The clostridial collagenases G and H are multidomain proteins. For collagen digestion, the domain arrangement is likely to play an important role in collagen binding and hydrolysis. In this study, the full-length collagenase H protein from Clostridium histolyticum was expressed in Escherichia coli and purified. The N-terminal amino acid of the purified protein was Ala31. The expressed protein showed enzymatic activity against azocoll as a substrate. To investigate the role of Ca\(^{2+}\) in providing structural stability to the full-length collagenase H, biophysical measurements were conducted using the recombinant protein. Size exclusion chromatography revealed that the Ca\(^{2+}\) chelation by EGTA induced interdomain conformational changes. Dynamic light scattering measurements showed an increase in the percent polydispersity as the Ca\(^{2+}\) was chelated, suggesting an increase in protein flexibility. In addition to these conformational changes, differential scanning fluorimetry measurements revealed that the thermostability was decreased by Ca\(^{2+}\) chelation, in comparison with the thermal melting point (\(T_m\)). The melting point changed from 54 to 49°C by the addition of excess Ca\(^{2+}\). These results indicated that the interdomain flexbility and the domain arrangement of full-length collagenase H are reversibly regulated by Ca\(^{2+}\).

Collagens are an important component of animal tissues, which are composed of many different cell types. Collagenases specifically break down collagen, which has a tight triple helical structure and is resistant to most proteases. Clostridium histolyticum collagenases G and H (ColG and ColH) can easily digest collagens, regardless of their types and sizes (27). Therefore, they are used extensively as a clinical tool in the nonsurgical treatment of Dupuytren’s disease (1, 13) and in the isolation of cells from tissues and organs, such as in the preparation of pancreatic islet cells for transplantation (5, 6, 15).

ColG and ColH require the presence of Ca\(^{2+}\) for their collagenolytic activity (2–4). The putative cleavage sites of ColG and ColH are the end portions and central portion of collagen, respectively (12, 20). The different cleavage sites may arise from the distinct domain architectures, that is, the different numbers of collagen binding domains (CBD) and polycystic kidney disease-like (PKD-like) domains. ColG and ColH consist of a collagenase module containing two domains, one or two copies of PKD-like domain, and one or two copies of CBD. A schematic diagram of the domain architectures of ColG and ColH is shown in Fig. 1. The percent sequence identity between these collagenases, calculated by ClustalW (26), is 41% for the entire sequences. The identities compared among individual domain types are 48% for the collagenase modules, 40 to 50% for the PKD-like domains, and 29 to 34% for the CBDs. The sequential alignments for the individual domains are shown in Fig. 2.

Structure determinations have been performed for the individual domains over the past few years. The crystal structures of the CBDs were solved as the apo- and holo-forms for ColG (29) and the holo-form for ColH (Protein Data Bank [PDB] entries 3IQX and 3JQW). The CBD structures revealed that two Ca\(^{2+}\) ions bind to the N-terminal linker region of the CBD and provide conformational stability to the CBD. Ca\(^{2+}\) binding is considered to have a critical role in enhancing interactions with collagen (29). Furthermore, the collagen triple-helix binding mode was proposed by a nuclear magnetic resonance titration and small-angle X-ray scattering study (22). Recently, the structures of the collagenase module and PKD-like domain of ColG have been reported (9, 10). The collagenase module adopted a unique architecture, and a collagenolysis mechanism was proposed based on the structural information for the individual domains of ColG. In the PKD-like domain, Ca\(^{2+}\) was expected to bind to the N-terminal region of the domain as well as the CBD (9).

The cloning of the colG and colH genes facilitated the development of collagenase expression systems. The Bacillus subtilis expression system produced ColH as the secreted form, but the transformation efficiency was low (17). The C. perfringens expression system alleviated the translational problems and expressed ColH as the secreted form (25). The Escherichia coli expression system by Ducka et al. provided the effective production of the mature and various truncated forms of ColG, ColH, and ColT (from Clostridium tetani) (8).

Rapid progress has been made in clarifying the molecular mechanisms of the individual domains for both the function and structure. However, no structural studies of the entire collagenase molecule, as a multidomain protein, have been performed. Understanding the interplay of the domains is important, and structural information about the full-length protein should provide valuable insight about its function. Here, we report the biophysicochemical properties of the full-length protein with and without Ca\(^{2+}\) chelation, and compare these results with those obtained from various truncated forms of the protein. These full-length collagenase molecules were expressed in E. coli and purified. The structure determinations revealed that the domains of the full-length collagenase molecules have unique architectures from those of the shorter molecules. The percent sequence identity between these collagenases, calculated by ClustalW (26), is 41% for the entire sequences. The identities compared among individual domain types are 48% for the collagenase modules, 40 to 50% for the PKD-like domains, and 29 to 34% for the CBDs. The sequential alignments for the individual domains are shown in Fig. 2.
cal characterization of full-length ColH, expressed in E. coli. In the presence and absence of Ca\(^{2+}\), ColH underwent reversible conformational changes. The Ca\(^{2+}\)-binding form of ColH was thermally more stable, suggesting that Ca\(^{2+}\) plays an important role in structural stability, rather than just the binding of the CBD to collagen. The binding of Ca\(^{2+}\) reduced the flexibility of the full-length ColH, implying that Ca\(^{2+}\) regulates the structural stability among the five domains in solution. Hydrophobic interactions also con-

FIG 1 Schematic diagram of the domain architectures of ColG and ColH. Pre, signal peptide; Pro, putative prodomain; PKD, polycystic kidney disease-like domain; CBD, collagen binding domain. The collagenase module is composed of two functional domains, an activator domain, and a peptidase domain.

(A)

(B)

(C)

FIG 2 Amino acid sequence alignment of the peptidase domains, PKD-like domains, and CBDs of ColG and ColH. Secondary-structure elements are indicated by looped lines (α-helix) and arrows (β-strand). Conserved amino acids between ColG and ColH are in bold type. (A) Peptidase domain (around the active site). Amino acids highlighted in the boxes are crucial for Zn\(^{2+}\) binding. (B) PKD-like domains. Putative Ca\(^{2+}\)-binding sites are highlighted in a gray background. (C) CBD. The Ca\(^{2+}\)-binding sites observed in the crystal structure (PDB entries 1NQD and 3JQW) are highlighted in a gray background. Secondary-structure elements are considered the holo-form.
MATERIALS AND METHODS

Expression vector preparation. C. histolyticum genomic DNA was prepared as described by Yoshihara et al. (30). The open reading frame (ORF) encoding ColH (GenBank accession number D29981.1) in the genomic DNA was amplified with the primers ColHN-Ndel (5'-GGGGGAAAGC-CTATGAAAAAGGACTG-3'; the lowercase letter shows an inserted base, and the underline shows the Ndel site) and ColHC-Xbal (5'-AAAAATAT TAGCATAGCTACGTACCTC-3'; lowercase letters show substituted bases, and the underline shows the Xbal site). The obtained PCR fragment was digested with Ndel and EcoRI or with EcoRI and Xbal. The Ndel-EcoRI fragment (1,605 bp) and the EcoRI-Xbal fragment (1,464 bp) were recovered as insert DNAs. The DNA fragments were inserted into pCold3-DNA (TaKaRa, Japan) digested with Ndel and Xbal. Furthermore, to introduce the His tag for the protein purification process, two phosphorylated oligonucleotides, 5'-pCTAGACATCATCATCTATCATCATTAG-3' and 5'-pCTAGACTAATGATGATGATGT-3', were mixed, denatured at 95°C for 5 min, and then left at room temperature. The annealed DNA fragment was inserted into the Xbal site of the plasmid described above. As the result of the insertion, the C-terminal amino acid sequence of ColH was changed from 1018SVGR* to 1018SVSRHHHHH*. The constructed plasmid vector was designated pCold3-ColH-His.

Enzymatic activity assay. E. coli strain BLR, transformed with pCold3-ColH-His by chemically induced transformation, was cultured at 37°C in LB broth containing 50 µg/ml ampicillin to an optical density at 660 nm (OD660) of 0.5. Isopropyl-β-D-thiogalactopyranoside (IPTG) (Nacalai Tesque, Japan) was then added to the broth (1 mM final concentration), and the E. coli culture was grown at 15°C for 24 h. The culture was centrifuged at 8,000 × g for 5 min at 4°C. The supernatant was used as a crude enzyme solution. The collagenase assay was performed by the improved method of Chavira et al. (7). A 200-µl aliquot of 0.5 M Tris-HCl (pH 7.0) and 200 µl of a 10-µg/ml azocoll suspension were mixed on ice. A 200-µl portion of the culture supernatant was added, and the mixture was immediately incubated at 30°C. After precisely 30 min, the mixture was cooled to 0°C on ice and centrifuged at 4°C for 10 min. The A540 of the supernatant was then measured. The culture broth of E. coli transformed with pCold3 was used as a blank. One unit of collagenase was defined as the difference in absorbance at 518 nm of the soluble chromogenic peptides released for 1 min.

Protein expression and purification. A 3-liter LB culture of E. coli strain BL21(DE3) harboring the expression plasmid was incubated at 37°C and induced by the addition of IPTG to a final concentration of 1 mM. After induction, the culture was incubated at 20°C for 24 h. The supernatant of the culture was obtained by centrifugation at 6,000 × g for 20 min at 4°C and then applied to a Ni-nitrilotriacetic acid (Ni-NTA) Superflow column (Qiagen, Germany) preequilibrated with 20 mM HEPES (pH 8.0), 1 M NaCl, and 1 mM CaCl2. The sample was eluted with a buffer containing 500 mM imidazole and concentrated by ultrafiltration membrane. The concentrated sample was diluted and then applied to an ion-exchange column (MonoQ; GE Healthcare, United Kingdom). The peak fractions were pooled and applied to a size exclusion chromatography column (Superdex 200; GE Healthcare) preequilibrated with 20 mM HEPES (pH 8.0), 100 mM NaCl, and 1 mM CaCl2. The protein was concentrated to the final concentration of 10 mg/ml. Protein concentrations were determined by the bicinchoninic acid (BCA) method (24), using a BCA protein assay kit (Thermo Fisher Scientific Inc., MA) and bovine serum albumin as a standard. The N-terminal sequence of the protein was analyzed by a protein sequencer, PPSQ-33A (Shimadzu, Japan).

Size exclusion chromatography analysis. Size exclusion chromatography was performed on a Superdex 200 column (10/300GL; GE Healthcare) at a flow rate of 0.3 ml/min at 25°C. The column was equilibrated with buffers composed of 20 mM HEPES (pH 8.0), 100 mM NaCl, and 1 mM CaCl2, and various concentrations (0, 0.25, 0.5, 0.75, 1, 1.5, and 10 mM) of EGTA. ColH (5 mg/ml) was prepared in the same buffer before application to the column. To estimate the molecular size, a gel filtration calibration kit (GE Healthcare) was used for molecular mass standards: ferritin, 440.0 kDa; aldolase, 158.0 kDa; conalbumin, 75.0 kDa; ovalbumin, 43.0 kDa; carbonic anhydrase, 29.0 kDa; RNase A, 13.7 kDa; and aprotinin, 6.5 kDa.

RESULTS

Protein preparation and assay. The protein sample was prepared from E. coli as the secreted form. The typical yield of protein purified from a 1-liter culture was 2 mg. The ColH protein was purified to homogeneity, as determined by an SDS-PAGE analysis (Fig. 3). The N-terminal amino acid sequence of the purified protein was analyzed. The first five residues were found to be AVDKN, consistent with sequence numbers 31 to 35 of ColH. This N-terminal sequence is identical to those obtained in the previously reported expression systems, in which E. coli DH5α, Bacillus subtilis DB104, and Clostridium perfringens were used as hosts (17, 25). The enzymatic activity of the culture broth of E. coli bearing pCold3-ColH-His was 16 U/liter.

Overall conformational change induced by calcium ions. To examine the effect of Ca2+ on full-length ColH, the apparent molecular mass of the protein was determined by size exclusion chromatography in the presence of various EGTA concentrations. The elution peak position shifted to faster positions at EGTA concentrations higher than 1 mM (Fig. 4). The apparent molecular mass was approximately 82 kDa without EGTA. In the presence of EGTA (>1 mM), the apparent molecular mass increased to 119 kDa. The calculated molecular mass is 114 kDa as a monomer. The
The apparent molecular mass of 82 kDa is significantly smaller than expected.

The hydrodynamic radius was measured in the presence and the absence of EGTA by dynamic light scattering (DLS). The estimated molecular masses and the percent polydispersity were 108 kDa and 31.6% (without EGTA) and 116 kDa and 38.4% (with EGTA), respectively (Table 1). These results showed that ColH exists as a monomer in solution and that the particle size distribution increases by the addition of EGTA.

We further investigated the structural conservation of ColH by far-UV CD measurements in the presence and absence of 2 mM EGTA. The two CD spectra were essentially identical. Thus, the secondary-structure elements and their arrangement in ColH were basically maintained in the presence and absence of Ca\(^{2+}\).

**Enhancement of thermostability by calcium ions.** Philomnathan et al. showed that Ca\(^{2+}\) contributes to the thermostability of the CBD of ColG, by differential scanning calorimetry (DSC) (23). In this study, the thermostabilizing effect of Ca\(^{2+}\) toward ColH was measured by differential scanning fluorimetry (DSF) (11, 21, 28). Fluorescent dyes, including SYPRO orange and 1-anilinonaphthalene-8-sulfonic acid (ANS), bind to hydrophobic regions of proteins and show intense fluorescence. During the thermal unfolding process, hydrophobic regions of the protein are exposed to the solvent, resulting in the stimulation of dye binding and an increase in the fluorescence intensity.

The thermostability was evaluated from the change in the \(T_m\) (Fig. 5). The \(T_m\) of ColH with calcium was 54°C. In the presence of >1 mM EGTA, the \(T_m\) decreased to 49°C, suggesting destabilization. Subsequently, the protein solution was mixed with 1 mM EGTA before the DSF measurement, and then Ca\(^{2+}\) was added to the solution. As a result, the \(T_m\) was observed to recover from 49 to 54°C, as the Ca\(^{2+}\) concentration in the solution was increased. Hence, the addition of >1 mM Ca\(^{2+}\) led to the full recovery of the \(T_m\). This result indicated that the presence of Ca\(^{2+}\) reversibly contributes to the thermostability of ColH.

As mentioned above, fluorescent dyes, such as SYPRO orange, are a useful tool to probe the existence of hydrophobic regions exposed to the solvent. In this study, a fluorescence analysis was conducted to assess the contribution of the hydrophobic interactions to the protein stability. Calmodulin, a known Ca\(^{2+}\)-binding protein, exposes its hydrophobic surfaces upon Ca\(^{2+}\) binding to stabilize protein-protein interactions (16, 18). We evaluated whether ColH also undergoes such a transition in the presence of Ca\(^{2+}\). The fluorescence intensities of ColH were very weak at EGTA concentrations of less than 1 mM. However, the fluorescence intensity dramatically increased at EGTA concentrations of >1 mM (Fig. 6). These results indicated that hydrophobic regions of CoH are exposed by calcium chelation.

**DISCUSSION**

The structures of the CBDs have been solved for both ColG and ColH, and they harbor two calcium binding sites (29). The PKD-like domain structure has been reported for ColG (9). Although calcium binding was not observed in the crystal structure, a putative calcium binding site was proposed to exist near the N-terminal region of the domain. ColH contains a collagenase module (activator and peptidase domains), two PKD-like domains, and one CBD. Therefore, at least four calcium binding sites are expected to be present in this protein. All four binding sites are located near linker regions between domains. According to previous reports (4, 20), ColH can accept approximately five Ca\(^{2+}\). It is possible that ColH has an additional Ca\(^{2+}\)-binding site, perhaps in the linker regions for which structural data are missing.

**TABLE 1** Hydrodynamic radius, estimated molecular mass, and percent polydispersity

<table>
<thead>
<tr>
<th>EGTA</th>
<th>Hydrodynamic radius (nm)</th>
<th>Molecular mass (kDa)</th>
<th>% Polydispersity</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>4.40 (±0.02)</td>
<td>108</td>
<td>31.6</td>
</tr>
<tr>
<td>+</td>
<td>4.54 (±0.04)</td>
<td>116</td>
<td>38.4</td>
</tr>
</tbody>
</table>

![FIG 3 SDS-PAGE of purified ColH (0.8 μg). M, molecular mass markers.](image1)

![FIG 4 Size exclusion chromatography analysis of ColH with various concentrations of EGTA. The inset shows the size chromatography. The dashed line indicates the theoretical molecular mass of ColH (114 kDa).](image2)
The size exclusion chromatography of full-length ColH revealed that the apparent molecular mass (82 kDa) was ~30% smaller than the calculated value (114 kDa). In the presence of EGTA (>1 mM), the apparent molecular mass was determined to be 119 kDa, which is comparable to the calculated value. The DLS measurements indicated that the percent polydispersity increased as EGTA was added, indicating an enhancement in the flexibility of the protein. No significant difference in the secondary structures was observed for the protein in the presence and absence of EGTA, according to the CD data. These results, therefore, suggest that Ca$^{2+}$ influences the interdomain conformation. The domain linker between the second PKD-like domain and CBD possesses two calcium binding sites (29). For the CBD of ColG, significant changes in the apparent molecular mass by Ca$^{2+}$ chelation were observed in size exclusion chromatography and DLS (19, 29). However, no significant change in the hydrodynamic radius of ColH was observed in the DLS measurements. The conformational changes of ColH by calcium chelation may reduce the number of interactions with the matrix in the size exclusion chromatography column. Spectrometric measurements with the fluorescent dye indicated that hydrophobic interactions also contribute to the interdomain conformation of ColH. Remarkable conformational changes have been suggested for the N-terminal region of the CBD of ColG; however, the ANS fluorescence intensity in the presence and absence of Ca$^{2+}$ did not change (23). Although the regions participating in the hydrophobic interactions have not been identified so far, they may be different from the linker region at the N terminus of the CBD.

Calcium ions were also found to influence the thermostability of full-length ColH. DSF measurements demonstrated that the $T_m$ decreased following calcium chelation. The difference in the $T_m$ was 5°C between the stabilized (with Ca$^{2+}$, 54°C) and destabilized (with EGTA, 49°C) states. The higher $T_m$ was reversibly restored by the addition of excess Ca$^{2+}$. Among the domains in ColG and ColH, the thermostability of the second CBD of ColG was previously evaluated by DSC (23). The $T_m$ of the CBD apo-form and a holo-form were determined to be 74 and 93°C, respectively. DSF and DSC are known to yield equivalent values as $T_m$ estimations (14). Considering the sequence conservation between ColH and ColG, it can be assumed that the CBD of ColH also has relatively high $T_m$. Consequently, the heat destabilization process in the initial stages of protein unfolding does not occur within the CBD of ColH. Although it is challenging to identify the initial region that is thermally destabilized, the peptidase domain and the following linker region are considered likely possibilities. The

### FIG 5
Thermal shift assay of ColH by differential scanning fluorimetry. (A) Effect of Ca$^{2+}$ on the thermostability of ColH. (B) Recovery of the thermostability by the addition of CaCl$_2$.

<table>
<thead>
<tr>
<th>EGTA (mM)</th>
<th>0</th>
<th>0.25</th>
<th>0.5</th>
<th>0.75</th>
<th>1</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_m$ (°C)</td>
<td>54</td>
<td>53</td>
<td>53</td>
<td>52</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CaCl$_2$ (mM)</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_m$ (°C)</td>
<td>54</td>
<td>54</td>
</tr>
</tbody>
</table>

### FIG 6
Fluorescence analyses of ColH with various concentrations of EGTA. The fluorescence intensity of ColH (2 μM) was measured after the addition of EGTA (0 to 10 mM) (circles). The fluorescence intensity recovery point following the addition of 1 mM CaCl$_2$ is indicated by the triangle. The inset shows the fluorescence spectra of the sample solutions with 0 mM EGTA, 1 mM EGTA, and 1 mM EGTA plus 1 mM CaCl$_2$. 

<table>
<thead>
<tr>
<th>EGTA concentration (mM)</th>
<th>0</th>
<th>0.1</th>
<th>0.5</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluorescence Intensity at $A_{max}$</td>
<td>8,000</td>
<td>7,000</td>
<td>6,000</td>
<td>5,000</td>
</tr>
</tbody>
</table>

The fluorescence intensity of ColH was measured after the addition of EGTA (0 to 10 mM) (circles). The fluorescence intensity recovery point following the addition of 1 mM CaCl$_2$ is indicated by the triangle. The inset shows the fluorescence spectra of the sample solutions with 0 mM EGTA, 1 mM EGTA, and 1 mM EGTA plus 1 mM CaCl$_2$. 

August 2012 Volume 78 Number 16 aem.asm.org 5843
structure of the PKD-like domain could not be exactly located in the crystal structure of CoG (Tyr119-Asn880) (10), and thus the collagenase module and the PKD-like domain may be flexibly linked.

In this study, we focused on full-length ColH as a multidomain structure and discussed the influence of interdomain flexibility on the arrangement of these individual domains. The results showed that the interdomain flexibility is predominantly and reversibly maintained by Ca$^{2+}$. The locations of the calcium binding sites are not known, except for the CBD. Nevertheless, the results presented herein indicate that the control of domain linker flexibility by Ca$^{2+}$ plays an important structural role, and particular domain arrangement control mechanisms by Ca$^{2+}$ may be commonly adopted by other collagenases, such as CoG.

ACKNOWLEDGMENTS
We acknowledge the support of the Biomedical Research Core of Tohoku University, Graduate School of Medicine. Part of this research was supported by the Coordination, Support and Training Program for Translational Research, from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES