cytotoxicity. This cytotoxicity was completely inhibited by GalNAc, suggesting that it was brought about by the binding of CEL-I to the specific carbohydrate chains on the target cell surface. Unexpectedly, the cytotoxicity of rCEL-I was considerably weak, although it showed very similar carbohydrate-binding specificity, as demonstrated by the hemagglutination inhibition assay. This might be caused by the lower binding affinity of rCEL-I arising from the decreased affinity for Ca^{2+}, as mentioned above.

The carbohydrate-recognition mechanism of C-type lectin has been extensively studied in the case of MBP. This protein basically recognizes mannose at its 3- and 4-hydroxyl groups, which form coordinate bonds with a Ca^{2+} ion in the carbohydrate-binding site as well as hydrogen bonds with amino acid residues in the vicinity (*29*). In addition to wild-type MBP, several mutants, whose amino acid residues were altered by site-directed mutagenesis, have been constructed (*30*). For the GalNAc-binding mutant, a histidine residue making van der Waals contact with the acetoamide group of GalNAc was introduced in addition to replacement of several amino acid residues and insertion of a glycine-rich sequence (*19*), based on comparison with rat hepatic lectin RHL-1 (*31*), which binds to GalNAc 60-fold stronger than to galactose. These observations suggest that the specificity of various C-type lectins could be changed by replacing only a limited number of amino acid residues around the carbohydrate-binding site. Since CEL-I exhibits much higher GalNAc-specificity (1,000-fold higher affinity for GalNAc than for galactose) than RHL-1 does, it should provide valuable clues for enhancing the specificity for particular carbohydrates using protein engineering techniques. With the data on the X-ray crystallographic structure of CEL-I, it should become possible to develop artificial lectins with novel carbohydrate-binding specificities in the future.

This work was supported by a Grant-in-Aid for Scientific Research (14560073 to T.H.) from the Japan Society for the Promotion of Science.

REFERENCES

1. Dodd, R.B. and Drickamer, K. (2001) Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. *Glycobiology* **11**, 71R–79R
2. Kilpatrick, D.C. (2002) Animal lectins: a historical introduction and overview. *Biochim. Biophys. Acta* **1572**, 187–197

3. Drickamer, K. (1988) Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **263**, 9557–9560
4. Day, A.J. (1994) The C-type carbohydrate recognition domain (CRD) superfamily. *Biochem. Soc. Trans* **22**, 83–88
5. Hatakeyama, T., Kohzaki, H., Nagatomo, H., and Yamasaki, N. (1994) Purification and characterization of four Ca^{2+}-dependent lectins from the marine invertebrate, *Cucumaria echinata*. *J. Biochem.* **116**, 209–214
6. Hatakeyama, T., Nagatomo, H., and Yamasaki, N. (1995) Interaction of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* with the erythrocyte membrane. *J. Biol. Chem.* **270**, 3560–3564
7. Hatakeyama, T., Furukawa, M., Nagatomo, H., Yamasaki, N., and Mori, T. (1996) Oligomerization of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* induced by the binding of carbohydrate ligands. *J. Biol. Chem.* **271**, 16915–16920
8. Oda, T., Tsuru, M., Hatakeyama, T., Nagatomo, H., Muramatsu, T., and Yamasaki, N. (1997) Temperature- and pH-dependent cytotoxic effect of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* on various cell lines. *J. Biochem.* **121**, 560–567
9. Hatakeyama, T., Matsuo, N., Shiba, K., Nishinohara, S., Yamasaki, N., Sugawara, H., and Aoyagi, H. (2002) Amino acid sequence and carbohydrate-binding analysis of the *N*-acetyl-D-galactosamine-specific C-type lectin, CEL-I, from the Holothuroidea, *Cucumaria echinata*. *Biosci. Biotechnol. Biochem.* **66**, 157–163
10. Ng, K.K., Drickamer, K., and Weis, W.I. (1996) Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. *J. Biol. Chem.* **271**, 663–674
11. Feinberg, H., Torgersen, D., Drickamer, K., and Weis, W.I. (2000) Mechanism of pH-dependent *N*-acetylgalactosamine binding by a functional mimic of the hepatocyte asialoglycoprotein receptor. *J. Biol. Chem.* **275**, 35176–35184
12. Poget, S.F., Freund, S.M., Howard, M.J., and Bycroft, M. (2001) The ligand-binding loops in the tunicate C-type lectin TC14 are rigid. *Biochemistry* **40**, 10966–10972
13. Feinberg, H., Mitchell, D.A., Drickamer, K., and Weis, W.I. (2001) Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. *Science* **294**, 2163–2166
14. Ng, K.K., Kolatkar, A.R., Park-Snyder, S., Feinberg, H., Clark, D.A., Drickamer, K., and Weis, W.I. (2002) Orientation of bound ligands in mannose-binding proteins. Implications for multivalent ligand recognition. *J. Biol. Chem.* **277**, 16088–16095
15. Shrive, A.K., Tharia, H.A., Strong, P., Kishore, U., Burns, I., Rizkallah, P.J., Reid, K.B., and Greenhough, T.J. (2003) High-resolution structural insights into ligand binding and immune cell recognition by human lung surfactant protein D. *J. Mol. Biol.* **331**, 509–523
16. Kozutsumi, Y., Kawasaki, T., and Yamashina, I. (1980) Isolation and characterization of a mannan-binding protein from rabbit serum. *Biochem. Biophys. Res. Commun.* **95**, 658–664
17. Drickamer, K., Dordal, M.S., and Reynolds, L. (1986) Mannose-binding proteins isolated from rat liver contain carbohydrate-recognition domains linked to collagenous tails. Complete primary structures and homology with pulmonary surfactant apoprotein. *J. Biol. Chem.* **261**, 6878–6887
18. Thiel, S., Vorup-Jensen, T., Stover, C.M., Schwaeble, W., Laursen, S.B., Poulsen, K., Willis, A.C., Eggleton, P., Hansen, S., Holmskov, U., Reid, K.B., and Jensenius, J.C. (1997) A second serine protease associated with mannan-binding lectin that activates complement. *Nature* **386**, 506–510
19. Kolatkar, A.R., Leung, A.K., Isecke, R., Brossmer, R., Drickamer, K., and Weis, W.I. (1998) Mechanism of N-acetylgalactosamine binding to a C-type animal lectin carbohydrate-recognition domain. *J. Biol. Chem.* **273**, 19502–19508
20. Suzuki, T., Takagi, T., Furukohri, T., Kawamura, K., and Nakauchi, M. (1990) A calcium-dependent galactose-binding lectin from the tunicate *Polyandrocarpa misakiensis*. Isolation, characterization, and amino acid sequence. *J. Biol. Chem.* **265**, 1274–1281
21. Poget, S.F., Legge, G.B., Proctor, M.R., Butler, P.J., Bycroft, M., and Williams, R.L. (1999) The structure of a tunicate C-type lectin from *Polyandrocarpa misakiensis* complexed with D-galactose. *J. Mol. Biol.* **290**, 867–879
22. Hatakeyama, T., Matsuo, N., Aoyagi, H., Sugawara, H., Uchida, T., Kurisu, G., and Kusunoki, M. (2002) Crystallization and preliminary crystallographic study of an invertebrate C-type lectin, CEL-I, from the marine invertebrate *Cucumaria echinata*. *Acta Crystallog. Sec. D* **58**, 143–144
23. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode in *Methods Enzymol.* (Carter, Jr., C.W. and Sweet, R.M., eds.,) Vol. **276**, pp. 307–326, Academic Press, New York
24. Mosmann, T. (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* **65**, 55–63
25. Matthews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497
26. Stanley, W.S., Peters, B.P., Blake, D.A., Yep, D., Chu, E.H., and Goldstein, I.J. (1979) Interaction of wild-type and variant mouse 3T3 cells with lectins from *Bandeiraea simplicifolia* seeds. *Proc. Natl Acad. Sci. USA* **76**, 303–307
27. Kim, M., Rao, M.V., Tweardy, D.J., Prakash, M., Galili, U., and Gorelik, E. (1993) Lectin-induced apoptosis of tumour cells. *Glycobiology* **3**, 447–453
28. Ohba, H., Bakalova, R., and Muraki, M. (2003) Cytoagglutination and cytotoxicity of wheat germ agglutinin isolectins against normal lymphocytes and cultured leukemic cell lines–relationship between structure and biological activity. *Biochim. Biophys. Acta* **1619**, 144–150
29. Weis, W.I., Drickamer, K., and Hendrickson, W.A. (1992) Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **360**, 127–134
30. Iobst, S.T. and Drickamer, K. (1996) Selective sugar binding to the carbohydrate recognition domains of the rat hepatic and macrophage asialoglycoprotein receptors. *J. Biol. Chem.* **271**, 6686–6693
31. Drickamer, K., Mamon, J.F., Binns, G., and Leung, J.O. (1984) Primary structure of the rat liver asialoglycoprotein receptor. Structural evidence for multiple polypeptide species. *J. Biol. Chem.* **259**, 770–778