LysM Domains from *Pteris ryukyuensis* Chitinase-A

A STABILITY STUDY AND CHARACTERIZATION OF THE CHITIN-BINDING SITE

Received for publication, August 27, 2007, and in revised form, December 14, 2007 Published, JBC Papers in Press, December 14, 2007, DOI 10.1074/jbc.M707156200

Takayuki Ohnuma1, Shoko Onaga5, Katsuyoshi Murata1, Toki Taira1, and Etsuko Katoh4,1

From the 4Division of Plant Sciences, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan and the 5Department of Bioscience and Biotechnology, Faculty of Agriculture, Ryukyu University, Nishihara cho, Okinawa 903-0213, Japan

The LysM domain probably binds peptidoglycans, but how it does so has yet to be described. For this report, we measured the thermal stabilities of recombinant LysM domains derived from *Pteris ryukyuensis* chitinase-A (PrChi-A) and monitored their binding to N-acetylglucosamine oligomers ([GlcNAc]n) using differential scanning calorimetry, isothermal titration calorimetry, and NMR spectroscopy. We thereby characterized certain of the domains' functional and structural features. We observed that the domains are very resistant to thermal denaturation and that this resistance depends on the presence of disulfide bonds. We also show that the stoichiometry of ([GlcNAc]2)/LysM domain binding is 1:1. ([GlcNAc]2)5 titration experiments, monitored by NMR spectroscopy, allowed us to identify the domain residues that are critical for ([GlcNAc]2) binding. The binding site is a shallow groove formed by the N-terminal part of helix 1, the loop between strand 1 and helix 1, the C-terminal part of helix 2, and the loop between helix 2 and strand 2. Furthermore, mutagenesis experiments reiterate the critical involvement of Tyr72 in ([GlcNAc]2)/LysM domain binding. Ours is the first report describing the physical structure of a LysM oligosaccharide-binding site based on experimental data.

Carbohydrates are often the signals that initiate biochemical processes. For example, as a response to pathogenic attacks, plant defense proteins first recognize and then attack the carbohydrate components of the pathogens' cell walls (1). As these defense proteins bind carbohydrates (2), via their noncatalytic carbohydrate-binding modules, their binding mechanisms are physiologically interesting. Many carbohydrate-binding modules have been structurally and/or functionally characterized (3) and, based on these characteristics, are currently grouped into 49 families in the Carbohydrate-Active Enzyme (CAZy) data base (available on the World Wide Web).

The LysM domain, which probably binds chitin, a β-(1→4)-linked N-acetylglucosamine oligosaccharide ([GlcNAc]n), was originally identified as a component of bacterial lysins. This domain is found in many of the enzymes involved in cell wall degradation and is also present in other proteins that are associated with bacterial cell walls (4–8). Basal level resistance by plants against certain pathogens also appears to involve the recognition of chitin oligosaccharides and related compounds. Recently, a chitin oligosaccharide elicitor receptor was purified, and it contains a LysM domain within its extracellular domain (9). Additionally, certain symbiotic relationships between leguminous plants and rhizobial bacteria appear to be mediated by LysM domain/chitin oligosaccharide interactions. Recent studies have identified a class of proteins that recognize Nod factors, which are the lipochitin oligosaccharide signal molecules secreted by symbiotic bacteria (10–12). These plant proteins are members of a serine/threonine receptor kinase family and contain extracellular LysM domains. After reviewing the available literature on LysM domains, it appears to us that its function is to bind peptidoglycans, although presently, it is not officially recognized as a carbohydrate-binding module, since it is not included in the CAZy data base.

Although the LysM domain is found in many proteins, as described above, only three tertiary structures for it have been reported. Bateman and Bycroft (13) determined the three-dimensional solution structure of the LysM domain found in the *Escherichia coli* membrane-bound lytic murein transglycosylase D (MltD; Protein Data Bank code 1EOG). Recently, the crystal structure of *Bacillus subtilis* YkUD, a protein containing one LysM domain was reported (Protein Data Bank code 1Y7M) (14). The structure of a LysM domain from the human hypothetical protein SB145 has been deposited in the Protein Data Bank (Protein Data Bank code 2DJP). These domains have a βααβ fold with the two helices packing against one side of the two-stranded antiparallel β-sheet.

Although the canonical three-dimensional LysM domain structure is now known, how a LysM domain physically interacts with an oligosaccharide is not. Bateman and Bycroft (13) reported that the loop between strand 1 and helix 1 lies at the end of a shallow groove on the surface of the domain and suggested that this region might be the ligand binding site. They based their suggestion both on an examination of the MltD solution structure and on an alignment of its sequence with other LysM domains (13). Recently, Mulder et al. (15) performed surface analyses and docking calculations using models (built by homology from the MltD structure) of the three LysM domains of the *Medicago truncatula* Nod factor perception protein with the goal of predicting the most favorable binding modes between chitin oligosaccharides and Nod factors. Their consensus interaction model primarily involves van der Waals interactions between a bound tetrasaccharide and the residues

---

*This work was supported by a grant from the Ministry of Agriculture, Forestry, and Fisheries of Japan (Green Technology Project IP4003). 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.

1 To whom correspondence should be addressed. Tel./Fax: 81-29-838-7910; E-mail: ekatoh@nias.affrc.go.jp.
of the long loop between helix 2 and strand 2. Neither of the aforementioned reports is based on an experimentally defined oligosaccharide-LysM complex, so it is not necessarily surprising that the two suggested binding sites differ significantly. Obviously, the oligosaccharide binding sites for LysM domains need to be experimentally defined.

Recently, we cloned the gene (prchIA) for a novel chitinase from Pteris ryukyuensis.\(^2\) It codes for two N-terminal LysM domains and a family 18-chitinase catalytic domain at its C terminus. For the study reported herein, we expressed and purified two recombinant LysM domains (LysM single and LysM tandem) derived from this chitinase. Using calorimetric and NMR techniques, we determined their thermostabilities and their binding affinities to various GlcNAc oligomers and characterized the carboxyhydrate binding site.

**EXPERIMENTAL PROCEDURES**

**Cloning and Purification of Recombinant LysM Domains—** Nucleotide sequences encoding PrChi-A LysM domains were inserted into pET-22b vectors (Novagen, Darmstadt, Germany). Fragments of prchIA cDNAs (AB247328; GenBank\(^\text{TM}\)) that encode LysM domains were amplified using PCR. LysM single (Cys\(^{60}\)-Lys\(^{107}\)) was constructed using the primers P1 (5'GGAATTCCATATGTGACAACATACACAATC-3') and P2 (5'-GGCGGATCTCTTATGGAATGCAACACAG-3'). The NdeI and BamHI recognition sites are positioned at the underlined and double underlined sequences, respectively. LysM tandem (Asp\(^{1}\)-Lys\(^{107}\)) was constructed using the primers P3 (5'-GGAAATTCATATGGATGTCACAAATCACAGCT-3') and P4. A PrChi-A deletion mutant \(\Delta 1\text{LysM} \) that consists of residues Cys\(^{60}\)-Ser\(^{395}\) (i.e. it lacks the LysM domain nearest the N-terminus) was also constructed. This mutant was constructed using the primers P1 and P4 (5'-GGCGGATCTCTTATGGAATGCAACACAG-3').

(The NdeI and BamHI recognition sites are positioned at the underlined and double underlined sequences, respectively.) LysM tandem (Asp\(^{1}\)-Lys\(^{107}\)) was constructed using the primers P3 (5'-GGAAATTCATATGGATGTCACAAATCACAGCT-3') and P4. A PrChi-A deletion mutant \(\Delta 1\text{LysM} \) that consists of residues Cys\(^{60}\)-Ser\(^{395}\) (i.e. it lacks the LysM domain nearest the N-terminus) was also constructed. This mutant was constructed using the primers P1 and P4 (5'-GGCGGATCTCTTATGGAATGCAACACAG-3'). A gene coding only for the catalytic domain (CatD; Lys\(^{125}\)-Ser\(^{395}\)) was also constructed. This mutant was constructed using the primers P5 (5'-GGAAATTCATATGGAATGTCACAAATCACAGCT-3') and P4. The purified PCR products were digested with NdeI and BamHI and then ligated into pET-22b. Site-directed mutagenesis (a Tyr\(^{72}\) → Ala mutation, Y72A) was accomplished using a QuickChange site-directed mutagenesis kit (Stratagene, La Jolla, CA). The primer used for this construct was P5 (5'-GGCGGATCTCTTATGGAATGCAACACAGCT-3'). The mutation was introduced into both the LysM single and \(\Delta 1\text{LysM} \) constructs.

All clones were transformed into E. coli BL21(DE3) cells. The bacteria were grown at 37 °C in Luria-Bertani broth for nonlabeled protein expression and in M9 minimal medium for uniformly\(^{15}\)N- and \(^{13}\)N/\(^{15}\)C-labeled protein expression. Ampicillin (50 μg/ml) was included in all broths. Protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside at a final concentration of 1 mM. Cells were then grown for 24 h at 18 °C, centrifuged at 5,000 × g for 15 min, and frozen.

Frozen pellets were thawed on ice, suspended in 20 mM sodium phosphate, pH 7.0, and lysed by sonication. Insoluble material was then removed by centrifugation at 27,000 × g for 30 min. The supernatants were dialyzed against 10 mM sodium acetate, pH 4.0, and precipitated. For LysM tandem, the sample was then concentrated and purified over Hi-Load Superdex 30 pg 26/60 (GE Healthcare Ltd.) in 20 mM sodium phosphate, pH 7.0, 100 mM NaCl. For LysM single, the supernatant was loaded onto a column of Toyoscreen Hexyl-650C (Tosoh Corp., Tokyo, Japan) resin. After washing the protein-loaded resin with 50 mM sodium acetate, pH 5.0, 2.5 mM ammonium sulfate, protein was eluted using a linear 2.5 to 0 mM ammonium sulfate gradient. The fraction containing LysM single was concentrated and purified over HiLoad Superdex 30 pg 26/60 in 20 mM sodium phosphate buffer, pH 7.0, 100 mM NaCl.

**Calorimetric Studies—** Thermal denaturations of both isolated LysM domains and LysM tandem in the presence of the reducing reagent β-mercaptoethanol (β-ME)\(^3\) or tris(2-carboxylethyl)phosphine (TCEP), each at a concentration of 5 mM, or in the presence of a 20-fold molar excess of (GlcNAc)\(_n\), were monitored by differential scanning calorimetry (DSC) using a MicroCal VP-DSC calorimeter (MicroCal, Northampton, MA). DSC thermograms were obtained from 20 to 130 °C at scan rates of 1 °C/min with protein concentrations of 1 mg/ml in 20 mM sodium phosphate, pH 7.0. Base-line corrections were performed by subtracting a buffer thermogram, obtained under identical conditions, from experimental thermograms. The data were analyzed using MicroCal Origin 7.0 (Origin-Lab Corp., Northampton, MA), assuming a two-state unfolding model.

Isothermal titration calorimetry (ITC) experiments were performed with a MicroCal VP-ITC calorimeter. The instrument’s design and its operation have been described in detail elsewhere (16). The LysM samples used in these experiments were prepared by first dialyzing them for about 16 h against 20 mM sodium phosphate, pH 7.0. (GlcNAc)\(_n\) was dissolved in the same buffer preparations.

For a given experiment, 7-μl aliquots of 9.65–31.9 mM (GlcNAc)\(_n\), in 20 mM sodium phosphate, pH 7.0, were added via a 250-μl syringe to a sample cell that contained 1.4482 ml of a stirred (310 rpm) 178–364 μM LysM solution equilibrated at 30 °C. After each injection, the heat released was measured. Prior to experimentation, the instrument containing water in the sample and reference cells was equilibrated overnight. Stable base lines were defined as those with root mean square noise levels of less than 5 ncal s\(^{-1}\). The heat of dilution caused by an injection of (GlcNAc)\(_n\) was measured under identical buffer, injection, and temperature conditions but by adding ligand to a sample solution that lacked protein. The heat of dilution was subtracted from the heat change that occurred when protein was present. Nonlinear fitting of the data was performed using MicroCal Origin 7.0 (Origin-Lab Corp.). The parameters that were varied to minimize the S.D. of the fit to the experimental data were the binding constant \(K_b\), the enthalpy change \(\Delta H\), and the number of binding sites per protein molecule (stoichiometry; \(n\)). The derived values for \(K_b\), \(\Delta H\), and \(n\) at 30 °C were

---

\(^{2}\) S. Onaga and T. Taira, submitted for publication.

\(^{3}\) The abbreviations used are: β-ME, β-mercaptoethanol; TCEP, tris(2-carboxylethyl)phosphine; DSC, differential scanning calorimetry; ITC, isothermal titration calorimetry.
then used to calculate the changes in free energy (ΔG) and entropy (ΔS).

NMR Spectroscopy—All NMR spectra were recorded at 30 °C using a Bruker AV600 spectrometer equipped with a 5-mm inverse triple resonance cryoprobe head with three-axis gradient coils. 1H, 13C, and 15N sequential resonance assignments were determined using data obtained from the following two-dimensional double resonance and three-dimensional double and triple resonance through-bond correlation experiments (17–19): two-dimensional 1H-15N HSQC, three-dimensional 15N-separated HOHAHA-HSQC, three-dimensional HNHA; three-dimensional HNCO, three-dimensional CBCA(CO)HN, and three-dimensional HNCA CB. All spectra were processed using NMRPipe software (20) and were analyzed using Sparky software (21). 1H, 13C, and 15N chemical shifts were referenced to HDO (4.64 ppm at 30 °C), indirectly to TSP (13C) (22), and to liquid ammonia (15N), respectively (23).

Two-dimensional 1H-15N HSQC spectra were recorded for 0.25 mM 15N-labeled LysM single in 20 mM sodium phosphate, pH 7.0, 100 mM NaCl, with and without 0.3 mM (GlcNAc)5. The amide 1H, 15N cross-peaks of the complex’s HSQC spectrum were identified using assignments obtained from the three-dimensional triple resonance experiments, three-dimensional CBCA(CO)HN and three-dimensional HNCA CB, of a solution containing 0.6 mM 15N/13C-labeled LysM and 0.72 mM (GlcNAc)5.

Homology Modeling—Multiple alignments of LysM domains were performed both manually and with the ClustalX program (24). The homology modeling program COMPOSER (25), which is part of the Sybyl 7.1 software package (SYBYL Tripos Associates, St. Louis, MO), was used to build the model LysM single structure. The structurally conserved regions (defined by COMPOSER) were built by homology using the NMR solution structure of the E. coli MltD LysM domain (13). Those peptide fragments, found in the Sybyl library of tertiary structures, that are most sequentially similar to the LysM domain loops, were used to model the loops. The loop geometries were optimized using the Tripos force field (26). The stereochemistries of the resulting models were checked using the PROCHECK program (27). Connolly surfaces were calculated using the MOLCAD program (28).

Oligosaccharide Docking—Oligosaccharide docking was performed using SurfDock-2.1 (29). Atomic coordinates for (GlcNAc)2 and (GlcNAc)5 were derived from the coordinates found for GlcNac in the cyclooxygenase-2/indomethacin/GlcNAc complex (PDB code 4COX) (30) and were then used for docking. The binding cavity was centered within a region defined by those residues whose NMR resonances were perturbed by (GlcNAc)5 binding. Surfex docking was performed using the “whole molecule” approach, with a maximum of 100 simultaneously considered posed fragments at each stage of the incremental construction process and default settings for all other parameters.

Chitin-binding Assay—The ability of LysM single and its Y72A mutant to bind chitin was examined using a column, packed with powdered chitin (0.5 × 10 cm), as described by Yamagami and Funatsu (31). Fifty μg of each domain were dissolved in 10 mM sodium phosphate, pH 7.0, and loaded onto the chitin column equilibrated with the same buffer. After washing the chitin with the same buffer to remove unadsorbed protein, the chitin was washed 1) with 1 M NaCl in the same buffer, 2) with 0.1 M acetic acid, 3) with 1 M acetic acid, and 4) with 5 M acetic acid.

Antifungal Activity—The antifungal activities of PrChi-A mutants were tested by measuring their abilities to inhibit Tri- choderma viride growth. A fungus inoculum was placed at the center of a Petri dish containing potato dextrose agar. Wells were subsequently punched into the agar at distances of 15 mm from the center of the Petri dish. Ten-μl samples, containing 200 or 400 pmol of protein in distilled water, were added to the wells. The plates were incubated for 24 h at room temperature and then photographed.

RESULTS

Cloning, Expression, and Purification of PrChi-A LysM Domains—PrChi-A contains two N-terminal LysM domains and a C-terminal class-IIIb chitinase-like catalytic domain (Fig. 1A). Initially, for the calorimetric and NMR experiments, three LysM recombinant genes were constructed: LysM tandem (Asp1–Lys107 and Cys60–Lys107). Although LysM single (Asp1–Lys107) and LysM single (Cys60–Lys107) have similar sequences (about 95% sequence identity; Fig. 1B), of the two constructs, only LysM single (Cys60–Lys107), hereafter known as LysM single, was produced in amounts (~8 mg/liter culture) sufficient for NMR spectroscopy and calorimetry. Sufficient quantities (~100 mg/liter culture) of LysM tandem (Asp1–Lys107), hereafter known as LysM tandem, were also produced. By the criterion of SDS-PAGE (Fig. 1C), the purities of both proteins are greater than 99%. LysM tandem and LysM single were used for the calorimetric experiments, and LysM single was used for the NMR experiments.

Characterizing the Thermostability of the LysM Domain by DSC—The energetics of a macromolecular conformational transition (e.g. protein denaturation) can be derived from the transition’s partial heat capacity’s dependence on temperature, which is obtainable from DSC measurements (32, 33). Fig. 2A shows the dependence of the Cp of unfolding on temperature for 1.0 mg/ml solutions of LysM tandem (black line) and LysM single (red line) in 20 mM sodium phosphate, pH 7.0. The LysM tandem and single domains denature at high temperatures, with thermal transition (Tm) values of 90.0 and 96.5 °C, respectively. Attempts to refold the domains by reducing the temperature failed, indicating that the thermal denaturations were completely irreversible for both domains. When heat denaturation causes protein aggregation, the higher temperature side of the corresponding DSC curve sharply decreases as a consequence of exothermic effects. At least in the case of the LysM tandem domain, irreversible denaturation is possibly caused by aggregation.

It is likely that the four cysteines found per domain form two disulfides. To determine if cysteine thiols are present in LysM tandem, LysM single, and/or LysM single Y72A, the proteins were assayed with the sulfhydryl-modifying reagent, 5,5′-dithiobis(2-nitrobenzoic acid) (34). No free thiols were detected in the proteins, suggesting that all cysteines participate in disul-
The Chitin-binding Site for a LysM Domain

A PST rich linker region

![Diagram of LysM domain and linker region]

B

| LysM tandem | LysM single | Mature protein (395 aa) |
|-------------|-------------|------------------------|
| First site  | Site 1      | Site 2                 |
| PrChi-A LysM 1st | 1           | 100                    |
| PrChi-A LysM 2nd | 59          | 90                     |
| MltD (180G) | 1           | 36                     |
| Chitinase V | 295         | 48(57)                 |
| Chitinase 1st | 105         |                        |
| Chitinase 2nd | 41          |                        |
| TPA inf: chitinase 18-10 1st | 283         |                        |
| TPA inf: chitinase 18-10 2nd | 348         |                        |
| N-acetylglucosamyl-L-alanine amidase | 160       |                        |
| Autolysin    | 631         |                        |
| Bacterial cell wall hydrolase | 565        |                        |
| CEBP 1st    | 85          |                        |
| CEBP 2nd    | 149         |                        |
| LysM 1st    | 99          |                        |
| LysM 2nd    | 154         |                        |
| NF5 1st     | 104         |                        |
| NF5 2nd     | 167         |                        |
| NF5 1st     | 52          |                        |
| NF5 2nd     | 124         |                        |

C

![Diagram of temperature dependence of Cp]

FIGURE 1. A, the PrChi-A domain order and the amino acid sequence of its LysM domains and linker region. The LysM domains are shown in red; the catalytic domain is shown in blue; and the linker regions, which are rich in Pro, Ser, and Thr residues, are shown in yellow. B, alignment of LysM amino acid sequences. Identities %

| Sequence      | Identities % |
|---------------|--------------|
| PrChi-A LysM 1st | 100          |
| PrChi-A LysM 2nd | 90           |
| MltD (180G) | 36           |
| Chitinase V | 48(57)       |
| Chitinase 1st | 43(61)       |
| Chitinase 2nd | 39(61)       |
| TPA inf: chitinase 18-10 1st | 34(52) |
| TPA inf: chitinase 18-10 2nd | 43(64) |
| N-acetylglucosamyl-L-alanine amidase | 41(64) |
| Autolysin    | 25(32)      |
| Bacterial cell wall hydrolase | 23(39) |
| CEBP 1st    | 131         |
| CEBP 2nd    | 192         |
| LysM 1st    | 153         |
| LysM 2nd    | 140         |
| NF5 1st     | 190         |
| NF5 2nd     | 81          |
| NF5 1st     | 150         |

FIGURE 2. A, plots of the temperature dependence of Cp (with base-line corrections) for the temperature-induced unfolding of LysM single (red line) and LysM tandem (black line). Protein concentrations of 1 mg/ml in 20 mM sodium phosphate, pH 7.0, were used. B, a plot of the temperature dependence of Cp when LysM tandem was unfolded as in A but with or without a reducing agent (β-ME or TCEP) present. C, the temperature dependence of Cp when LysM tandem was unfolded as in A without an oligosaccharide present (black line) or with (GlcNAc)4 (green line), (GlcNAc)3 (blue line), or (GlcNAc)2 (red line) present. The oligosaccharides were each present at a 20-fold molar excess.

fide bridges (data not shown). (Other indirect evidence for the presence of disulfide bonds is discussed below.) To determine whether such likely disulfide bonds contribute significantly to domain stability, the DSC experiment using LysM tandem was repeated, but this time in the presence of 5 mM β-ME or 5 mM TCEP (Fig. 2B). Since TCEP is more stable and a more effective reducing reagent than β-ME (35) and since reducing reagents sometimes produce spurious calorimetric data, using two different reducing reagents seemed prudent. For LysM tandem, in the presence of β-ME, its Tm value is 77.4 °C, and in the presence of TCEP, its Tm value is 68.4 °C. Apparently, there are disulfides present in LysM tandem that remarkably stabilize its structure.

The effects of GlcNAc oligomers on the thermal stability of LysM tandem were also studied. Fig. 2C shows the temperature dependence of Cp for uncomplexed LysM tandem (black line) and LysM tandem in the presence of (GlcNAc)3 (green line), (GlcNAc)4 (blue line), and (GlcNAc)5 (red line) after base-line subtraction. The Tm values are somewhat affected by the presence of the oligosaccharides: 89.9 °C for the isolated domain and for the complexes, 90.2 °C in the presence of (GlcNAc)3, 90.9 °C in the presence of (GlcNAc)4, and 91.6 °C in the presence of (GlcNAc)5. The Tm values for LysM tandem unfolding increased slightly as the length of the GlcNAc oligomer increased.
The Chitin-binding Site for a LysM Domain

Thermodynamics of (GlcNAc)ₙ/LysM Binding Monitored by ITC—ITC is the most direct method for measuring the heat of binding at constant temperature. A ligand is titrated into a solution containing its binding partner with the result that the complex binding constant ($K_b$), the enthalpy change for binding ($\Delta H$), and the stoichiometry of the complex (n) can be derived (32, 36). ITC measurements of (GlcNAc)ₙ (n = 3, 4, and 5) binding to the LysM constructs were measured at 30 °C. Fig. 3 shows the results for an ITC experiment when 7-µl portions of 9.65 mM (GlcNAc)₃ were titrated into a 1.4482-ml solution containing 129.8 µM LysM single in 20 mM phosphate, pH 7.0. In Fig. 3A, the area within each trough is a measure of the heat released upon the addition of (GlcNAc)₃. Fig. 3B is a plot of the heat released, normalized to 1 mol of oligosaccharide, versus the molar ratio of LysM single to a GlcNAc oligomer. The theoretical “best fit” binding curves are also shown in Fig. 3B. No matter which GlcNAc oligomer was used, when n was free to vary during fitting, the data were best fit with a stoichiometry of approximately 1 for the LysM single complex data and a stoichiometry of approximately 2 for the LysM tandem complex data (data not shown).

The derived thermodynamic values, $K_b$, $\Delta H$, $\Delta S$, and $\Delta G$, for the experiments involving (GlcNAc)₄ and (GlcNAc)₅ binding to the domains are summarized in Table 1. Since the $K_b$ values for the binding of both LysM constructs with (GlcNAc)₃ are very small, the other thermodynamic parameters, $\Delta H$, $\Delta S$, and $\Delta G$, could not be calculated. The binding affinities of the LysM domain for the tetrameric and pentameric GlcNAc oligomers are approximately in the millimolar range. The binding process is enthalpically driven, and the entropy loss opposes binding.

**TABLE 1**

| Ligand | n  | $K_b$ | $\Delta H$ | $\Delta S$ | $\Delta G$ |
|--------|----|-------|------------|------------|------------|
|        |    | mol⁻¹ | kcal/mol   | kcal/mol   | kcal/mol   |
| LysM single |     |       | -8.7       | -11.7      | -5.2       |
| (GlcNAc)₃ | 1.1 | 5,310 |            |            |            |
| (GlcNAc)₄ | 1.1 | 1,450 | -8.7       | -14        | -4.5       |
| LysM tandem |   |       | -7.8       | -9.2       | -5.0       |
| (GlcNAc)₅ | 2.3 | 3,740 |            |            |            |
| (GlcNAc)₆ | 2.3 | 1,180 | -7.2       | -9.7       | -4.4       |
would be Cys\textsuperscript{60} with Cys\textsuperscript{104} and Cys\textsuperscript{71} with Cys\textsuperscript{94}. These disulfide bonds were added to the model before refinement, and their geometries were optimized using the Tripos force field. The overall quality of the resulting model was verified by the PROCHECK program. Fig. 4A displays the model of the PrChi-A LysM single domain, and Fig. 4B displays the template structure, which is the MltD LysM domain. The modeled PrChi-A LysM single has a βαβ fold, with a geometry that is almost same as that of the MltD LysM domain.

Identification of the Residues of LysM Single That Interact with (GlcNAc)_5—To identify which residues of LysM single interact with (GlcNAc)_5, we recorded \(^1\)H-\(^{15}\)N HSQC spectra at 30 °C for LysM single in the presence and absence of (GlcNAc)_5. These experiments identified the protein amide \(^{15}\)N and \(^1\)H chemical shifts that are affected by (GlcNAc)_5 and thereby identify the residues that interact directly or indirectly with (GlcNAc)_5 (Fig. 5A). Whether (GlcNAc)_5 was present or not, the LysM single spectra contain well dispersed, intense signals, indicating well defined molecular structures. Certain LysM single resonances shift dramatically when (GlcNAc)_5 is present, presumably because their corresponding protein residues interact with (GlcNAc)_5 in a LysM-(GlcNAc)_5 complex. The backbone amide resonances of the \(^1\)H-\(^{15}\)N HSQC spectra of LysM single in the presence and absence of (GlcNAc)_5 were assigned using the correlations obtained from the CBCA-(CO)NH and HNCA CB spectra. The cross-peaks of a CBCA-(CO)NH spectrum can be used to correlate a residue’s \(^{15}\)N, HN amide resonances to the Cα and Cβ resonances of the preceding residue; a HNCA CB spectrum provides the means to correlate intraresidue \(^{15}\)N, HN, Cα and \(^{15}\)N, HN, Cβ resonances.

The residue assignments are given in Fig. 5A. The cross-peaks of five residues, Ser\textsuperscript{81}, Glu\textsuperscript{86}, Ala\textsuperscript{90}, Cys\textsuperscript{94}, and Asn\textsuperscript{95}, displayed significant chemical shift perturbations when (GlcNAc)_5 was present (Fig. 5A and B). The amide signals of five other residues, Gly\textsuperscript{68}, Thr\textsuperscript{70}, Cys\textsuperscript{71}, Tyr\textsuperscript{72}, and Ile\textsuperscript{74}, were undetectable when the oligosaccharide was present, probably because of extreme peak broadening, which would be indicative of an intermediate binding exchange rate compared with the NMR time scale. The residues with chemical shifts most affected by (GlcNAc)_5 are color-coded in the model of LysM single (Fig. 6A). Gly\textsuperscript{68} is located at the loop that resides between strand 1 and helix 1, and residues Thr\textsuperscript{70}–Ile\textsuperscript{74} are positioned near the N-terminal of helix 1. Glu\textsuperscript{86} is located at the C-terminal of helix 2, and Ala\textsuperscript{90}, Asn\textsuperscript{94}, and Ala\textsuperscript{95} are part of the loop between helix 2 and strand 2. These residues are all found on one side of the molecule and form a long shallow groove at the molecular surface. (Evidence for the groove is given below.)
Docking of (GlcNAc)_n onto the PrChi-A LysM Single Model—Accessible surface calculations on PrChi-A LysM single were performed with the MOLCAD program. The results of these calculations allowed us to identify a few cavities that might be GlcNAc binding sites. Most notably, 1) there is a large cavity that is formed by residues whose NMR resonances are not affected by oligosaccharide binding, but 2) a cavity associated with those residues whose resonance are affected by oligosaccharide binding could not be readily discerned. We therefore performed a second accessible surface area calculation within the region defined by the residues whose resonances are affected by (GlcNAc)_5 binding. A shallow groove delineated by those residues could be discerned, and it was used as the LysM docking surface. Docking simulations were then performed with Surfex-Dock. Docking was initially performed using (GlcNAc)_2 and the model of LysM single. The results of this simulation identified many possible docking sites (data not shown). We then constructed a model of (GlcNAc)_4 using the (GlcNAc)_2 coordinates and docked (GlcNAc)_4 to LysM single. The binding modes of (GlcNAc)_4 for all docking runs are very similar to each other for the portion of the oligosaccharide found to dock at the N-terminal region of helix 1 and the loop between strand 2 and helix 2 (lower region of the molecule in Fig. 6B). However, the binding modes for the oligosaccharide were much more diverse at the top of the groove (Fig. 6B).

Mutational Analysis—Because the binding of carbohydrates to proteins is known to often be mediated by aromatic residues (3), we postulated that mutation of Tyr72, which is one of the residues whose NMR amide signals became undetectable in the presence of (GlcNAc)_5, might disrupt LysM chitin binding and enzymatic antifungal activity of PrChi-A. To establish if the conformations of the wild-type and mutant forms of LysM single are the same, we compared their 1H-15N HSQC spectra. The amide signals of the mutant are as well dispersed as those of the wild type spectrum, and in many cases the cross-peaks of the two spectra overlap (Fig. 7A). Furthermore, the thermostability of LysM single Y72A, as measured by DSC, is the same as that of the wild type. Both the shapes of the thermograms and the T_m values of the two proteins coincide (data not shown). Therefore, LysM single Y72A and the wild type domain have virtually the same fold. Having eliminated the possibility that loss of activity upon mutation could be attributed to a conformational change, we then applied LysM single and LysM single Y72A to the chitin column to evaluate the chitin binding ability of these proteins by comparing their relative elution positions. As shown in Fig. 7B, LysM single was retained on the column, demonstrating its chitin binding ability. Although it was eluted as separate peaks, LysM single was mainly eluted from the column by 1 M acetic acid. On the other hand, approximately
half of the applied Y72A mutant passed through the column, and almost all of the rest of it eluted with 0.1 M acetic acid, indicating that Tyr72 in LysM single is responsible for chitin binding of the LysM domain. We then postulated that since mutation of Tyr72 to Ala disrupts LysM chitin-binding, it might also disrupt the enzymatic antifungal activity of PrChi-A, which is the case (Fig. 7C). The antifungal activity of Δ1LysM-Y72A, Δ1LysM, and CatD (The constructs are shown in the figure.) The test samples (200 or 400 pmol of a purified chitinase in 10 μl of distilled water) were individually placed into wells positioned around a field of T. viride. D, an isothermal microcalorimetric profile of a titration of a 9.65 mM (GlcNAc)₅ solution into a 129.8 μM LysM single Y72A solution. The area of each trough corresponds to the heat generated after the addition of 7 μl of (GlcNAc)₅ to 1.4482 ml of protein solution.

**DISCUSSION**

There is intense interest in LysM domain proteins because of their important roles in chitin oligosaccharide recognition and their widespread occurrence. In higher plants, numerous genes for LysM domain-containing proteins have been identified, although there is only limited information concerning their three-dimensional structures. Therefore, little is known about their structure-function relationships, specifically chitin oligosaccharide binding modes.

**Disulfide Bonds Structurally Stabilize the PrChi-A Domain—** We have shown, using DSC experiments (Fig. 2), that LysM single and LysM tandem are thermostable (i.e. have high Tₘ values). DSC measurements were also made on LysM tandem with the reducing reagents, β-ME and TCEP (35), present (Fig. 2B). The Tₘ value of LysM tandem decreased in the presence of...
The Chitin-binding Site for a LysM Domain

the reducing agents. Our LysM constructs have four cysteine residues per domain. Other chitinases have cysteine residues in their LysM domains, and the positions of these cysteines are conserved with respect to those of PrChi-A (Fig. 1). We assayed for free cysteines within the recombinant domain constructs using 5,5′-dithiobis(2-nitrobenzoic acid) and found none. This finding suggests that all cysteines within the domains form disulfides. Further, it is known that the 13C chemical shifts of protein disulfides and of cysteines found in random coiled structures differ (i.e. Ca at 57.68 ppm and Cβ at 38.38 ppm for disulfides; and Ca at 58.8 ppm and Cβ at 29.75 ppm for cysteines) (38). The difference is especially large for the Cβ chemical shifts. We found that the chemical shift values for the Cβ resonances of Cys69, Cys71, Cys94, and Cys104 of LysM single are 41.54, 41.06, 42.40, and 38.96 ppm, respectively. They are very similar to typical values found for disulfides. Therefore, although we have not directly demonstrated the existence of disulfide bonds in the LysM domains, it is probable that they exist and contribute significantly to the stability of the domains.

The LysM Domain Interacts with (GlcNAc)₅—The LysM domain is found in various enzymes that participate in bacterial cell wall degradation (13). Recently, a chitin oligosaccharide elicitor receptor has been purified, and it contains an extracellular LysM domain (9). Furthermore, a molecular recognition site for lipochitin oligosaccharide signal molecules (Nod factors), which are secreted by symbiotic bacteria, has been characterized recently, and it is a LysM domain (10–12). Based on the results of these reports and others, it appears that the function of LysM domains is to bind peptidoglycans, although how they do so has not been determined. The results of our calorimetric experiments directly show that the PrChi-A LysM domain binds (GlcNAc)₅ (n = 3–5). The thermostability of LysM tandem increases in the presence of (GlcNAc)₅ (n = 3–5) (Fig. 2C). Furthermore, the results of the ITC experiments, which directly measured the binding affinities of the LysM domains for (GlcNAc)₅ (Fig. 3), show that the binding affinities increase as the number of GlcNAc repeats increases. These results suggest that the LysM domain recognizes the central portion of a GlcNAc oligomer rather than its termini. Using the NMR titration and modeling data, PrChi-A LysM single appears to have a shallow groove formed by a cluster of hydrophobic residues and is the oligosaccharide binding site.

Identification of Residues in LysM Single That Are Critical for (GlcNAc)₅ Binding—Although the LysM domain has been identified in many proteins, there is no experimental information about its oligosaccharide binding site. Bateman and Bycroft (13) reported that the loop between strand 1 and helix 1 lies at the end of a shallow groove on the domain surface and suggested that this region is the best candidate for a ligand binding site. Recently, Mulder et al. (15) predicted the most favored binding modes between chitin oligosaccharides and three Nod factor domains. They proposed a consensus model whereby tetrasaccharide interact mostly through van der Waals bonds with the amino acids of the long loop between helix 2 and strand 2. The LysM binding site proposed by Bateman and Bycroft (13) and by Mulder et al. (15) differ, which is not surprising, since neither is derived from experimental data. The interpretation of our NMR experimental data indicates that the (GlcNAc)ₙ recognition site is made out of two LysM regions, one being the N-terminal part of helix 1 and the loop between helix 1 and strand 1 (site 1) and the other being the C-terminal portion of helix 2 and the loop between helix 2 and strand 2 (site 2). These two sites form a long continuous shallow groove on the molecular surface. Ours is the first description of a LysM oligosaccharide-binding site that incorporates experimental knowledge of the binding site’s residues.

As discussed above, the N-acetyl-chitooligosaccharide recognition site can be subdivided in two. After calculating the degree of sequence identity between the LysM site 1 of PrChi-A and each of the other proteins listed in Fig. 1B and averaging the results, it is notable that although residues Gly68–Ile73 are well conserved in the enzymes (averaging about 78% sequence identity), they are less well conserved in the receptor proteins (averaging about 50% sequence identity). On the other hand, we find a larger variation for both the type of residue at a given position in site 2 and the length of site 2. Therefore, site 1 may be the signature sequential feature that distinguishes between the LysM domains of enzymes and receptors. We also found that Tyr72 of PrChi-A is critically involved in (GlcNAc)₅ binding. An aromatic residue, either a tyrosine or a tryptophan, is found in many of the LysM domains that are a part of a chitinase. Therefore, an aromatic residue at the position that corresponds to residue 72 of PrChi-A may be a signature feature of chitinase LysM domains. Although LysM domains of other enzymes, such as MltD, do not have an aromatic residue at the position corresponding to Tyr72, they do have conserved or conservatively replaced residues (e.g. Asp69 and Ile73) that correspond to PrChiA LysM domains whose chemical shifts are affected by (GlcNAc)₅ binding. Therefore, LysM domains associated with such enzymes may also interact with GlcNAc oligomers in the same manner as does the PrChiA LysM domain.

In summary, we clarified the functional and structural features of the LysM domain. PrChi-A LysM domains are very thermostable, which is probably conferred in large part by disulfides. Extrapolating from our data, we expect that, in general, the stoichiometry of LysM domains and bound carbohydrates is 1:1. We also conclude that LysM domains bind chitin oligosaccharides via a shallow groove composed of the N-terminal part of helix 1, the C-terminal part of helix 2, and the loops between strand 1 and helix 1 and between helix 2 and strand 2.

Acknowledgments—We thank Dr. Katsuhide Yutani (RIKEN SPring-8 Center, Harima Institute) for useful advice about calorimetry and for critically reading the manuscript. We thank Toshikazu Kobayakawa, Toshiro Kimura, and Atsushi Midorikawa (World Fusion Co., Ltd.) for helpful discussions.

REFERENCES
1. Gabius, H. J., Siebert, H. C., Andre, S., Jimenez-Barbero, J., and Rudiger, H. (2004) ChemBioChem. 5, 740–764
2. Rudinger, H.-J. G. (2001) Glycoconj. J. 18, 589–613
3. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Biochem. J. 382, 769–781
4. Jerse, A. E., Yu, J., Tall, B. D., and Kaper, J. B. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 7839–7843
5. Birkeland, N. K. (1994) Can J. Microbiol. 40, 658–665
6. Frankel, G., Lider, O., Hershkoviz, R., Mould, A. P., Kachalsky, S. G.,
The Chitin-binding Site for a LysM Domain

Candy, D. C., Cahalon, L., Humphries, M. J., and Dougan, G. (1996) J. Biol. Chem. 271, 20359–20364

7. Pellegrini, M., Marcotte, E. M., Thompson, M. J., Eisenberg, D., and Yeates, T. O. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 4285–4288

8. Ponting, C. P., Aravind, L., Schultz, J., Bork, P., and Koonin, E. V. (1999) J. Mol. Biol. 289, 729–745

9. Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., Minami, E., and Shibuya, N. (2006) Proc. Natl. Acad. Sci. U. S. A. 103, 11086–11091

10. Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., Sato, N., Nakamura, Y., Tabata, S., Sandal, N., and Stougaard, J. (2003) Nature 425, 585–592

11. Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., and Stougaard, J. (2003) Nature 425, 637–640

12. Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003) Science 302, 630–633

13. Bateman, A., and Bycroft, M. (2000) J. Mol. Biol. 299, 1113–1119

14. Bielnicki, J., Devedjiev, Y., Derewenda, U., Dauter, Z., Joachimiak, A., and Derewenda, Z. S. (2006) Proteins 62, 144–151

15. Mulder, I., Lefebvre, B., Cullimore, I., and Imberty, A. (2006) Glycobiology 16, 801–809

16. Wiseman, T. W., S. Brandts, J. F., and Lin, L.-N. (1989) Anal. Biochem. 179, 131–137

17. Bax, A., and Grezesiek, S. (1993) Acc. Chem. Res. 26, 131–138

18. Clore, G. M., and Gronenborn, A. M. (1991) Science 252, 1390–1399

19. Clore, G. M., and Gronenborn, A. M. (1998) Trends Biotech. 16, 22–34

20. Delaglio, F., Grzesie, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biol. NMR 6, 277–293

21. Goddard, T. D., and Kneller D. G., SPARKY 3, University of California, San Francisco

22. Live, D. H., Davis, D. G., Agosta, W. C., and Cowburn, D. (1984) J. Am. Chem. Soc. 106, 1934–1941

23. Wüthrich, K., Billeter, M., and Braun, W. (1983) J. Mol. Biol. 169, 949–961

24. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) Nucleic Acids Res. 24, 4876–4882

25. Topham, C. M., Thomas, P., Overington J. P., Johnson M. S., Eisenmenger F, and Blundell T. L. (1990) Biochem. Soc. Symp. 57, 1–9

26. Clark, M., Cramer, R. D. I., and van den Opdenbosch, N. (1989) J. Comput. Chem. 10, 982–1012

27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 782–791

28. Waldherr-Teschner, M., Goetzte, T., Heiden, W., Knoblauch, M., Vollhardt H., and Buckmann, J. (1992) in Advances in Scientific Visualization, (Hin, A. T. S., ed) pp. 58–67, Springer-Verlag New York Inc., New York

29. Jain, A. N. (2007) J. Comput. Aided Mol. Des. 21, 283–306

30. Kurumbail, R. G., Stevens, A. M., Gierse, J. K., McDonald, J. L., Stegeman, R. A., Pak, J. Y., Gildehaus, D., Miyashiro, J. M., Penning, T. D., Seibert, K., Isakson, P. C., and Stalling, W. C. (1996) Nature 384, 644–648

31. Yamagami, T., and Funatsu, G. (1996) Biosci. Biotechnol. Biochem. 60, 1081–1086

32. Jelesarov, I., and Bosshard, H. R. (1999) J. Mol. Recognit. 12, 3–18

33. Remmel, R. L., and Gombotz, W. R. (2000) BioPharm. 13, 36–45

34. Ellman, G. L. (1959) Arch. Biochem. Biophys. 82, 70–77

35. Han, J. C., and Han, G. Y. (1994) Anal. Biochem. 220, 5–10

36. Pierce, M. M., Raman, C. S., and Nall, B. T. (1999) Methods 19, 213–221

37. Finn, R. D., Benjamin Schuster-Böckler, J. M., Griffiths-Jones, S., Hollich, V., Lassman, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. L., and Bateman, A. (2006) Nucleic Acids Res. 34, D247–D251

38. Wang, Y., and Jardetzky, O. (2002) Protein Sci. 11, 852–861