THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 43, Issue of October 22, pp. 30818–30825, 1999
Printed in U.S.A.

Crystal Structure of Chitosanase from *Bacillus circulans* MH-K1 at 1.6-Å Resolution and Its Substrate Recognition Mechanism*

(Received for publication, March 2, 1999, and in revised form, July 21, 1999)

Jun-ichi Saito‡, Akiko Kita‡, Yoshiki Higuchi‡, Yoshiho Nagata§, Akikazu Ando§, and Kunio Miki‡¶

From the ‡Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan and the §Department of Biotechnology, Graduate School of Science and Technology, Chiba University, Matsudo-city, 271-8510, Japan

Chitosanase from *Bacillus circulans* MH-K1 is a 29-kDa extracellular protein composed of 259 amino acids. The crystal structure of chitosanase from *B. circulans* MH-K1 has been determined by multiwavelength anomalous diffraction method and refined to crystallographic R = 19.2% (R_{free} = 23.5%) for the diffraction data at 1.6-Å resolution collected by synchrotron radiation. The enzyme has two globular upper and lower domains, which generate the active site cleft for the substrate binding. The overall molecular folding is similar to chitosanase from *Streptomyces* sp. N174, although there is only 20% identity at the amino acid sequence level between both chitosanases. However, there are three regions in which the topology is remarkably different. In addition, the disulfide bridge between Cys^{50} and Cys^{124} joins the β1 strand and the α7 helix, which is not conserved among other chitosanases. The orientation of two backbone helices, which connect the two domains, is also different and is responsible for the differences in size and shape of the active site cleft in these two chitosanases. This structural difference in the active site cleft is the reason why the enzymes specifically recognize different substrates and catalyze different types of chitosan degradation.

Chitin is an abundant biopolymer of β-(1,4)-linked GlcNAc, which is hydrolyzed by chitinase (EC 3.2.1.14). Chitinase is one of the key enzymes in plant defense systems against fungal infection (1, 2). It is widely distributed in microorganisms and plants, and the primary structures of these molecules have been reported for many sources. They are classified on the basis of amino acid sequence similarities into either family 18 or 19 among the 72 glycosyl hydrolase families (3–5). The three-dimensional structures of 25 of the 72 families of glycosyl hydrolases have already been determined (6). Barley chitinase in family 19 is similar in its three-dimensional structure to the well studied glycosyl hydrolase, lysozyme, but differs in its substrate specificity (7).

On the other hand, a different enzyme, chitosanase (EC 3.2.1.132), which is a member of glycosyl hydrolase family 46, hydrolyzes chitosan, a polymer of GlcN produced by partial (over 60%) or full deacetylation of chitin. Most chitosanases are found in microorganisms (8–11), and a few are found in plants (12–15). The complete amino acid sequences have been reported for procaryotic chitosanases from *Bacillus circulans* MH-K1 (MH-K1 chitosanase)[1] (16), *Streptomyces* sp. N174 (N174 chitosanase) (17), and *Nocardioides* sp. N106 (N106 chitosanase) (18) and for a eucaryotic chitosanase from *Fusarium solani* f. sp. *phaseoli* SUF386 (19). Recently, chitosanase genes were found by genomic analysis of chlorella virus PBCV-1 (20), chlorella virus CVK2 (21), and *Bacillus subtilis* (22). Furthermore, three additional chitosanase genes from *Nocardioides* sp. K-01, *Amycolatopsis* sp. CsO-2, and *Pseudomonas* sp. A-01 have been cloned.[2] No sequence similarities were found between these chitosanases and any previously reported chitinases, although they hydrolyze substrates the chemical structures of which differ only at the acetyl group on the C-2 atom of the sugar. It is important to understand the substrate recognition and catalytic mechanisms of chitosanase and chitinase based on their three-dimensional structures. While the crystal structures of chitinases from *Hordeum vulgare* (23–25), *Hevea brasiliensis* (26) and *Serratia marcescens* (27) have been determined, the structural information of the chitosanases is available only for the crystal structure of N174 chitosanase (28).

Microbial chitosanases are classified into three subclasses based on the specificity of the cleavage positions for partially acetylated chitosan (Table I) (18). Subclass I chitosanases such as N174 chitosanase can split GlcN-GlcN and GlcNAc-GlcN linkages (29), whereas *Bacillus* sp. no. 7-M chitosanase in subclass II can cleave only GlcN-GlcN linkages (30). On the other hand, subclass III chitosanases such as MH-K1 chitosanase can split both GlcN-GlcN and GlcN-GlcNAc linkages (31). This selectivity at the cleavage position of the substrates might be controlled by rigid substrate recognition by chitosanases in these subclasses. To clarify the mechanism of their highly selective enzymatic reaction, comparison of the precise three-dimensional structures of these chitosanases is required. Here, we report the crystal structure at 1.6-Å resolution of MH-K1 chitosanase, which belongs to subclass III. We compared the structure of this chitosanase with that of N174 chitosanase (28)

* This work was supported in part by the Japan Society for the Promotion of Science (JSPS) Research for the Future Program Grant 97L00501 (to K. M.) and by Grant-in-Aid for JSPS fellows 6585 (to J. S.) from the Ministry of Education, Science, Sports and Culture. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1QGI) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ A member of the Sakabe Project of TARA (Tsukuba Advanced Research Alliance), University of Tsukuba. To whom correspondence should be addressed. Tel.: 81-75-753-4029; Fax: 81-75-753-4032; E-mail: miki@kuchem.kyoto-u.ac.jp.

[1] The abbreviations used are: MH-K1 chitosanase, chitosanase from *B. circulans* MH-K1; N174 chitosanase, chitosanase from *Streptomyces* sp. N174; N106 chitosanase, chitosanase from *Nocardioides* sp. N106; MAD, multiwavelength anomalous diffraction.

[2] A. Ando, unpublished results.

and will discuss here the reaction mechanism and specificity of substrate recognition of this enzyme.

EXPERIMENTAL PROCEDURES

Crystallization—The recombinant MH-K1 chitosanase was purified and crystallized as previously reported (32). Crystallization was performed at 20 °C by the sitting drop vapor diffusion method using ammonium sulfate as the precipitant. Crystals belong to the orthorhombic space group $P2_12_12$ with unit-cell dimensions of a = 43.3 Å, b = 128.0 Å, and c = 57.7 Å.

Data Collection and Processing—Intensity data on native crystals were collected using synchrotron radiation with a wavelength of 1.00 Å at BL-6A of Photon Factory (KEK, Tsukuba, Japan). Four full sets were measured around two different rotation axes in which two sets were measured mainly for weak reflections at higher resolution with longer exposure time (20 s/degree), and the other two were measured for strong reflections at lower resolution with shorter exposure time (4 s/degree) to avoid saturation of these reflection intensities. A screenless Weissenberg camera for macromolecular crystals was used with a 0.1-mm aperture collimator and a cylindrical cassette of radius 286.5 mm for high resolution data (1.6 Å) and 429.7 mm for low resolution data (2.5 Å) (33). For the multiwavelength anomalous diffraction (MAD) method, intensity data of the K_2PtCl_4 derivative were collected using one crystal at BL-18B of Photon Factory at four wavelengths (1.0000 Å for remote 1, 1.0721 Å for edge, 1.0722 Å for peak, and 1.0728 Å for remote 2) determined from a fluorescence scan. All data were collected at room temperature. Diffraction intensities were recorded on 200 × 400-mm imaging plates (Fuji Photo Film, Co., Ltd.) and read on a Fuji BAS 2000 scanner (34). The intensity data were processed using the program *DENZO* (35) and merged using the program *SCALEPACK* (Table II).

Phase Determination—Many heavy atom compounds were extensively screened, but only the K_2PtCl_4 derivative was found to be effective. This derivative was prepared by soaking crystals with 0.1 mM K_2PtCl_4 in 81% (w/v) saturated ammonium sulfate in Tris-HCl buffer for 24 h. The refinement of heavy atom parameters and calculation of MAD phases were performed with the program *MLPHARE* in the *CCP4* package (36) (Table II). MAD phases were initially calculated at 4.0-Å resolution, and the obtained electron density map was improved by solvent flattening and histogram mapping using the program *DM* in the *CCP4* package (37, 38).

Model Building and Refinement—The C-α chain was traced on the MAD phased electron density map using the program *TURBO-FRODO* (39). Phases calculated from the partially constructed model were combined with the MAD phases using the programs *SFALL* and *SIGMAA* in the *CCP4* package (38, 40). The map was again improved by solvent flattening, and phases were extended to 3.0-Å resolution. The molecular envelope was re-estimated from the partial backbone model with the aid of the molecular model of N174 chitosanase (Protein Data Bank code 1CHK) for comparison of the whole structure using the program *MAMA* in the *CCP4* package (41). After 80% of the total polyalanine model had been constructed, this model was refined against the merged native data at 5.0–1.8-Å resolution. All refinements were performed using the *X-PLOR* program (42). This refined model was subjected to molecular dynamics and simulated annealing refinement with slow cooling from 3000 to 300 K at 5.0–1.6-Å resolution. After repeated manual rebuilding and fitting the model into $2F_o - F_c$ and $F_o - F_c$ maps, positional and individual atomic B factor refinements were carried out. The stereochemistry of the final model was analyzed using the program *PROCHECK* (43).

TABLE I
Subclasses of microbial chitosanases

○, GlcNAc; ●, GlcN; ⊖ and -●, reducing end. Three subclasses of microbial chitosanases are classified on the basis of the degradation products of partially acetylated chitosan substrate. The degradation products of N174 chitosanase are GlcN in both nonreducing and reducing ends or GlcN in nonreducing end and GlcNAc in reducing end. Therefore, N174 chitosanase can split GlcN-GlcN and GlcNAc-GlcN linkages between subsites D and E of six sugar binding sites, subsites A–F, where F is the reducing end in a similar manner to the subsites in lysozyme. On the other hand, the degradation products of MH-K1 chitosanase are GlcN in both nonreducing and reducing ends or GlcN in reducing end and GlcNAc in nonreducing end. Therefore, MH-K1 chitosanase can split GlcN-GlcN and GlcN-GlcNAc linkages.

	Subclass I (*Streptomyces* sp. N174)	Subclass II (*Bacillus* sp. No. 7-M)	Subclass III (*B. circulans* MH-K1)
Degradation products	●●- ●●●- ●⊖ ●●⊖ ●○⊖ ●●○⊖ ●○○⊖ ●○○○⊖	●●- ●●●- ●●●●- ●○●●●- ●●○●●- ●○●●●●- ●●○●●●-	●●- ●●●- ●○●●- ●●○●●- ○●●●-
Subsite	D-E	D-E	D-E
Cleavage points	GlcN-GlcN GlcNAc-GlcN	GlcN-GlcN	GlcN-GlcN GlcN-GlcNAc

TABLE II
Statistics for data collection and MAD phasing

	Native					K_2PtCl_4 derivative			
	Native 1	Native 2	Native 3	Native 4	Merged native	Remote 1	Edge	Peak	Remote 2
Data collection									
Rotation axis	c^*	a^*	c^*	a^*		a^*	a^*	a^*	a^*
Wavelength (Å)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0721	1.0722	1.0728
Resolution (Å)	1.60	1.60	2.50	2.50	1.60	2.50	3.00	3.00	3.00
Outer shell (Å)	1.63–1.60	1.63–1.60	2.54–2.50	2.54–2.50	1.63–1.60	2.59–2.50	3.11–3.00	3.11–3.00	3.11–3.00
No. of observed reflections	137,671	141,857	66,529	67,615	462,029	41,369	24,165	24,289	24,330
No. of unique reflections	29,340	30,230	8,880	8,913	40,769	16,641	10,196	10,267	10,231
R_{merge}[a]	0.044	0.044	0.078	0.047	0.085	0.041	0.039	0.039	0.038
Outer shell	0.586	0.728	0.277	0.163	0.656	0.198	0.112	0.090	0.092
Completeness (%)	68.4	70.4	76.0	75.4	95.1	76.3	80.6	81.2	80.9
Outer shell (%)	50.5	52.4	52.4	52.4	83.7	55.7	59.7	61.2	59.8
Phase determination									
R_{cullis} (50–4.0 Å)[b]							0.88	0.84	0.84
Phasing power (50–4.0 Å)[c]									
Acentric							0.87	1.02	1.01
Centric							0.71	0.79	0.77

[a] $R_{merge} = \Sigma|I(h) - \langle I(h)\rangle|/\Sigma I(h)$, where $\langle I(h)\rangle$ is the mean value of the reflection h for all measurements of $I(h)$.

[b] $R_{cullis} = \Sigma\|F_{P\lambda i} - F_{P\lambda 0}| - |F_{\lambda i}(\text{calc})\|/\Sigma|F_{P\lambda i} - F_{P\lambda 0}|$, where $F_{P\lambda i}$ is the structure factor of the data at λi, $F_{P\lambda 0}$ is the structure factor of the data collected at 1.0000 Å, and $F\lambda i(\text{calc})$ is the calculated contribution of anomalous scatterer.

[c] Phasing power = $\langle F_h\rangle/E$, where $\langle F_h\rangle$ is the mean of calculated heavy atom structure factor amplitude and E is the root mean square lack of closure error.

RESULTS AND DISCUSSION

Structure Determination—The crystal structure of MH-K1 chitosanase was solved by MAD phasing of the K_2PtCl_4 derivative. Although the space group could not be unambiguously determined between $P2_12_12$ and $P222_1$ from the results of the preliminary diffraction experiments (32), the MAD phasing gave a reasonable solution only when the former space group was employed. The electron density with initial MAD phases at 4.0-Å resolution enabled us to incorporate six α helices (α3, α7, α8, α9, α10, and α12) corresponding to 67 of the total of 259 amino acid residues. Subsequently, phases calculated from these portions were combined with the MAD phases, and then phases were improved and extended to 3.0-Å resolution by solvent flattening. Despite repeated combination and improvement of the phases, residues 89–120, corresponding to the platinum binding site and a few loop regions between α helices, were not well fitted to the electron density map. Finally, the polyalanine model with 80% of the total structure was constructed, which was used for refinement against the merged native data at 5.0–1.8-Å resolution. At this stage, side chains of the α helices could be easily identified in the $2F_o - F_c$ map, and α helix and β sheet were identified in the region of residues 89–120 in $F_o - F_c$ maps. During the course of refinement, the free R factor (44) dropped from 47.7 to 23.5% for 5% of the total reflections. The crystallographic R factor for the final model, including 150 water molecules and an SO_4^{2-} ion lying in the crystallographic 2-fold axis, was 19.2% at 1.6-Å resolution (Table III). Fig. 1 shows the final $2F_o - F_c$ map in the highly hydrophobic core region. The stereochemistry of the final model was reasonable, with no residues lying at unfavorable regions in the Ramachandran plot. The root mean square deviations from standard values were 0.009 Å for bond lengths and 1.375° for angles (Table III).

Overall Structure—The overall molecular structure of MH-K1 chitosanase is shown in Fig. 2, which shows 14 α helices and 5 β strands. The overall folding of MH-K1 chitosanase is similar to that of N174 chitosanase (28) except for three regions described below, the longest dimensions of the both molecules being 57 and 55 Å, respectively (Fig. 3*a*). The amino acid sequences were aligned between MH-K1 and N174 chitosanases on the basis of their conformational comparison of the secondary structure (Fig. 4). Despite the folding similarity,

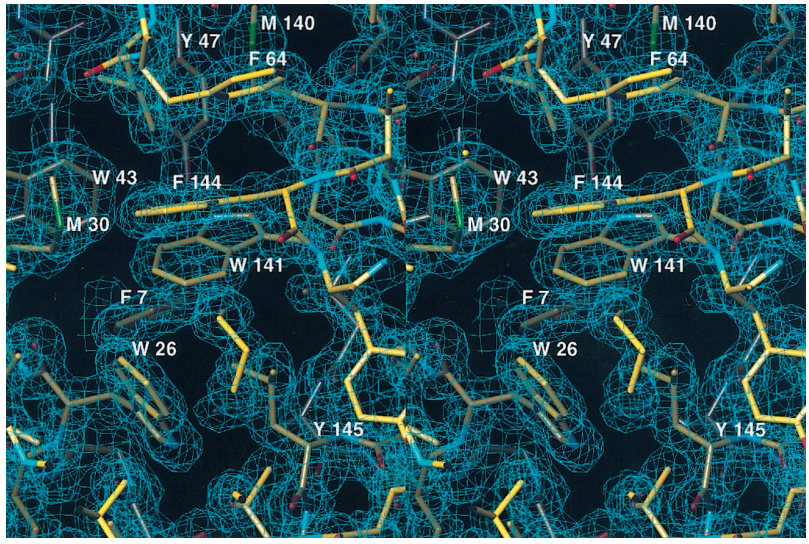

FIG. 1. **Stereo view of the final $2F_o - F_c$ map at 1.6-Å resolution.** This highly hydrophobic core region includes α4 and α8 helices and a β sheet loop between β2 and β3 strands (residues Phe[7], Trp[26], Met[30], Trp[43], Tyr[47], Phe[64], Met[140], Trp[141], Phe[144], and Tyr[145]). This *figure* was drawn by *TURBO-FRODO* (39).

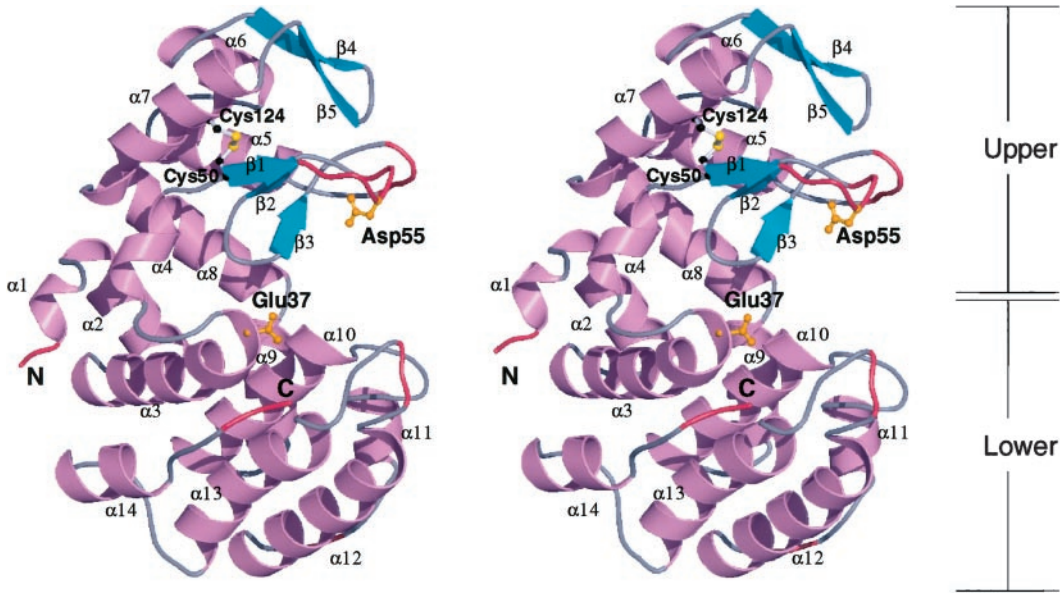

FIG. 2. **Stereo view of the overall three-dimensional structure of MH-K1 chitosanase illustrated as *ribbon diagrams*.** The α helices and β strands are indicated in *violet* and *blue*, respectively, and two domains are labeled. Catalytic residues (Glu[37] and Asp[55] in *orange*) and a disulfide bridge (Cys[50] and Cys[124] in *yellow*) are shown as *ball-and-stick models*. The loop regions with high temperature factors ($B \geq 30$ Å^2) are indicated in *red*. This *figure* was drawn by *MOLSCRIPT* (51) and *Raster 3D* (52, 53).

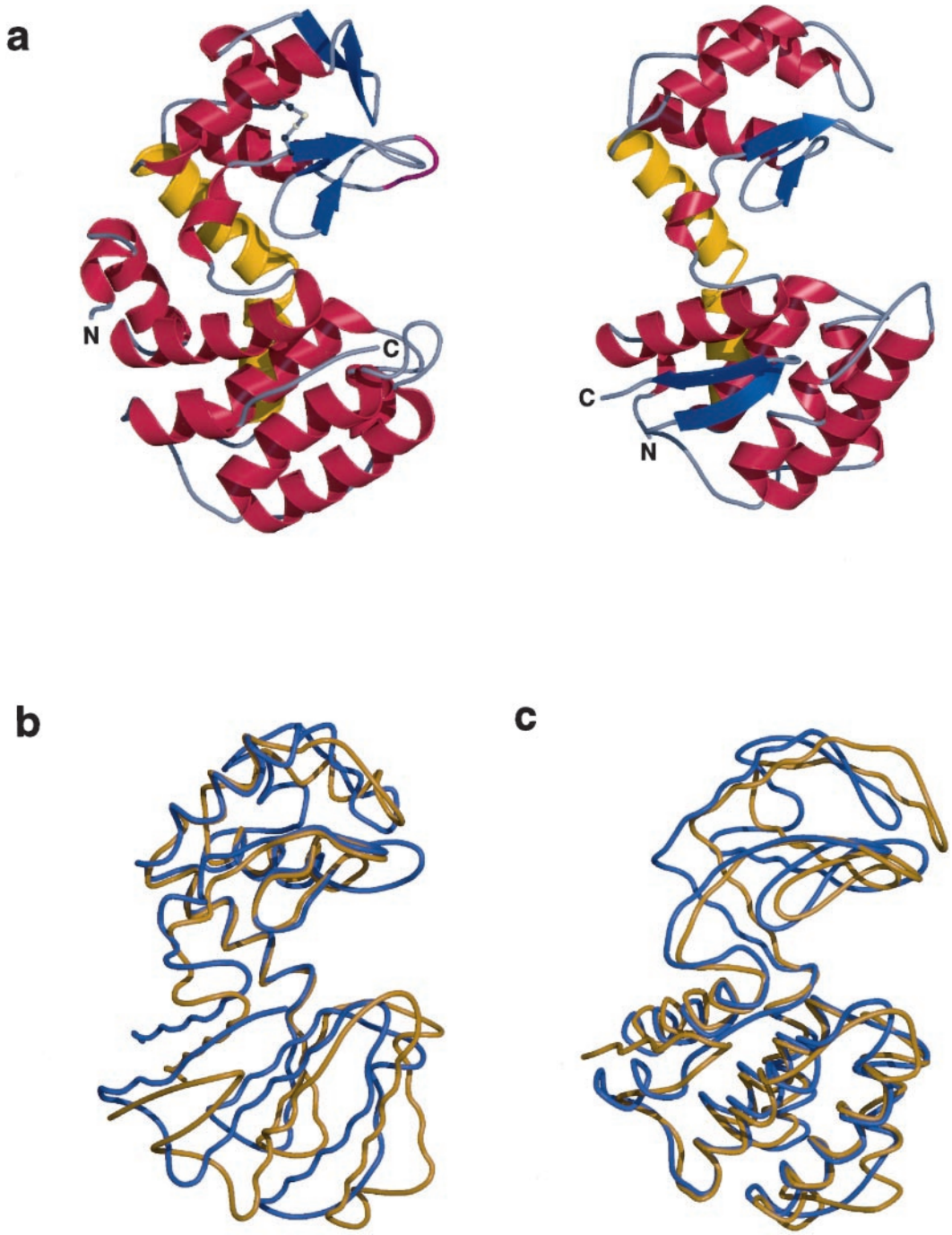

FIG. 3. **Comparison of the structures of MH-K1 and N174 chitosanases.** *a*, the overall structures of MH-K1 (*left*) and N174 (*right*) chitosanases are illustrated as *ribbon diagrams*. The backbone helices are shown in *yellow*. The protruding roof of the cleft in MH-K1 chitosanase (*left*) is shown in *pink*. *b* and *c*, superposition of the C-α trace models of MH-K1 (*blue*) and N174 chitosanases (*yellow*). The C-α atoms of the upper (47–68, 77–88, 120–130, and 134–148 in the sequence of MH-K1 chitosanase) and lower (23–40, 147–159, 164–178, 199–216, and 232–239) domains are superimposed in *b* and *c*, respectively. Residues 1–21 (MH-K1 chitosanase) and 1–6 (N174 chitosanase) have been omitted for clarity. This *figure* was drawn by *MOLSCRIPT* (51) and *Raster 3D* (52, 53).

there was only 20% identity in the amino acid sequences of both chitosanases (18). While there was a significantly conserved segment in the N-terminal region (residues 19–69), its sequence alignment from the secondary structure was difficult in the region of residues 90–119. Three marked differences in the molecular structures were observed between MH-K1 and N174 chitosanases. First, there were two additional helices (α1 and α2) in the N-terminal region of MH-K1 chitosanase, which were 16 residues longer than that of N174 chitosanase. Second, there were two β strands (β4 and β5) following to the α6 helix in the top region of the upper domain, whereas there was only an α5 helix in N174 chitosanase. They were located in the unaligned region described above (residues 90–119 in MH-K1 chitosanase and residues 70–93 in N174 chitosanase). Third, the secondary structures were completely different at the C-terminal regions of both chitosanases, a helical structure (α14) in MH-K1 chitosanase and two β sheets (β4 and β5) in N174 chitosanase. In addition, the disulfide bridge between Cys^{50} and Cys^{124} joined the β1 strand and the α7 helix only in MH-K1 chitosanase, which is not conserved among other chitosanases.

Domain Structure—The molecule was shown to be composed of two globular domains, the upper and lower domains. Two

backbone helices shown in *yellow* in Fig. 3*a*, the α8 and α9 helices in MH-K1 chitosanase (α7 and α8 helices in N174 chitosanase), connect the upper and lower domains. Val^{147} and Tyr^{148} (Val^{121} and Tyr^{122} in N174 chitosanase) localized between the two helices are conserved between MH-K1 and N174 chitosanases. Although the lengths of the two helices in both chitosanases were almost the same, the angles formed by the two helices, Asp^{133}–Val^{147} and Val^{147}–Gly^{160} in MH-K1 chitosanase, Asp^{107}–Val^{121} and Val^{121}–Gly^{134} in N174 chitosanase, were different, being approximately 110 and 130°, respectively. In other words, the cleft formed by the upper and lower domains of MH-K1 chitosanase is more open than that of N174 chitosanase. The relative orientations of the upper and lower domains were slightly different. This difference in relative orientation of the upper and lower domains affects the size and shape of cleft. In superposition of the overall structures between MH-K1 and N174 chitosanases (45), the root mean square deviation for the corresponding 129 C-α atoms was 2.03 Å (MH-K1: residues 23–40, 47–68, 77–88, 120–130, 134–147, 149–159, 164–178, 199–216 and 232–239; N174: 8–25, 32–53, 57–68, 94–104, 108–121, 123–133, 137–151, 179–196 and 210–217). Nevertheless, the domain structures of MH-K1 and N174 chitosanases were very similar in each domain, and individual root mean square deviations for the upper and lower domains were 0.98 Å (Fig. 3*b*) and 1.35 Å (Fig. 3*c*), respectively. Although the sequence identity between MH-K1 and N174 chitosanases was only ~20%, the superimposed C-α atoms of each domain showed a high degree of similarity in both secondary and tertiary structures.

TABLE III
Refinement statistics

Parameter	Value
No. of reflections ($I \geq 2\sigma(I)$)	38,353
Resolution range (Å)	5.0–1.6
No. of nonhydrogen protein atoms	2037
No. of ions (SO_4^{2-})	1 on 2-fold axis
No. of water molecules	150
R factor[a] (%)	19.2
R_{free}[b] (%)	23.5
Root mean square deviation from ideal geometry	
Bond lengths (Å)	0.009
Bond angles (degrees)	1.375

[a] *R* factor = $\Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.
[b] The R_{free} value was calculated from a set of reflections (5%) randomly, which were omitted from structure refinement.

Substrate Recognition—A highly conserved sequence segment was found in the N-terminal region of the procaryotic MH-K1, N174, and N106 chitosanases (18, 46). A site-directed

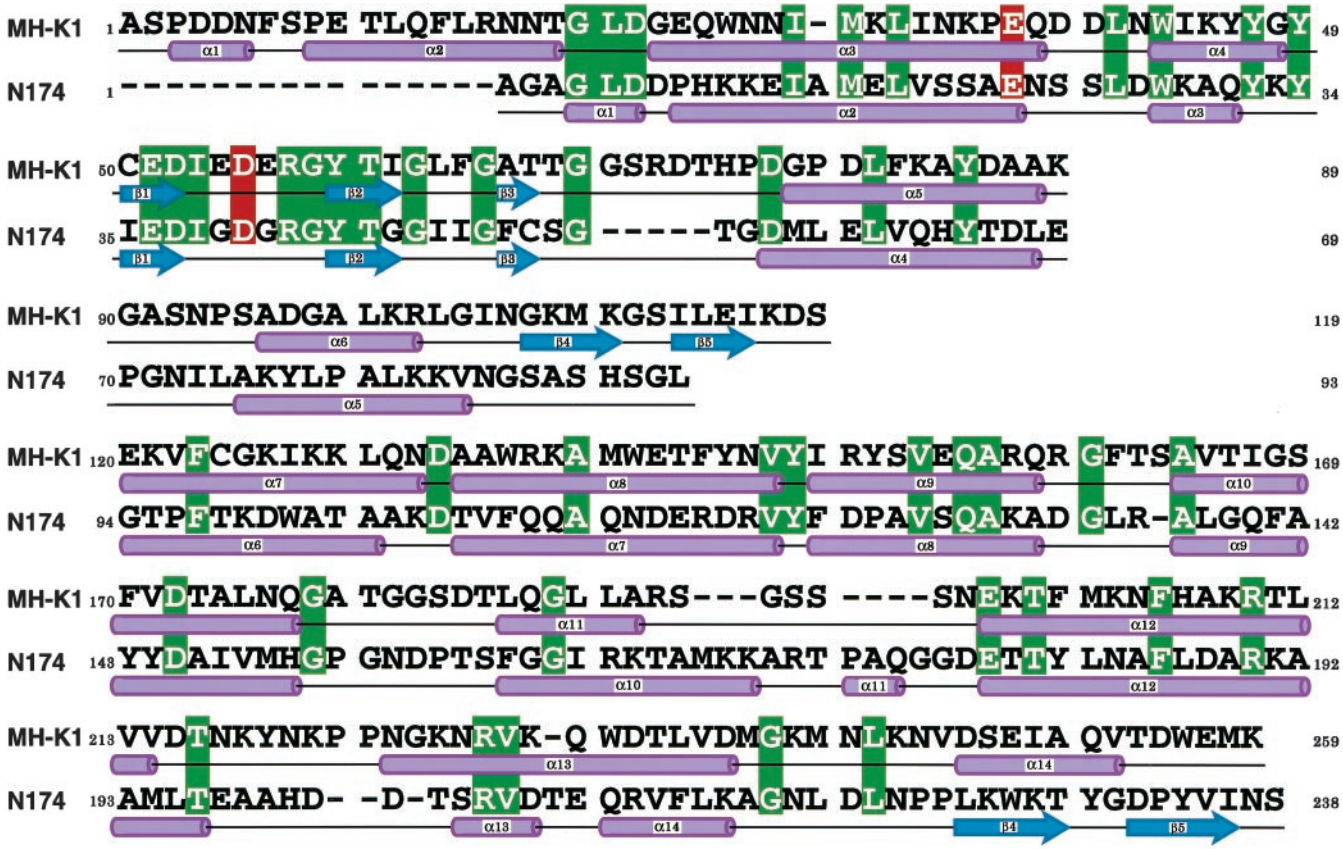

FIG. 4. **Amino acid sequence alignment of MH-K1 (GenBank™ accession number D10624) and N174 (GenBank™ accession number L40408) chitosanases based on their secondary structures.** Residues 90–119 in MH-K1 chitosanase and 70–83 in N174 chitosanase were not aligned because the structural and positional differences were too large. Identical amino acid residues are *shaded* in *green*, and the catalytic residues are *shaded* in *red*. *Below* the sequence, the secondary structure elements are represented by *cylinders* and *arrows* for α helices and β strands, respectively. Amino acid residues 77–83 (RW**P**GP**L**S, uncorrected amino acids are *bold*), 99 (D), and 159–160 (HA) in the previously reported sequence (16) disagreed with those used in the present crystal structure determination as 77–82 (DG**P**D**L**F, one residue omitted), 98 (G), and 158–160 (QRG, one residue additional), respectively. The sequence shown in this *figure* could be corrected on the basis of model fitting to the 1.6-Å resolution electron density map, where the shapes of the side chains were clearly identified. This correction was supported by the amino acid sequence of chitosanase from *B. ehimensis* sp. EAG1 (90% identity with MH-K1 chitosanase) (deposited in GenBank™, accession number AB008788 (Akiyama, K., Fujita, T., Kuroshima, K., Sakane, T., Yokota, A., and Takata, R.). Most of the reported amino acid sequence of MH-K1 chitosanase was determined from the protein sequence, but several regions were partially deduced from the nucleotide sequence (16). The amino acids corrected in the present study were located in the regions deduced from the nucleotide sequence.

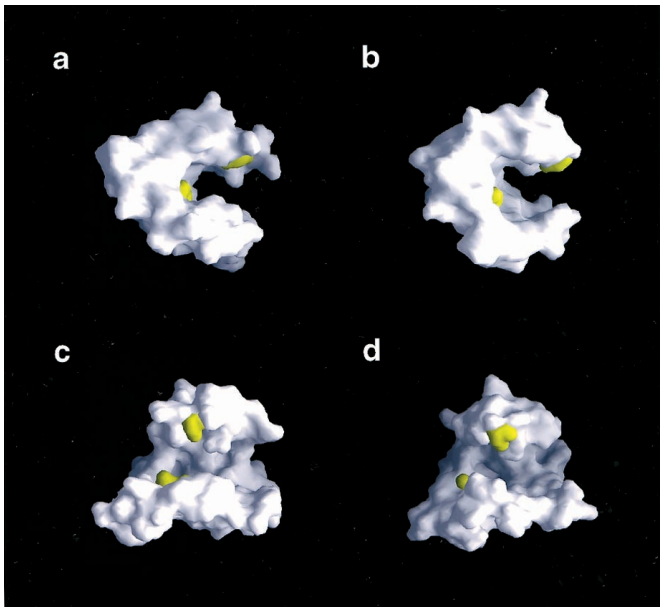

FIG. 5. **Comparison of the molecular surfaces of the substrate binding cleft of MH-K1 (*a* and *c*) and N174 chitosanases (*b* and *d*).** The viewpoint of *a* and *b* is the same as that of Fig. 3*a*. The viewpoint of *c* and *d*, which directly shows the cleft opening space, is rotated 90° around a vertical line from that of *a* and *b*. The catalytic residues in the cleft are indicated in *yellowish green*. This *figure* was drawn by *GRASP* (48).

mutagenesis study revealed that both Glu^{22} and Asp^{40} localized within the conservative N-terminal region in N174 chitosanase are essential for catalytic activity (47). These two residues are conserved in MH-K1 chitosanase (Glu^{37} and Asp^{55}) as shown in Fig. 4, which suggests that these two are also catalytic residues. This was also supported by mutagenesis of Asp^{55} of this chitosanase.[2]

To understand their substrate specificities, the surface electrostatic potentials of MH-K1 and N174 chitosanases were calculated using the program *GRASP* (48). There were no marked differences in electrostatic distribution at the potential substrate binding cleft, where the electrostatic potentials were negative in both chitosanases. However, there was a significant difference in the shape of the substrate binding cleft (Fig. 5). There was an insertion region (residues 70–74) after β3 strand in the upper domain of MH-K1 chitosanase, which formed a protrusion at the roof of the cleft (shown in *pink* in Fig. 3*a*). On the other hand, the shape of the cleft created by the lower domain was also different. The C-terminal region of MH-K1 chitosanase was composed of α14 helix, whereas there were two β strands (β4 and β5) in N174 chitosanase. The α14 helix extended toward the cleft, which covered a hollow observed in N174 chitosanase and provided the flat base of the cleft in MH-K1 chitosanase. Consequently, the cleft in MH-K1 chitosanase was a little smaller than that in N174 chitosanase. In addition, the relative orientation of the upper and lower domains was slightly different in these two chitosanases as discussed above. These structural and orientational differences affecting the substrate binding cleft should account for the differences in substrate recognition between MH-K1 and N174 chitosanases. To examine the substrate recognition mechanism, a substrate analogue (chitosan hexamer; hexa-β-(1,4)-D-glucosamine, $GlcN_6$) was adjusted into the cleft of MH-K1 chitosanase (Fig. 6). This binding model was constructed on the basis of the structure of human lysozyme complexed with tetrasaccharide (49) and the speculative chitosan hexamer model of N174 chitosanase (28). The interaction between the cleft and the substrate analogue was specified only at three subsites, C,

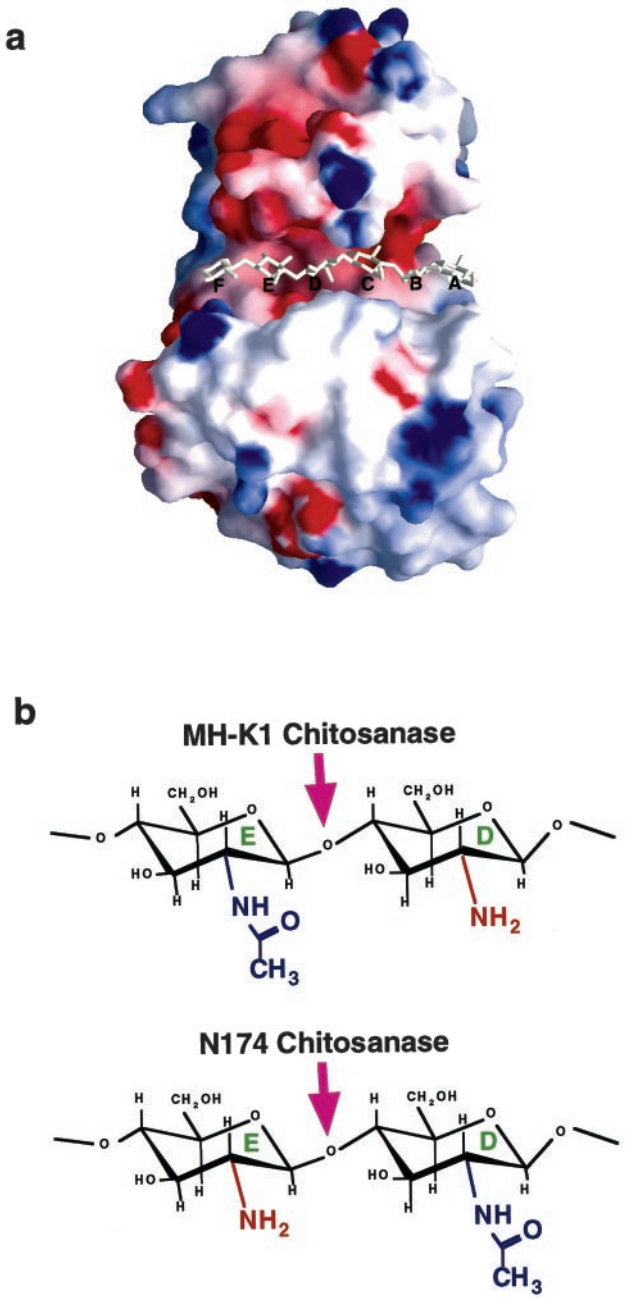

FIG. 6. **Binding of substrate in MH-K1 chitosanase.** *a*, a binding model with a substrate analogue (chitosan hexamer; hexa-β-(1,4)-D-glucosamine, $GlcN_6$) in the active site. The subsites in the binding site of the cleft (six sugar binding sites) A–F are labeled, where F is the reducing end of the sugar. *Panel a* was drawn using the program *GRASP* (48). *b*, the substrate cleavage positions for the partially acetylated chitosan in MH-K1 and N174 chitosanases.

D, and E, among the six sugar binding sites, whereas the positions of A, B, and F sugars were more speculative due to the loose interaction between the cleft and sugars. In subsites A–F, the F sugar was assigned as the reducing end in a similar manner to the subsites in lysozyme. The substrate was cleaved between D and E sugars by two catalytic residues, Glu^{37} and Asp^{55}. The glycosyl hydrolases are classified into those two showing mechanisms of action due to the substrate binding site structures of β-glycosidases; the average distance between the two catalytic residues is 9.5 Å in "inverting enzymes," whereas it is 5.3 Å in "retaining enzymes" (50). MH-K1 chitosanase belongs to the inverting enzymes because the distance between oxygen atoms of Glu^{37} and Asp^{55} was 10.9 Å. In the binding

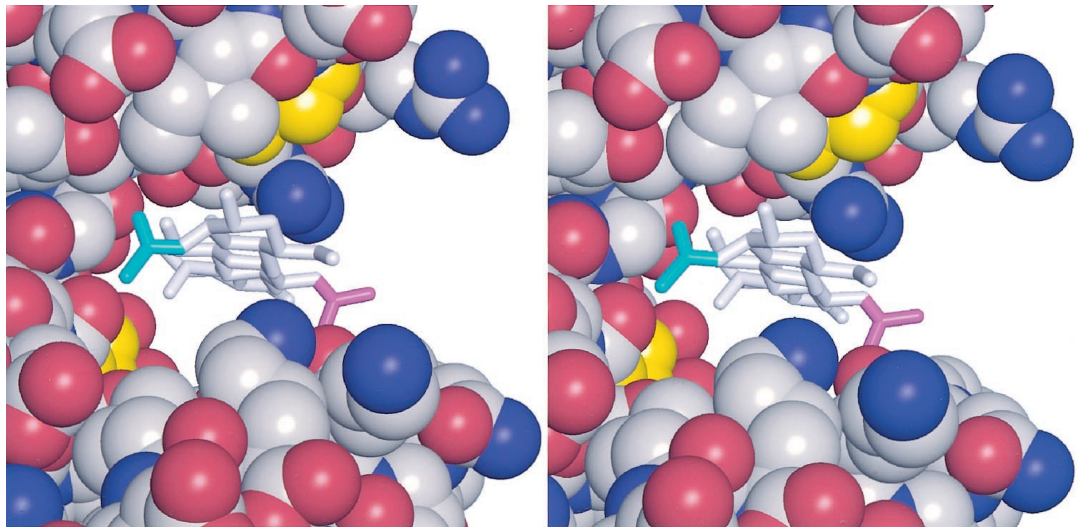

FIG. 7. **Stereo view showing substrate binding to the cleft of the molecular surface of MH-K1 chitosanase.** The chitosanase molecule is shown as a space-filling model with the catalytic residues (Glu^{37} and Asp^{55}) in *yellow*. The artificial combined substrate of GlcN-GlcNAc-GlcNAc (GlcN-GlcN-GlcNAc + GlcN-GlcNAc-GlcN) shown as a *stick model* is bound in the substrate binding cleft. The acetyl groups on the C-2 atom in the D sugar (GlcN-GlcNAc-GlcN) are shown in *pink*, whereas those in the E sugar (GlcN-GlcN-GlcNAc) are shown in *cyan*. This *figure* was drawn by *MOLSCRIPT* (51).

model with chitosan hexamer, these two catalytic residues have to move to make close contact with the linkage between D and E sugars. Glu^{37} fixed on the long central helix may act as a general acid, while Asp^{55} located on the β sheet loop between $\beta1$ and $\beta2$ may act as a general base to polarize the attacking water molecule. The temperature factors of this β sheet loop and the insertion loop between $\beta3$ and $\alpha5$ were significantly higher than those of the other regions (Fig. 2). These loops with high temperature factors may be structurally flexible so that they can make suitable contacts and for efficient recognition with the substrates.

MH-K1 chitosanase (subclass III) can split the GlcN-GlcNAc linkage (corresponding to the D-E sugar linkage) but cannot split the GlcNAc-GlcN linkage (Fig. 6). On the other hand, N174 chitosanase (subclass I) cannot split GlcN-GlcNAc but can cleave GlcNAc-GlcN. The environments around the acetyl group of GlcNAc in the substrate were investigated on the basis of the binding model of MH-K1 chitosanase with chitosan hexamer. In Fig. 7, an artificial substrate model (GlcN-GlcNAc-GlcNAc, corresponding to C-D-E sugars), where the acetyl groups are located on both D and E sugars of the chitosan hexamer, is accommodated in the cleft of MH-K1 chitosanase. In this model, two catalytic residues (Glu^{37} and Asp^{55}) are located near the cleavage bond of the substrate. The acetyl group on the C-2 atom of the D sugar (*pink*) is too close to the atoms at the base of the cleft, causing a significant steric hindrance. However, the acetyl group on the C-2 atom of the E sugar (*cyan*) could be located at a suitable depth in the cleft without any steric hindrance. Based on this substrate binding model, we concluded that MH-K1 chitosanase can accommodate only the GlcN-GlcNAc motif at the D-E subsite of its substrate binding cleft; *i.e.* it cannot bind the GlcNAc-GlcN motif. This may be the reason why MH-K1 chitosanase cannot cleave the GlcNAc-GlcN linkage but the GlcN-GlcNAc linkage. This model shows that the size and shape of the cleft are such that the substrate sugar with the acetyl groups at positions suitable for the specific cleavage reaction can be accommodated in the active site, which affords reaction specificity for substrate recognition of this chitosanase.

Acknowledgments—We thank Drs. N. Sakabe, N. Watanabe, and M. Suzuki of the Photon Factory for kind help in intensity data collection and processing, which were performed under the approval of the Photon Factory Advisory Committee (Proposal 95G048).

REFERENCES

1. Verburg, J. G. & Huynh, Q. K. (1991) *Plant Physiol.* **95,** 450–455
2. Huynh, Q. K., Hironaka, C. M., Levine, E. B., Smith, C. E., Borgmeyer, J. R. & Shah, D. M. (1992) *J. Biol. Chem.* **267,** 6635–6640
3. Henrissat, B. (1991) *Biochem. J.* **280,** 309–316
4. Henrissat, B. & Bairoch, A. (1993) *Biochem. J.* **293,** 781–788
5. Henrissat, B. & Bairoch, A. (1996) *Biochem. J.* **316,** 695–696
6. Davies, G. & Henrissat, B. (1995) *Structure* **3,** 853–859
7. Holm, L. & Sander, C. (1994) *FEBS Lett.* **340,** 129–132
8. Yabuki, M., Uchiyama, A., Suzuki, K., Ando, A. & Fujii, T. (1988) *J. Gen. Appl. Microbiol.* **34,** 255–270
9. Okajima, S., Ando, A., Shinoyama, H. & Fujii, T. (1994) *J. Ferment. Bioeng.* **77,** 617–620
10. Okajima, S., Kinouchi, T., Mikami, Y. & Ando, A. (1995) *J. Gen. Appl. Microbiol.* **41,** 351–357
11. Boucher, I., Dupuy, A., Vidal, P., Neugebauer, W. A. & Brzezinski, R. (1992) *Appl. Microbiol. Biotechnol.* **38,** 188–193
12. Ouakfaoui, S. E. & Asselin, A. (1992) *Phytochemistry* **31,** 1513–1518
13. Ouakfaoui, S. E. & Asselin, A. (1992) *Plant Sci.* **85,** 33–41
14. Sharma, P., Børja, D., Stougaard, P. & Lönneborg, A. (1993) *Physiol. Mol. Plant Pathol.* **43,** 57–67
15. Osswald, W. F., Shapiro, J. P., Doostdar, H., McDonald, R. E., Niedz, R. P., Nairn, C. J. & Mayer, R. T. (1994) *Plant Cell Physiol.* **35,** 811–820
16. Ando, A., Noguchi, K., Yanagi, M., Shinoyama, H., Kagawa, Y., Hirata, H., Yabuki, M. & Fujii, T. (1992) *J. Gen. Appl. Microbiol.* **38,** 135–144
17. Masson, J.-Y., Denis F. & Brzezinski, R. (1994) *Gene* (*Amst.*) **140,** 103–107
18. Masson, J.-Y., Boucher, I., Neugebauer, W. A., Ramotar, D. & Brzezinski, R. (1995) *Microbiology* **141,** 2629–2635
19. Shimosaka, M., Kumehara, M., Zhang, X.-Y., Nogawa, M. & Okazaki, M. (1996) *J. Ferment. Bioeng.* **82,** 426–431
20. Lu, Z., Li, Y., Que, Q., Kutish, G. F., Rock, D. L. & Von Etten, J. L. (1996) *Virology* **216,** 102–123
21. Yamada, T., Hiramatsu, S., Songsri, P. & Fujie, M. (1997) *Virology* **230,** 361–368
22. Parro, V., Román, M. S., Galindo, I., Purnelle, B., Bolotin, A., Sorokin, A. & Mellado, R. P. (1997) *Microbiology* **143,** 1321–1326
23. Hart, P. J., Monzingo, A. F., Ready, M. P., Ernst, S. R. & Robertus, J. D. (1993) *J. Mol. Biol.* **229,** 189–193
24. Hart, P. J., Pfluger, H. D., Monzingo, A. F., Hollis, T. & Robertus, J. D. (1995) *J. Mol. Biol.* **248,** 402–413
25. Song, H. K. & Suh, S. W. (1996) *Acta Crystallogr. Sec. D* **52,** 289–298
26. Scheltinga, A. C. T. v., Kalk, K. H., Beintema, J. J. & Dijkstra, B. W. (1994) *Structure* **2,** 1181–1189
27. Perrakis, A., Tews, I., Dauter, Z., Oppenheim, A. B., Chet, I., Wilson, K. S. & Vorgias, C. E. (1994) *Structure* **2,** 1169–1180
28. Marcotte, E. M., Monzingo, A. F., Ernst, S. R., Brzezinski, R. & Robertus, J. D. (1996) *Nat. Struct. Biol.* **3,** 155–162
29. Fukamizo, T., Honda, Y., Goto, S., Boucher, I. & Brzezinski, R. (1995) *Biochem. J.* **311,** 377–383
30. Izume, M., Nagae, S., Kawagishi, H., Mitsutomi, M. & Ohtakara, A. (1992) *Biosci. Biotech. Biochem.* **56,** 448–453
31. Mitsutomi, M., Ueda, M., Arai, M., Ando, A. & Watanabe, T. (1996) *Chitin Enzymol.* **2,** 273–284

32. Saito, J., Kita, A., Nagata, Y., Ando, A. & Miki, K. (1995) *Acta Crystallogr. Sec. D* **51,** 856–857
33. Sakabe, N., Ikemizu, S., Sakabe, K., Higashi, T., Nakagawa, A., Watanabe, N., Adachi, S. & Sasaki, K. (1995) *Rev. Sci. Instrum.* **66,** 1276–1281
34. Miyahara, J., Takahashi, K., Amemiya, Y., Kamiya, N. & Satow, Y. (1986) *Nucl. Instr. Methods* **A246,** 572–578
35. Otwinowski, Z. & Minor, M. (1997) *Methods Enzymol.* **276,** 307–326
36. Collaborative Computational Project 4 (1994) *Acta Crystallogr. Sec. D* **50,** 760–763
37. Wang B. C. (1985) *Methods Enzymol.* **115,** 90–112
38. Zhang K. Y. J. & Main, P. (1990) *Acta Crystallogr Sec. A* **46,** 377–381
39. Roussel, A. & Cambillau, C., (1992) *TURBO-FRODO*, Biographics, LCCMB, Marseille, France
40. Read, R. J., (1986) *Acta Crystallogr. Sec. A* **42,** 140–149
41. Kleywegt, G. J. & Jones, T. A., (1993) *ESF/CCP4 Newsletter* **28,** 56–59
42. Brünger, A. T. (1993) *XPLOR Version 3.1 Manual*, Yale University
43. Laskowski, P. V., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1994) *J. Appl. Crystallogr.* **26,** 283–291
44. Brünger, A. T. (1992) *Nature* **355,** 472–475
45. Kleywegt, G. J. & Jones, T. A., (1994) *ESF/CCP4 Newsletter* **31,** 9–14
46. Fukamizo, T. & Brzezinski, R. (1998) *Biochem. Cell Biol.* **75,** 687–696
47. Boucher, I., Fukamizo, T., Honda, Y., Willick, G. E., Neugebauer, W. A. & Brzezinski, R. (1995) *J. Biol. Chem.* **270,** 31077–31082
48. Nicholls, A., Bharadwaj, R. & Honig, B. (1993) *Biophys. J.* **64,** 166–167
49. Song, H., Inaka, K., Maenaka, K. & Matsushima, M. (1994) *J. Mol. Biol.* **244,** 522–540
50. McCarter, J. D. & Withers, S. G. (1994) *Curr. Opin. Struct. Biol.* **4,** 885–892
51. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
52. Bacon, D. J. & Anderson, W. F. (1988) *J. Mol. Graphics* **6,** 219–220
53. Merritt, E. A. & Murphy, M. E. P. (1994) *Acta Crystallogr. Sec. D* **50,** 869–873