

博士學位論文

Structure-function relationships of lysin motifs

derived from an alga, *Volvox carteri*

平成 30 年 1 月 10 日

近畿大学大学院

農学研究科 バイオサイエンス専攻

北奥 喜仁

ABBREVIATIONS

(GlcNAc)_n	β-1,4 linked N-acetyl-D-Glucosamines with DP = n
A₂₈₀	Absorbance at 280 nm
BMRB	Biological Magnetic Resonance Bank
CBM	Carbohydrate-Binding Module
CD	Circular Dichroism
GH	Glycoside Hydrolase
HIC	Hydrophobic Interaction column Chromatography
HSQC	Heteronuclear Single Quantum Coherence transfer
IC₅₀	Inhibition Constant at 50 % inhibition activity
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
LB	Luria-bertani Broth
NMR	Nuclear Magnetic Resonance spectroscopy
O.D.₆₀₀	Optical Density at 600 nm
DP	Degree of polymerization
RMSD	Root Mean Square Deviation
SDS-PAGE	Sodium Dodecyl Sulfate Poly-Acrylamide Gel Electrophoresis
SEC	Size-Exclusion column Chromatography

CONTENTS

CHAPTER I	Introduction	1
CHAPTER II	LysM from an alga, <i>Volvox carteri</i>	12
CHAPTER III	Structure-function relationship of LysM1 and LysM2 from <i>VcChi</i>	28
CHAPTER IV	Concluding remarks	51

REFERENCES

ACKNOWLEDGEMENTS

Chapter I

Introduction

Chitin

Chitin is a secondary most abundant polysaccharide next to cellulose, and a linear polysaccharide, which is composed of *N*-acetylglucosamine (GlcNAc) linked by β -1,4-glycosidic linkages (Fig.I-1). Chitin and the *N*-deacetylated form, chitosan, naturally occur in exoskeletons of crustaceans, cuticle of insects, fungal cell wall and alga. Chitin and modified chitinous materials play roles of structural scaffold of tissue cells and support the structure against mechanical stresses.

Crystalline chitin found in nature takes different three forms having same chemical structures, but having different packing in the crystals called α -, β - and γ -chitin (Jang et al. 2004). α -chitin, which is mainly found in exoskeleton of crustaceans, has anti-parallel arrangement of single chitin chains (Carlstrom 1957). β -chitin, which is found, for example, in pen of *Loligo*, has parallel arrangement of single chitin chains (Blackwell 1969). γ -chitin is relatively minor form of chitin, which can have been found in cocoon fiber of *Ptinus* beetle and stomach of *Loligo* (Jang et al. 2004). γ -chitin is composed of two parallel and one anti-parallel chitin chain. Anti-parallel arrangement of α -chitin gives rise to the close packing of each chain and α -chitin as the most stable structure among these three.

Since fungal cell walls contain chitin to some extents, the degradation products, chitin-oligosaccharides, play roles on plant defense responses against pathogenic fungi (Kim et al. 2000). Similar interactions have been reported for plant-bacteria interactions through the cell wall component, peptidoglycan, and the degradation products (Willmann et al. 2011). Peptidoglycan is a co-polymer of *N*-acetylglucosamine and *N*-acetylmuramic acid cross-linked by oligopeptides. Cell wall fragments released by constitutively expressed lytic enzymes are perceived by receptors initiating intracellular signaling cascade and bringing about the change in expression level of genes, ion flux, accumulation of callose and etc (Vidhyasekaran 2014). Taken together, chitin and the derivatives intermediate the processes on plant-microbe interactions. Recently, Chitin- and chito-oligosaccharides, which are oligomer of chitin and chitosan, respectively, are reported to help wound healing and to have antimicrobial activity. As such, these oligosaccharides are gaining interests attribute to the potential for medical applications.

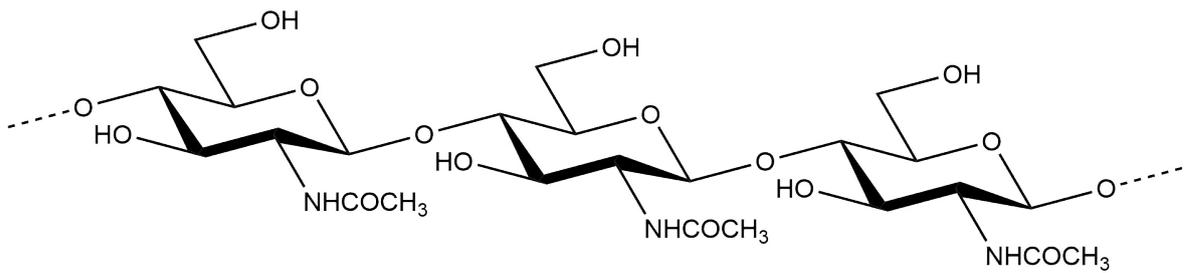


Figure I-1. Chemical structure of chitin.

Chitinases

Chitinases (E.C. 3.2.1.14) are hydrolytic enzymes, which degrade chitin. Chitinases can be classified into two families, glycoside hydrolases family 18 (GH18) and family 19 (GH19), based on their amino acid sequences (CAZy: <http://www.cazy.org/Glycoside-Hydrolases.html>). GH18 chitinases are commonly found in wide range of organisms from bacteria to vertebrates. On the other hand, GH19 chitinases are mainly found in plants and only limited species of bacteria. GH18 and GH19 do not share any sequence similarity and the three-dimensional (3D) structures are completely different from each other.

Two groups reported the first 3D structures of GH18 chitinases in the journal, “Structure”, in 1994 as contiguous two reports (Perrakis et al. 1994; Terwisscha van Scheltinga et al. 1994). One of them was from a gram-negative bacterium, *Serratia marcescens*, and the other was from a rubber tree, *Hevea brasiliensis*. These reports showed that GH18 chitinases have $(\beta/\alpha)_8$ -barrel fold (Fig. I-2). As clearly seen from the figure I-2, GH18 chitinases are varied in domain composition. Some of them have an additional ligand-binding domain, and the others do not. Some of them have an insertion domain composed of α/β secondary-structural units, and the others do not. Ohnuma *et al.* suggested that the variation in the domain composition is derived from the natural selection during the molecular evolution processes (Ohnuma T. et al. 2014).

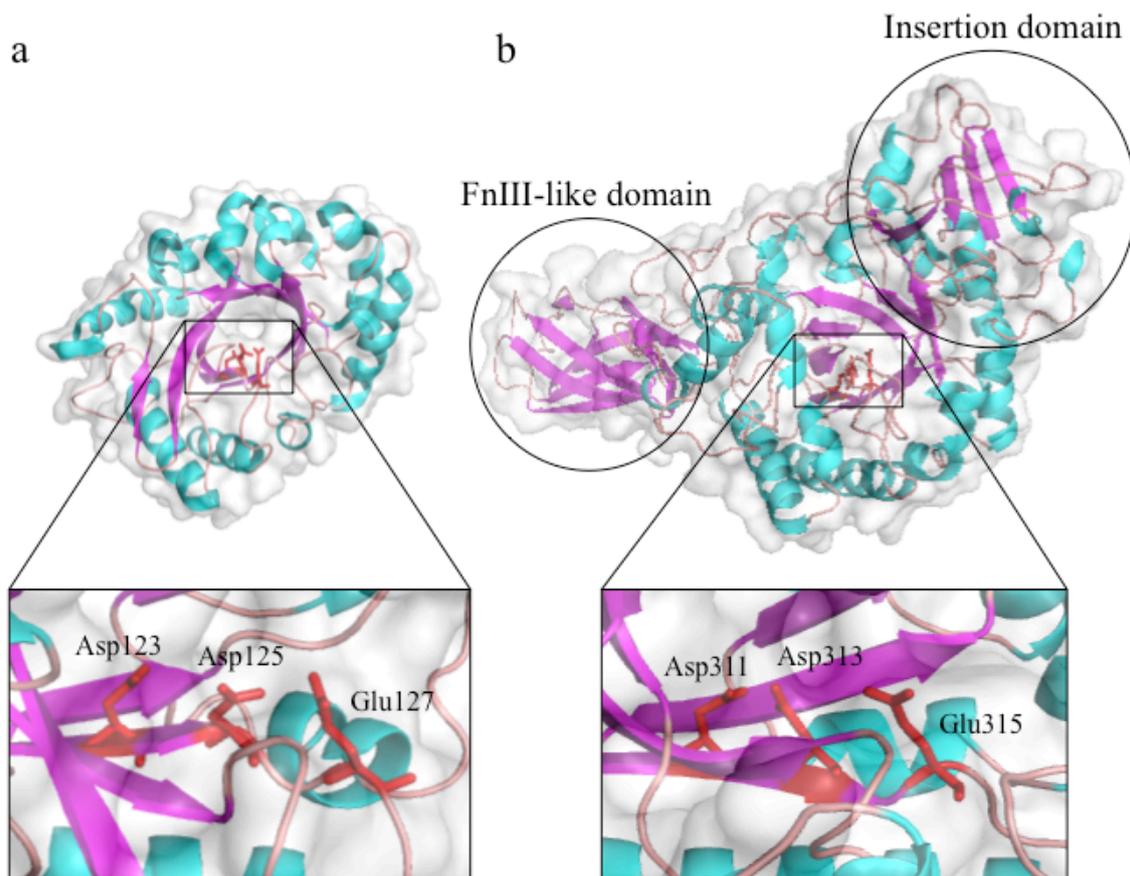


Figure I-2. Crystal structures of hevamine and SmChiA. β -strands are colored as magenta, α -helices are as cyan and loop regions are colored as pink. A catalytically important DxDxE motif is indicated by red and the side chains are depicted as stick model.

The first 3D structure of GH19 chitinases was reported in 1993 and was from barley, *Hordeum vulgare*. This report showed that GH19 chitinases have α -rich lysozyme-like fold (Fig. I-3) (Hart et al. 1993). The structure of this barley chitinase was composed of two lobes containing six loop structures. Based on the nomenclature of the previous reports, these loop structures can be denoted as loop I-V and C-terminal loop (Taira et al. 2011). Since chitinases lacking some of these loop structures had been reported, the barley chitinase was designated as a “loopful” chitinase in the previous reports (Ohnuma et al. 2013). The loop structures have been reported to participate in substrate binding. However, the functionalities of each loop structure have not yet been fully understood.

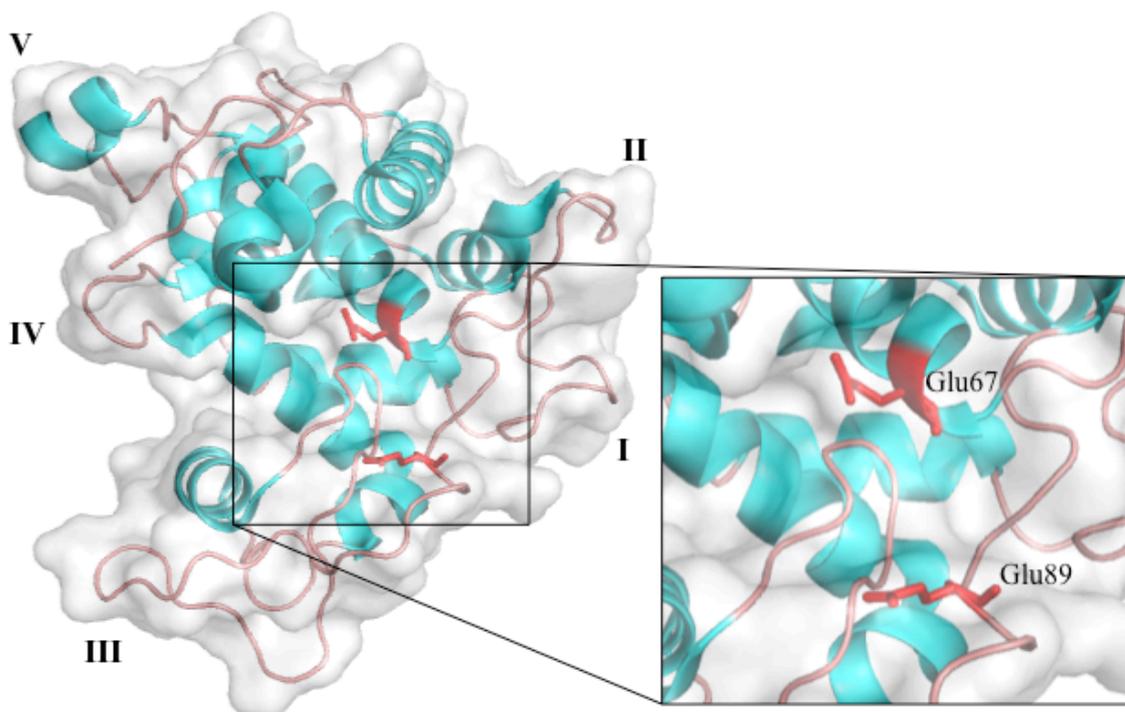


Figure I-3. The first GH19 chitinase structure (PDB ID: 2BAA). The loop structures (loop I-V) are indicated by roman numerals. Catalytic acid and base residues are colored as red and the stick models of the side chains are depicted.

Carbohydrate binding modules

Carbohydrate binding modules, CBMs, are classified into 83 families in CAZY database (<http://www.cazy.org/>) to date, based on the amino acid sequence similarity. CBMs can also be classified into three types, i.e. Type A, B and C, based on the shape of binding site (Boraston et al. 2004). Type A CBMs have flat ligand-binding site, which is mainly composed of aromatic side chains and are specific to insoluble crystalline ligands. Prevalence of aromatic side chain in ligand binding site is common for carbohydrate binding proteins. However, the platform-like architecture is characteristic for this type of CBMs. This planer architecture of the Type A CBMs is thought to be complementary to the planer architecture of crystalline polysaccharides to which the Type A CBMs are specific. Since aromatic side chains comprising carbohydrate-binding surface are thought to form CH- π stacking interactions with pyranose rings of sugar moieties of crystalline polysaccharides, complementarity of the surface architecture is advantageous to simultaneously form two or more interactions

(Ramírez-Gualito et al. 2009). This rather promiscuous binding mode sometimes enables sliding of ligand upon degradation of substrate by the attached catalytic domain.

Type B CBMs have binding sites, which can accommodate a single glycan chain (Boraston et al. 2002). Hence, the Type B CBMs have cavities to accommodate glycan chain, which are often described as grooves or clefts. Type B CBMs generally show significant affinity to oligosaccharides with degree of polymerization (DP) of four or larger, but show weak or negligible affinity to oligosaccharides with DP of three or shorter. In contrast to the cases of Type A CBMs, binding-sites of Type B CBMs are generally composed of several ligand-binding subsites, where individual subsites accommodate monosaccharide units of polymeric ligands. Thus, the binding affinity elevates as DP become longer. Aromatic side chains also play important roles in ligand binding, as in the case of Type A CBMs. In contrast to Type A CBMs, however, a key role of direct hydrogen bonds upon ligand binding is characteristic for Type B CBMs.

Type C CBMs lack extended binding site and accommodate only mono-, di- and tri-saccharides (Shinya et al. 2013). Distinction between the Type B and C is subtle as some CBM has both characteristics. As in Type B CBMs, a hydrogen-bonding network plays a key role on binding but seems to be more intensive.

Lysin motif

In 1986, Garvey and coworkers reported C-terminal repetitive sequence in a viral lysin (Garvey et al. 1986). Nowadays, this sequence is known to act as a binding module and reported to exist in various kinds of organisms ranging from bacteria to human (Buist et al. 2008). This protein module is called lysin motif or its abbreviation, LysM, as many of them had been found in lytic enzymes.

Bateman and Bycroft reported the first solution structure of LysM of bacterial lysin (MltD) using NMR in 2000, revealing that LysM has a core structure comprising ca. 40 amino acids and folded into antiparallel two stranded β -sheet connected to two α -helices on the opposite side, so called $\beta\alpha\alpha\beta$ -fold (Bateman and Bycroft 2000). Six years later, Bielnicki and coworkers reported the first crystal structure and verified that the $\beta\alpha\alpha\beta$ -fold is shared among LysMs (Bielnicki et al. 2006). Ohnuma and coworkers analyzed the thermodynamic character and the binding site of chitin oligosaccharides on LysM of a chitinase derived from a fern, *Pteris ryukyuensis* and identified the binding

cleft located on the α -helical surface based on NMR chemical shift perturbation and docking simulation (Ohnuma *et al.* 2008). Liu *et al.* reported crucial evidence for binding mode of LysM by showing crystal structure of LysM of Chitin Elicitor Receptor Kinase (CERK1) in complex with (GlcNAc)₄ (Liu *et al.* 2012). In this complex structure, binding cleft extends from the loop structure connecting β -sheet and α -helices to the end of α 1. As Ohnuma *et al.* showed the LysMs acting as subdomains in Carbohydrate-Active enZymes (CAZymes), LysM is classified as CBM50 in CAZY database.

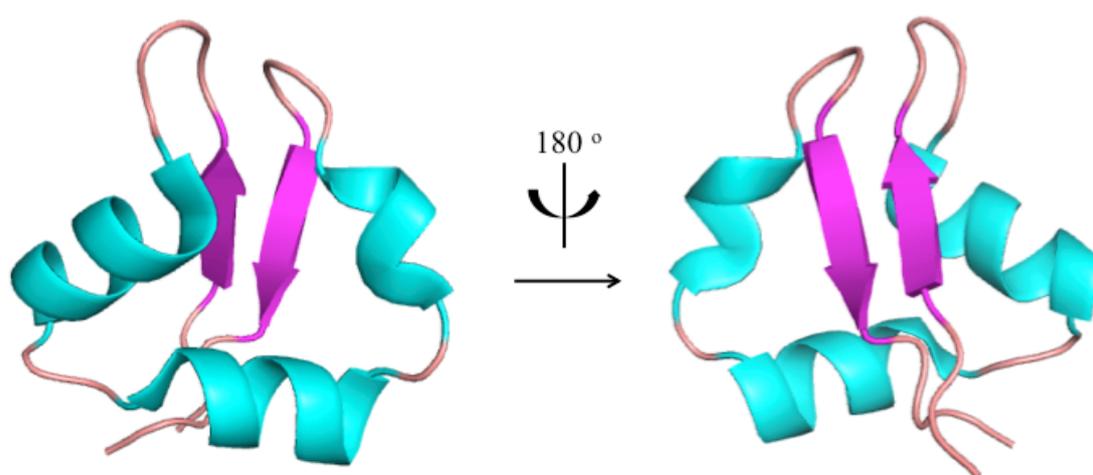


Figure I-4. The first structure of LysM determined by NMR. β -strands are colored as magenta, α -helices are as cyan and loop regions are colored as pink. The right figure is a 180° rotated image of the left.

Plant chitinases

Based on the structures and domain compositions, plant chitinases were classified by a different system from CAZY. According to the review of Taira (Taira 2010), plant chitinases are classified into five classes and two additional subclasses, class I-V and class IIL, class IIIb, respectively. Class I, II, IIL and IV belong to the GH19 family, though Class III, IIIb and V belong to the GH18 family.

Class I and II chitinases have homologous catalytic domain, though Class I chitinases have additional CBM18 domain. In analogy to Class I and II chitinases, Class

IV and Class IIL chitinases also share analogous catalytic domains, but have different domain structures. Class IV chitinases have additional CBM18 domain like Class I chitinases, while Class IIL chitinases are composed of a single domain like Class II chitinases.

Class III and IIIb chitinases share similar domain composition, but have poor sequence similarities to each other (<20 %). Class IIIb chitinases lack disulfide bridges conserved in Class III chitinases, but have an inserted loop structure comprising substrate-binding site. Class V chitinases lack disulfide bridges but have insertion domain mentioned above in the section of “Chitinases”.

LysM containing plant chitinase, PrChiA

PrChiA is a chitinase derived from a fern, *Pteris ryukyuensis*. Onaga *et al.* first reported the cloning of the gene, construction of the heterologous expression system in *E.coli* and characterization of the recombinant PrChiA in 2008. They showed that the chitinase is composed of N-terminal two LysM domains (PrLysM1, PrLysM2) linked in tandem (PrLysM tandem) and C-terminal Class IIIb catalytic domain (PrChiA cat). Each domain is linked by proline, serine and threonine-rich linker region. Degradation pattern of chitin oligosaccharide by PrChiA was characteristic to the Class IIIb chitinases, i.e. typical Class III cleaved (GlcNAc)₅ mainly into (GlcNAc)₄ and GlcNAc though PrChiA cleaved (GlcNAc)₅ mainly into (GlcNAc)₃ and (GlcNAc)₂ as was reported for previous studies about PrChiA (Onaga and Taira 2008). Sequence analysis of PrChiA also showed strong homology to Class IIIb chitinases, but showed weak homology to Class III chitinases. They also performed LysM truncation experiments and revealed the role of LysM domains of PrChiA in this chitinase action. Truncation of LysM domains decreased not only chitin-binding affinity, but also chitin-degradation ability of PrChiA. Truncation of LysM and mutation in catalytic center got rid of antifungal activity of PrChiA indicating that the LysM domain and catalytic activity is essential for exhibiting the antifungal activity.

Ohnuma *et al.* cloned the genes of PrLysM2 and PrLysM tandem, constructed the heterologous expression system in *E.coli* as single domain proteins and characterized based on spectroscopic and thermodynamic methods, revealing that the binding mode and thermodynamics of PrLysM2 toward chitin oligosaccharides. Although several

reports about PrChiA had been elucidating enzymatic character and the binding mode of attached LysM domain, structure-function relationship of PrChiA cat have not yet been clarified. Hence, we reported the structure and function of PrChiA cat at molecular level.

Firstly, we determined the crystal structure of PrChiA cat and compared the structure with that of a Class III chitinase from *hevea brasiliensis*, hevamine. Comparison of structures of these two chitinases is shown in the Fig. I-5. Detailed inspection revealed the differences in orientations of $\alpha 1$, $\alpha 5$ and $\alpha 7$ and also in the length of $\beta 6$. The most striking difference of PrChiA was the extruded loop structure and an additional antiparallel β -sheet structure supporting the loop structure, which was not found in Class III chitinases. Chitin oligosaccharide degradation experiments showed that PrChiA cat degrade chitin oligosaccharides in a similar manner to those of the other Class IIIb chitinases, which were different from the degradation profiles of Class III chitinases suggesting that the substrate recognition mode of Class IIIb chitinases are clearly different from those of Class III chitinases. Thermal unfolding experiments in the presence and absence of chitin oligosaccharides and titration experiments by isothermal titration calorimetry (ITC) using inactive mutants of PrChiA cat showed it can bind to (GlcNAc)₃ and longer, but not to (GlcNAc)₂. Thermodynamic analysis of binding showed that the favorable enthalpy drove the binding to (GlcNAc)₃ and (GlcNAc)₄ accompanied with a smaller entropy loss. We also analyzed binding of allosamidin, a well-known GH18 inhibitor, to PrChiA cat based on ITC titration experiment, inhibition assay and docking simulation. These results showed that PrChiA bound allosamidin with relatively weak affinity and was comparable to chitin oligosaccharides. Inhibition assay of allosamidin upon degradation of (GlcNAc)₄ showed more than 10-fold higher IC₅₀ value (172±16 μ M) than typical GH18 chitinases (0.04-10 μ M). Docking simulation showed that the binding site of PrChiA to allosamidin was similar to hevamine's binding site (Fig. I-6). The extruding loop region, which is absent in hevamine, seems to make a wall on the binding site. This may interfered the correct binding to allosamidin and resulted in weaker binding affinity than typical GH18 chitinases.

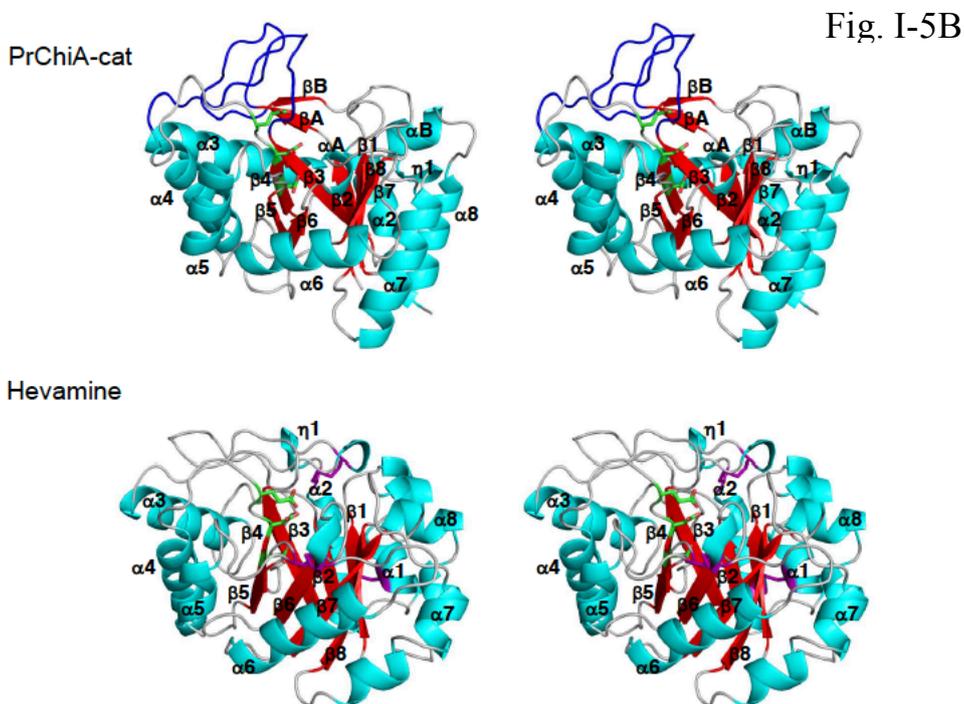
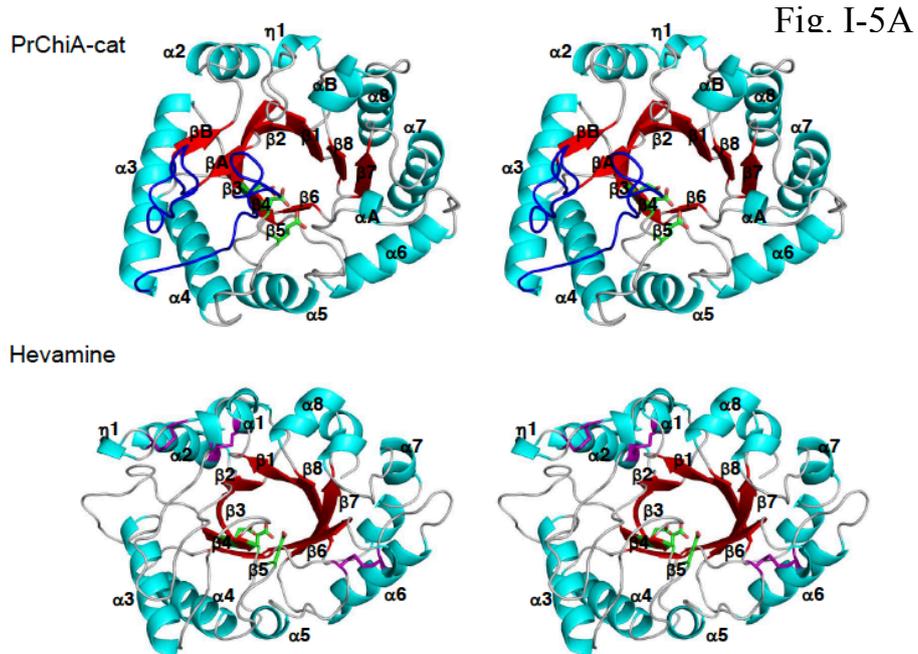


Figure I-5. Stereo views of the crystal structures of PrChiA-cat and hevamine. A: views from the top of the $(\beta/\alpha)_8$ barrel. α -Helices and a 3_{10} helix are colored in cyan, while β -strands are colored in red. The catalytic motif DxDxE (Asp243-Asp245-Glu247) is highlighted by a stick model colored in green. The disulfide bonds found in hevamine are colored in purple. B: Views from the side of the $(\beta/\alpha)_8$ barrel. Loop structures found only in PrChiA are colored in blue.

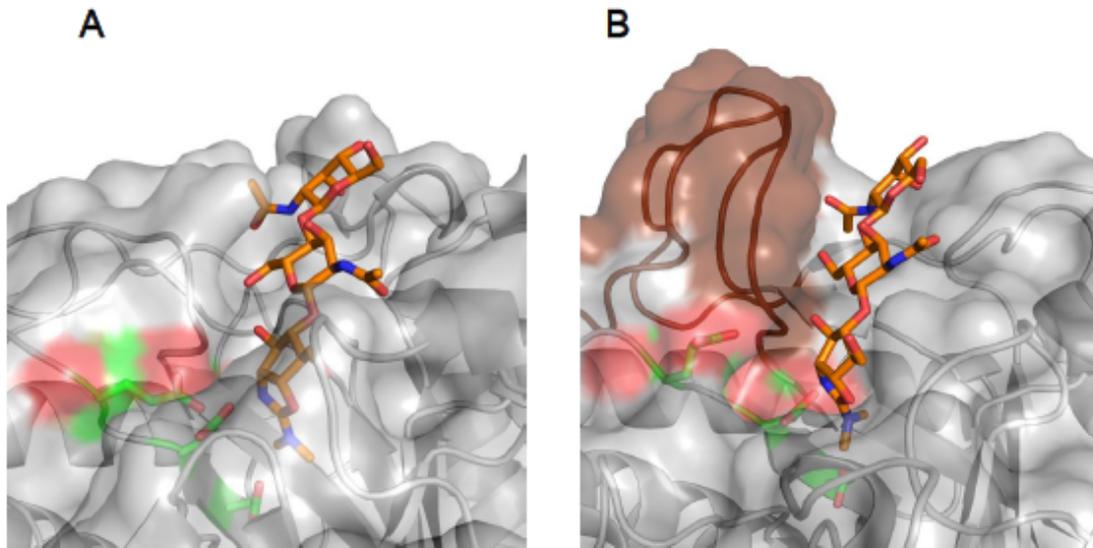


Figure I-6. A: Crystal structure of hevamine in complex with allosamidin (PDB code: 1LLO). B: Modeled structure of PrChiA-cat in complex with allosamidin. A docking simulation was performed using AUTODOC software.

Volvox carteri

Volvox carteri is a spherical multicellular green alga comprising many small *Chlamydomonas*-like biflagellate somatic cells with few large non-motile reproductive cells called gonidia at surface area and filled with Extra Cellular Matrix, ECM (Kirk 2004). Somatic cells form monolayer surface and face the biflagellate side out and cooperatively work to swim with a characteristic rolling motion. Gonidia just lie under the layer of somatic cells and the layer comprising somatic cells and gonidia is called cellular zone in the ECM. *V. carteri* reproduces in an asexual reproduction cycle in haploid and a sexual life cycle is normally not utilized. Single cycle of asexual reproduction process takes 48 h and is controlled by light and dark periods. On the other hand, sexual life cycle is used only for making dormant, diploid zygotes, which can overcome harsh dryness in summer hot. Sexual reproduction process is under control of the sex-inducing pheromone, which is one of the most competent pheromone working at sub-femto molar (Starr and Jaenicke 1974).

In 1986, Kirk *et al.* reported the sexual reproduction is elicited by a heat shock (Kirk *et al.* 1986). Although the molecular mechanism of the induction of the sexual

pheromone is still unclear, the mechanism of pheromone induced activation process of the sexual reproduction was suggested by Sumper *et al.* and is as follows; Positively charged pheromone is attracted by the negatively charged ECM and accumulated on the surface of the spheroid. Somatic cells exposed on the spheroids detect pheromone molecule and induce secretion of a protein called pherophorin II, which have 42 kDa fragment in the sequence and homologous to pheromone. Secreted pherophorin molecule is proteolytically cleaved and releases the homologous region to the pheromone into ECM. This process further amplifies the signal and the accumulated signaling molecules induce cleavage of gonidia and initiate development of reproductive sexual cells.

A chitinase derived from *V. carteri*, VcChi

Amon and coworkers cloned the genes of two sex-inducing pheromone inducible proteins in 1998 (Amon et al. 1998). Both clones were multi-modular and had N-terminal repetitive sequences homologous to chitin-binding proteins. One of these two clones showed homology to protease family. On the other hand, the other clone didn't show any homology to any other protein families, but the N-terminal repetitive sequence showed similarity to repetitive sequences attached to chitinases and lysozymes including the lysozyme from *Bacillus* phage ϕ 29, i.e. LysM. This fact led them to perform chitinase assay and showed the activity on glycol chitin.

There had been no reports following result above and no detailed analysis of this chitinase. Hence, we decided to study on this protein. Initially, we tried to characterize the N-terminal LysM domains in depth. Based on previous reports about LysMs from proteins with different functionalities, functionality of VcChi was discussed.

Chapter II

LysM from an alga, *Volvox carteri*

Introduction

LysM is a carbohydrate-binding module family, which can bind to saccharides containing GlcNAc as the component, such as chitin, peptidoglycan, their oligomers, lipo-chitooligosaccharides and etc. LysMs are composed of ca. 50 amino acid residues having canonical $\beta\alpha\beta$ fold comprising anti-parallel 2-fold β -sheet on one side and triangular two α -helices and loop on the other side. The first complex structure of LysM binding to chitin oligomer was reported on 2012. (T. Liu *et al.* 2012). Chitin tetramer was accommodated by the shallow cleft on the α -helical surface.

LysMs have been found in various organisms ranging from bacteria to human. Since LysMs are specific to chitin and the related compounds, which are found as a structural element of cell walls, functionalities of LysM attached proteins involved in degradation of bacterial and fungal cell walls and perception of molecular signals have also been reported so far.

Amon *et al.* reported that *Volvox carteri* secretes a chitinase containing LysMs (VcChi) into ECM (Amon *et al.* 1998). We decided to characterize the LysMs attached to the VcChi (Fig. II-1). In this study, functionality of second LysM from the N-terminus (VcLysM2) was discussed.

(A) Chitinase from *Vovox carteri*



(B) Amino acid sequences of VcLysM1 and VcLysM2

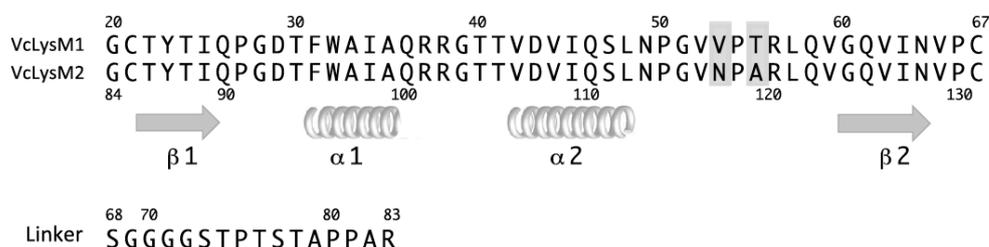


Fig. II-1 Domain organization of a chitinase from *Volvox carteri* (a) and amino acid sequences of LysMs and the linker region (b). Secondary structures of VcLysM2 revealed by the crystal structure are aligned as $\beta 1$, $\alpha 1$, $\alpha 2$, and $\beta 2$ below the sequence. Amino acids shaded in grey are amino acid substitutions between VcLysM1 and VcLysM2.

Materials and Methods

Plasmid construction

VcChi LysM2 gene sequence was optimized and the gene was synthesized by the gene synthesis service of Thermo Fisher Scientific Inc. Yokohama, GeneArt, for *E. coli* expression system. The optimized gene fragment was amplified by polymerase chain reaction using forward primer (5'-ATGGGCTGCACATATAACAATTCAACCG-3') and reverse primer (5'-TTAACACGGCACGTTGATCACCTGG-3'). The amplified gene fragment was inserted into pETBlue-1 AccepTor vector (Novagen) by TA-cloning according to the instruction manual. The recombinant plasmid was designated pET-VcLysM2.

Site-directed mutagenesis

Mutation at Trp96 to Ala was introduced by Quikchange site-directed mutagenesis kit (Stratagene). Forward (5'-GGGTGACACATTCGCGGCGATCGCGCAG-3') and reverse (5'-CTGCGCGATCGCCGCGAATGAGACACCC-3') (mutation sites should be underlined) primers were used to amplify the mutated pET-VcLysM2 gene. The mutated sequence was verified by sequencing. Mutated plasmid was named

pET-VcLysM2 W96A.

Protein expression and purification

SHuffle T7 (NEB) harboring *placI* and pETBlue-1 containing VcLysM2 gene (pET-VcLysM2) and the mutant were used to express VcLysM2. Cell culture was prepared in 1 L of Luria-Bertani Broth (LB) media containing 50 µg/mL ampicillin sodium. Cells were grown at 37 °C until it reaches $O.D._{600} \approx 0.6$. Then, protein expression was induced by adding 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and the cell culture was shaken at 25 °C for 24 h. Cell culture was collected and centrifuged at 6,500 rpm at 4 °C for 15 min. Cell pellets were re-suspended in 20 mM Tris-HCl, pH 7.5 and lysed by sonication. Cell lysate was disrupted by sonication and the supernatant was used for further purification.

Protein solution was dialyzed against 10 mM sodium acetate, pH 4.0 for overnight. The precipitation was removed by centrifuge at 12,000 rpm and 4 °C for 15 min. pH of the supernatant was adjusted to 5.0 by dialyzing against 10 mM sodium acetate, pH 5.0. To perform hydrophobic interaction column chromatography, 1.5 M ammonium sulfate was directly dissolved in the protein solution. The sample was applied to TOYOPEARL Butyl-650 M (TOSOH) equilibrated with the dialysis buffer with the same concentration of ammonium sulfate to the sample. Protein adsorbed column was washed twice with the equilibration buffer. Protein adsorbed on the column was eluted by the dialysis buffer without salt. Eluted protein was fractionated into 2 mL fractions. Protein amounts of fractions were estimated based on A_{280} . Fractions containing protein were collected and applied to size-exclusion column chromatography (SEC). Hiprep Sephacryl S-100 HR column equilibrated with 10 mM sodium acetate, pH 5.0, 0.1 M sodium chloride was used for SEC. After loading the sample, protein was eluted isocratically by the equilibration buffer. Purity of the protein was verified by tricine SDS-PAGE.

Crystallization and NMR sample was prepared using non-labeled and labeled M9 media instead of LB. Purification was performed in a similar manner to that used for other protein preparations.

Crystallization and data collection

Purified VcLysM2 was dialyzed twice with distilled water. Sample was concentrated to 6.0 mg/mL by Amicon Ultra-15 (Merck Millipore). Crystallization condition was screened at 480 conditions from Crystallization screens (Hampton research). Crystallization drops were composed 1 μ Ls of protein sample and reservoir solution. Droplets were set up on each stage of 96-well crystallization plate (Crystalquick plate, Greiner bio-one) and equilibrated against 90 μ L of reservoir solution at 293 K. X-ray diffraction experiments were performed at the beamline, BL-17A of Photon Factory (Tsukuba, Japan). Crystal grew in 0.2 M sodium citrate and 20 % PEG 3,350 was cryoprotected with 0.08 M sodium citrate, 10 % PEG 3,350 and 20 % ethylene glycol and mounted within a nylon loop. Mounted crystal was set on the goniometer, frozen by flash cooling and irradiated by X-ray with 0.98 Å wavelength. Diffraction data set was collected with PILATUS3 S 6M (DECTRIS) under the 0.1 s exposure of X-ray beam during 0.1 ° oscillation of crystal. Diffraction images were analyzed by HKL2000 (Otwinowski and Minor 1997). Statistics of diffraction data were summarized in Table II-1.

Determination and refinement of crystal structure

Molecular replacement was exploited to make the initial phase estimation using Phaser (McCoy et al. 2007) in CCP4 program suite (Collaborative Computational Project, Number 4 1994). A-chain of crystal structure of a LysM domain of PrChiA (PDB ID, 4PXV) was used as the search model. Refinement was performed using REFMAC (Murshudov et al. 1997) and Phenix.refine (Afonine et al. 2012). Molecular model was modified each step of refinement using COOT (Emsley and Cowtan 2004). Finally, $R_{\text{work}}/R_{\text{free}}$ reached to 13.9/18.0 for the structure of VcLysM2 with the highest resolution of 1.20 Å. On the Ramachandran plot, dihedral angles of 97.9, 2.1 and 0 % of the residues were at favored, allowed and disallowed regions, respectively. Refined structure was deposited in PDB under the PDB code, 5K2L. Refinement statistics are summarized in Table II-1.

Table II-1 Data collection and refinement statistics.

Data collection	
Space group	<i>P4₁2₁2</i>
Cell parameters	
<i>a, b, c</i> (Å)	29.1, 29.1, 94.3
<i>α, β, γ</i> (°)	90, 90, 90
Wavelength (Å)	0.98
Resolution (Å)	50.0–1.20 (1.22–1.20)
<i>R</i> _{merge}	0.073 (0.118)
<i>I</i> / <i>σI</i>	73.0 (27.7)
Completeness (%)	99.5 (98.7)
Redundancy	23.5 (22.0)
Refinement	
Resolution (Å)	1.20
No. of reflections	12,739
<i>R</i> _{work} / <i>R</i> _{free} (%)	13.9/18.0
No. of atoms	
Protein	403
Ligand/ion	4
Water	89
RMS deviation	
Bond length (Å)	0.0259
Bond angles (°)	1.9986

Sequential assignment of backbone and partial side chain resonances

0.2 mM of ¹³C, ¹⁵N-labeled VcChi LysM2 dissolved in 20 mM sodium acetate, pH 5.0 (90 % H₂O/10 % D₂O) were used to obtain NMR spectra. Two-dimensional (1H-15N HSQC) and three-dimensional (HNCO, HNCACO, HNCACB, CBCACONH) spectra were measured at 300 K on Bruker AVANCE III 500 MHz spectrometer equipped with a triple-resonance pulsed-field-gradient cryo-probe head under the control of TopSpin 2.1. Raw FID data were processed on NMRPipe (Delaglio et al. 1995) and the resulting spectra were analyzed by Sparky. Assignment data for VcLysM2 were deposited in the BMRB database as the accession code, 12002.

Chemical shift perturbation upon binding to (GlcNAc)_n (n = 3-6)

(GlcNAc)₃ (0-4.8 mM), (GlcNAc)₄ (0-0.7 mM), (GlcNAc)₅ (0-0.6 mM) and (GlcNAc)₆ (0-0.5 mM) were titrated into ¹⁵N-labeled VcLysM2 (62 μM) dissolved in the same buffer used for resonance assignment. Chemical shift perturbation was observed on ¹H-¹⁵N HSQC spectra and the degree of signal migration, Δδ, for each residue was calculated by the following equation,

$$\Delta\delta = |\Delta\delta_H| + |\Delta\delta_N/5|$$

where Δδ_H and Δδ_N represent chemical shift migration on ¹H- and ¹⁵N-axis, respectively. Amino acid residues corresponding to the signals migrated more than Δδ = 0.25 or broaden and hard to be followed were mapped on the surface model of the crystal structure of VcLysM2 to estimate the ligand binding sites.

Thermodynamics of VcLysM2 and W96A mutant

VcLysM2 (130 μM) and VcLysM2-W96A (112 μM) dissolved in 20 mM sodium acetate, pH 5.0 were used for titration experiments using iTTC200 (MicroCal) controlled by the operating software. Experimental data were analyzed by Origin 7. (GlcNAc)_n (n = 2-6) dissolved in the dialysis buffer used to prepare protein samples were titrated into the protein solution in the sample cell (202.8 μL). Energy applied to keep constant the temperature of cells in the adiabatic jacket was recorded every five seconds and were plotted against the experimental time (μcal/sec). Thermogram obtained was baseline-subtracted, and the individual responses were integrated and plotted against the molar ratio of ligand to protein. Using One Set of the Sites model defined in the Origin software, thermodynamic parameters were optimized by fitting theoretical curves to the experimental data.

Results

Crystal structure of VcLysM2

Crystal structure of VcLysM2 was successfully determined at 1.2 Å resolution. One VcLysM2, eighty-nine water molecules and one ethyleneglycol molecule were found in the asymmetric unit. VcLysM2 adopted $\beta\alpha\alpha\beta$ fold canonical for LysM domains as is shown on the Fig. II-2. The first strand (β 1), the first helix (α 1), the second helix (α 2) and the second strand (β 2) were ranging from Thr86 to Ile89, Phe95 to Ala99, Val106 to Leu112 and Gly124 to Asn128, respectively. A disulfide bridge formed between Cys85 and Cys131 tied N- and C-termini. On the surface model of the structure, we found a shallow cleft comprising the loop between β 1 and α 1, α 2 and the loop between α 2 and β 2. Ethyleneglycol molecule was accommodated in a cavity on the shallow cleft and forming indirect hydrogen bonding to VcLysM2 molecule through water molecules.

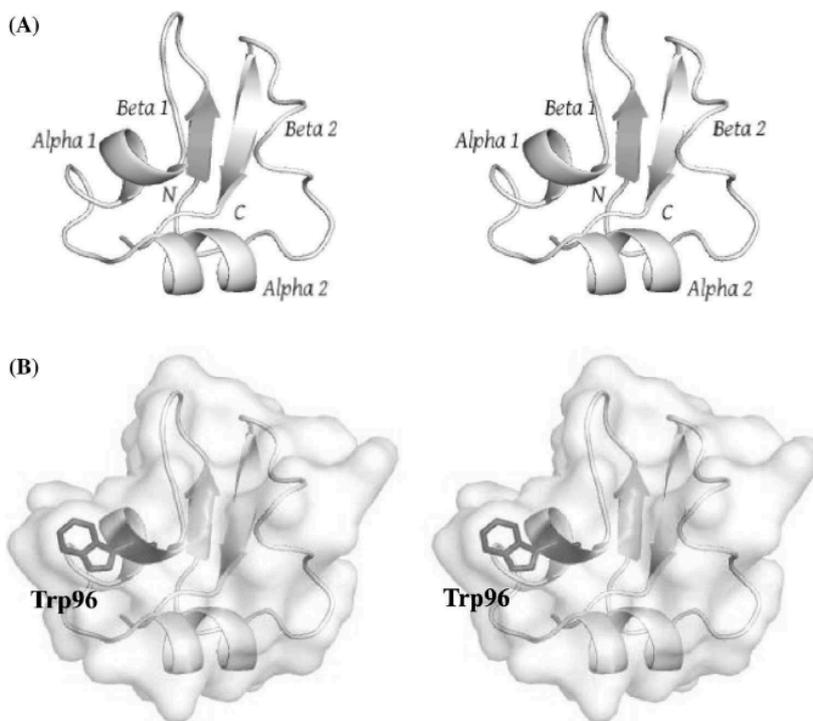


Fig. II-2 Stereo views of the crystal structure of VcLysM2. Ribbon representation of the main chain structure (a) and the surface model of the entire structure (b). The Trp96 side chain involved in GlcNAc residue binding is highlighted by a stick model.

To determine the (GlcNAc)_n interaction site, we attempted to co-crystallize VcLysM2 with ligand or to soak the crystal with ligand solution. However, the trials were unsuccessful. Since VcLysM2 molecules were tightly packed in the asymmetric unit, we concluded that the tight packing of VcLysM2 molecules resulted in the difficulties to obtain ligand bound crystals. Therefore, we tried to estimate the ligand-binding site by NMR.

Backbone resonance assignments

As is shown in Fig II-3 (red), the signals of VcLysM2 on ¹H-¹⁵N HSQC spectrum are fully separated. It helped us to almost completely assign backbone and C_β resonances. Finally, 90 % of the backbone ¹H_N and ¹⁵N, 100 % of the ¹³C_α, 100 % of the ¹³C_β and 98 % of the ¹³C' resonances were assigned excluding four proline residues and the N-terminal methionine residue. Corresponding assignments to ¹H-¹⁵N correlation signals are labeled nearby in Fig II-3.

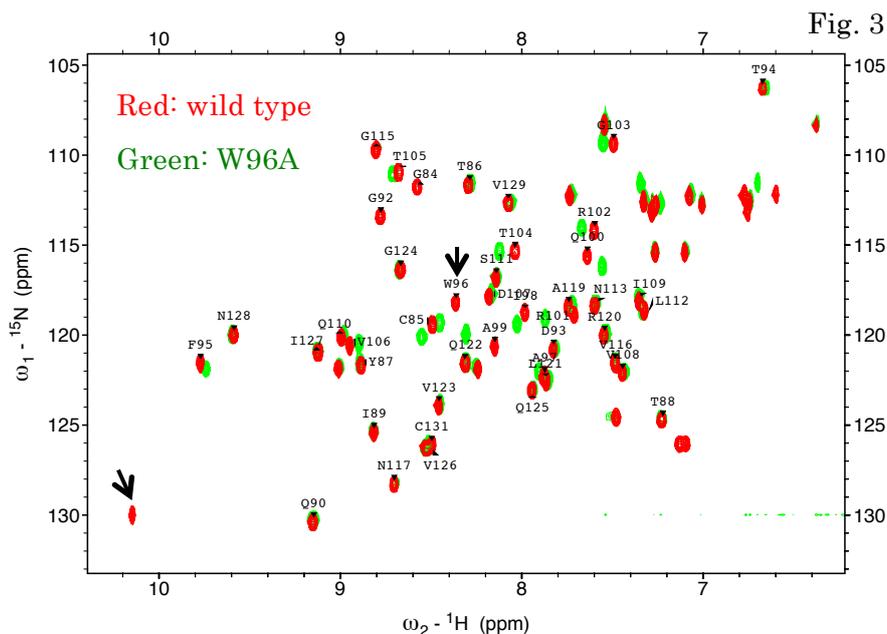


Fig. II-3 ¹H-¹⁵N HSQC spectra of ¹⁵N-labeled VcLysM2 (red) and VcLysM2-W96A (green). The main chain resonances of VcLysM2 were assigned as labeled in the spectrum. The resonances designated by arrows are the main chain and side chain resonances of Trp96 of VcLysM2.

(GlcNAc)₃₋₆ binding site

Upon titration of (GlcNAc)₃₋₆ solution into the protein solution, migration of backbone NH signals were observed. Thus the degrees of signal migrations, $\Delta\delta$, for each peak corresponding to residues were calculated and plotted against the residue number (Fig II-4 a-d). Significantly ($\Delta\delta > 0.25$) titrated residues were mapped on the crystal structures and are shown as the insets of each bar graph. Along (GlcNAc)₃ titration, the backbone NH resonance of Trp96 were broaden and beyond recognition during titration. Signals were gradually migrated suggesting that the exchange rate of binding was fast or intermediate. At saturated a point, backbone resonances of Gly92, Thr94, Phe95, Ile98, Leu121 and Gln122 were significantly titrated. When (GlcNAc)₄ was used for titration experiment instead of (GlcNAc)₃, backbone NH signals of Ala99, Ala119, Arg120 and Gly124 were significantly titrated in addition to the residues affected by (GlcNAc)₃. Additional line broadenings were observed for backbone NH resonances of Ile98 and Gln100. When (GlcNAc)₅ was used as a titrant, Asp93, Ala97 and Leu121 were further affected upon binding. In the case of (GlcNAc)₆, the result was almost identical to those obtained with (GlcNAc)₅, but further line-broadening was observed for backbone NH signal of Asp93. Overall, the binding sites of VcLysM2 to chitin oligosaccharides are located on the shallow binding groove formed by the loop connecting β 1 and α 1 and that connecting α 2 and β 2. Broadening of several residues may be due to the change in the dynamic state of the α 1-helical region.

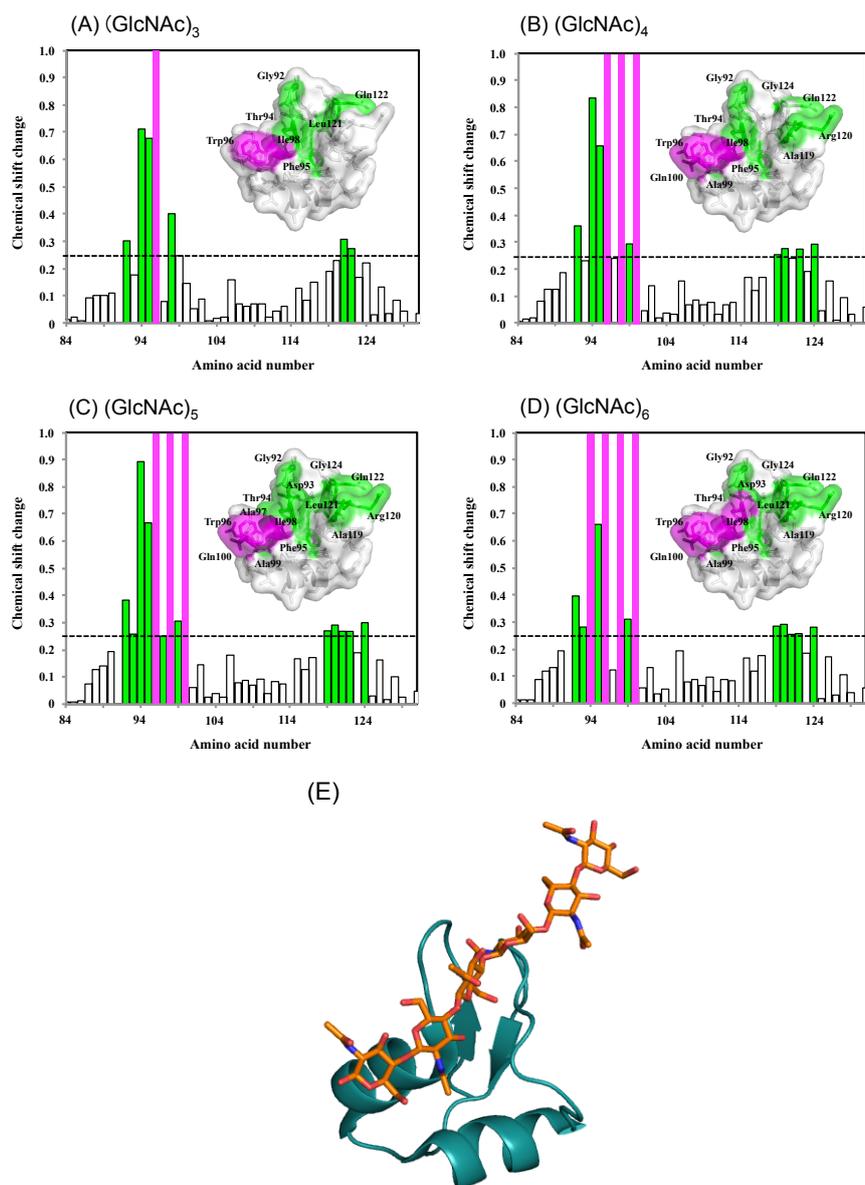


Fig. II-4 Ligand-binding site of VcLysM2 estimated by NMR-based titration experiments. Histograms of chemical shift changes (green) or line broadening (magenta) of the individual main chain resonances of VcLysM2 induced by the binding of (GlcNAc)₃ (a), (GlcNAc)₄ (b), (GlcNAc)₅ (c), and (GlcNAc)₆ (d) at the saturated conditions. Insets are the surface model of the crystal structure of VcLysM2, in which amino acid residues, the resonances of which were affected by (GlcNAc)_n binding, are highlighted in the same colors as in the histograms. (e) Crystal structure of (GlcNAc)₆-bound LysM from a putative NlpC/P60 D,L endopeptidase from *T. thermophilus* (PDB code, 4UZ2).

Effect of mutation of Trp96

Figure II-4e shows the crystal structure of LysM from putative D,L-endopeptidase, NlpC/P60, in complex with (GlcNAc)₆. Wong and co-workers reported that the Tyr28 of the LysM are interacting with GlcNAc residue at the reducing end through stacking interaction. Since VcLysM2 also has Trp96 at the similar position, we assume that the residue can be involved in the interaction. In order to investigate the contribution of Trp96 to the interaction with (GlcNAc)_n, we mutated Trp96 to alanine (VcLysM2-W96A) and characterized the mutant. As is shown in Fig. II-3, the backbone NH signal of Trp96 in the ¹H-¹⁵N HSQC spectrum of VcLysM2 (red) disappeared in the spectrum of VcLysM2-W96A (green) as predicted previously. Backbone NH signals of other residues around Trp96 were also shifted, but we assumed it was due to the lack of aromatic ring. Therefore, we assumed the mutation doesn't entirely affect the whole structure of the protein. We assigned the characteristic resonance at 10.15 ppm of the ¹H-axis and 130 ppm of the ¹⁵N-axis in ¹H-¹⁵N HSQC of VcLysM2 (Fig. II-3, red) as the indole NH of Trp96 since VcLysM2 contain only one Trp residue. This resonance is absent in the spectrum of VcLysM2-W96A (Fig. II-3, green) indicating that the assignment was correct. Titration of (GlcNAc)₃ to VcLysM2 shifted the side chain signal of Trp96 to higher field a bit (Fig. II-5a), but was markedly different response to the cases of (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ (Fig. II-5b-d), in which large migrations to lower field were observed. This indicate the major contribution of this side chain only to (GlcNAc)₄ or longer.

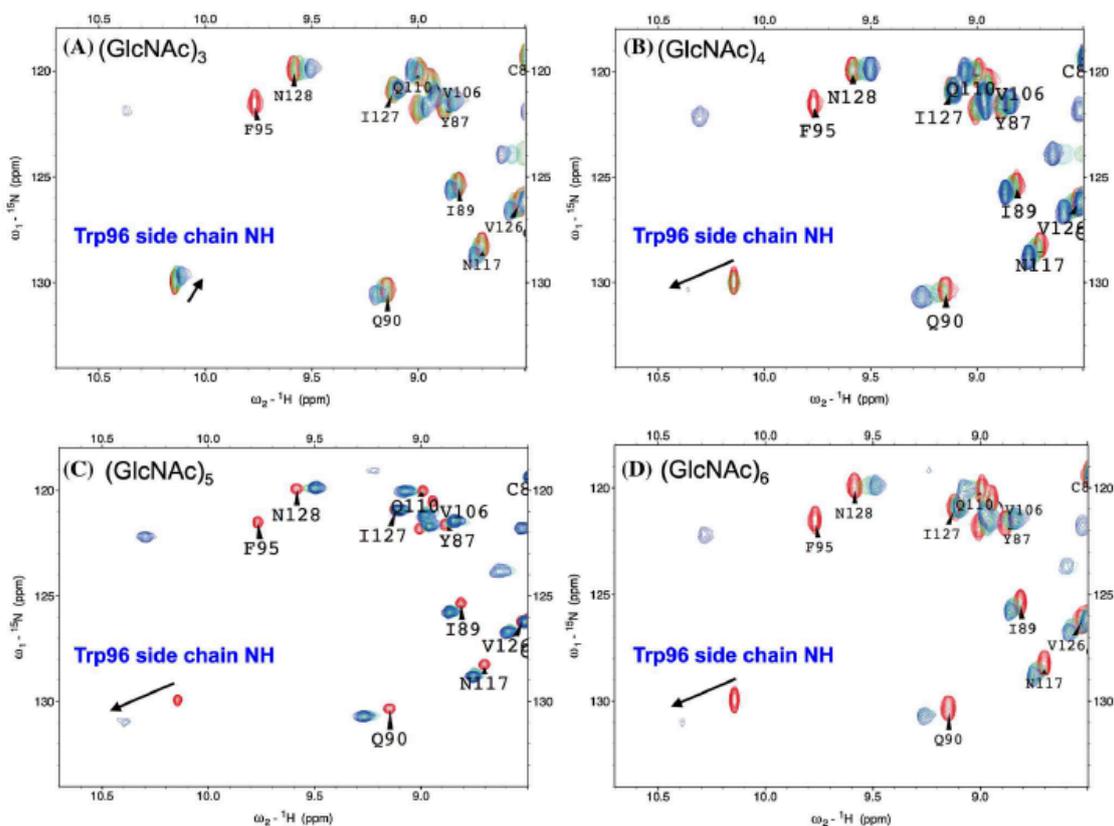


Fig. II-5 Overlays of ^1H - ^{15}N HSQC spectra of ^{15}N -labeled VcLysM2 in the absence or presence of (GlcNAc) $_3$ (a), (GlcNAc) $_4$ (b), (GlcNAc) $_5$ (c), or (GlcNAc) $_6$ (d). NMR titrations of the individual ligands were conducted in 20 mM sodium acetate pH 5.0. The molar ratios of (GlcNAc) $_n$:protein were 0:1 (red), 5:1 (green), 23:1 (cyan), and 96:1 (blue) for (GlcNAc) $_3$ (a); 0:1 (red), 1:1 (green), 4:1 (cyan), and 14:1 (blue) for (GlcNAc) $_4$ (B); 0:1 (red), 2:1 (green), 4:1 (cyan), and 11:1 (blue) for (GlcNAc) $_5$ (c); and 0:1 (red), 1:1 (green), 2:1 (cyan), and 12:1 (blue) for (GlcNAc) $_6$ (d)

Thermodynamics of (GlcNAc) $_3$ -6 binding to VcLysM2 and VcLysM2-W96A

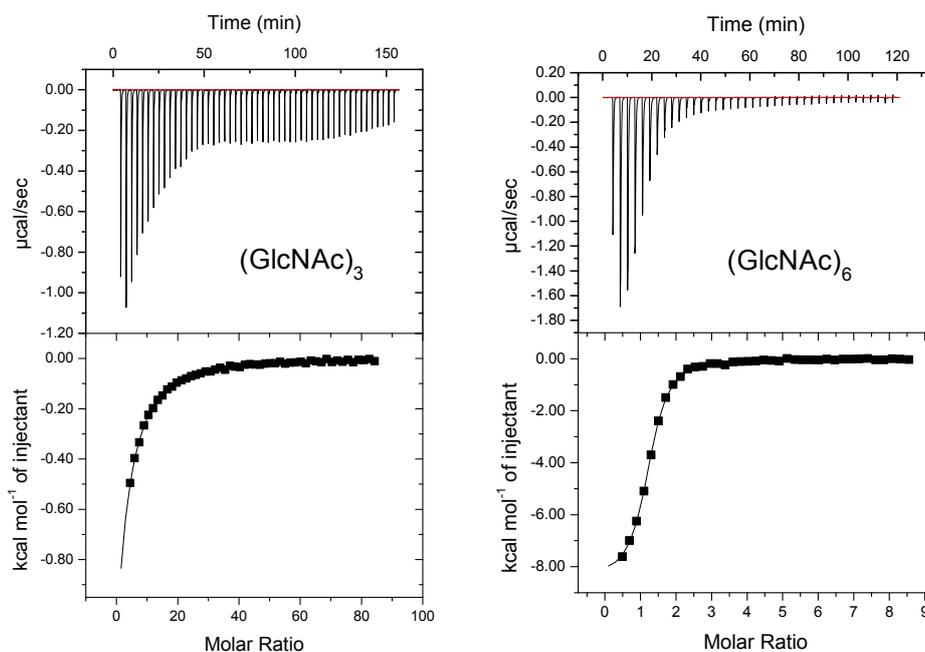
Fig. II-6 shows thermograms and binding isotherms of (GlcNAc) $_3$ -6 binding to VcLysM2 VcLysM2 and VcLysM2-W96A. All bindings were exothermic and the heat evolutions increase as the degrees of polymerization increase. (GlcNAc) $_2$ was also titrated, but thermogram didn't show any characteristic profile for bindings. Therefore, the thermodynamic parameters were analyzed only for (GlcNAc) $_3$ -6. Thermodynamic parameters were determined by data-fitting, based on One set of the sites model in the Origin software. Binding stoichiometry of weak bindings were fixed at one as was

observed in VcLysM2-(GlcNAc)₃ and VcLysM2-W96A-(GlcNAc)₃ and (GlcNAc)₆ binding data. Resulting thermodynamic parameters are summarized in the Table II-2. All in all, thermodynamic parameters were driven by favorable enthalpy changes accompanied with smaller unfavorable entropy changes. Although binding of VcLysM2 to (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ showed tendencies that longer oligomers have larger enthalpy gain accompanied with larger entropy loss, binding to (GlcNAc)₃ showed largest enthalpy and entropy contributions in all oligosaccharides tested. When we used VcLysM2-W96A instead of VcLysM2, smaller contributions of both enthalpy and entropy changes were observed among oligosaccharides tested. Resulting free energy change showed moderate difference in affinities between (GlcNAc)₃ and (GlcNAc)₆ with -1.0 kcal/mol for VcLysM2-W96A, whereas marked difference in affinities between these oligosaccharides by -2.9 kcal/mol for VcLysM2. This suggests that VcLysM2-W96A has shortened ligand-binding site.

Table II-2 Thermodynamic parameters obtained for chitin oligosaccharide binding to VcLysM2 and VcLysM2-W96A

	Ligand	<i>n</i>	K_{assoc} (1/M)	ΔH° (kcal/mol)	$-T\Delta S^\circ$ (kcal/mol)	ΔG° (kcal/mol)
VcLysM2	(GlcNAc) ₃	1.00	9.47×10^2	-8.8	4.7	-4.1
	(GlcNAc) ₆	1.23	1.24×10^5	-8.2	1.3	-6.9
VcLysM2-W96A	(GlcNAc) ₃	1.00	9.03×10^2	-4.7	0.7	-4.0
	(GlcNAc) ₆	1.00	5.08×10^3	-5.4	0.3	-5.1

VcLysM2



VcLysM2-W96A

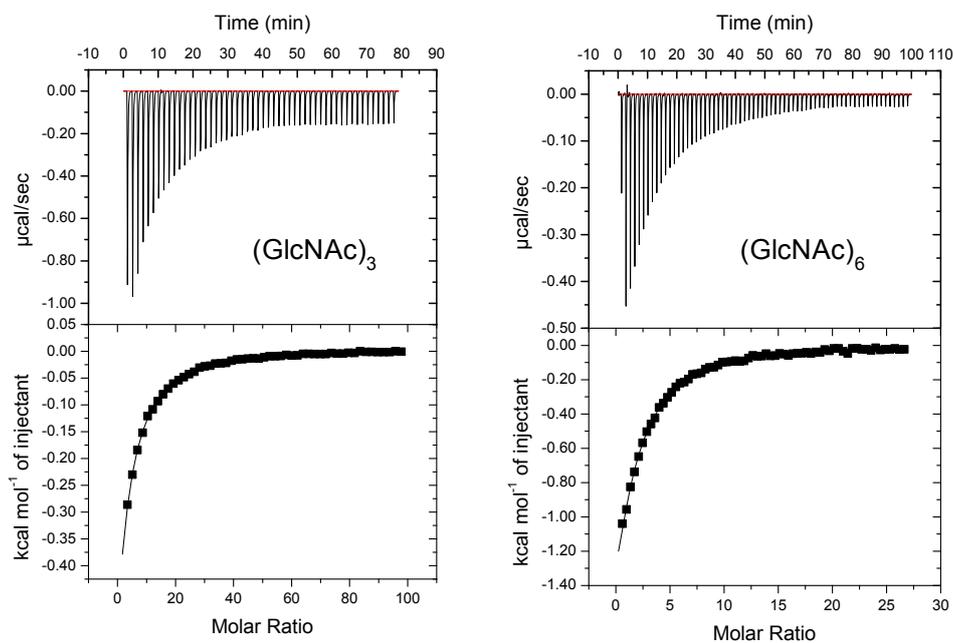


Fig. II-6 ITC thermograms and theoretical fits for the binding of (GlcNAc)₃ and (GlcNAc)₆ to VcLysM2 and VcLysM2-W96A. Titrations were conducted in 20 mM sodium acetate pH 5.0. Thermodynamic parameters obtained from individual experiments are listed in Table 1.

Discussion

Chitinases expressed in higher plants are induced by various biotic and abiotic stresses and thought to act in host-defensive reaction against pathogens. In the case of VcChi, the expression is induced by wounding and sex-inducing pheromone as was reported by Amon *et al.* Sex-inducing pheromone of *V. carteri* is induced by heat stresses and result in formation of dormant zygotes, which are considered to work for overcoming dryness of ponds in the summer heat (Kirk and Kirk 1986; Amon *et al.* 1998). Therefore, the VcChi expression seems be directly or indirectly induced by abiotic stresses. However, the expression of VcChi can also be induced by mechanical stresses caused by biotic stresses. VcChi doesn't show any similarities to chitinases from other higher plants. In this context, the structural and functional study of the enzyme is highly desired.

Boraston proposed a classification of CBMs into three types based on the carbohydrate-recognition mechanisms, i.e. Type A, Type B and Type C (Boraston *et al.* 2004). Type A CBMs have planer binding surface, which is mainly composed of aromatic residues and bind to insoluble ligands (Tormo *et al.* 1996). Type B CBMs have relatively opened binding cleft, which can accommodate various degrees of carbohydrates (Boraston *et al.* 2002), but not planer ligand binding site. Type C CBMs have relatively short binding site, which can accommodate only mono-, di, or trisaccharides (Shinya *et al.* 2013). VcLysM2 can be classified as Type B, based on the fact that the protein showed specificities toward various degrees of oligosaccharides with different affinities through the shallow binding cleft.

The disulfide bridge connecting the N- and C-termini was found in the crystal structure. Thermal denaturation experiments showed T_m more than 90 °C for VcLysM2 (data not shown). Thermal stability of LysM domains have been reported for PrChiA LysM2 containing two disulfide bridges and the melting temperature was at 89.9 °C (Ohnuma *et al.* 2008). One of the covalent cross-linkages bridges almost same position as VcLysM2 and the other bridges the beginning of $\alpha 1$ and the loop between $\alpha 2$ and $\beta 2$. VcLysM2 has Phe95 corresponding to the second disulfide bridge and the residue seems forming the hydrophobic core. This can potentially contribute to the thermal stability.

Mutational analysis of VcLysM2-W96A partially elucidated the ligand-recognition mechanism of VcLysM2. Aromatic side chains of amino acid residues are known to

have ability to form CH- π stacking interaction with the pyranose ring. The combination of crystal structure, NMR and ITC binding analyses showed that the stacking interaction is also the case for VcLysM2. Crystal structure showed that VcLysM2 has surface-exposed aromatic side chain on Trp96, which is able to interact with the ligand. Titration experiments using NMR showed that Trp96 residue form the binding site on VcLysM2. In the ITC experiments, binding affinities for (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ were significantly different from each other (0.9, 1.3 and 1.9 kcal/mol, respectively), though almost no difference in the binding affinities to (GlcNAc)₃ were observed between VcLysM2 and VcLysM2-W96A (0.1 kcal/mol). As was also suggested from signal migration of aromatic NH of Trp96 upon binding to (GlcNAc)₃, (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆, Trp96 specifically binds to (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ and forms a binding subsite at the end of the chitin-binding site. Similar effect was also observed for the LysM domains from a putative NlpC/P60 D,L-endopeptidase from *Thermus thermophiles* reported by Wong *et al.*, who clearly showed that the Tyr28 at the beginning of $\alpha 1$ are located at the edge of binding site and interacting with GlcNAc residue through CH- π stacking (Wong et al. 2015). Taken together, we concluded that Trp96 of VcLysM2 forms GlcNAc binding subsite at the edge of the binding site and interacting through CH- π interaction.

Chapter III

Structure-function relationship of LysM1 and LysM2 from VcChi

Introduction

Biochemical significance of the multi-modularity in proteins has been shown for many kinds of proteins and also for LysM-containing proteins. Mechanisms of taking advantages from the multi-modularity in proteins vary from protein to protein. Plant Receptor-like kinases (RLKs) and receptor-like proteins (RLPs) containing extracellular LysMs as sensor domain cooperatively bind to chitin fragment released from pathogenic microorganisms leading to defense-signaling cascade (Liu et al. 2012; Liu et al. 2016). Chitin elicitor-receptor kinase 1 (CERK1) and Chitin elicitor-binding protein (CEBiP) are one of the most famous RLK-RLP system that recognize chitin oligosaccharides derived from fungal cell wall through forming intermolecular dimeric structure (Liu et al. 2016). On the other hand, the opponents, fungi, secrete effector protein containing multiple LysMs, which bind to chitin oligosaccharides with much higher affinity than plant-receptor proteins, to escape from plant defense responses. Ecp6 from *Cladosporium fulvum* is one of such effector proteins comprising three LysMs, which can bind to chitin oligosaccharides with ultra-high (pM) affinity (Sánchez-Vallet, A. et al. 2013). To accomplish the high affinity, Ecp6 forms an intramolecular dimerization using two of three LysMs in the protein. To fit the catalytically active domain to its substrate, bacterial lysin harboring an endopeptidase domain, NlpC/P60 domain, adopts intermolecular dimerization of LysMs to attach the catalytic domain to peptide fragment of peptidoglycan (Wong et al. 2014).

Although the reports on multi-modular LysM containing proteins described above revealed the structural assembly and binding mode on these proteins, the contributions of individual LysMs have not yet been clarified in detail. We then cloned the genes of VcLysM1, VcLysM2 and VcLysMs linked in tandem, purified the proteins, and analyzed the binding modes of individual domains.

Materials and Methods

Plasmid construction

pET-VcLysM2 used for the experiments on the chapter II were also used for the experiments in this chapter. Plasmid containing VcLysM1 was constructed by double mutations of Asn117 and Ala119 of VcLysM2 into asparagine and alanine, respectively, and the genes encoding intermediate proteins between VcLysM1 and VcLysM2, VcLysM2-N117V and VcLysM2-A119T, were also constructed by introducing mutations into pET-VcLysM2. Mutations were introduced by QuikChange[®] Site-directed mutagenesis kit (Stratagene) using the primers on the table III-1. A gene encoding VcLysM tandem was amplified by PCR using the primers listed on the table III-1 and inserted into pETBlue-1 AccepTor vector (Novagen) by TA-cloning.

Table III-1. Primers used for site-directed mutagenesis.

Primers	Primer sequences (5'-3')
N117V Fwd.	GAATCCTGGCGTT <u>GTT</u> CCGGCACGCCTGC
N117V Rev.	GCAGGCGTGCCGGA <u>ACA</u> ACGCCAGGATTC
A119T Fwd.	GGCGTT <u>GTT</u> CCG <u>ACA</u> CGCCTGCAGG
A119T Rev.	CCTGCAGGCGT <u>GTC</u> CGGA <u>ACA</u> ACGCC
N117V/A119T Fwd.	CTGGCGTTAATCCG <u>ACA</u> CGCCTGCAGG
N117V/A119T Rev.	CCTGCAGGCGT <u>GTC</u> CGGATTAACGCCAG
VcLysM tandem Fwd.	ATGGGTTGTACCTATACCATTTCAGCCTG
VcLysM tandem Rev.	TTAACACGGCACGTTGATCACCTGG

*Single and double underline correspond to N117V and A119T mutations, respectively.

Preparation of protein samples

SHuffle T7 express (NEB) harboring pET-VcLysM1, VcLysM2 N117V, VcLysM1 A119T, VcLysM2 and VcLysM tandem were used to express corresponding genes. Purification procedures were almost the same as the procedure for VcLysM2 described on the chapter II, but 1.5 M ammonium sulfate was used for VcLysM tandem on hydrophobic interaction column chromatography instead.

All experiments were performed in 20 mM Sodium acetate, pH 5.0.

Determination of protein concentration

Protein concentrations were determined by absorbance at 280 nm. Molar extinction coefficient, ϵ_M , was calculated based on the amino acid sequences. Equation used for the calculation is as follows:

$$\epsilon_M = 5,500(\#\text{tryptophan}) + 1,490(\#\text{tyrosin}) + 125(\#\text{cystine}) \text{ (Pace et al. 1995).}$$

Isothermal titration experiments

An iTC₂₀₀ system (GE Healthcare) was used for titration experiments. The standard software from the supplier was used to operate the equipment. Ligand solution filled in the syringe was titrated into protein solution in the sample cell for all experiments. Raw data were processed and analyzed by Origin software. Protein and ligand concentrations used for experiments are listed on the table III-2.

Table III-2. Samples used for ITC experiments

Protein	Concentration	Ligand	Concentration
VcLysM1	115 μ M	(GlcNAc) ₃	46.0 mM
		(GlcNAc) ₄	11.5 mM
		(GlcNAc) ₅	4.6 mM
		(GlcNAc) ₆	4.6 mM
VcLysM1 V53N	97.0 μ M	(GlcNAc) ₃	52.0 mM
		(GlcNAc) ₄	13.0 mM
		(GlcNAc) ₅	4.5 mM
		(GlcNAc) ₆	3.8 mM
VcLysM1 T55A	130 μ M	(GlcNAc) ₃	52.0 mM
		(GlcNAc) ₄	26.0 mM
		(GlcNAc) ₅	16.0 mM
		(GlcNAc) ₆	5.2 mM
VcLysM2	94.0 μ M	(GlcNAc) ₄	7.0 mM
		(GlcNAc) ₅	4.1 mM
VcLysM tandem	127 μ M	(GlcNAc) ₃	48.0 mM
		(GlcNAc) ₄	24.0 mM
		(GlcNAc) ₅	15.0 mM
		(GlcNAc) ₆	3.6 mM

During titration experiments, sample solution was kept stirred by the syringe tip and the stirring speed was at 1,000 rpm. Data points were collected by titrations every 5 seconds and plotted against the experimental time. Peaks for each titration on the thermogram were integrated and plotted against the molar ratio of ligand to protein (Fig. III-1 and Fig. III-13). Based on the experimental data for VcLysMs, non-linear data fittings were performed by assuming one set of the sites model defined in the Origin software to obtain the thermodynamic parameters for binding of (GlcNAc)_n. A sequential model was used to fit the data for VcLysM tandem. Resulting thermodynamic parameters are summarized on the table III-3 and table III-6.

Circular dichroism of VcLysMs

Circular dichroism spectra of VcLysM1, VcLysM1 V53N, VcLysM1 T55A at 200-260 nm were recorded by a J-720 (JASCO) equipped with a PTC-423L temperature controller under operation of the Spectra manager software (JASCO). All measurements were conducted with 2 mm quartz cuvette and the temperature was kept at 298 K. Ellipticities of samples at every 0.1 nm were recorded and normalized as molar ellipticity, θ (deg cm² dmol⁻¹).

Acquisitions of NMR spectra

Bruker UltraShield 500 Plus magnet equipped with a Avance III spectrometer and a triple-resonance pulse-field-gradient cryoprobe head was used to record spectra. All equipments were operated by Bruker TopSpin 2.1. Finally, one-dimensional (¹H), two-dimensional (¹H-¹⁵N HSQC, ¹H-¹³C HSQC, NOESY, DQF-COSY) and three-dimensional (HNCO, HNCACO, CBCACONH, HNCACB, HCCH-TOCSY, CCH-TOCSY, ¹⁵N-edited TOCSY, ¹⁵N-edited NOESY, ¹³C-edited NOESY) spectra were recorded were collected at 300 K. one-dimensional spectra were processed and analyzed by TopSpin and two-dimensional and three-dimensional spectra were processed by NMRPipe (Delaglio et al. 1995) and analyzed by Sparky (Goddard et al.; <https://www.cgl.ucsf.edu/home/sparky/>).

H-D exchange experiments

^1H - ^{15}N HSQC spectra were recorded 12 min to 24 hours after exchanging normal buffer with deuterated buffer. Buffer was exchanged by concentration by Vivaspin[®] 6-3K and dilution with the deuterated buffer. H-D exchange was observed based on the disappearance of HSQC signals.

Structural calculation by CYANA

Three-dimensional structure of VcLysM1 and VcLysM2 were calculated based on distance constraints obtained from NOE experiments and dihedral angle constraints generated by TALOS+ (Shen et al. 2009) using CYANA (Güntert and Buchner 2015). After several cycles of refinements, distance constraints for hydrogen bonding derived from H-D experiments are combined. Finally, 687 NOE-based distance, 216 dihedral angle, 32 hydrogen bonding based distance and 3 disulfide bonds-based distance restraints were used for structural calculation of VcLysM1, and 721 NOE-based distance, 216 dihedral angle, 32 hydrogen bonds-based distance and 3 disulfide bonds-based distance restraints were used for structural calculation of VcLysM2. Among 100 conformers calculated, 20 conformers were chosen as the representative structure based on the target function. The representative structures are shown on the Fig. III-5.

Chemical shift perturbation experiments of VcLysM1 upon titration of (GlcNAc)₆

To define the ligand binding site of VcLysM1, totally 100 μL of 0.5 mM (GlcNAc)₆ were titrated into 500 μL of 50 μM VcLysM1 and the chemical shift migration was observed for each ^1H - ^{15}N resonance. Combined chemical shift migration, $\Delta\delta$, was calculated as follows:

$$\Delta\delta = |\Delta\delta_{\text{H}}|/5 + |\Delta\delta_{\text{HN}}|.$$

$\Delta\delta$ s were plotted against the corresponding residue numbers (Fig. III-7). Amino acid residues representing the ^1H - ^{15}N resonances with $0.25 < \Delta\delta$ were mapped on the solution structure of VcLysM1. Amino acid residues of VcLysM2 selected by the same criteria

based on the results described in chapter II were also mapped on the solution structure of VcLysM2.

4, 8, 16, 28, 48, 72, 100 μL of 30 mM $(\text{GlcNAc})_3$ was repeatedly titrated into VcLysM tandem and $\Delta\delta$ were calculated for Gly28/92, Gly51/115 and Gly60/124. Peak position was determined by fitting Gaussian to the partially overlapped peaks using built-in integration script of Sparky. $\Delta\delta$ were plotted against the free ligand concentration and the association constants, K_a s, and corresponding ΔG° s were determined by fitting regression curves.

Docking simulation of VcLysM1 to $(\text{GlcNAc})_6$

To estimate the binding mode of VcLysM1 to $(\text{GlcNAc})_6$, docking simulation was performed using HADDOCK 2.2 web server (Dominguez, C. *et al.* 2003 and de Vries, S.J. *et al.* 2007). Top 20 structures of the solution structure of VcLysM1 and $(\text{GlcNAc})_6$ structure prepared by combination of GLYCAM (Woods group 2005-2017: <http://glycam.org>) and PRODRG (Schüttelkopf, A.W. *et al.* 2004) were used for the simulation. Gly28, Asp29, Thr30, Trp32, Gln36, Val53, Thr55, Arg56, Gln58 and Gly60 of VcLysM1 were selected as the active residues based on the $(\text{GlcNAc})_6$ titration experiment using NMR and accessible surface area calculated by ASA-view (<http://www.abren.net/asaview/>). Passive residues were selected automatically. Simulated complex structure is shown on the Fig. III-9.

Results

Binding thermodynamics of VcLysMs binding to $(\text{GlcNAc})_{2-6}$

$(\text{GlcNAc})_2$, $(\text{GlcNAc})_3$, $(\text{GlcNAc})_4$, $(\text{GlcNAc})_5$ and $(\text{GlcNAc})_6$ solutions were titrated into VcLysM1, VcLysM2-N117V, and VcLysM2-A119T solutions, and $(\text{GlcNAc})_4$ and $(\text{GlcNAc})_5$ were titrated into VcLysM2. Profiles of heat evolutions were satisfactorily obtained to determine the thermodynamic parameters for $(\text{GlcNAc})_3$, $(\text{GlcNAc})_4$, $(\text{GlcNAc})_5$ and $(\text{GlcNAc})_6$ titration into all proteins (Fig. III-1). Since no significant heat evolution was observed for $(\text{GlcNAc})_2$ (data not shown), thermodynamic parameters were determined only for $(\text{GlcNAc})_3$, $(\text{GlcNAc})_4$,

(GlcNAc)₅ and (GlcNAc)₆ and are summarized on the table III-3. Due to the poor data points at lower molar ratio, which are important to correctly determine the thermodynamic parameters (Turnbull and Daranas 2003), the data for (GlcNAc)₃ were inaccurate when compared with longer chain oligosaccharides. Therefore, discussion will be based only on the thermodynamic parameters for (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆.

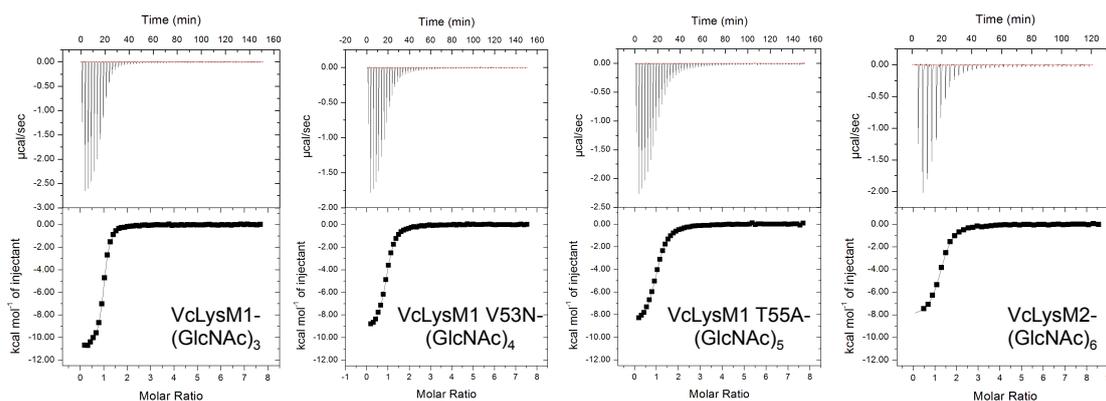


Figure III-1. Comparison of Chitin oligosaccharide binding experiments of VcLysMs. Upper panels are thermograms and the lowers are binding isotherms of binding of VcLysM1, VcLysM1 V53N, VcLysM1 T55A and VcLysM2 to chitin hexamer from left to right.

Table III-3. Thermodynamics of VcLysM1-(GlcNAc)₃₋₆ binding experiments.

Protein	Ligand	N	K_s (M^{-1})	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
VcLysM1	(GlcNAc) ₃	1	3.94×10^3	-8.67	3.75	-4.92
	(GlcNAc) ₄	1.02	5.30×10^4	-9.21	2.76	-6.45
	(GlcNAc) ₅	1.03	2.10×10^5	-9.93	2.66	-7.27
	(GlcNAc) ₆	0.957	4.67×10^5	-11.0	3.22	-7.78
VcLysM1 V53N	(GlcNAc) ₃	1	1.92×10^3	-7.17	2.69	-4.48
	(GlcNAc) ₄	1	2.67×10^4	-7.97	1.93	-6.04
	(GlcNAc) ₅	0.968	1.05×10^5	-8.68	1.81	-6.86
VcLysM1 T55A	(GlcNAc) ₆	0.904	2.24×10^5	-9.43	2.13	-7.30
	(GlcNAc) ₃	1	9.05×10^2	-7.61	3.58	-4.03
	(GlcNAc) ₄	1	1.03×10^4	-8.71	3.22	-5.49
VcLysM2*	(GlcNAc) ₅	1.04	4.16×10^4	-8.96	2.66	-6.30
	(GlcNAc) ₆	0.996	1.07×10^5	-9.06	2.20	-6.86
	(GlcNAc) ₃	1	9.47×10^2	-8.76	4.68	-4.08
VcLysM2*	(GlcNAc) ₄	1.01	1.13×10^4	-5.95	0.423	-5.52
	(GlcNAc) ₅	0.866	3.39×10^4	-7.51	1.15	-6.19
	(GlcNAc) ₆	1.23	1.24×10^5	-8.24	1.28	-6.95

*Thermodynamics of (GlcNAc)₃ and (GlcNAc)₆ are from Kitaoku *et al.* (2017)

All in all, binding stoichiometry converged to 1 indicating the binding of VcLysMs to chitin oligosaccharides tested were 1:1 interaction. Binding reactions were driven by enthalpy accompanied with smaller and unfavorable entropy and both favorable and unfavorable contributions increased as polymerization degree become larger. Binding constants decreased in the order of VcLysM1, VcLysM2-A119T, VcLysM2-N117V and VcLysM2 indicating that the mutants have intermediate affinity between VcLysM1 and VcLysM2. In most cases, VcLysM1 showed largest favorable contributions of enthalpy accompanied with largest unfavorable contributions of entropy among these proteins. VcLysM2-N117V also showed large contributions of favorable enthalpy changes, but unfavorable contributions of entropy resulted in the affinity rather similar to VcLysM2. In comparison with VcLysM2-N117V, A119T mutant showed rather modest contributions of unfavorable entropy changes and resulted in second smallest ΔG° values among the proteins used.

To examine more details about the effects of mutations, changes in ΔH° and $T\Delta S^\circ$ for individual mutations were calculated and shown on the Fig. III-2. Mutations in VcLysM1 at both Val53 and Thr55 resulted in unfavorable contributions to binding thermodynamics. When ΔH° and $T\Delta S^\circ$ are compared between (GlcNAc)₃, (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆, effects on mutation at Val53 are larger than the mutation at Thr55 for (GlcNAc)_{4,5}, but are smaller for (GlcNAc)₆, indicating certain gap in the binding modes between (GlcNAc)_{4,5} and (GlcNAc)₆.

Entropy changes for (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ were plotted against enthalpy changes on the Fig III-3. No entropy-enthalpy compensation (EEC) was observed between (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆. However, EEC was observed between VcLysM1 and VcLysM1 V53N for each oligosaccharide suggesting that the interactions contributing to the binding to VcLysM1 and VcLysM2-A119T are similar from each other.

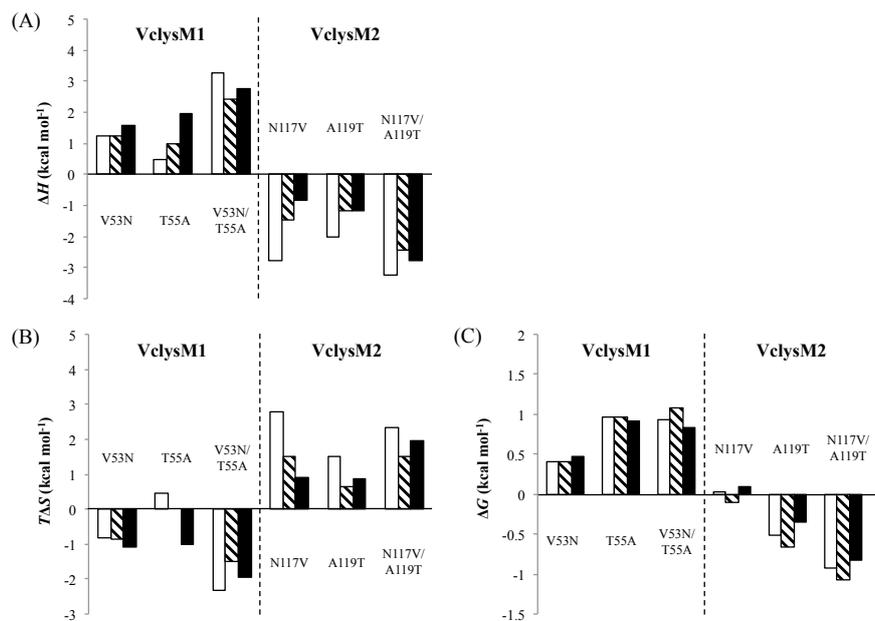


Figure III-2. Effects of mutations on two mutation sites, 53/117 and 55/119. Difference in affinities are shown as blank, striped and black bar representing differences in binding affinities against (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆, respectively. From left to right, these bars represent effects of mutations at the first (Val53), second (Thr55) and both on VcLysM1 and inverse of them on VcLysM2.

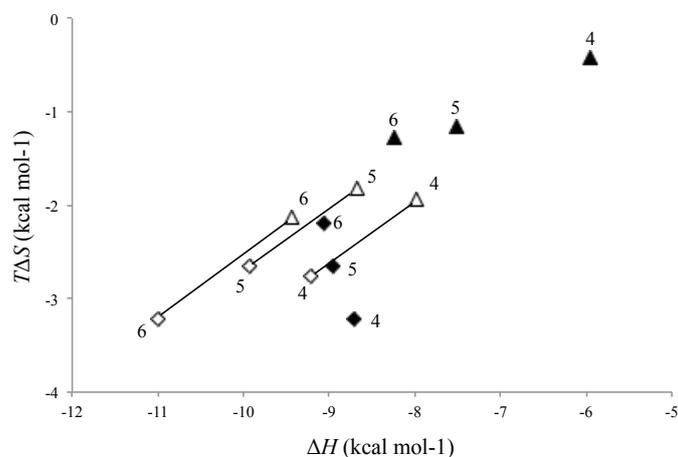


Figure III-3. EEC of VcLysMs-(GlcNAc)₃₋₆ binding. Symbols: open rectangle; VcLysM1, open triangle; VcLysM1 V53N, closed rectangle; VcLysM1 T55A; closed triangle; VcLysM2. Numbers Solid lines represent EEC between VcLysM1 and VcLysM1 V53N upon binding to (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆

Circular dichroism VcLysMs

CD spectra were compared between VcLysMs ranging from 200-260 nm (Fig III-4). All proteins showed peaks around 207 nm and 222 nm, which are characteristic to α -helical structures, as was speculated from $\beta\alpha\alpha\beta$ structure conserved for LysMs. Sharp increases at the shorter edges of wavelength indicate correct folding of proteins. Spectra of VcLysM1 and VcLysM2 are deviated from each other and those of the two intermediate mutants were those in between of VcLysM1 and VcLysM2, suggesting that the two mutated residues may be structural factors affecting the binding affinities. Since structural deviations were suggested from CD spectra, we decided to determine the solution structures of VcLysM1 and VcLysM2 to get insights into the structure-function relationship.

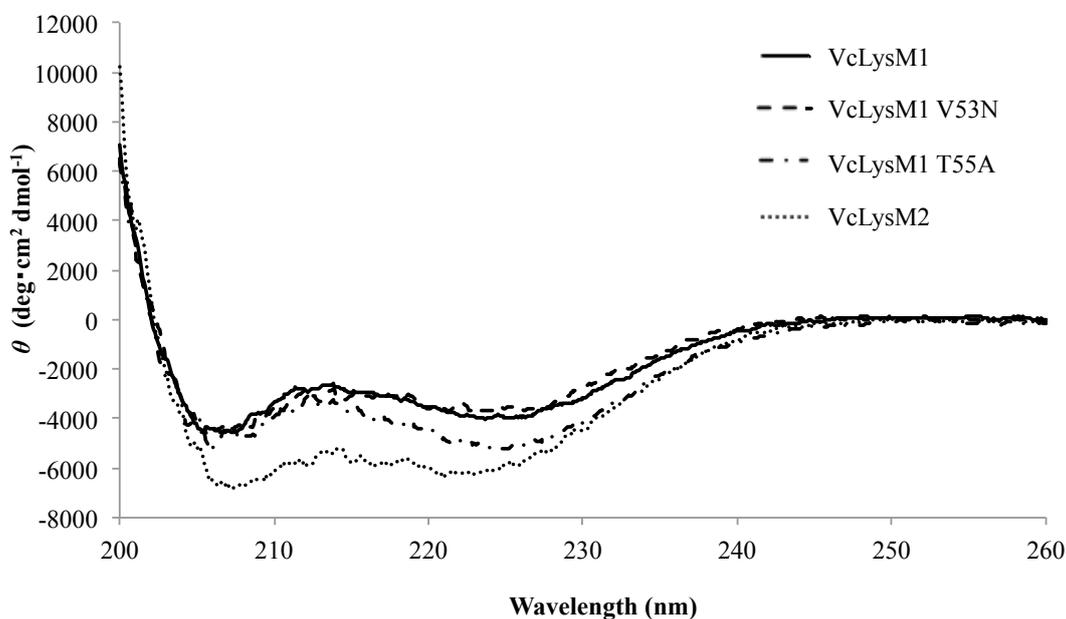


Figure III-4. Circular dichroism of VcLysMs. Molar ellipticity ranging from 200 to 260 nm were plotted against the wavelength. Solid (—), dashed (---), dashed dotted (- · -) and dotted (·····) lines correspond to VcLysM1, VcLysM1 V53N, VcLysM1 T55A and VcLysM2, respectively.

Solution structures of VcLysM1 and VcLysM2

Solution structures of VcLysM1 and VcLysM2 were successfully determined. Statistics of top 20 structures with lowest target function is shown on the table III-4. As we can see on the Fig III-5, well-converged structures were obtained for both LysMs and the main chain RMSD of VcLysM1 and VcLysM2 were 0.51 and 0.53 Å, respectively. Structural validation with PSVS (Protein Structure Validation Suite) showed that 90.0 %, 10.0 %, 0.0 % and 0.0 % of ϕ - ψ angles of VcLysM1 and 89.1 %, 10.8 %, 0.0 % and 0.0 % of ϕ - ψ angles of VcLysM2 were in most favorable, additionally allowed, generously allowed and disallowed region, respectively. Although all of criteria based on softwares in PSVS showed acceptable Z-scores ($-5 < Z$) for VcLysM1 (-2.57, -0.79, -1.49, -3.49 and -0.39 for Verify3D, ProsaII, Procheck (ϕ - ψ), Procheck (all dihedral angles) and Molprobitry Clashscore) and VcLysM2 (-1.93, -0.66, -1.85, -3.96 and -1.55, respectively for the same softwares for VcLysM1), Z-scores for both proteins were deviated from the recommended value ($-3 < Z$). I assumed that these deviations were simply due to the dynamic character of proteins since these criteria are based on crystal structures with relatively high resolution (< 1.80 Å) as was also mentioned by Bhattacharya and coworkers. Since the other criteria showed good proximity to the mean values of each data set, these structures were used for further analyzes.

Superposition of solution structures of VcLysM1 and VcLysM2 in Fig. 5 showed high structural homology between these two structures (RMSD=0.53 Å). Backbone traces well overlapped except the loop region between α_2 and β_2 , where Val53 and Thr55 are located. I also could see the effect of deviation in ^1H - ^{15}N HSQC spectra, in which ^1H - ^{15}N signals of only the residues on this loop region were affected (Fig. III-6).

Table III-4. Structural statistics of best 20 structures.

	VcLysM1	VcLysM2
NMR distance and dihedral restrains		
Total unambiguous NOE	687	721
Hydrogen-bond restraints	32	32
Dihedral angle restraints		
Phi constraints	108	108
Psi constraints	108	108
Structural statistics		
Pairwise RMSDs (mean \pm SD)		
Heavy atoms (\AA)	0.93 ± 0.10	0.90 ± 0.14
Backbone atoms (\AA)	0.51 ± 0.10	0.53 ± 0.16
Ramachandran statistics (%)		
Residues in most favorable regions	90.0	89.2
Residues in additionally allowed regions	10.0	10.8
Residues in generously allowed regions	0.0	0.0
Residues in disallowed regions	0.0	0.0

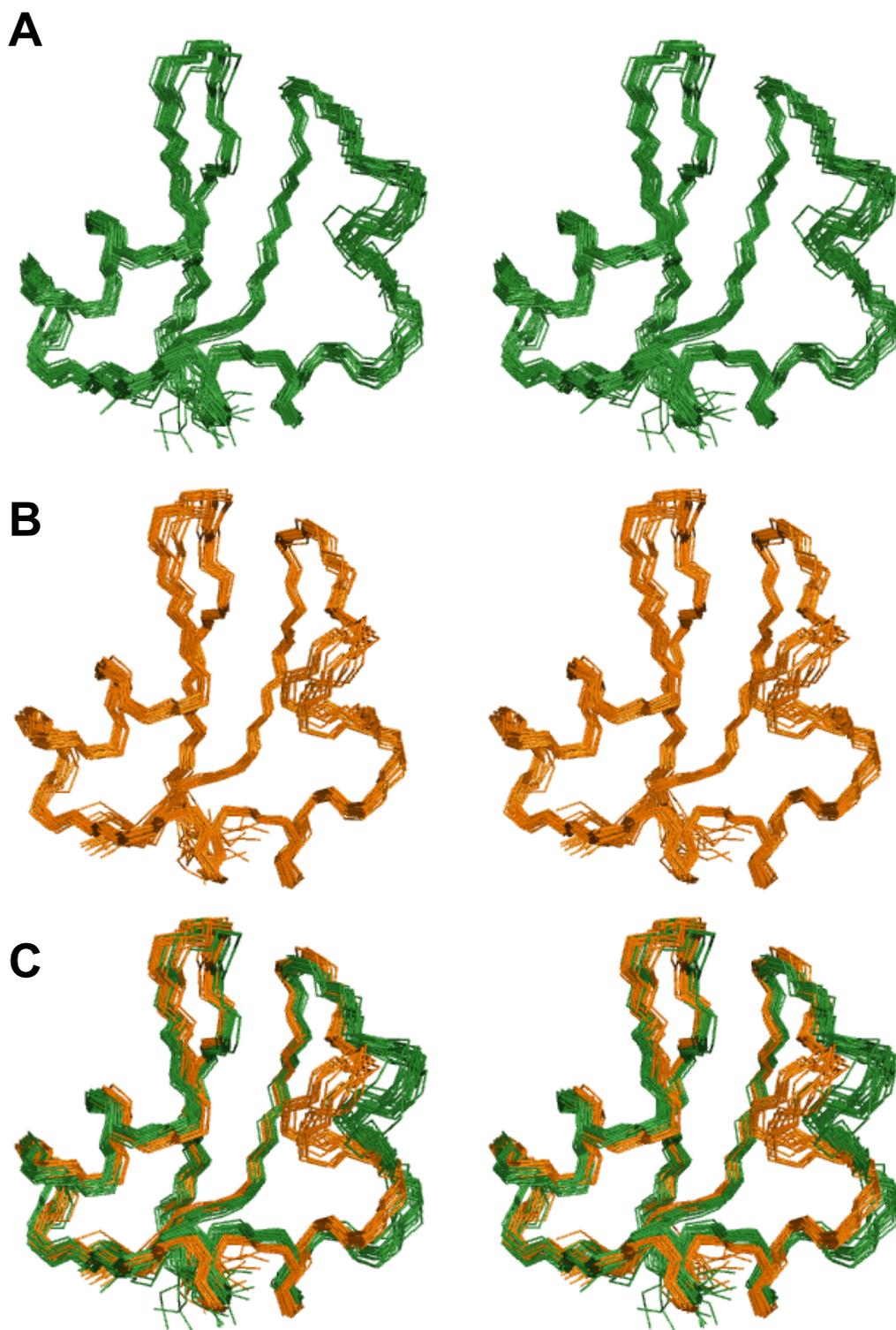


Figure III-5. Solution structures of VcLysMs. Superposition of top 20 structures of VcLysM1 and VcLysM2 are depicted as line models on PyMOL. A, Stereo view of solution structure of VcLysM1 (*green*). B, Stereo view of Solution structure of VcLysM2 (*orange*). C, Superposition of A and B.

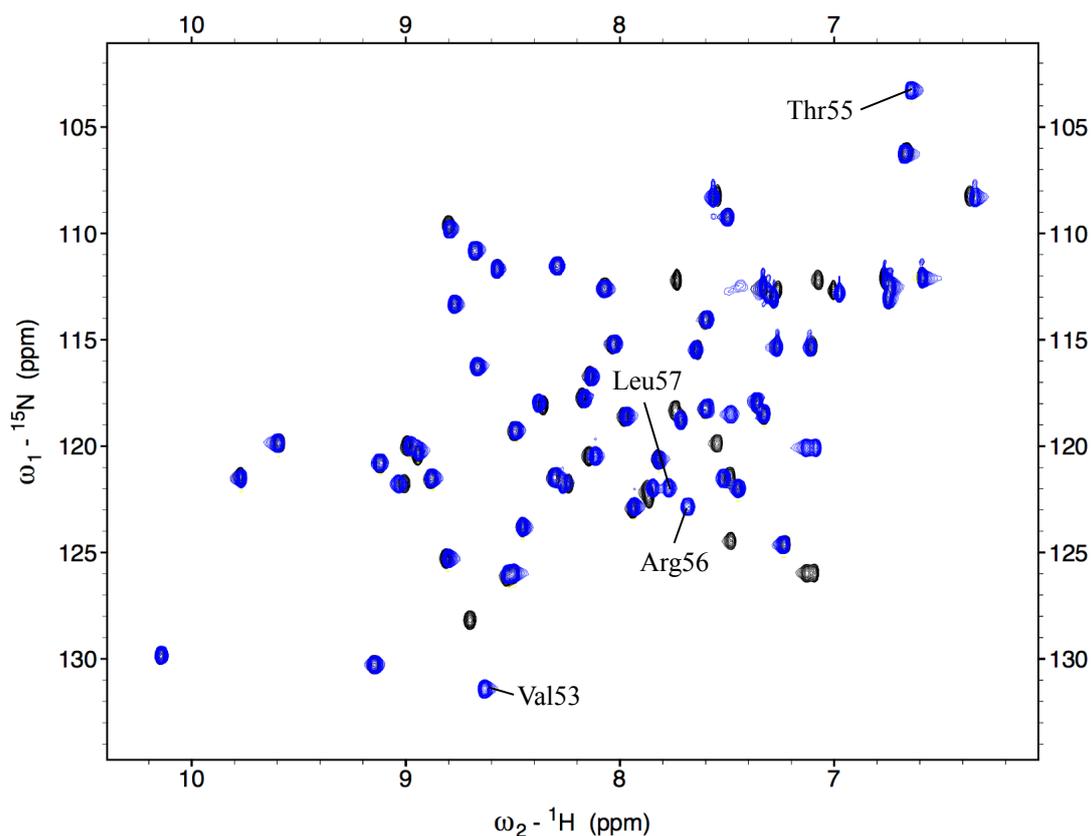


Figure III-6. Overlaid ${}^1\text{H}$ - ${}^{15}\text{N}$ HSQC spectra of VcLysM1 (*blue*) and VcLysM2 (*black*).

Chemical shift perturbation of VcLysM1 upon binding to $(\text{GlcNAc})_6$

$\Delta\delta$ values plotted against the residue numbers are shown in Fig. III-7. Residues with $0.25 < \Delta\delta$ were located on the loop region between $\beta 1$ and $\alpha 1$, the beginning of $\alpha 1$ and the loop region between $\alpha 2$ and $\beta 2$. Strongly affected residues of VcLysM1 and VcLysM2 were mapped onto the solution structures and compared to each other (Fig. III-8). Binding sites are conserved and are located on the shallow cleft on the α -helical surface. Almost the same set of residues were affected upon binding to $(\text{GlcNAc})_6$ except Val53. To see more details of the effect of deviation in the loop structure, binding mode of VcLysM1 to $(\text{GlcNAc})_6$ was simulated.

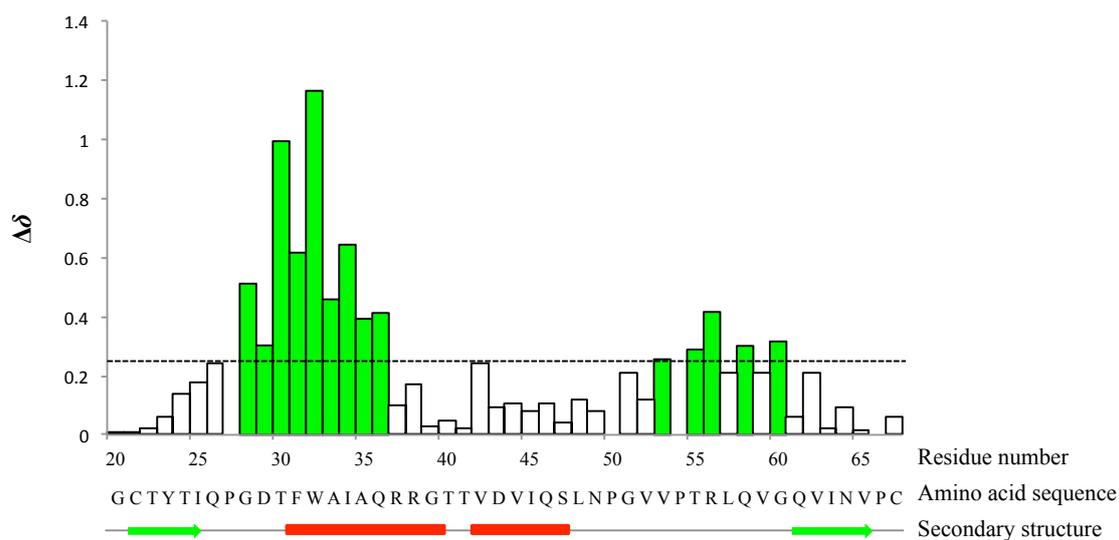


Figure III-7. Chemical shift migration plotted against amino acid residue number. Chemical shift migrations ($\Delta\delta$) for each residue calculated based on the equation described in the text were plotted against the residue number. Amino acid sequence and secondary structure corresponding to the residue numbers were also shown below. Green bars and red bars are corresponding to elongated and helical structures, respectively.

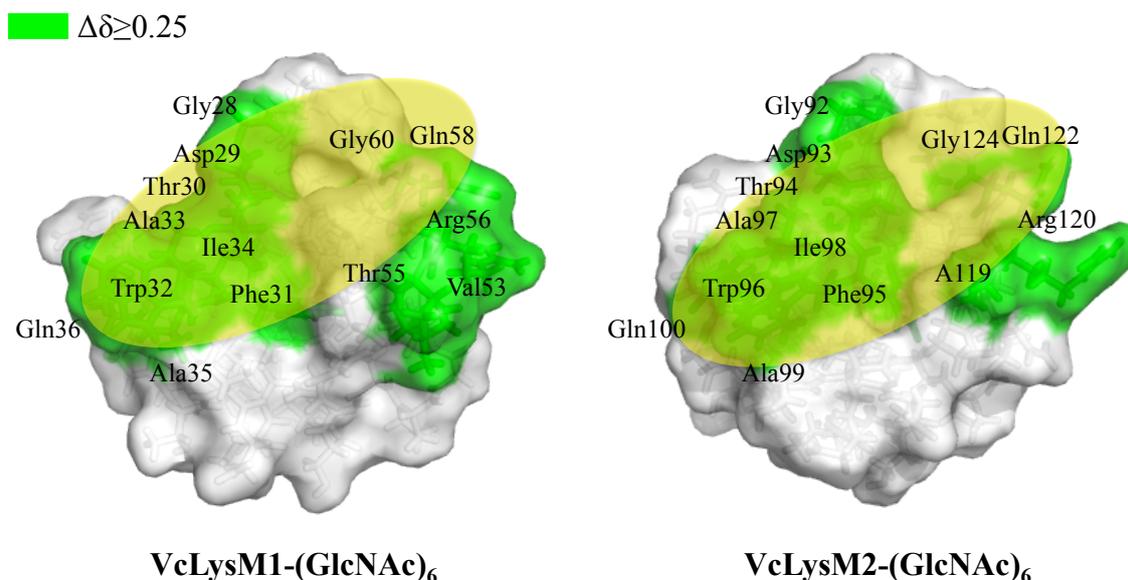


Figure III-8. Mapping of perturbed amino acid residues on titration experiments using NMR. Amino acid residues perturbed significantly ($\Delta\delta \geq 0.25$) upon (GlcNAc)₆ titration experiments were mapped (green) on the solution structures.

Binding mode of VcLysM1 to (GlcNAc)₆

(GlcNAc)₆ was docked onto the solution structure of VcLysM1 using HADDOCK2.2. Based on the HADDOCK score, which is an RMSD in the weighted sum of various energy terms, 6 clusters were generated. One of these clusters having lowest HADDOCK score was used and the most feasible structure was chosen by visual inspection. The docked structure is presented as PyMol image (Fig.III-9) and LigPlot+ diagram (Fig.III-10). Individual GlcNAc residues were named GlcNAc1-6 from the reducing end, respectively. Hydrogen bonds between VcLysM1 and (GlcNAc)₆ are depicted as dashed lines.

GlcNAc1 seems to interact with VcLysM1 by only one hydrogen bond between O6 of GlcNAc1 and N_ε of Trp32 of VcLysM1 and major part of the residue was stuck out from the binding cleft of VcLysM1 indicating that amino acid residues around GlcNAc1 doesn't form subsite for accommodating sugar residue. Pyranose ring of GlcNAc2 seems to be stacked with the aromatic ring of Trp32 as was also described for VcLysM2 in the figure II-5. The middle two GlcNAc residues, GlcNAc3 and GlcNAc4, interacted with main chain amide group of VcLysM1 by hydrogen bonds. Especially, acetyl group of GlcNAc3 contributed to the interaction by two hydrogen bonds suggesting importance of *N*-acetyl group of the sugar residue located at this subsite for the VcLysM1-(GlcNA)_n interaction. *N*-acetyl group of GlcNAc5 also interacted with main chain polar atoms of VcLysM1 by two hydrogen bonds. Furthermore, methyl group of this *N*-acetyl group seems to be accommodated by the hydrophobic cavity of VcLysM1 indicating that this subsite is also specific to sugar moiety with *N*-acetyl group. Although no hydrogen bond was formed between GlcNAc6 and VcLysM1, most parts of the residue were accommodated in the binding cleft. Thr55 participated in the binding to GlcNAc5, but the interaction was formed with the main chain carbonyl. It was still possible for Thr55 to form indirect hydrogen bond mediated by a water molecule, but Val53 was located outside of the (GlcNAc)₆-binding region.

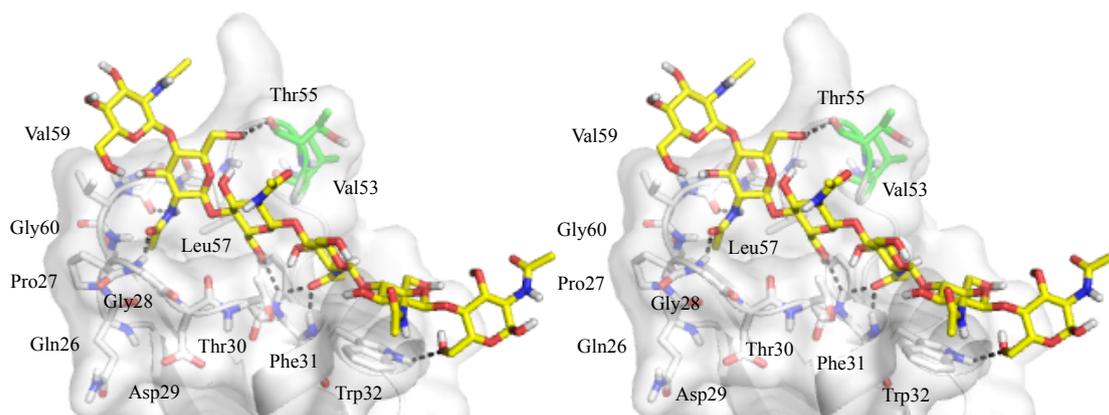


Figure III-9. Stereo views of simulated VcLysM1-(GlcNAc)₆ complex structure. Carbon atoms of interacting residues indicated by LigPlot⁺ and Val53 are shown as white stick. Carbon atoms of (GlcNAc)₆ are shown by yellow stick. Hydrogen, nitrogen and oxygen atoms are colored white, blue and red, respectively. Dashed lines represent hydrogen bonds between VcLysM1 and (GlcNAc)₆.

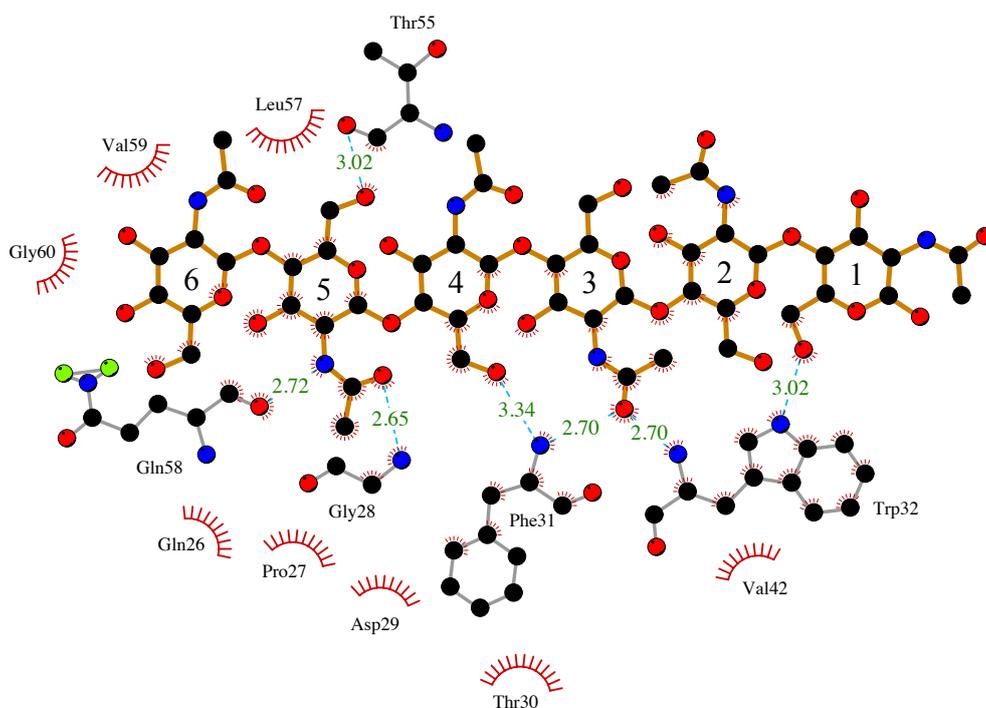


Figure III-10. LigPlot⁺ description of the interactions participating in VcLysM1-(GlcNAc)₆ binding. Hydrogen bonds (dashed line) and hydrophobic interactions are depicted. Numerals listed on the hydrogen bonds indicate the distances between the corresponding atoms.

Effect of formation of tandem linkage on VcLysMs

^1H - ^{15}N HSQC spectra of VcLysM tandem are overlaid with the spectra of VcLysM1 and VcLysM2 (Fig. III-11). Assignment of VcLysM tandem is indicated in the figure. Assignment was almost completed except Gly-Gly sequence in the linker region as resonances of C_α , C_β and C_γ of these two residues were completely overlapped indicating these two glycine residues are in random coil structure.

Most signals of VcLysM tandem were overlapped with the signals for VcLysM1 and VcLysM2 except the linker and neighboring residues indicating that the linkage between the VcLysM1 and VcLysM2 affected only the linker region and its nearest neighbor.

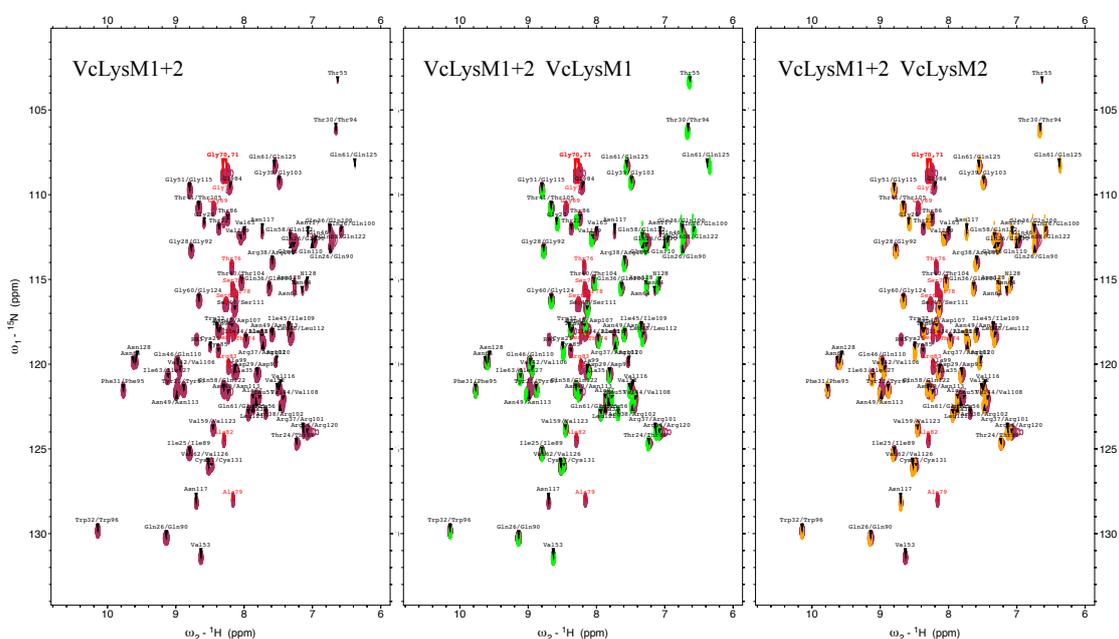


Figure III-11. Overlay of the ^1H - ^{15}N HSQC spectrum of VcLysM tandem (*purple*) with those of VcLysM1 (*green*) and VcLysM2 (*orange*). The assignment of VcLysM tandem was indicated on each peak.

Effect of tandem-linkage formation upon binding properties of (GlcNAc)₃₋₆

To determine the binding affinities on each domain of VcLysM tandem, i.e. VcLysM1 and VcLysM2 linked in tandem, (GlcNAc)₃ was titrated, chemical shift migration was calculated for each ligand concentration and binding constants were determined based on the signals derived from each domain by non-linear curve fitting procedure (Fig. III-12 and Table III-5).

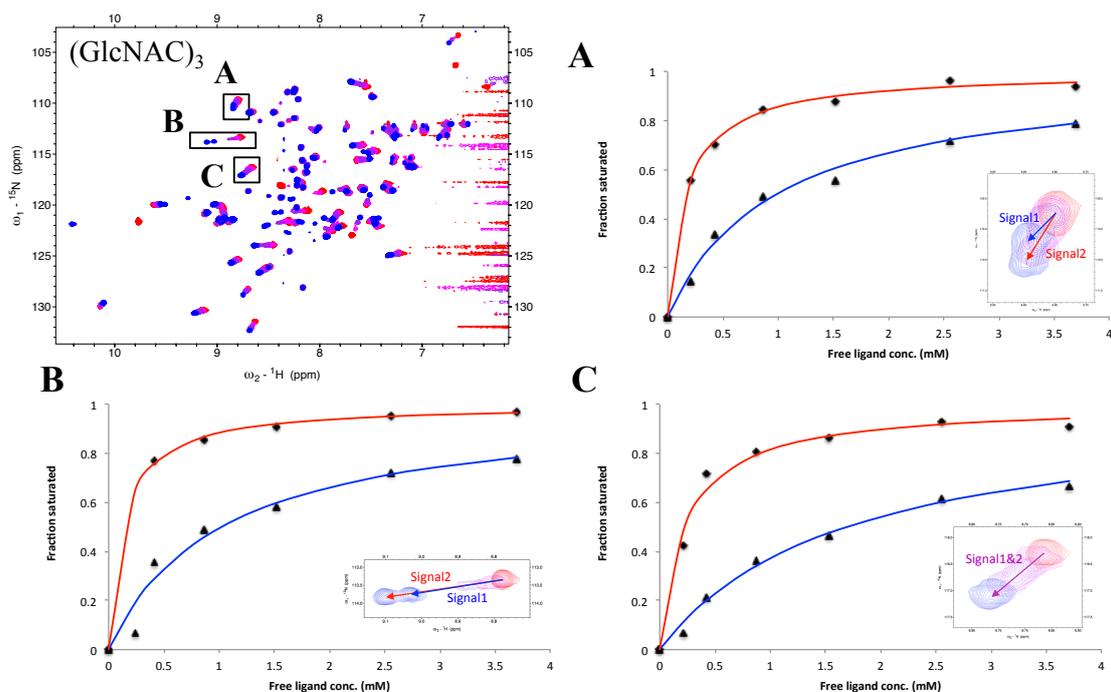


Figure III-12. Chemical shift perturbation experiments of VcLysM tandem upon binding to (GlcNAc)₃. Red and blue lines in binding curves and the insets corresponds to binding curves and chemical shift migration of resonances derived from VcLysM1 and VcLysM2, respectively.

Table III-5. Binding constants and corresponding Gibbs free energy changes of VcLysM tandem-(GlcNAc)₃ binding.

	Gly28	Gly92	Gly51	Gly115	Gly60	Gly124
K_a (M ⁻¹)	7.57×10^3	9.80×10^2	4.40×10^3	5.93×10^2	5.96×10^3	1.02×10^3
ΔG (kcal mol ⁻¹)	- 5.29	- 4.08	- 4.97	- 3.78	- 5.15	- 4.10

Since binding curves derived from chemical shift perturbation experiments often depend on the residue selected and binding affinities for (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ were too high to determine the binding affinity by the experiments, (GlcNAc)₃₋₆ titration experiments were also performed using ITC (Fig. III-13). The thermodynamic parameters were calculated by fitting regression curves assuming sequential binding mode as this model showed better fitness to the experimental data than Two Sets of the Sites model in the Origin software and is summarized on the table III-6.

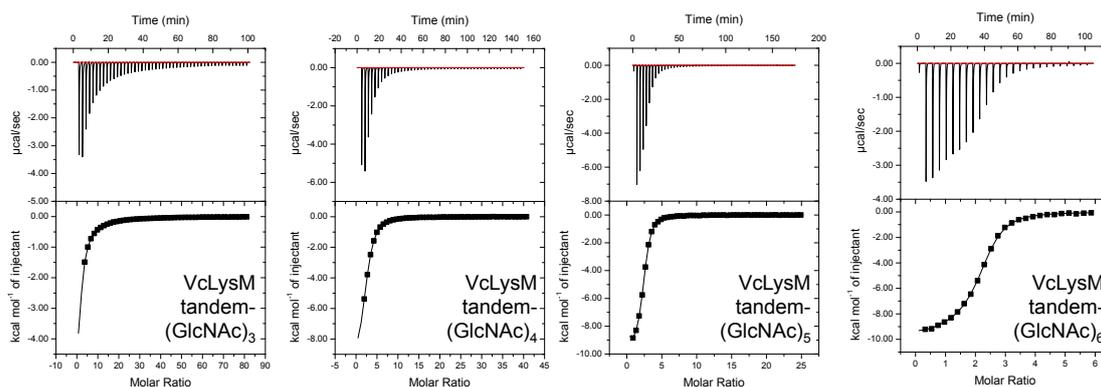


Figure III-13. Thermograms and corresponding binding isotherms of VcLysM tandem-(GlcNAc)₃₋₆ titration experiments.

Table III-6. Thermodynamics parameters of VcLysM tandem-(GlcNAc)₃₋₆ binding experiments.

		(GlcNAc) ₃	(GlcNAc) ₄	(GlcNAc) ₅	(GlcNAc) ₆
Site 1	<i>N</i>	1.00	1.00	1.00	1.00
	<i>K_a</i> (M ⁻¹)	6.60 x 10 ³	9.84 x 10 ³	1.55 x 10 ⁴	2.43 x 10 ⁴
	ΔH (kcal mol ⁻¹)	- 9.68	- 14.8	- 11.9	- 12.3
	$T\Delta S$ (kcal mol ⁻¹)	- 4.47	- 9.33	- 6.14	- 6.29
	ΔG (kcal mol ⁻¹)	- 5.21	- 5.44	- 5.71	- 5.98
<hr/>					
Site 2	<i>N</i>	1.00	1.00	1.00	1.00
	<i>K_a</i> (M ⁻¹)	7.89 x 10 ²	8.59 x 10 ³	5.33 x 10 ⁴	9.76 x 10 ⁴
	ΔH (kcal mol ⁻¹)	- 9.46	- 7.64	- 9.85	- 7.60
	$T\Delta S$ (kcal mol ⁻¹)	- 5.51	- 2.27	- 3.40	- 0.796
	ΔG (kcal mol ⁻¹)	- 3.95	- 5.36	- 6.44	- 6.80

To assign the two datasets, Site 1 and Site 2, to the VcLysM1/VcLysM2, the data obtained from chemical shift perturbation experiments (Fig. III-12 and Table III-5) were used for comparison. Calculated ΔG° for three pairs of residues, Gly28/Gly92, Gly51/Gly115 and Gly60/Gly124, were comparable to the values given by the ITC experiment; hence, the Site 1 and Site 2 are most likely to correspond to the binding sites on VcLysM1 and VcLysM2, respectively. When we compare the ΔG of VcLysM tandem with VcLysM singles, small decrease for VcLysM1 (-0.3 kcal/mol) and small increase in VcLysM2 (0.1 kcal/mol) was observed.

We calculated the thermodynamic parameters for (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ based on ITC experiments. Apparently, thermodynamic parameters for Site 1 were more exothermic than Site 2 suggesting that Site 1 and Site 2 correspond to VcLysM1 and VcLysM2, respectively. When we referred the chemical shift perturbation experiments upon binding to (GlcNAc)₆ using NMR, the binding curves drawn by plotting $\Delta\delta$ against ligand concentration without normalization resulted in sigmoidal patterns for both resonances for VcLysM1 and VcLysM2 (data not shown) indicating that the binding affinities for individual domains strongly affected each other, although it is hard to discuss based on precise affinities for individual domains.

Discussion

Although we couldn't obtain crystal structure of VcLysMs in complex with chitin oligosaccharides, docking simulation provided deeper insights into the binding mode of VcLysM1. Five sugar residues of (GlcNAc)₆ were accommodated in the binding cleft, but one of the GlcNAc residues, GlcNAc1, poorly interacted with VcLysM1 indicating that the ligand binding site of VcLysM1 is composed of five subsites excluding the subsite for GlcNAc1.

When we see more details on the interaction between VcLysM1 and (GlcNAc)₆, Val53 is unlikely to directly interact with (GlcNAc)₆, but likely to support the binding site from behind the ligand binding cleft. Effect of mutation of Val53 was also limited in minor change in the binding thermodynamics in comparison with the mutation of Thr55. However, VcLysM1 and the V53N mutant showed EEC upon binding to chitin oligosaccharides (Fig. III-3). We assumed that this EEC phenomenon was due to the similarity in the interactions between these two proteins and was derived from change in the loop structure supported by Val53. On the other hand, effect of mutation in Thr55 was greater than that of V53 mutation and thermodynamic parameters of T55A mutant upon binding to (GlcNAc)₄, (GlcNAc)₅ and (GlcNAc)₆ were dissimilar to those of VcLysM1 nor VcLysM2 suggesting that there was certain difference in the interactions. Since there was no direct interaction by the side chain of Thr55, we assume that this change in thermodynamic parameters was due to the change in the loop structure or indirect hydrogen-bonding interaction through water molecule.

Since the binding site on VcLysM1 seems to be composed of several polar groups and single aromatic residue stacking through the side chain and the shape of the binding site was rather open, but not flat indicating that VcLysM1 belongs to Type B CBMs. Specificity on the catalytic domain can severely be affected by the specificity of attached CBM. As such, natural substrate of VcChi can be speculated to be amorphous or region close to an amorphous polysaccharide chain.

When we compared the spectra of VcLysM tandem and VcLysM singles, spectra were almost completely overlapped except the residues neighboring to the linker region. Therefore, there was no direct interaction between LysMs linked in tandem. In addition to this observation, there was almost no effect on binding affinities to (GlcNAc)₃ indicating that there is neither intermolecular nor intramolecular interaction, which was observed for previous reports (Sánchez-Vallet et al. 2013; Wong et al. 2015; Liu et al.

2016). In the case of binding to (GlcNAc)₄ and longer oligosaccharides, however, the interactions seem to be affected by the formation of tandem structure as was mentioned above (Table III-6). Patra *et al.* (2015) reported negative cooperativity of dimeric LysM domain connected to a lectin domain through steric hindrance (Patra et al. 2016). Therefore we concluded that decrease in the affinity in one LysM and sigmoidal pattern in the binding curves were simply due to steric hindrance of oligosaccharides bound to another LysM.

Chapter IV

Concluding remarks

For binding experiments between CBMs and carbohydrates, there are uncertainties in the results obtained. For example, even if macroscopic experiments show binding of a CBM to the ligand, these results may be derived from non-specific adsorption or specific binding to only limited part of the ligand. In this meaning, we have to perform experiments with homogeneous samples. Using modern techniques or compromise with heterogeneous samples with such uncertainties. To get insights into the binding mode to heterogeneous ligands, we have to carefully analyze the experimental results obtained using homogeneous ligands homologous to such ligands. Since carbohydrates have completely different physical properties depending on the chemical composition, the glycosidic linkages and the degree of polymerization, polysaccharides are composed of several parts with various kinds of properties. Hence it is hard to examine the binding mode of CBMs on such polysaccharides. Since structure of CBM is complementary to the ligand and the specificity can partially be speculated based on the structure, binding modes of CBMs have sometimes been speculated from the structures (Boraston et al. 2004).

LysMs had been reported to bind to chitinous materials, such as chitin, chitin oligosaccharides, lipo-chitooligosaccharides, peptidoglycan and its fragments, in which each sugar unit is normally linked by β -1,4 glycosidic linkage (Ohnuma et al. 2008; Liu et al. 2012; Mesnage et al. 2014; Bozsoki et al. 2017). However, the (speculated) target molecules for LysMs vary from oligosaccharides to complex polysaccharides and, in some cases, molecular recognition mechanisms have not yet been clarified. Especially for LysMs acting on polymeric substances, molecular recognition mechanisms had hardly been clarified due to the difficulties mentioned above. Several reports have showed that LysMs or LysM containing proteins can also bind to insoluble materials, but the information are limited.

In the present study, we determined the binding site of VcLysMs and the binding site was found to be located on the cleft on the α -helical surface, as reported for LysMs derived from other proteins. Tryptophan residue located on the edge of the binding site seems to be important for comprising a binding subsite for $(\text{GlcNAc})_n$ with $n=4$ and longer indicating that LysMs containing aromatic residue at this position potentially

target tetramer and longer ligands.

Individual VcLysMs can potentially act on n=4 or longer oligomers; however, the two VcLysMs linked in tandem negatively influenced the affinity to such oligomers. Thus, linking in tandem of the two LysMs is unlikely suitable for such a functionality. I assume that the negative influences may have been resulting from the flexible linker and the LysMs, which do not interact with each other. It seems rather like these LysMs work on polymeric complex substances like peptidoglycans. This is like LysMs acting on peptidoglycan (Wong et al. 2015). Actually, hydroxyl group on C2 of GlcNAc2, GlcNAc4, GlcNAc6, which are replaced by lactate in peptidoglycan, in the simulated structure of VcLysM1 in complex with (GlcNAc)₆ are stuck out into solvent and not interacting with VcLysM1 indicating the potential binding ability of VcLysM1 to peptidoglycan. Although Onaga and coworkers showed that LysMs in chitinase could bind to insoluble crystalline chitin, it is hard to think that these LysMs actually bind to crystalline region of the ligand, but rather likely to bind to amorphous region of such chitinous material (Onaga and Taira 2008). They also reported that presence of LysMs in chitinase is crucial for the antifungal activity. However, the chemical composition and the orientations of components in fungal cell walls have not yet been clarified at molecular level; therefore, the relationship between the LysM structure and antifungal activity is unclear at present.

Taken together, although the natural ligand for VcLysMs is still unknown, it appears to be polysaccharide with amorphous region, such as cell walls of microorganisms. Although there are only a few reports about pathogenesis and symbiosis of bacteria in algal cell, VcChi may play a role in eliminating the microorganisms in the ECM of *V. carteri* to prepare for long-term survival of the zygote (Ramanan et al. 2016).

LysM1 and LysM2 showed difference in the binding affinities. The difference in the binding affinity was derived from the deviation in the loop structure between $\alpha 2$ and $\beta 2$. It is hard to discuss based on the relatively small difference in the binding affinity, but the deviation in the structure might reflect difference in the ligand structure since fungal cell walls are modified by deacetylation and other kinds of glycans, such as β -1,3/1,6 glycans (Latgé 2007). Figure IV-1 depicts environments where VcChi may work in. Further study about the structure of such cell walls will answer this question.

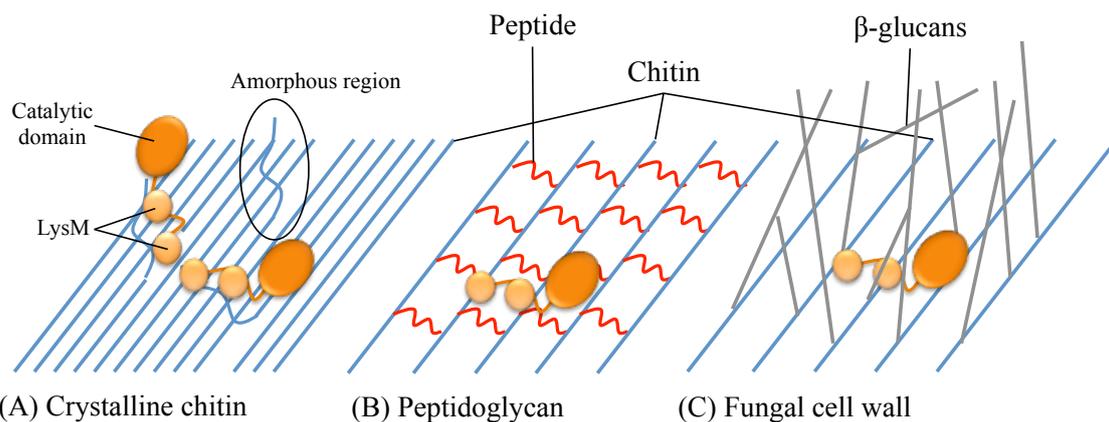


Fig IV-1. Candidates of substrates for VcChi. (a) Amorphous region or crystalline region next amorphous region. (b) Glycosidic linkages of peptidoglycan. (c) Polymeric chitin in fungal cell wall. Polymeric chitin is represented as blue lines. Peptide is represented as red lines. β -glucans are represented as gray lines. VcChi is represented as orange robes linked by orange linker region.

REFERENCES

- Afonine P V, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68:352–67 . doi: 10.1107/S0907444912001308
- Amon P, Haas E, Sumper M (1998) The Sex-Inducing Pheromone and Wounding Trigger the Same Set of Genes in the Multicellular Green Alga *Volvox*. *Plant Cell* 10:781–789 . doi: 10.1105/tpc.10.5.781
- Bateman A, Bycroft M (2000) The structure of a LysM domain from *E. coli* membrane-bound lytic murein transglycosylase D (MltD) 1 Edited by P. E. Wight. *J Mol Biol* 299:1113–1119 . doi: 10.1006/jmbi.2000.3778
- Bielnicki J, Devedjiev Y, Derewenda U, Dauter Z, Joachimiak A, Derewenda ZS (2006) *B. subtilis* ykuD protein at 2.0 Å resolution: Insights into the structure and function of a novel, ubiquitous family of bacterial enzymes. *Proteins Struct Funct Genet* 62:144–151 . doi: 10.1002/prot.20702
- Blackwell J (1969) Structure of β -chitin or parallel chain systems of poly- β -(1→4)-N-acetyl-D-glucosamine. *Biopolymers* 7:281–298 . doi: 10.1002/bip.1969.360070302
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* 382:769–781 . doi: 10.1042/BJ20040892
- Boraston AB, Nurizzo D, Notenboom V, Ducros V, Rose DR, Kilburn DG, Davies GJ (2002) Differential Oligosaccharide Recognition by Evolutionarily-related β -1,4 and β -1,3 Glucan-binding Modules. *J Mol Biol* 319:1143–1156 . doi: 10.1016/S0022-2836(02)00374-1

- Bozsoki Z, Cheng J, Feng F, Gysel K, Vinther M, Andersen KR, Oldroyd G, Blaise M, Radutoiu S, Stougaard J (2017) Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. *Proc Natl Acad Sci* 201706795 . doi: 10.1073/pnas.1706795114
- Buist G, Steen A, Kok J, Kuipers OP (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol Microbiol* 68:838–847 . doi: 10.1111/j.1365-2958.2008.06211.x
- Carlstrom D (1957) THE CRYSTAL STRUCTURE OF alpha-CHITIN (POLY-N-ACETYL-D-GLUCOSAMINE). *J Cell Biol* 3:669–683 . doi: 10.1083/jcb.3.5.669
- Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr Sect D Biol Crystallogr* 50:760–763 . doi: 10.1107/S0907444994003112
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–93
- Emsley P, Cowtan K (2004) *Coot* : model-building tools for molecular graphics. *Acta Crystallogr Sect D Biol Crystallogr* 60:2126–2132 . doi: 10.1107/S0907444904019158
- Garvey KJ, Saedi MS, Ito J (1986) Nucleotide sequence of Bacillus phage phi 29 genes 14 and 15: homology of gene 15 with other phage lysozymes. *Nucleic Acids Res* 14:10001–8
- Goddard, T. D. Kneller, D. G. Sparky 3, University of California, San Francisco

- Güntert P, Buchner L (2015) Combined automated NOE assignment and structure calculation with CYANA. *J Biomol NMR* 62:453–471 . doi: 10.1007/s10858-015-9924-9
- Hart PJ, Monzingo AF, Ready MP, Ernst SR, Robertus JD (1993) Crystal Structure of an Endochitinase from *Hordeum vulgare* L. Seeds. *J Mol Biol* 229:189–193 . doi: 10.1006/JMBI.1993.1017
- Jang MK, Kong BG, Jeong Y Il, Lee CH, Nah JW (2004) Physicochemical characterization of α -chitin, β -chitin, and γ -chitin separated from natural resources. *J Polym Sci Part A Polym Chem* 42:3423–3432 . doi: 10.1002/pola.20176
- Kim YS, Lee JH, Yoon GM, Cho HS, Park SW, Suh MC, Choi D, Ha HJ, Liu JR, Pai HS (2000) CHRK1, a chitinase-related receptor-like kinase in tobacco. *Plant Physiol* 123:905–915 . doi: 10.1104/pp.123.3.905
- Kirk DL (2004) Volvox. *Curr Biol* 14:R599-600 . doi: 10.1017/CBO9780511529740
- Kirk DL, Kirk MM (1986) Heat shock elicits production of sexual inducer in Volvox. *Sci. (New York, NY)* 231:51–54
- Latgé JP (2007) The cell wall: A carbohydrate armour for the fungal cell. *Mol Microbiol* 66:279–290 . doi: 10.1111/j.1365-2958.2007.05872.x
- Liu S, Wang J, Han Z, Gong X, Zhang H, Chai J (2016) Molecular Mechanism for Fungal Cell Wall Recognition by Rice Chitin Receptor OsCEBiP. *Structure* 24:1192–1200 . doi: 10.1016/j.str.2016.04.014
- Liu T, Liu Z, Song C, Hu Y, Han Z, She J, Fan F, Wang J, Jin C, Chang J, Zhou J-M, Chai J (2012) Chitin-Induced Dimerization Activates a Plant Immune Receptor. *Science (80-)* 336:1160–1164 . doi: 10.1126/science.1218867

- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674 . doi: 10.1107/S0021889807021206
- Mesnager S, Dellarole M, Baxter NJ, Rouget J-B, Dimitrov JD, Wang N, Fujimoto Y, Hounslow AM, Lacroix-Desmazes S, Fukase K, Foster SJ, Williamson MP (2014) Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nat Commun* 5: . doi: 10.1038/ncomms5269
- Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr Sect D Biol Crystallogr* 53:240–255 . doi: 10.1107/S0907444996012255
- Ohnuma T, Onaga S, Murata K, Taira T, Katoh E (2008) LysM domains from *Pteris ryukyuensis* chitinase-A: A stability study and characterization of the chitin-binding site. *J Biol Chem* 283:5178–5187 . doi: 10.1074/jbc.M707156200
- Ohnuma T, Umemoto N, Kondo K, Numata T, Fukamizo T (2013) Complete subsite mapping of a “loopful” GH19 chitinase from rye seeds based on its crystal structure. *FEBS Lett* 587:2691–2697 . doi: 10.1016/j.febslet.2013.07.008
- Ohnuma T. TT (2014) **植物キチナーゼの構造と機能, その利用. 応用糖質科学** 4:121–126
- Onaga S, Taira T (2008) A new type of plant chitinase containing LysM domains from a fern (*Pteris ryukyuensis*): Roles of LysM domains in chitin binding and antifungal activity. *Glycobiology* 18:414–423 . doi: 10.1093/glycob/cwn018
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326 . doi: 10.1016/S0076-6879(97)76066-X

- Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* 4:2411–2423 . doi: 10.1002/pro.5560041120
- Patra D, Mishra P, Vijayan M, Surolia A (2016) Negative Cooperativity and High Affinity in Chitooligosaccharide Binding by a *Mycobacterium smegmatis* Protein Containing LysM and Lectin Domains. *Biochemistry* 55:49–61 . doi: 10.1021/acs.biochem.5b00841
- Perrakis A, Tews I, Dauter Z, Oppenheim AB, Chet I, Wilson KS, Vorgias CE (1994) Crystal structure of a bacterial chitinase at 2.3 Å resolution. *Structure* 2:1169–1180 . doi: 10.1016/S0969-2126(94)00119-7
- Ramanan R, Kim BH, Cho DH, Oh HM, Kim HS (2016) Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnol Adv* 34:14–29 . doi: 10.1016/j.biotechadv.2015.12.003
- Ramírez-Gualito K, Alonso-Ríos R, Quiroz-García B, Rojas-Aguilar A, Díaz D, Jiménez-Barbero J, Cuevas G (2009) Enthalpic nature of the CH/ π interaction involved in the recognition of carbohydrates by aromatic compounds, confirmed by a novel interplay of NMR, calorimetry, and theoretical calculations. *J Am Chem Soc* 131:18129–18138 . doi: 10.1021/ja903950t
- Sánchez-Vallet A, Saleem-Batcha R, Kombrink A, Hansen G, Valkenburg DJ, Thomma BPHJ, Mesters JR (2013) Fungal effector Ecp6 outcompetes host immune receptor for chitin binding through intrachain LysM dimerization. *Elife* 2013:1–16 . doi: 10.7554/eLife.00790
- Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44:213–223 . doi: 10.1007/s10858-009-9333-z

- Shinya S, Ohnuma T, Yamashiro R, Kimoto H, Kusaoke H, Anbazhagan P, Juffer AH, Fukamizo T (2013) The first identification of carbohydrate binding modules specific to chitosan. *J Biol Chem* 288:30042–30053 . doi: 10.1074/jbc.M113.503243
- Starr RC, Jaenicke L (1974) Purification and characterization of the hormone initiating sexual morphogenesis in *Volvox carteri* f. *nagariensis* Iyengar. *Proc Natl Acad Sci U S A* 71:1050–1054 . doi: 10.1073/pnas.71.4.1050
- Taira T (2010) Structures and Antifungal Activity of Plant Chitinases. *J Appl Glycosci* 57:167–176 . doi: 10.5458/jag.57.167
- Taira T, Mahoe Y, Kawamoto N, Onaga S, Iwasaki H, Ohnuma T, Fukamizo T (2011) Cloning and characterization of a small family 19 chitinase from moss (*Bryum coronatum*). *Glycobiology* 21:644–654 . doi: 10.1093/glycob/cwq212
- Terwisscha van Scheltinga AC, Kalk KH, Beintema JJ, Dijkstra BW (1994) Crystal structures of hevamine, a plant defence protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure* 2:1181–1189 . doi: 10.1016/S0969-2126(94)00120-0
- Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y, Steitz TA (1996) Crystal structure of a bacterial {family-III} cellulose-binding domain: a general mechanism for attachment to cellulose. *{Embo} J* 15:5739–5751 . doi: 8918451
- Turnbull WB, Daranas AH (2003) On the Value of c: Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? doi: 10.1021/JA036166S
- Vidhyasekaran P (2014) *PAMP Signals in Plant Innate Immunity*. Springer

Willmann R, Lajunen HM, Erbs G, Newman M-A, Kolb D, Tsuda K, Katagiri F, Fliegmann J, Bono J-J, Cullimore J V, Jehle AK, Götz F, Kulik A, Molinaro A, Lipka V, Gust AA, Nürnberger T (2011) Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc Natl Acad Sci U S A* 108:19824–9 . doi: 10.1073/pnas.1112862108

Wong JEMM, Alsarraf HMAB, Kaspersen JD, Pedersen JS, Stougaard J, Thirup S, Blaise M (2014) Cooperative binding of LysM domains determines the carbohydrate affinity of a bacterial endopeptidase protein. *FEBS J* 281:1196–1208 . doi: 10.1111/febs.12698

Wong JEMM, Midtgaard SR, Gysel K, Thygesen MB, Sorensen KK, Jensen KJ, Stougaard J, Thirup S, Blaise M (2015) An intermolecular binding mechanism involving multiple LysM domains mediates carbohydrate recognition by an endopeptidase. *Acta Crystallogr Sect D Biol Crystallogr* 71:592–605 . doi: 10.1107/S139900471402793X

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my old supervisor, Prof. Tamo Fukamizo, School of biochemistry in Suranaree University of technology. Without his kind guidance, supports and critical review, I couldn't finish this work.

I also would like to express my gratitude to my supervisor, Prof. Shigeru Shigeoka, faculty of agriculture of Kindai University. If he didn't take over the supervision from Prof. Tamo Fukamizo, I couldn't continue my work in Kindai University.

I would like to express my appreciation to the members of my thesis committee, especially Prof. Ryutaro Utsumi and Prof. Takashi Kitayama, Faculty of Agriculture, Kindai University, for giving advice and reviewing this thesis.

I greatly appreciate suggestions and discussions given by Assoc. Prof. Takayuki Ohnuma, faculty of agriculture of Kindai University. His sever criticism gave me deeper insight into experimental results.

I wish to thank Prof. Finn Lillelund Aachmann, Norwegian University of science and technology, as he took supervision for the major part of the last year of my doctoral degree. His kind supports in Norway and encouragement helped me to keep hanging in there.

I would like to show appreciation to Prof. Kjell Morten Vårum, Norwegian University of science and technology. His kind guidance for my attitude as a researcher helped me to keep clear mind through this work. I hope that his soul may rest in peace.

I appreciate Prof. Toki Taira, faculty of agriculture, University of the Ryukyus. He supplied the gene of plant chitinases and gave me inspirations on my work.

I want to thank Assistant Prof. Shigenori Nishimura, School of Life and environmental sciences of Osaka Prefecture University, for his kind supports on determining solution structures and related discussions.

I appreciate Dr. Tomoyuki Numata, Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology, as he instructed me about X-ray crystal structure analyses.

I wish to thank members of Biomolecular chemistry laboratory and coworkers in Norwegian biopolymer group. They helped me a lot through my doctoral degree.

Finally, I would like to thank my family for kind supports in many aspects.

I acknowledge the financial supports by Research Fellowship for Young Scientists from Japan Society for the Promotion of Science (16J10483), Fellowship of Kyoto International Forum for Environmental and Energy and The Scandinavia·Japan Sasakawa Foundation (grant no. 15-17).