Substrate Recognition by the Collagen-binding Domain of
Clostridium histolyticum Class I Collagenase*

Received for publication, April 24, 2000, and in revised form, December 7, 2000
Published, JBC Papers in Press, December 19, 2000, DOI 10.1074/jbc.M003450200

Osamu Matsushita‡‡, Takaki Koide§§, Ryoji Kobayashi‡‡, Kazuhiro Nagata¶¶, and Akinobu Okabe‡‡

From the Departments of §§Microbiology and ¶¶Chemistry, Faculty of Medicine, Kagawa Medical University, Miki-cho, Kita-gun, Kagawa 761-0793, Japan, ¶¶Core Research for Evolutional Science and Technology, Japan Science and Technology Corporation, Kawauchi, Saitama 332-0012, Japan, and the §Department of Molecular and Cellular Biology, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8397, Japan

Clostridium histolyticum type I collagenase (ColG) has a segmental structure, S1+S2+S3a+S3b. S3a and S3b bound to insoluble collagen, but S2 did not, thus indicating that S3 forms a collagen-binding domain (CBD). Because S3a+S3b showed the most efficient binding to substrate, cooperative binding by both domains was suggested for the enzyme. Monomeric (S3b) and tandem (S3a+S3b) CBDs bound to telopeptide, which contains only the collagenous region. However, they did not bind to telopeptides immobilized on Sepharose beads. These results suggested that the binding site(s) for the CBD is(are) present in the collagenous region. The CBD bound to immobilized collagenous peptides, (Pro-Hyp-Gly)n and (Pro-Pro-Gly)n, only when n is large enough to allow the peptides to have a triple-helical conformation. They did not bind to various peptides with similar amino acid sequences or to gelatin, which lacks a triple-helical conformation. The CBD did not bind to immobilized Glc-Gal disaccharide, which is attached to the side chains of hydroxylysine residues in the collagenous region. These observations suggested that the CBD specifically recognizes the triple-helical conformation made by three polypeptide chains in the collagenous region.

Collagens are the major components of the extracellular matrix. They are the most abundant proteins in mammals, constituting a quarter of their total weight. Collagens are not only structural proteins with a high tensile strength but also affect cell differentiation, migration, and attachment. Nineteen different types are known to date, and type I collagen is the major species in higher vertebrates. In the endoplasmic reticulum of fibroblasts, collagen precursor peptides are hydroxylated by various dioxygenases (prolyl hydroxylase and a lysyl hydroxylase), glycosylated at the hydroxylysine residues, and folded into a triple helix. HSP47 associates with procollagen during this modification and/or folding processes and is assumed to function as a collagen-specific molecular chaperone (1). The precursor is secreted into the extracellular space, and its N- and C-terminal propeptides are removed by procollagen peptidases. Cleavage triggers the arrangement of collagen monomers into a staggered array called fibrils, which are insoluble macromolecular assemblies having lengths up to a few hundred micrometers. One can observe 67-nm-wide cross striations (overlap and gap zones) on the fibrils by electron microscopy, which are formed by the regular arrangement of collagen monomers. Each collagen monomer consists of collagenous and noncollagenous regions. The former have a regular amino acid sequence, which is a repeat (338 times in type I collagen) of the Gly-X-Y triplet, forming a long (300 nm in type I collagen) triple-helical domain. Proline and hydroxyproline residues are most common in the X and Y positions, respectively, and this combination stabilizes the helix to the highest extent (2). The latter regions lack the sequence and conformation characteristic of the former, and are named telopeptides, because they are located at the N and C termini of the collagenous region.

Clostridium histolyticum is a strictly anaerobic Gram-positive bacterium that is one of the causative agents of gas gangrene. The pathogen produces a variety of collagenases in large quantity, which efficiently degrade collagen in connective tissue. Although the enzyme has been widely used in biological experiments, its detailed biochemical characteristics have not been fully understood (3). Previously, we have shown that C. histolyticum has two divergent collagenase genes (colG and colH) in its chromosome, which encode two distinct enzymes; class I collagenase (ColG) and class II collagenase (ColH) (4, 5). Their deduced amino acid sequences showed no significant similarity to those of eukaryotic matrix metalloproteases. Comparison of the ColG and ColH sequences revealed their segmental structure; S1+S2+S3a+S3b for ColG and S1+S2a+S2b+S3 for ColH as shown in Fig. 1A (see below). A truncated ColH consisting of only S1 showed full hydrolytic activity on a peptide substrate (6). When a consensus motif for zinc proteases, HEXXH, present in S1 was altered, catalytic activity was drastically reduced (7). Thus, S1 was assumed to contain the catalytic domain. However, an isolated C-terminal

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

8761
fragment, S2b+S3 from ColH, bound to lyophilized insoluble collagen to the same extent as full-length ColH (6). This observation lead us to conclude that this C-terminal fragment contains the collagen-binding domain (CBD). Considering the complex macromolecular structure of collagen fibrils, however, the structure of the substrate molecule/moiety recognized by CBD is not fully understood. It could be even a disparate molecule closely associated with collagen fibrils.

For applications, we have shown that the CBD can be used to anchor peptide-signaling molecules to collagen-containing tissues (8). Fusion proteins between growth factors and S2b+S3 bound not only to insoluble collagen in vitro but to connective tissue in vivo. A fusion protein carrying basic fibroblast growth factor joined to the N terminus of S2b+S3 remained in type I collagen-rich tissue when injected subcutaneously into nude mice and exerted a growth-promoting effect at the sites of injection much longer than basic fibroblast growth factor alone. To apply the findings more widely to drug delivery systems, it is essential to diminish the antigenicity of the CBD. One straightforward approach is to synthesize mimics of the CBD by rational drug design, based on its molecular structure. To achieve this objective, detailed information is needed, i.e. the minimal region for collagen binding, the minimal receptor moiety, and the structure of the bimolecular complex. Furthermore, elucidation of the molecular mechanism of collagen binding provides insights to understand the biosynthesis and degradation of extracellular matrix proteins. Thus, we carried out binding experiments using various C-terminal fragments of class I collagenase (ColG) and subcomponents or mimics of human type I collagen.

**EXPERIMENTAL PROCEDURES**

**DNA Manipulations and Sequencing**—Restriction endonucleases were purchased from Takara Shuzo Co. (Kyoto, Japan), Toyobo (Osaka, Japan), and New England BioLabs (Beverly, MA). The DNA ligation kit was a product of Takara Shuzo. All recombinant DNA procedures were carried out as described by Sambrook et al. (9). All constructs were sequenced to confirm the reading frame on an automated nucleotide sequencer (model ABI PRISM 377, PerkinElmer Life Sciences, Foster City, CA). A Thermo Sequenase fluorescence-labeled primer cycle sequencing kit with 7-deaza-dGTP (Amersham Pharmacia Biotech, Uppsala, Sweden) were used for sequencing the GST fusion constructs.

**Purification of GST Fusion Proteins**—The segmental boundaries of *C. histolyticum* class I collagenase (ColG) were determined from an alignment of the amino acid sequences of three clostridial collagenases (4). Nucleotide fragments encoding the C-terminal segments were obtained by PCR using either pCHG5 or pCHG51 (4) plasmid DNA as the template and the following pairs of synthetic primers: 5'-CCCGGAC-GATATATGGGATATAG-3' and 5'-CTCGAGTTAATCCTCTTCTT-TATTCTA-3' for segment 2 (S2, Thr677-Asp772); and 5'-CCGGGAC-AACACTTACTAA-3' (S3a-5' primer) and 5'-CTCGAGCTATTTC-3' for segment 3a (S3a, Thr773-Lys896); and 5'-CCGGGACGAGAATGTTGAAA-3' and 5'-CTCGAGCTATTTC-3' for segment 3b (S3b, Asp897-Lys1008). Two hundred picomoles of the cleavage products was separated for segment 3 (S3a, Thr773-Lys1008) was prepared by PCR using *C. histolyticum* genomic DNA and a pair of oligonucleotide primers: 5'-CCCCCAGCTTTGCTAGTACGG-3' and 5'-GGATCTTATTTATCCTTACT-3'. The amplified fragment was inserted into pT7Blue T-vector DNA. A recombinant plasmid carrying an insert with the correct nucleotide sequence was chosen. To express the collagenase, a pET-11a vector (Novagen, Madison, WI) was employed. A unique EcoRI site in the vector was deleted by the fill-in reaction using Klenow enzyme prior to use (pET-11aARD). The colG fragment was inserted into the vector between NheI and BamHI sites, and E. coli BL21-CodonPlus(DE3)-RIL (Stratagene, La Jolla, CA) was transformed with the resultant plasmid (pEG101). The cells were grown in 2 liter of LB broth supplemented with 100 μg/ml ampicillin and 30 μg/ml chloramphenicol to an optical density at 600 nm of 0.7. Isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 1 mM, and the cells were grown for more 2 hours. Cells were harvested by centrifugation at 7000 × g for 15 min at 4 °C, suspended in 40 ml of CB buffer, and disrupted two times in a French pressure cell at 10,000 p.s.i. Ammonium sulfate was added to the cleared lysate to 60% saturation, and the precipitate was dissolved in 7 ml of 50 mM Tris-HCl buffer (pH 7.5). The sample (approximately 8.5 ml) was applied to a size exclusion chromatography column (Superose 12 10/30, Amersham Pharmacia Biotech) pre-equilibrated with 10 mM Na-HEPES (pH 7.4), 150 mM NaCl, 1 mM CaCl2.

A glutamate residue in the putative catalytic center, Ser472 (HEY477), was mutated to a glutamine by site-directed mutagenesis. PCR was carried using pCHG4 (4) plasmid DNA, and two mutagenic primers;
5'-GAAGAATTTGTAGCATACATATCATCATTCTTAC3' and 5'-TTGTAATATGTAATATGTTACAAAACTTTC3'. After the mutation was confirmed by nucleotide sequencing, the insert EcoRI fragment (0.80 kilobase) was used to replace the corresponding fragment in pEG101. The mutant collagenase, rColG(E414Q), was produced purified in the same manner as the wild type.

Specific activity against insoluble collagen was measured as described previously (7). One unit of enzyme activity equals 1 μmol of l-leucine equivalents liberated from collagen in 5 h at 37 °C under the specified conditions. The N-terminal amino acid sequence up to the 20th residue was determined as described above. The molecular mass was determined by MALDI/TOF MS.

Collagen Binding Assay—Binding of the CBD to type I collagen was determined as follows: 10 μg of collagen (acid-soluble collagen, Sigma type I) was incubated in 100 μl of protein mixture at room temperature for 30 min. The basal mixture (100 μl) contained 5 μg of bovine serum albumin (Sigma), 5 μg of chicken ovalbumin (Sigma), 5 μg of porcine myoglobin ( Sigma), 5.42 μg (0.2 ml) of S5a3b (ColG), and 2.76 μg (0.2 nmol) of S3b (ColG) in CB buffer in CB buffer. The water phase was examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). When necessary, the salt concentration in CB buffer was modified (0, 0.1, 0.5, or 1.0 M). To alter the pH of the buffer, Tris-HCl was replaced with 0.1 M BisTris-HCl (pH 6.0 or 7.0) or Tris-HCl (pH 7.0, 8.0, or 9.0). The effect of divalent cations was examined by replacing CaCl2 with 0.1M EGTA plus 0.1M CaCl2.

Semiquantitative measurements of binding were carried out as described previously (6), where the amount of the unbound fraction protein was quantitated by GST assay using 1-chlore-2,4-dinitrobenzene as the substrate.

Gel-filtration Studies—Size exclusion chromatography experiments were performed at room temperature on a high performance liquid chromatography (HPLC) system equipped with a Superdex 75 column (1 × 30 cm, Amersham Pharmacia Biotech) at a flow rate of 0.5 ml/min. S3b (20 μg) was dissolved in 50 μl of CB buffer or the same buffer containing 1 mM EGTA instead of Ca2+, and chromatographed. The following proteins were used as molecular mass standards: γ-globulins, 158 kDa; bovine serum albumin (BSA), 68 kDa; chicken ovalbumin, 44 kDa; and cytochrome c, 12.5 kDa. The measurements were carried out in triplicate.

Measurement of Hydrodynamic Radius—The hydrodynamic radius of S3b was measured at 20 °C using a dynamic laser light-scattering instrument (model DynaPro 99, Protein Solutions, Charlottesville, VA). S3b was dissolved in CB buffer or CB buffer containing 1 mM EGTA instead of Ca2+ at a concentration of 5.66 mg/ml. The samples were filtered through a 0.22-μm filter before analysis.

Binding to Atelocollagen—Bovine skin atelocollagen was purchased from Koken (Tokyo, Japan). To remove the water-soluble fraction, atelocollagen was treated with water-soluble carbodiimide prior to use. The peptide was covalently immobilized to a biosensor chip by the supplier’s standard procedure, and purified by reversed phase chromatography on a Shimadzu HPLC system using a Tosoh ODS-80TM column. The purity of the final peptide fractions was greater than 95% estimated by the same HPLC system. Identity of the peptides was confirmed by amino acid sequencing using an Applied Biosystems 492 protein sequencer. Collagenous peptides, (PPO)10, (PPO)10, and (PPO)10 were purchased from Peptide Institute, Inc. (Osaka, Japan). The synthesis and purification of the other collagenous peptides were described elsewhere (12).

Coupling of the Peptides to Sepharose Beads—Peptides were dried under vacuum before weighing to prepare solutions in 0.2 M NaHCO3, 1 mM NaCl, pH 8.3 (for peptide solution). The protein was dissolved in water at 4 °C for 48 h prior to use. To prepare 150 μl of the peptide-coupled gel beads, 43 mg of CNBr-activated Sepharose 4B (Amersham Pharmacia Biotech) powder was suspended in 1 ml of 1 mM HCl and incubated for 5 min. The suspension was centrifuged at 500 × g for 1 min to remove the water phase. This washing was repeated a total of three times. The gel was washed once with 1 ml of distilled water in the same way, and was suspended in 500 μl of the peptide solution. The suspension was incubated on a rotary shaker for 2 h. The peptide solution was removed by centrifugation, and the gel was incubated in 1 ml of 0.1 M Tris-HCl (pH 8.0) over night on a rotary shaker to block any remaining active groups. The gel was sequentially washed with 1 ml each of 0.1 M NaHCO3, 0.5 M NaCl (pH 8.3), Coupling buffer, 0.1 M sodium acetate, 0.5 M NaCl (pH 4.5); and Coupling buffer. CB buffer was added to the gel to obtain the final suspension (1 ml). All the above procedures were carried out at room temperature (25 ± 1 °C), and the suspension was stored at 4 °C. Gelatin-Sepharose 4B was purchased from Amersham Pharmacia Biotech.

Coupling of a Disaccharide to Sepharose Beads—Epoxide-activated Sepharose 6B powder (Amersham Pharmacia Biotech) was rehydrated and washed with distilled water as described by the supplier. 2-O-α-D-glucopyranosyl-(1→3)-O-α-D-glucopyranosyl-(1→2)myoglobin (Sigma), 5.42 mg (0.2 m mo l) was obtained from Toronto Research Chemicals (North York, ON, Canada) and dissolved in 150 μl of 0.1 M NaOH prior to use. The swollen gel beads (75 μl) were mixed with all the carbohydrate solution. After the suspension was incubated at 37 °C for 16 h with gentle mixing, the beads were washed with 300 μl of 0.1 M NaOH and subsequently with 400 μl of 1 M ethanolumine-HCl (pH 8.0). The beads were suspended in 400 μl of 1 M NaOH (pH 8.0) for 10 h to block any remaining active groups. The gel was sequentially washed alternatively with 0.4 ml each of 0.1 M sodium acetate, 0.5 M NaCl (pH 4.5) and 0.1 M Tris-HCl, 0.5 M NaCl (pH 8.0) three times. The beads were washed with 0.4 ml of CB buffer three times and used for the binding assay.

Examination of the Binding to Substrate-coupled Sepharose—A suspension containing 50 μl of Sepharose beads coupled with peptides, or disaccharide was added to an 0.5-ml microcentrifuge tube. The gel was centrifuged at 500 × g for 1 min and washed once with 0.4 ml of CB buffer. Then it was suspended in 100 μl of the above protein mixture. The suspension was incubated at 25 °C for 30 min on a rotary shaker, and the water phase was collected by centrifugation at 500 × g for 1 min. The sample (30 μl) was mixed with 4 × SDS reducing dye (10 μl), boiled for 3 min, and analyzed by SDS-PAGE using 15% SDS-polyacrylamide gels (13). To examine the bound protein, the gel beads were twice washed with 0.4 ml of CB buffer and suspended in 50 μl of 2× SDS reducing dye. The suspension was boiled for 3 min to solubilize the bound protein(s). After centrifugation, the water phase (40 μl) was analyzed by SDS-PAGE.

Quantitative Analysis of the Binding to a Peptide Substrate—Binding of S3b, S3a3b, and rColG(E414Q) to a collagenous peptide was measured by surface plasmon resonance using a BIACORE apparatus (Biacore, Uppsala, Sweden). A peptide, G(POG)8, was dissolved in 10 mM sodium acetate (pH 6.0), and the solution was incubated at 4 °C for 48 h prior to use. The peptide was covalently immobilized to a biosensor chip CM5 through the α-amino group following the supplier’s protocol. To minimize interaction between multiple ligands and a single protein molecule ("bridging"), the peptide was used at a low concentration (0.1 mM). To achieve a low ligand density (4 RU/5), resonance was measured in 10 mM Na-HEPES (pH 7.4), 150 mM NaCl, 1 mM CaCl2, 0.005% Tween 20 with a flow rate of 0.1 M HCl. Values for apparent dissociation constants, Kd(app), were calculated from equilibrium binding data at eight protein concentrations (100 nM to 300 μM) for S3b, and at eight protein concentrations (30 nM to 100 μM) for S3a3b. Data were directly fit using the following equation by least square method,

\[ CRU = K_a \times CRU_{\text{max}} \times \frac{[\text{protein}]}{1 + K_a \times [\text{protein}]} + \text{background} \] (Eq. 1)

where CRU is the response at equilibrium considered for bulk refractive index errors using a sham-coupled flow cell blocked with ethanolamine, [protein] is the analyte concentration, and K_a is the association constant. Kd(app) is the reciprocal of K_a.
Collagen-binding Domain of Clostridial Collagenase

RESULTS

Purification of the C-terminal Fragments—Various C-terminal segments (S2, S3a, S3b, and S3a3b) of *C. histolyticum* class I collagenase (ColG) were expressed as GST fusion proteins (Fig. 1B). The C-terminal fragments were cleaved off from the N-terminal carrier by thrombin, and their N-terminal amino acid sequences were determined. For each fragment, the initial 20 amino acid residues coincided with the expected sequence. Although binding of S3b and S3a3b increased at 5 M urea, the latter proteins all gave similar curves (Fig. 2A). Because all the proteins seemed to be unfolded in the presence of 9 M urea, CD spectra were also recorded in this condition (Fig. 2, B–E). In the absence of urea, the spectrum was characteristic for the respective protein, but the uniqueness was lost in the presence of urea. The molar ellipticity profile of S3a3b (Fig. 2E) correlated well to the averaged pattern for S3a (Fig. 2C) and S3b (Fig. 2D). Binding of monomeric (S3b) and tandem (S3a3b) S3 to type I collagen were measured in various buffers. Neither addition of sodium chloride up to 1 M (Fig. 3A) or alteration of the pH within a range between 6.0 and 9.0 (Fig. 3B) affected binding significantly. However, addition of 1 mM EGTA instead of Ca2+ inhibited binding significantly (Fig. 3C). The effect of the chelator was reversed by the addition of a 1 mM excess of Ca2+. The apparent molecular mass of S3b was measured by size exclusion chromatography. The value was 11.4 ± 2.12 kDa (mean ± S.D. of three experiments) in the presence of 1 mM Ca2+ and 16.0 ± 0.08 kDa in the presence of 1 mM EGTA. The hydrodynamic radius (R_H) of S3b was 1.93 nm in the presence of 1 mM Ca2+ as measured by dynamic laser light scattering. This size corresponds to about 14.2 kDa, based on the standard size/mass relationship for globular proteins. When Ca2+ was replaced with 1 mM EGTA, the R_H value and estimated molecular mass were 2.16 nm and 18.7 kDa, respectively.

Binding to Tropocollagen Segments—When collagen is treated with pepsin, the nonhelical regions (telopeptides) are digested, giving the triple-helical region named atelocollagen. Binding to commercially available atelocollagen was measured to determine which region is recognized by S3. The binding assay was carried out in the same condition as with collagen using the basal protein mixture. Although binding of S3b and S3a3b was observed, significant amounts of atelocollagen α-chains were observed in the water phase (data not shown). Thus, we pretreated atelocollagen with water-soluble carbodiimide to cross-link the α-chains to prevent their release into the water phase. This substrate showed binding to S3b and S3a3b, and negligible amounts of α-chains were observed in the water phase (Fig. 4A).

Various subcomponents of human type I collagen were coupled to Sepharose beads, and the binding of S3b and S3a3b to these gels was examined (Fig. 4, B and C). When the proteins were incubated with the gel beads coupled with a synthetic peptide corresponding to one of the four telopeptides present in human type I collagen, S3b and S3a3b remained in the water.

**TABLE I**

| Fusion protein (fragment) | Collagen binding |
|--------------------------|-----------------|
| GST-S2 (Thr<sup>677</sup>-Asp<sup>772</sup>) | 4.1 ± 2.6 |
| GST-S3a (Thr<sup>777</sup>-Lys<sup>896</sup>) | 79.5 ± 0.93 |
| GST-S3b (Asp<sup>894</sup>-Lys<sup>1009</sup>) | 97.2 ± 0.29 |
| GST-S3a3b (Thr<sup>777</sup>-Lys<sup>1008</sup>) | 99.7 ± 0.26 |
| GST | 1.1 ± 1.8 |

Collagen-binding activities of fusion proteins carrying various C-terminal segments of class I collagenase

Five milligrams of washed insoluble collagen were mixed with 100 pmol of fusion protein in 60 μl of CB buffer and incubated at room temperature for 30 min. The collagen-binding activity of each fusion protein was determined by calculating the difference in the GST activity in supernatants incubated with and without collagen. The values represent the average of triplicate trials plus deviations.
When the proteins were incubated with beads coupled with a collagenous peptide, (prolyl-hydroxyprolyl-glycine)_{10} \[(POG)_{10}\], S3b and S3a3b disappeared from the water phase (lane 6F). The bound fraction was recovered by boiling the affinity gel in SDS sample buffer (lane 6B). Although a small amount of BSA was present in this fraction, a similar amount of the protein was also visible in the fraction bound to Tris-blocked Sepharose beads used as a control (lane 5B). A disaccharide, 2-O-a-glucopyranosyl-D-galactopyranose, is present in the triple-helical region at the hydroxylysine residues. We coupled this disaccharide to the Sepharose beads, and S3 did not bind (Fig. 4C).

To examine if S3 specifically binds to the collagenous peptide, cleared lysates were prepared from uninduced E. coli cells, which express GST-S3a3b or GST-S3b. Each lysate containing 1 mg of protein in CB buffer was incubated at 25 ± 1 °C for 30 min with 50 μl of (POG)_{10}-Sepharose beads, and the bound fractions were examined by SDS-PAGE. In each fraction, one distinct band was present that seemed to correspond to the respective GST fusion protein (Fig. 4D). Because each band showed a strong signal by Western blotting and immunoanalysis using affinity-purified anti-S3b rabbit antibody (data not shown), it was concluded that the bands were GST-S3a3b and GST-S3b, respectively.

Finally, the binding of S3a3b and S3b to the collagenous peptide, (POG)_{10}, were examined in the same set of buffers as used for the collagen binding assay shown in Fig. 3, which was carried out in the buffers with varying salt concentrations, varying pH values, and varying Ca^{2+} concentrations. Their binding to this peptide was similar to those against type I collagen (Fig. 5).

**Binding to Various Peptides**—Various substrates related to (POG)_{10} were immobilized on Sepharose beads, and their binding activity to S3a3b and S3b was examined. First, the number of repeats of the POG triplets was altered (Fig. 6A). The proteins bound to (POG)_{n}-Sepharose to a similar extent as to (POG)_{10}-Sepharose, but they bound poorly to (POG)_{5}-Sepharose. The better binding to the longer peptides was also true for prolyl-prolyl-glycine (PPG). Binding reached a maximum level when the number of repeats was 8 (Fig. 6B). The proteins did not bind to Sepharose beads coupled with any of the following 24-mer peptides; (PG)_{12}, (PPPG)_{6}, all-D-(PPG)_{8}, (APG)_{8}, or (PAG)_{8} (Fig. 6C, lanes 1–5) or to gelatin-Sepharose beads (Fig. 6C, lane 6).

**Estimation of Dissociation Constants**—The time course of binding between S3b and a collagenous peptide was determined by surface plasmon resonance (Fig. 7). The reaction was fast enough to reach equilibrium within about 30 s. The dissociation constant \(K_D\) and corrected maximum response cRU_{max} were calculated by plotting the equilibrium values. \(K_D\) and cRU_{max} values for S3b were 5.47 × 10^{-5} M and 2.34 × 10^{3} RU, respectively. Binding of S3a3b resulted in similar sensorgram profiles, although the response was larger than that of S3b at the same concentration (data not shown). \(K_D\) and cRU_{max} values for S3a3b were 6.30 × 10^{-5} M and 3.41 × 10^{3} RU, respectively.
Collagen-binding Domain of Clostridial Collagenase

Binding of the Recombinant ColG to the Peptide Substrate—Full-length ColG was purified from a recombinant E. coli cell lysate (Fig. 8A, lane 1). Because the synthetic peptides recognized by the CBD were expected to be hydrolyzed by this full-length enzyme, we also prepared an enzyme with a conservative mutation in the putative catalytic center located in the catalytic domain (S1) (Fig. 8A, lane 2). N-terminal amino acid sequences of the wild type and the mutated enzyme were identical and matched the deduced sequence with the exception of the removal of the initial methionine residue. Two extra residues, Ala-Ser, were present at their N termini, compared with mature ColG. The molecular masses measured by MALDI-TOF MS were 114,208 and 114,181 Da for the wild type and the mutated enzyme, respectively. The observed values coincided well with the respective calculated values, 114,120.8 and 114,119.8 Da. The specific activity of recombinant collagenase against insoluble collagen was 881 ± 6145 units/mg of protein, comparable to the value obtained for the authentic enzyme (826 ± 42 unit/mg of protein). The mutated enzyme, rColG(E414Q), had a low level of hydrolytic activity (34.7 ± 9.6 unit/mg of protein) against insoluble collagen. The latter enzyme bound to (POG)10-Sepharose beads but not to Tris-blocked beads (Fig. 8B). Positive binding of the mutated enzyme to G(POG)8 peptide was also observed by surface plasmon resonance. In this binding profile (Fig. 8C), however, a slow binding reaction was present in addition to the fast binding reaction. Due to this complexity, we could not estimate the $K_D$ value for binding between full-length enzyme and the collagenous peptide.

DISCUSSION

We previously reported that a C-terminal fragment (S2b+S3) of C. histolyticum class II collagenase (ColH) binds to...
insoluble type I collagen (6). Recently, we cloned the structural gene for the class I enzyme (ColG) from the same strain, and found that S3 was tandemly repeated at its C terminus (4). Comparison of the predicted amino acid sequences of these divergent collagenases allowed us to precisely predict the boundary between each segment. Based on these results, we produced various C-terminal segments of ColG as GST fusion proteins. The results of semi-quantitative binding experiments indicated that the S3 monomer is a functional collagen-binding domain (CBD). Tandem S3 (S3a3b) bound to insoluble collagen more efficiently than S3b, suggesting cooperative binding of the two functional domains to this macromolecular substrate.

The fusion proteins were cleaved by thrombin, and the C-terminal fragments were purified by affinity chromatography. Collagen-binding domain peptides, S3a, S3b, and S3a3b, all gave a β-sheet CD spectrum. By using a k2d server available on the Web (14, 15), the percentages of secondary structure were estimated as follows: α-helix 5%, β-sheet 47%, random 48% for S3a; α-helix 9%, β-sheet 44%, random 48% for S3b; α-helix 5%, β-sheet 47%, random 48% for S3a3b. These percentages agree with a secondary structure prediction carried out by using the PredictProtein (available on the Web) (16). The CD spectrum of the tandem domain (S3a3b) resembles the average of the S3a and S3b spectra, which suggests that the secondary structure of each monomer is conserved in the tandem domain. When treated with 9 M urea, all of these spectra became characteristic of non-native proteins, suggesting that they are unfolded in this solvent.

Neither addition of salt up to 1 M or alteration of the pH between 6 and 9 affected binding significantly, suggesting that ionic interactions do not play a major role in the interaction between the CBD and collagen. A plasma glycoprotein, von Willebrand factor, mediates platelet adhesion by binding both to collagen in a damaged blood vessel and to a glycoprotein on the platelet membrane. The A3 domain interacts with collagen in the perivascular tissue of damaged cell walls. Based on a crystal structure analysis, Huizinga et al. (17) suggested that binding is achieved primarily through ionic interactions between negative charged residues on the domain and positively charged residues in collagen. Elucidation of the precise mode of CBD binding by structural methods is essential to understand its mechanism.

The molecular mass of S3b estimated by size exclusion chromatography and the value estimated by laser light scattering were the same in the presence and absence of Ca\textsuperscript{2+}. It was suggested that S3b is a monomer in both conditions. The reduction of the Stokes radius and hydrodynamic radius in the presence of Ca\textsuperscript{2+} was consistent with observations for various Ca\textsuperscript{2+}-binding proteins, e.g. troponin C or calmodulin (18–20). Because ionic interactions are not likely to play a major role in the binding, the enhanced binding in the presence of Ca\textsuperscript{2+} may be due to a conformational change through exposure of a hy-
The efficient binding of the CBD to collagen in the presence of Ca^{2+} seems to be one of the reasons why collagenase requires the cation for full activity (23).

C-terminal fragments of class I collagenase (S3b and S3a3b) bound to various types of collagen. Because their amino acid sequences are more diverse in the noncollagenous region (telopeptides) than in the collagenous region, it was hypothesized that the binding site(s) is(are) present in the collagenous region. This working hypothesis was supported by the facts that CBD binds to atelocollagen but not to synthetic telopeptides. In type I collagen, the collagenous region consists of three polypeptide chains each of which contains 338 Gly-X-Y peptide units. Although various residues are present at the X and Y positions, we initially used the representative oligopeptide (POG)_{10} to mimic collagen, because this triplet is the most frequent in type I collagen. Alternatively, the CBD might recognize a disaccharide, 2-O-a-glucopyranosyl-D-galactopyranosyl, group attached to hydroxylysine residues in the polypeptides. Thus, we also examined binding to immobilized Glc-Gal.

Among various triplets, Gly-Pro-Hyp is known to contribute the most to the formation of a triple helix. The thermal stability of triple helices has been reported for (POG)_{n} peptides of various length. The melting temperatures ($T_{m}$) are 2 °C when $n = 5$, 44.5 °C when $n = 8$ (24), and 61 °C when $n = 10$ (25). Because we carried out all the binding experiments at 25 ± 1 °C, it was assumed that most of the (POG)_{5} molecules are in a random conformation, whereas most of the (POG)_{8} and (POG)_{10} molecules are triple-helical. Binding was observed when the peptides are supposed to be triple-helical. To see if these results are due to steric hindrance, we immobilized the (POG)_{5} and (POG)_{10} peptides on NHS-Sepharose beads, which have a 9-atom spacer arm, and examined binding of CBD and obtained essentially the same results (data not shown). The length of the collagenous peptides estimated from x-ray crystallography are longer than the measured diameter of S3b (3.86 nm), also suggesting that CBD recognizes the conformation

---

**Fig. 6. Binding of S3 to various peptide-coupled beads.** Two hundred picomoles each of S3a3b and S3b, and 5 µg each of BSA, ovalbumin, and myoglobin in 100 µl were incubated at 25 ± 1 °C with 50 µl of various peptide-coupled Sepharose beads. For each binding experiment, the following three samples are shown: C, negative control incubated in the absence of the beads; F, free fraction in the water phase; B, fraction bound to the beads. A, (POG)$_n$, B, (PPG)$_n$, C, lanes 1, (PG)$_{12}$; lanes 2, (PPPG)$_{6}$; lanes 3, all-D-(PPG)$_8$; lanes 4, (APG)$_8$; lanes 5, (PAG)$_8$; lanes 6, gelatin-Sepharose.

---

2 T. Toyoshima, O. Matsushita, J. Minami, N. Nishi, A. Okabe, and F. Itano, manuscript submitted.
generated only when the number of triplet reiteration is large enough (26). Another series of oligopeptides, (PPG)_n, were known to have a triple-helical conformation when n is sufficiently large (27). When n increased, the binding of S3 increased, and it seemed to reach a maximum when n = 8. To exclude the possibility that binding is due to nonspecific hydrophobic interactions, binding to various peptides that have similar sequences of the same length (24 residues) without the proper triple-helical conformation was examined. Two peptides, (Pro-Gly)_{12} and (Pro-Pro-Pro-Gly)_6, lack triplets and, hence, the triple-helical conformation despite their similar hydrophobicity. The enantiomer of the original peptide, (Pro-Pro-Gly)_6, should have a triple-helical conformation but in the reverse orientation. Two similar peptides, (APG)_{8}, and (PAG)_{8} were plotted against protein concentration, and fitted to the equation described under “Experimental Procedures” by the least square method to calculate K_d and RU_{max} values. Dissociation constant values measured by surface plasmon resonance do not correspond to values measured in solution, being affected by chemical modification and/or immobilization by one of the reactants.

Fig. 7. Estimation of K_d value by surface plasmon resonance. A collagenous peptide, G(POG)_8, was immobilized on the surface of thin gold membrane at a low density. S3b solution with varying concentration (100 nM to 300 μM) was injected at a time of 100 s, and the time course of the binding was observed. Inset, equilibrium response values were plotted against protein concentration, and fitted to the equation described under “Experimental Procedures” by the least square method to calculate K_d and RU_{max} values. Dissociation constant values measured by surface plasmon resonance do not correspond to values measured in solution, being affected by chemical modification and/or immobilization by one of the reactants.

Fig. 8. Binding of a recombinant ColG to collagenous peptides. A, purified recombinant ColG proteins (2 μg each). M, molecular mass markers; lane 1, wild type; lane 2, mutated collagenase, rColG(E414Q), in which a glutamate residue in the putative catalytic center (Glu^{414}) was replaced with a glutamine residue. The samples were subjected to SDS-PAGE on a 12.5% polyacrylamide gel. Numbers on the left are molecular masses (in kDa) of the markers. B, binding of rColG(E414Q) to ligand-coupled Sepharose beads. M, molecular mass markers; C, control incubated in the absence of the beads; F, free fraction remained in the water phase; B, fraction bound to the beads. Lanes 1, Tris-blocked; lanes 2, (POG)_8. C, sensorgram of the binding of rColG(E414Q) to a collagenous peptide, G(POG)_8. The concentration of the analyte ranges from 1 nM to 30 μM. Slow binding reaction overlapped to the fast binding reaction.

fast to estimate association/dissociation rate constants (k_{on}/ k_{off}) by this method. This fast reaction might contribute to efficient recycling of the enzyme. Tandem S3 (S3a3b) did not show cooperative binding when the immobilized ligand density was low. “Bridging” might be necessary for cooperative binding by two independent domains. Alternatively, two domains might form a single continuous binding surface. Analysis using a sensor chip on which a longer (25 aa) peptide is immobilized should give more insight.

A recombinant wild type collagenase, rColG, showed full activity against insoluble collagen despite the two extra residues at its N terminus. This suggested that the structures necessary for substrate binding and hydrolysis are conserved in the recombinant enzyme. However, it is likely that the collagenous peptide ligands are hydrolyzed by its catalytic activity during the binding assay. Previously, we have reported a mutational analysis on the catalytic center of the C. histolyticum class II collagenase (ColH) (7). As for this enzyme, conservative mutations in the putative catalytic center, HEXXH, significantly affected its catalytic activity without affecting its binding activity (6, 7). Thus, it seems unlikely that the conservative mutation, E414Q, alters the global conformation of the homologous enzyme. As expected, the mutated full-length enzyme bound to a collagenous peptide immobilized on the beads. How-
ever, surface plasmon resonance analysis revealed that the enzyme has a slow binding mechanism in addition to the fast one governed by CBD. Because S2 did not bind significantly to collagen, gelatin, (POG)$_n$, or (POG)$_{10}$ (data not shown), it was suspected that the slow interaction occurs between the catalytic domain (S1) and the substrate.

A simple model peptide, (POG)$_n$ has an amino acid sequence like the collagenous region of various collagen molecules. We showed that the peptide possesses the structural information that allows specific binding of the CBD of a bacterial collagenase. Because actual collagen molecules have diversity in the amino acid residues at the X and Y positions, the CBD has a binding spectrum wide enough to allow this diversity. This binding spectrum would contribute to the wide substrate specificity of the bacterial enzymes, which is different from the cleavage of specific collagens by eukaryotic matrix metalloproteases (29). In the collagen biosynthetic process, a possible chaperone protein, HSP47, transiently interacts with procollagen molecules in the endoplasmic reticulum to regulate their folding and/or modification (30, 31). Koide et al. (12) showed that HSP47 interacts with (PPG)$_n$ peptides but not with (POG)$_n$ peptides. They also showed that prolyl 4-hydroxylase has the same substrate preference based on prolyl 4-hydroxylation at the Y position. Our results showed that the CBD of C. histolyticum class I collagenase lacks such selectivity. Elucidation of the underlying molecular mechanism would give insights into the evolution and structure-function relationships of various collagen-binding proteins. From a practical view point, the binding of CBD to atelocollagen is beneficial; i.e. it extends the clinical applications of the drug delivery system we proposed previously (8), because atelocollagen is one of the more widely used biomaterials in clinical fields. To apply the system more generally, however, it is necessary to replace the CBD with low molecular mimics to reduce or eliminate its antigenicity. To achieve these objectives we have begun attempts to make cocrystals between the CBD and collagenous peptides.

Acknowledgments—We thank Kayoko Yamashita and Yuki Taniguchi for their technical assistance. We also thank David B. Wilson (Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY) for invaluable discussion and assistance in preparing the manuscript.

REFERENCES

1. Nagata, K. (1996) Trends Biochem. Sci. 21, 22–26
2. Ramshaw, J. A., Shah, N. K., and Brodsky, B. (1998) J. Struct. Biol. 122, 86–91
3. Worthington, C. C. (1988) Worthington Enzyme Manual: Collagenase, Worthington Biochemical Co., Freehold, NJ
4. Matsushita, O., Jung, C.-M., Katayama, S., Minami, J., Takahashi, Y., and Okabe, A. (1999) J. Bacteriol. 181, 923–933
5. Yoshihara, K., Matsushita, O., Minami, J., and Okabe, A. (1994) J. Bacteriol. 176, 6489–6496
6. Matsushita, O., Jung, C.-M., Minami, J., Katayama, S., Nishi, N., and Okabe, A. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 7018–7023
7. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
8. Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4350–4354
9. Matsushita, O., Yoshihara, K., Katayama, S.-I., Minami, J., and Okabe, A. (1994) J. Bacteriol. 176, 149–156
10. Koide, T., Asada, S., and Nagata, K. (1999) J. Biol. Chem. 274, 34523–34526
11. Koide, T., Asada, S., and Nagata, K. (1999) J. Biol. Chem. 274, 34523–34526
12. Koide, T., Asada, S., and Nagata, K. (1999) J. Biol. Chem. 274, 34523–34526
13. Laemmli, U. K. (1970) Nature 227, 680–685
14. Andrade, M. A., Chacin, P., Merelo, J. J., and Morán, F. (1995) Protein Eng. 6, 383–390
15. Merelo, J. J., Andrade, M. A., Prieto, A., and Morán, F. (1994) Neurocomputing 6, 445–454
16. Rost, B. (1996) Methods Enzymol. 266, 525–539
17. Huizinga, E. G., Martijn van der Plas, R., Kroon, J., Sixma, J. J., and Gros, P. (1997) Structure 5, 1147–1156
18. Byers, D. M., and Kay, C. M. (1982) Biochemistry 21, 229–233
19. Klee, C. B., Crouch, T. H., and Richman, P. G. (1980) Annu. Rev. Biochem. 49, 489–515
20. McDonald, J. R., Walsh, M. P., McCubbin, W. D., Oikawa, K., and Kay, C. M. (1985) Biochem. J. 232, 569–575
21. LaPorte, D. C., Wierman, B. M., and Storm, D. R. (1980) Biochemistry 19, 3814–3819
22. Tanaka, T., and Hidaka, H. (1980) J. Biol. Chem. 255, 11078–11080
23. Bond, M. D., and Van Wart, H. E. (1984) Biochemistry 23, 3092–3099
24. Shah, N. K., Ramshaw, J. A., Kirkpatrick, A., Shah, C., and Brodsky, B. (1996) Biochemistry 35, 10262–10268
25. Sakakibara, S., Inouye, K., Shudo, K., Kishida, Y., Kohayashi, Y., and Prockop, D. J. (1973) Biochem. Biophys. Acts 303, 198–202
26. Bell, J., Eatun, M., Brodsky, B., and Berman, H. M. (1994) Science 266, 75–81
27. Sakakibara, S., Kishida, Y., Kikuchi, Y., Sakai, R., and Kakihara, K. (1998) Bull. Chem. Soc. Jpn. 41, 1273–1273
28. Burjanadze, T. V. (1979) Biopolymers 18, 931–938
29. Ott, J., Gabriel, D., Murphy, G., Knauper, V., Tominaga, Y., Nagase, H., Kroger, M., Tschesche, H., Bode, W., and Moroder, L. (2000) Chem. Biol. 7, 119–132
30. Nakai, A., Satoh, M., Hirayoshi, K., and Nagata, K. (1992) J. Cell Biol. 117, 903–914
31. Satoh, M., Hirayoshi, K., Yokota, S., Hosokawa, N., and Nagata, K. (1996) J. Cell Biol. 133, 469–483