Structure
Article

Structure-Function Analysis of a CVNH-LysM Lectin Expressed during Plant Infection by the Rice Blast Fungus *Magnaporthe oryzae*

Leonardus M.I. Koharudin,[1] Arturo R. Viscomi,[2] Barbara Montanini,[2] Michael J. Kershaw,[3] Nicholas J. Talbot,[3] Simone Ottonello,[2,*] and Angela M. Gronenborn[1,*]
[1]Department of Structural Biology, University of Pittsburgh School of Medicine, Biomedical Science Tower 3, 3501 Fifth Avenue, Pittsburgh, PA 15260, USA
[2]Dipartimento di Biochimica e Biologia Molecolare, Università di Parma, Viale G.P., Usberti 23/A, 43100 Parma, Italy
[3]School of Biosciences, University of Exeter, Exeter EX4 4QD, UK
*Correspondence: simone.ottonello@unipr.it (S.O.), amg100@pitt.edu (A.M.G.)
DOI 10.1016/j.str.2011.03.004

SUMMARY

The rice blast fungus *Magnaporthe oryzae's* genome encodes a hypothetical protein (MGG_03307) containing a type III CVNH lectin, in which a LysM domain is inserted between individual repeats of a single CVNH domain. At present, no structural or ligand binding data are available for any type III CVNH and functional studies in natural source organisms are scarce. Here, we report NMR solution structure and functional data on MGG_03307. The structure of the CVNH/LysM module revealed that intact and functionally competent CVNH and LysM domains are present. Using NMR titrations, carbohydrate specificities for both domains were determined, and it was found that each domain behaves as an isolated unit without any interdomain communication. Furthermore, live-cell imaging revealed a predominant localization of MGG_03307 within the appressorium, the specialized fungal cell for gaining entry into rice tissue. Our results suggest that MGG_03307 plays a role in the early stages of plant infection.

INTRODUCTION

The Cyanovirin-N homolog (CVNH) proteins are a recently described class of lectins whose biological roles are poorly understood. All CVNH proteins comprise tandem sequence repeats with amino acid identities ranging from 11% to 40% between the repeats (Koharudin et al., 2008; Percudani et al., 2005), including the parent protein CV-N (Bewley et al., 1998; Boyd et al., 1997). Based on their domain organization, CVNHs have been grouped into three categories: type I proteins contain either single or multiple CVNH domains, each composed of tandem-sequence repeats; type II are multidomain proteins which, in addition to the CVNH domain sequence, also contain an MS8 domain of unknown function; and type III proteins comprise interrupted CVNH multidomains, in which a LysM domain sequence is inserted between individual repeats of a single CVNH domain sequence (Percudani et al., 2005). All CVNH proteins share a common pseudosymmetric fold originally determined for CV-N (Bewley et al., 1998), and the structures of four homologs and several mutant type I proteins have been determined by NMR spectroscopy or X-ray crystallography (Barrientos et al., 2003; Fromme et al., 2007; Koharudin et al., 2009; Koharudin et al., 2008; Matei et al., 2008). CVNHs have previously been reported to be mannose-binding lectins and their specificity for certain linkages has made them useful as tools in glycobiology (Bewley, 2001; Botos et al., 2002; Koharudin et al., 2008; Shenoy et al., 2002). In addition, CV-N and other cyanobacterial lectins exhibit potent anti-viral activity, which is mediated by binding to high-mannose sugars on virus envelope glycoproteins (Barrientos and Gronenborn, 2005; Botos et al., 2002; Boyd et al., 1997; Williams et al., 2005). In contrast to the wealth of structural and sugar-binding data for type I CVNHs, no such information is available for type II and III CVNHs. Given their extensive sequence similarity, it is expected that the fold of the CVNH domains in the modular type II proteins is analogous to that of type I proteins. However, for type III proteins, it is not clear, a priori, whether the characteristic fold of the CVNH domain will be retained, given the interruption in the sequence by the insertion of a Lysin motif (LysM) domain.

The LysM domain was originally described in a lysozyme from the *Bacillus* phage phi 29 (Garvey et al., 1986; Saedi et al., 1987). Lysozymes degrade bacterial cell walls by hydrolyzing glycosidic bonds between N-acetylmuramic acid and N-acetyl-D-glucosamine (GlcNAc) in peptidoglycan (Buist et al., 2008; de Jonge and Thomma, 2009; Donovan, 2007). Subsequently, LysM domains were found in many modular bacterial and fungal enzymes involved in cell wall degradation. For example, chitinases, enzymes that break down polymeric GlcNAc, the main constituent of fungal cell walls (Buist et al., 2008; de Jonge and Thomma, 2009) as well as various hydrolases, involved in remodeling of cell wall peptidoglycans during bacterial cell division, contain LysM domains (Layec et al., 2009; Moll et al., 2010). In addition to the hydrolytic activity documented for bacterial LysM proteins, a morphogenetic role associated with endospore coat assembly and development has also been described (Costa et al., 2006).

In contrast to CVNH sequences that have only been found in bacteria, filamentous fungi and seedless plants so far, LysM

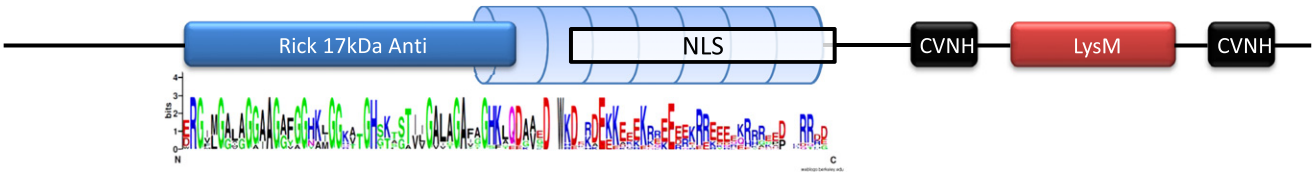

Figure 1. Domain Architecture of Type III CVNH Proteins

Conserved domains present in MGG_03307 and in nine additional type III CVNHs are shown. These include the N-terminal Rick 17 kDa domain (PF05433), the CVNH domain (PF08881) and the LysM domain (PF01476). Also conserved among the ten type III CVNH proteins is a predicted coiled-coil region (light-blue cylinder), comprising a putative nuclear localization signal (*NLS*). A logo representation of amino acid sequence conservation in the N-terminal portion of type III CVNHs is shown below the Rick 17kDa domain and the coiled-coil/NLS region of the proteins. See also Figure S1.

domains are also present in higher eukaryotes, including plants and humans (Buist et al., 2008; Ponting et al., 1999). For example, extracellular LysM domain proteins have been identified in plant cell surface receptors that mediate interaction between symbiotic bacteria and leguminous plants by recognizing lipid substituted-chitin oligosaccharides (Nod factors) released by Rhizobia (Knogge and Scheel, 2006; Zhang et al., 2007). In *Arabidopsis*, a LysM-containing receptor-like kinase was shown to be involved in defense signaling in response to fungal pathogens (Miya et al., 2007; Wan et al., 2008).

More than 400 putative LysM domain-containing proteins have been identified in fungi (de Jonge and Thomma, 2009) and most of them solely contain strings of LysMs. For example, type A LysM proteins, also called "LysM effectors," are secreted polypeptides that have been shown to be involved in the interaction between the fungus and its host (de Jonge and Thomma, 2009). They contain from one up to seven LysM domains and have been reported to play functional roles such as shielding of the fungal cell wall from carbohydrate degrading enzymes released by the invaded plant and suppressing host defense mechanisms via sequestration of oligosaccharide (e.g., chitin) breakdown products of fungal origin that otherwise trigger plant immunity (de Jonge and Thomma, 2009; de Jonge et al., 2010; van den Burg et al., 2006). LysM domains are also found together with enzyme domains such as chitinase and other hydrolases, or with additional carbohydrate-binding modules. In contrast, the presence of a putative LysM domain within another domain sequence has only been seen, so far, for the modular type III CVNHs (Percudani et al., 2005) that constitute one of the least populated groups of fungal LysM-containing proteins (de Jonge and Thomma, 2009). Based on sequence similarity alone, it had been suggested that the conformationally distinct, yet functionally related, LysM and CVNH domains may fold into their canonical structures and simultaneously bind oligo-N-acetylglucosamine (LysM) and oligo-mannose (CVNH) moieties of complex sugars (Percudani et al., 2005).

We observed that out of the ten ascomycete fungi that contain type III CVNHs, nine are plant disease-causing species. We therefore decided to carry out a structure-function study of a type III CVNH from a plant pathogenic fungus, determining the glycan-binding properties of the lectin and investigating its role in fungal pathogenicity. We decided to analyze a CVNH from the devastating rice blast fungus *Magnaporthe oryzae* (Wilson and Talbot, 2009). Rice blast disease causes substantial losses in yield of cultivated rice and is a considerable threat to global food supplies (Pennisi, 2010).

Here, we report the structure of the type III CVNH module (hereafter designated as MoCVNH-LysM) embedded in the hypothetical MGG_03307 protein from *M. oryzae*. The structure and sugar binding studies of the CVNH/LysM module of MoCVNH-LysM revealed intact and functionally competent CVNH and LysM domains. Carbohydrate specificities for both domains were identified using direct titrations of sugars into protein solutions, using ^{1}H-^{15}N HSQC spectroscopy to identify and follow binding. In vivo localization studies found MGG_03307 to be preferentially expressed and localized within the appressorium, a dome-shaped specialized structure that is used by the fungus to breach the leaf cuticle and invade the plant during infection. The results of our combined structural and in vivo localization studies provide direct insight into the function of MGG_03307.

RESULTS AND DISCUSSION

Domain Architecture and Cellular Localization of MGG_03307

A schematic diagram of the predicted domain architecture of type III CVNHs, based on the analysis for ten proteins, including MGG_03307, is displayed in Figure 1. In all proteins, a Rick 17 kDa domain (PF05433), followed by a predicted coiled-coil domain, bearing a putative nuclear localization signal (NLS) precedes the CVNH (PF08881) domains with their interposed LysM (PF01476) domain. The Rick domain resembles a Rickettsia 17 kDa surface antigen of unknown function. Although no nuclear function thus far has been ascribed to any type III CVNH, it is interesting to note that fusion with MGG_03307 restores the nuclear localization capacity of an NLS-lacking transcription factor in yeast (see Figure S1 available online), and that a nucleotide-binding site in a putative region for transcriptional regulation (GO0000166 and GO0006355, respectively) has been identified within the Rick 17 kDa domain of type III CVNHs (de Jonge and Thomma, 2009). Also worthy of note is the absence of any recognizable secretion signal and the presence of only one, poorly conserved cysteine in the LysM domain sequence of all type III CVNHs, compared to multiple Cys residues in most eukaryotic LysM effectors, with some cysteines essential for protein stability upon secretion in the host apoplast (de Jonge and Thomma, 2009). MGG_03307 and its homologs are therefore expected to be intracellular proteins.

To gain direct insight into MGG_03307 localization in its source organism, we generated N- and C-terminal GFP gene

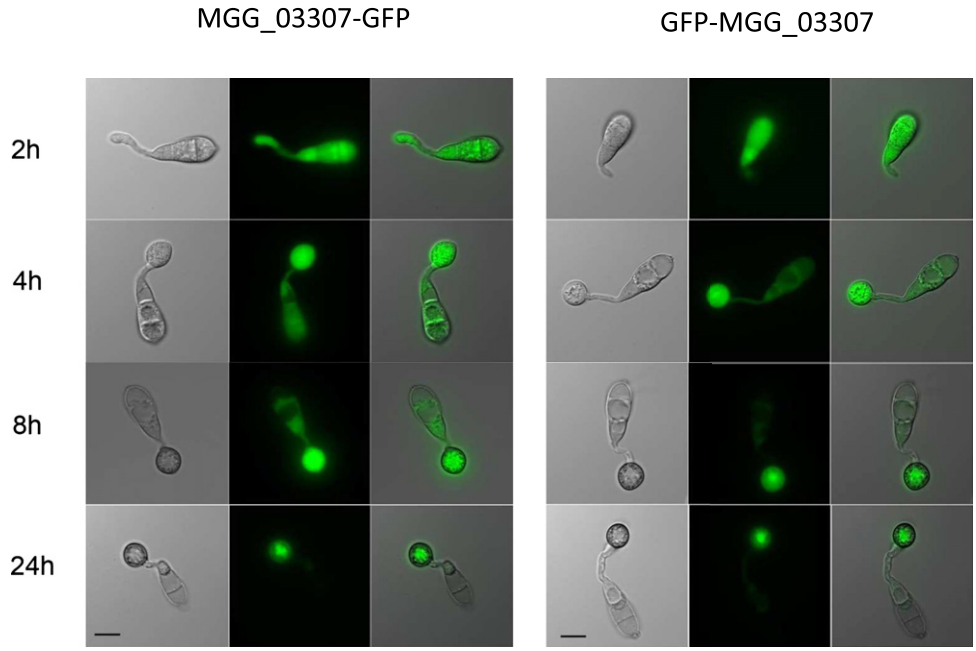

Figure 2. Expression and Cellular Localization of MGG_03307 during Infection-Related Development of *M. oryzae*
Conidia were harvested from Guy-11 transformants expressing N-terminal (*MGG_03307-GFP*) or C-terminal (*GFP-MGG_03307*) MGG_03307-SGFP fusion proteins, deposited onto glass coverslips, and observed by epifluorescence microscopy at the times indicated (scale bars, 10 μm).

fusions and expressed these under control of the natural *MGG_03307* promoter in *M. oryzae.* Transformants were selected, analyzed by DNA gel blot and then investigated by live cell imaging. Spores were inoculated onto hydrophobic surfaces to induce the development of appressoria (Wilson and Talbot, 2009). We observed MGG_03307-GFP fluorescence at an early stage (2 hr) of conidia germination, followed by a progressively increased accumulation of MGG_03307-GFP fluorescence during appressorium formation (Figure 2). Fluorescence was predominantly associated with the cytoplasm, although a more punctate distribution of fluorescence was observed in mature appressoria (Figure 2). Differential expression of MGG_03307 was confirmed by analysis of previously published transcriptional profiling data that revealed high level, specific expression of MGG_03307 in appressoria (Oh et al., 2008).

Isolation and Functional Validation of the Multidomain CVNH/LysM Module

We next focused on the CVNH/LysM dual-domain, the only functionally interpretable module in MGG_03307. As shown by the sequence alignment in Figure 3A, there is extensive similarity between the CVNH and the LysM domain sequences of MGG_03307 and the corresponding sequences of prototypic CVNH and LysM proteins of known structure. A notable difference is the presence of two glycine-rich, nine-residue-long linker sequences in MGG_03307 as opposed to a single, four-residue linker connecting the tandem repeats of canonical (type I) CVNHs (Figure 3A). One linker connects the end of the first CVNH repeat (residue 225) to the N-terminal residue of the LysM domain (residue 235), while the other joins the C terminus of the LysM domain (residue 282) to the first amino acid of the second CVNH repeat (residue 293). As for most CVNHs, there is extensive similarity between the CVNH repeats in MGG_03307 (~36%), and a substantial sequence conservation is also apparent between the individual CVNH repeats of MGG_03307 and those of other CVNHs, suggesting a structural and/or functional role in carbohydrate binding. A number of conserved hydrophobic amino acids are present in the core of the CVNH domain structures (e.g., L194, L212, F293, L309, and L326), but sequence similarity also extends to structurally important polar amino acids such as Ser183 and Ser296, whose corresponding residues in CV-N connect helical turns α1 and α3 to strands β9 and β6, and to amino acids within the sugar binding sites. High sequence similarity is also seen for the LysM domain (Figure 3A). Conserved residues are mainly located at the N terminus of helix α1 and the C terminus of helix α2 in the known LysM structures. In addition, the tetrapeptide GD(S/T)L, located in the loop between strand β1 and helix α1, is present in all the examined LysM sequences. In contrast, a significant sequence divergence is observed in the loop region between helix α2 and strand β2.

Sequence alignment also provided the rational framework for dissecting the C-terminal CVNH/LysM dual-domain module into its constituent units for structural and functional analysis. We selected residue G174 in the full-length protein as the N terminus of the LysM-domain construct and renumbered it as residue G1 of MoCVNH-LysM. The corresponding His-tagged recombinant protein was expressed in *E. coli* and purified to homogeneity by metal-affinity chromatography. A 174 amino acid protein, carrying seven additional N-terminal residues that remain after thrombin cleavage of the 6xHis purification tag, was produced.

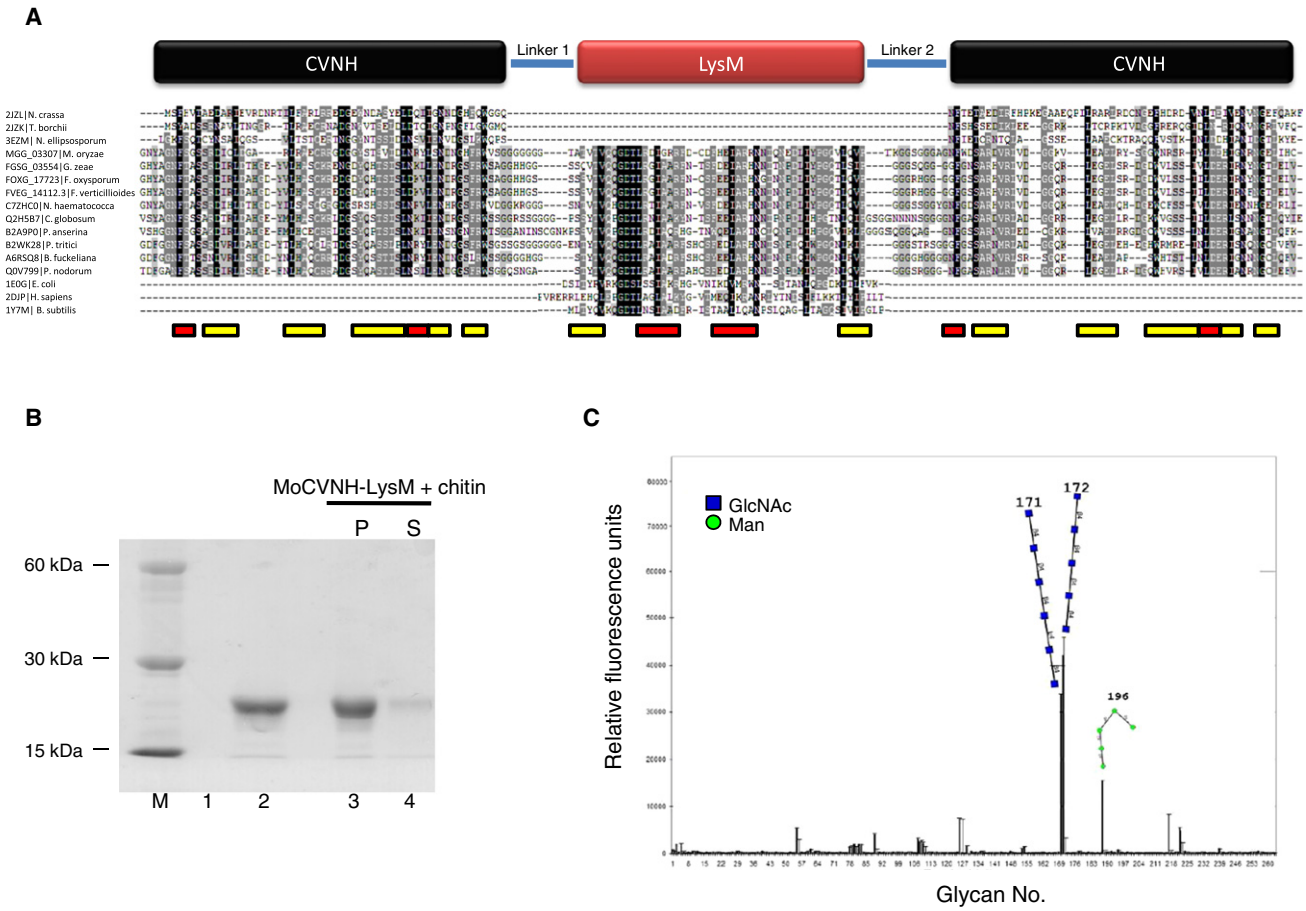

Figure 3. Sequence and Carbohydrate Binding Analysis of MoCVNH-LysM

(A) Alignment of the CVNH/LysM region of MGG_03307 with the corresponding regions of the other type III CVNHs. Also included in the alignment are the sequences of three CVNH (top) and three LysM (bottom) domains of known structure (red, helix; yellow, strand). Amino acid residues conserved in all sequences, in >70%, or in >60% of the sequences are shown on a black, gray, and light-gray background, respectively. The accession numbers and the source organisms for each sequence are listed on the left.

(B) Affinity capture and precipitation of MoCVNH-LysM by chitin-derivatized beads. Free chitin beads and MoCVNH-LysM protein are shown in lanes 1 and 2; MoCVNH-LysM recovered in the pellet (P) and supernatant (S) after incubation with chitin beads are shown in lanes 3 and 4. Samples were analyzed by SDS-polyacrylamide gel electrophoresis and Coomassie blue staining (see Experimental Procedures for details); molecular mass markers (*M*) are shown on the left.

(C) Glycan array analysis of MoCVNH-LysM. Relative fluorescence unit (RFU) values are the average (±SE) of six replicates, from which the maximum and the minimum RFUs were removed; antibody-mediated detection was used to probe MoCVNH-LysM binding. A list of the glycans present on the array (Printed Array version 2.0) can be found at http://www.functionalglycomics.org/static/index.shtml. See also Figure S3.

Importantly, as revealed by exploratory carbohydrate binding experiments, the recombinant MoCVNH-LysM protein bound chitin (Figure 3B), and specifically interacted with one mannose and two N-acetylglucosamine-containing oligosaccharides when challenged with more than 250 different carbohydrates via glycan array analysis (Figure 3C). MoCVNH-LysM therefore was a functionally competent lectin module, capable of selective sugar binding.

Structures of MoCVNH-LysM, the CVNH-LysM Domain of MGG_03307

The 3D solution structure of the MoCVNH-LysM was determined by NMR spectroscopy using uniformly ^{15}N- and $^{13}C/^{15}N$-labeled samples. As evidenced by the excellent chemical shift dispersion and line widths in the ^{1}H-^{15}N HSQC spectra (Figure 4), the protein is well folded. Complete backbone assignments were obtained and resonances of residues residing in the CVNH domain are labeled by residue name and number and are colored in blue and magenta for subdomains A and B, respectively in Figure 4. Pairs of resonances were observed for the seven N-terminal tag residues that remain after thrombin cleavage (Y_{-1}, G_{-3}, I_{-4}, and H_{-5}; colored in red in Figure 4), as well as for G1 and N2, indicating two distinct conformations for this region of the protein in solution. The remainder of the protein (residues Y3 to C167), however, exhibited no resonance doubling. The observed doubling of resonances is most likely caused by *cis*/trans proline isomerization at P_{-2} in the appended tag sequence. Careful inspection of NOESY cross-peaks associated with these residues did not provide structurally distinct NOEs from either conformation, suggesting that it exists mostly

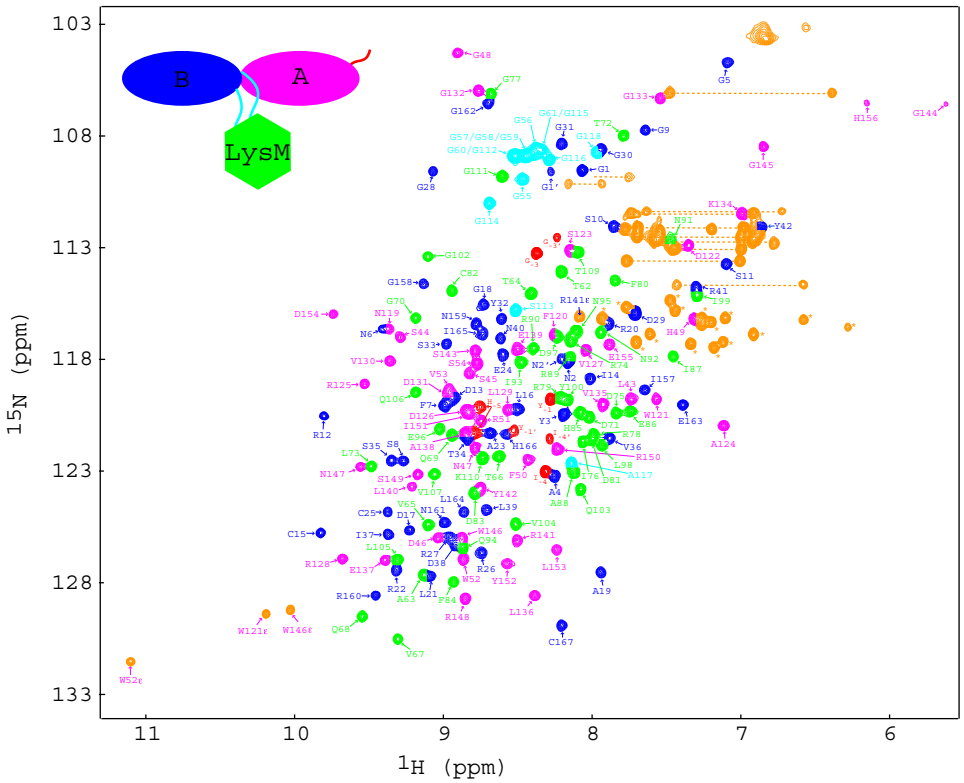

Figure 4. ^{1}H-^{15}N HSQC Spectra of MoCVNH-LysM

Assigned backbone amide resonances are labeled by single amino acid code and residue number for each individual domain; CVNH resonances from domains A and B are colored blue and magenta, respectively, and LysM resonances are colored green. The seven additional residues, appended at the N terminus that remain from the tag after thrombin cleavage are designated $G_{-7}S_{-6}H_{-5}I_{-4}G_{-3}P_{-2}Y_{-1}$ (red) (A). All expected amide resonances of MoCVNH-LysM are observed, including some resonances of the N-terminal appendix (H_{-5}, I_{-4}, G_{-3}, and Y_{-1}). Side-chain resonances of tryptophan, arginine, glutamine, and asparagines residues are colored orange.

as a random tail. This is also supported by the lack of amide signals for S_{-6}, and G_{-7}.

Solution structures of MoCVNH-LysM were calculated using the program CNS, incorporating a total of 2366 distance (NOE and H-bond) and 161 dihedral angle constraints (see Experimental Procedures for details). One thousand structures were calculated and a final ensemble of 25 conformers with the lowest restraint energies was chosen as the structural model of the protein. Best fit superpositions of the 25 conformers of MoCVNH-LysM are displayed in Figures 5A and 5B (C_{α} representation). Since two flexible linkers connect the two domains and no fixed orientation between them could be derived from the NMR data, conformers of each domain were superimposed separately. Average backbone and heavy atom rmsd values with respect to the mean coordinate position are 0.25 ± 0.03 and 0.90 ± 0.07 Å for the CVNH domain (residues 4–52 and 118–166; Figure 3A), and 0.21 ± 0.05 and 0.87 ± 0.07 Å for the LysM domain (residues 61–109; Figure 5B). All individual conformers in the ensembles exhibit excellent covalent geometry and no violation of experimental data (Table 1). The fact that the domains behave as two independent units, rather than a single one, is supported by several pieces of evidence: (1) no NOEs were observed between residues in the linker regions and either of the two domains, and (2) heteronuclear NOE relaxation data indicate that the two linker regions are strikingly flexible (with hetNOE values ranging from 0.13 to 0.51), compared with residues located in either domain. Note that numerous NOEs are present for residues between the two sequence repeats of the CVNH domain, creating a well-defined interface between the two halves of the CVNH domain.

The overall topology of MoCVNH-LysM is shown in ribbon representation using the lowest restraint energy model in the 25 conformer ensemble (Figure 5C). The architecture of the individual domains closely resembles those of previously determined structures of CVNH and LysM domains. The CVNH domain of MoCVNH-LysM (residues 1–52 for sequence repeat 1; residues 120–167 for sequence repeat 2) comprises two triple-stranded β sheets (one formed by strands β1, β2, and β3, the other formed by strands β6, β7, and β8), two β-hairpins (one formed by strands β9 and β10, the other formed by strands β4 and β5), and four 3_{10}-helical turns (α1, α2, α3, and α4). As seen in the superposition of the CV-N structure on the CVNH domain structure determined here (Figure 5D), only small differences between the two domains exist.

The LysM domain (residues 62–109), bisecting the two CVNH repeats, is connected to the canonical CVNH structure by two linkers (shown in orange in Figure 5C). It comprises two antiparallel β strands (β1 and β2), flanked by two helices (α1 and α2), and its overall fold is very similar to both prokaryotic and eukaryotic LysM structures available in the RCSB PDB database.

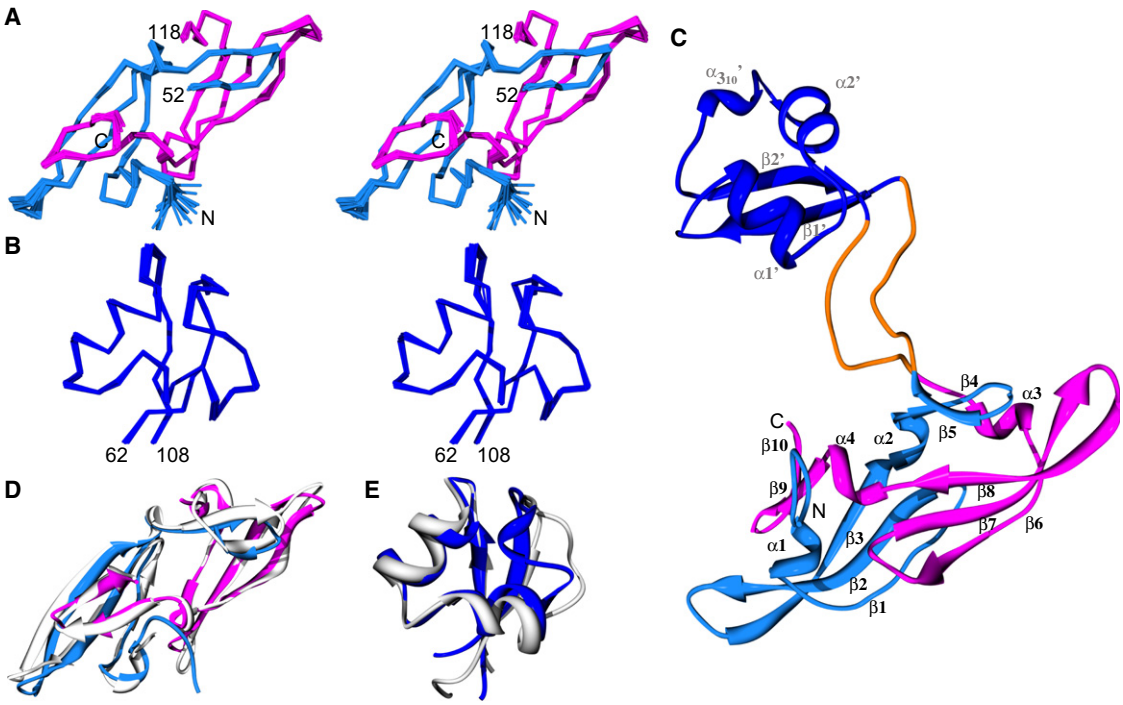

Figure 5. Overall Fold of MoCVNH-LysM

(A and B) Stereo views of the 25 conformer ensembles for the individual CVNH and LysM domains (C_α representation). For the CVNH domain (A), residues 1–52 (sequence repeat 1) and residues 120–167 (sequence repeat 2) are colored in light blue and magenta, respectively. Residues of the LysM domain are colored in dark blue (B).

(C) Ribbon representation of the full-length MoCVNH-LysM structure using the same color code as in (A) and (B). The two flexible linkers are colored orange.

(D and E) Display structural superpositions of the individual MoCVNH-LysM domains onto CV-N (gray) and the LysM domain of the *B. subtilis* YkuD (gray), respectively.

A superposition of the LysM domains of MoCVNH-LysM and of the *B. subtilis* YkuD protein is shown in Figure 5E. The only notable structural differences between these two LysM domains are located in the loops between strand β1 and helix α1 and between helix α2 and strand β2 (Figure 5E). This structural difference is accompanied by poor sequence conservation (see Figure 3A).

Manα(1-2)Man Binding: The CVNH Domain

It is well established that CV-N blocks HIV entry by binding to the Manα(1-2)Man epitopes on the D1 and D3 arms at the nonreducing end of high-mannose carbohydrates present on the viral envelope glycoprotein gp120 (Botos et al., 2002; Shenoy et al., 2002). In addition, all members of the CVNH family are able to recognize Manα(1-2)Man to differing extents (Koharudin et al., 2008). In order to test for Manα(1-2)Man binding by MoCVNH-LysM, we carried out NMR titration experiments and monitored chemical shift perturbations by ^{1}H-^{15}N HSQC spectroscopy. As revealed by these titrations, MoCVNH-LysM interacts with Manα(1-2)Man (Figure 6) solely through the CVNH domain, since only resonances from this domain, but not from the LysM domain were affected. Similar to CV-N and some CVNH homologs, two binding sites were identified: one on subdomain A and the other on subdomain B. The binding site on subdomain A is composed of residues located on the first 3_{10}-helix, α1, the loop connecting β strands 2 and 3, and the β-hairpin between strands 9 and 10, while that on subdomain B comprises residues from the third 3_{10}-helical turn, α3, the loop connecting β strands 7 and 8, and the β-hairpin between strands 4 and 5. Although binding is relatively week, Manα(1-2)Man has a slight preference for subdomain A (apparent K_d of 5.23 ± 0.42 mM compared to subdomain B (apparent K_d 15.13 ± 1.13 mM) (see insets in Figure 6).

Binding of Various N-Acetylglucosamine Oligomers and Derivatives to the LysM Domain

LysM domains are carbohydrate-binding modules that recognize the N-acetylglucosamine (N-GlcNAc) building blocks of complex sugars, such as chitin, a polymer of N-GlcNAc units connected by β-(1,4) linkages, and peptidoglycans, composed by alternating β-(1,4) linked N-GlcNAc and N-acetylmuramic acid (N-MurNAc) units (de Jonge and Thomma, 2009). Although classified as a lectin domain, carbohydrate binding studies with chitin-like N-GlcNAc oligomers have only been reported for the LysM domain of a chitinase-A protein (PrChi-A) (Ohnuma et al., 2008), with apparent binding affinities in the mM range for tetramers and pentamers of N-GlcNAc. Based on structural modeling, the sugar binding site was proposed to reside in a shallow groove lined by the N terminus of helix α1, the loop between α1 and β1, the C terminus of helix α2, and the loop between α2 and β2. No binding data is available for smaller N-GlcNAc oligomers, or for any other LysM domain containing protein. We therefore investigated binding by the LysM domain

Table 1. NMR Constraints and Refinement Statistics for MoCVNH-LysM

	174 aa[a] 25 structures[b]
NMR distance and dihedral constraints	
Distance constraints	2366
Total NOE	2236
Intra residue (|i-j| = 0)	429
Inter residue	
Sequential (|i-j| = 1)	614
Medium-range (1 < |i-j| < 5)	358
Long range (|i-j > 5)	835
Hydrogen bonds	130
Dihedral angle constraints	
phi	81
psi	80
Structure statistics	
Violations (mean and SD)	
Distance constraints (Å)	0.042 ± 0.001
Dihedral angle constraints (°)	0.710 ± 0.053
Max. dihedral angle violation (°)	5
Max. distance constraint violation (Å)	0.5
Deviations from idealized geometry	
Bond lengths (Å)	0.0048 ± 0.0001
Bond angles (°)	0.553 ± 0.008
Impropers (°)	0.534 ± 0.011
CVNH mean rmsd[c] (Å)	
Backbone atoms	0.25 ± 0.03
Heavy atoms	0.90 ± 0.07
LysM mean rmsd[c] (Å)	
Backbone atoms	0.21 ± 0.05
Heavy atoms	0.87 ± 0.07

[a] 167 residues from the main protein chain plus 7 residues of the His-tag cloning artifact after thrombin cleavage.
[b] 1000 structures were calculated and 25 structures with the lowest CNS energy were refined in explicit water.
[c] Mean rmsd was calculated for residues 4–52 and 118–166 and for residues 61–109 for the CVNH and LysM domains, respectively, of the 25 refined structure ensemble.

of MoCVNH-LysM to various N-GlcNAc-containing carbohydrates using NMR chemical shift perturbation as a readout.

Using monomeric N-GlcNAc and N-MurNAc as potential ligands in NMR titrations revealed no or extremely small perturbations in the ^{1}H-^{15}N HSQC spectra, even for a >250-fold molar excess of N-GlcNAc (Figure S2A), whereas detectable chemical shift differences were observed upon titration with the N-MurNAc monosaccharide (Figure S2B). At a molar protein:ligand ratio of ~1:250, R160, R148, Y152, D46, R150, C25, S149 resonances were affected. However, structural mapping of these resonances onto the structure of MoCVNH-LysM did not identify a single patch in the protein (see inset in Figure S2B), and the changes were not saturable upon further addition of N-MurNAc. Therefore, the observed spectral differences are most likely caused by nonspecific interactions between MoCVNH-LysM and the N-MurNAc monosaccharide.

Subsequent titrations were performed with N-GlcNAc-β-(1,4)-N-GlcNAc (or $(N\text{-GlcNAc})_2$) and N-GlcNAc-β-(1,4)-N-MurNAc (Figure 7). These disaccharides represent the simplest oligomeric units of chitin and of the carbohydrate component of peptidoglycan, respectively. The $(N\text{-GlcNAc})_2$ titration, although causing some resonance changes, failed to reach saturation, suggesting that the binding affinity for $(N\text{-GlcNAc})_2$ is at best in the millimolar range. In contrast, binding between the N-GlcNAc-β-(1,4)-N-MurNAc disaccharide and MoCVNH-LysM was saturable and in slow exchange on the chemical shift scale, with an apparent K_d of 3.93 ± 1.45 mM (see inset in Figure 7B). Although surprising, the observed slow exchange may be related to an overall conformational rearrangement in the LysM domain upon carbohydrate binding. Importantly, in both titrations, only resonances of residues belonging to the LysM, but not to the CVNH domain were affected, indicating that there is no communication between the two domains (see also below).

To characterize the interaction between MGG_03307 and higher oligomers of chitin-related cell wall carbohydrates, titrations were also carried out with $(N\text{-GlcNAc})_3$, $(N\text{-GlcNAc})_4$, $(N\text{-GlcNAc})_5$, and $(N\text{-GlcNAc})_6$. These additional titration experiments are displayed in Figure 8. ^{1}H-^{15}N HSQC spectra in the absence or presence of $(N\text{-GlcNAc})_3$, $(N\text{-GlcNAc})_4$, $(N\text{-GlcNAc})_5$, and $(N\text{-GlcNAc})_6$ are shown. Binding curves were determined on the basis of the intensity of five bound resonances (see insets), and the same resonances were used in all titrations with chitin-related glycans. The estimated binding constants increased as a function of N-GlcNAc multiplicity up to n = 6, with apparent K_d values of 560 ± 399, 100 ± 27, 28 ± 3, and 21 ± 4 μM for $(N\text{-GlcNAc})_3$, $(N\text{-GlcNAc})_4$, $(N\text{-GlcNAc})_5$, and $(N\text{-GlcNAc})_6$, respectively. Note that no significant difference was observed between $(N\text{-GlcNAc})_5$ and $(N\text{-GlcNAc})_6$ and only LysM domain resonances were perturbed upon interaction with N-GlcNAc oligomers. Since no appreciable change in line width was observed between free and glycan-bound MoCVNH-LysM, even in the case of $(N\text{-GlcNAc})_6$, we conclude that only a single $(N\text{-GlcNAc})_n$ glycan is bound by each LysM module, with optimal binding reached at the stage of the pentasaccharide. Equivalent titrations with N-GlcNAc-β-(1,4)-N-MurNAc oligomers larger than a disaccharide could not be performed because these compounds are not commercially available.

We next compared the chitin oligomer binding profile of the LysM domain of MoCVNH-LysM with that of PrChi-A LysM from a plant chitinase (Ohnuma et al., 2008). Although binding data were obtained using different methodologies, isothermal titration calorimetry (ITC) for PrChi-A LysM and NMR in the current work, it is interesting to note that the estimated binding affinities of PrChi-A LysM for $(N\text{-GlcNAc})_4$ and $(N\text{-GlcNAc})_5$ are about three orders of magnitude lower than the binding affinity values measured here for MoCVNH-LysM. Inspection of the amino acid sequence differences between these two LysM domains (Figure S3) reveals that they are located in the loop region between helix α2 and strand β2, the site of largest variation in all LysM sequences (Figure 3A). Resonances of amino acids in this loop were clearly affected upon interaction with $(N\text{-GlcNAc})_{4\text{-}5}$ oligomers and almost all amide resonances within the LysM domain of MoCVNH-LysM were strongly perturbed,

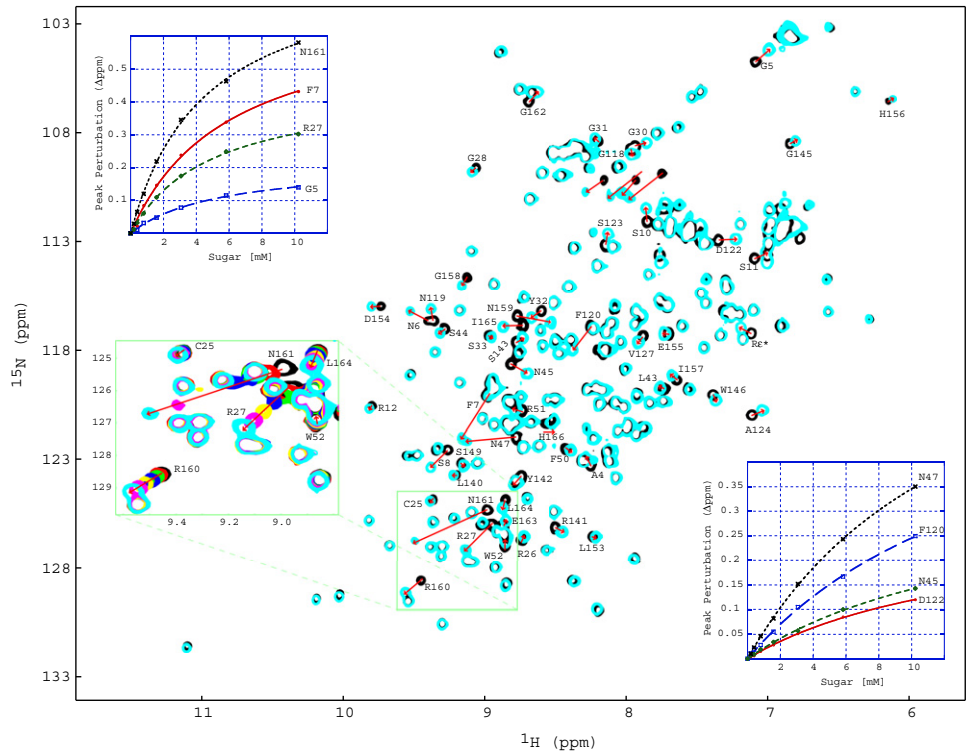

Figure 6. Interaction of MoCVNH-LysM with Manα(1-2)Man

Superposition of ^{1}H-^{15}N HSQC spectra without (black) and with excess Manα(1-2)Man (cyan) for the final point in the titration (~51 molar excess). Affected resonances are labeled by amino acid type and number. Apparent binding constants were extracted from binding curves using resonances of four residues in each subdomain (insets). An enlarged view is provided as an inset to illustrate selected chemical shift changes for the interaction between MoCVNH-LysM and Manα(1-2)Man (fast exchange). Apparent dissociation constants (K_d) of 5.23 ± 0.42 and 15.13 ± 1.13 mM were extracted from the binding isotherms for subdomains A and B, respectively. As indicated in the insets, Kd and standard deviation values were calculated based on four data points for each binding site.

different from the data reported for PrChi-A LysM. This finding is consistent with a much tighter chitin binding by MoCVNH-LysM and suggests larger $(N\text{-GlcNAc})_{4\text{-}5}$-induced conformational changes, compared to PrChi-A. Strongly affected residues in MoCVNH-LysM include those located in the loop region between helix α1 and stand β1, in helix α1, and in the loop region between helix α2 and strand β2. Interestingly, the largest changes are seen within the conserved GD(S/T)L motif, suggesting that this region may also critically contribute to carbohydrate binding strength and specificity.

Further comparison was carried out between the MoCVNH-LysM domain and that of Ecp6, a secreted protein from *Cladosporium fulvum*, a fungal pathogen that causes leaf mold of tomato (de Jonge et al., 2010). Interestingly, Ecp6 also binds to $(N\text{-GlcNAc})_{4\text{-}6}$ in the low micromolar range (6.4–11.5 μM) as detected by ITC (de Jonge et al., 2010). Importantly, however, is that the reported binding data were for the full-length protein with three LysM domains per Ecp6 molecule.

Conclusions

The NMR solution structure of a type III CVNH/LysM dual-lectin protein was determined. It comprised canonical structures for the individual domains that are flexibly connected via two long linkers that emerge from the middle of the bilobal CVNH fold. This structural independence is paralleled by the functional independence (i.e., distinct carbohydrate binding specificity) of the two domains. Although both domains are carbohydrate-interaction modules, each specific glycan only affects the cognate recognition module. Therefore, these functionally decoupled domains represent a clear example of protein evolution in which shuffling, acquisition, and structural intermingling of functionally related, but otherwise distinct domains allow for new proteins with multiple functions to emerge.

MoCVNH-LysM binds oligo-mannose and oligo-N-GlcNAc carbohydrates with millimolar and micromolar affinities, respectively. The latter are the building blocks of chitin, a major constituent of fungal cell walls. Binding of the peptidoglycan component N-GlcNAc-β-(1,4)-N-MurNAc by MoCVNH-LysM was also observed, with an apparent affinity in the low millimolar range. However, the presence of peptidoglycans has not been documented in fungi or plants, and therefore this interaction may not be functionally important. Our observation that binding of $(N\text{-GlcNAc})_{4\text{-}5}$ by MoCVNH-LysM is much tighter than the value reported for the PrChi-A LysM from a plant chitinase, but similar to that of fungal Epc6, strongly suggests that a chitin-related and oligomannose-containing polysaccharide may be the physiological ligand of this dual-domain lectin. In this context it is important to note that chitin-based polysaccharides are structural constituents not only of the outer cell wall, but also of internal septa, such as those associated with spore formation. In addition, it may be worth pursuing further mutational studies to assess the in vivo recognition of the carbohydrate in the plant.

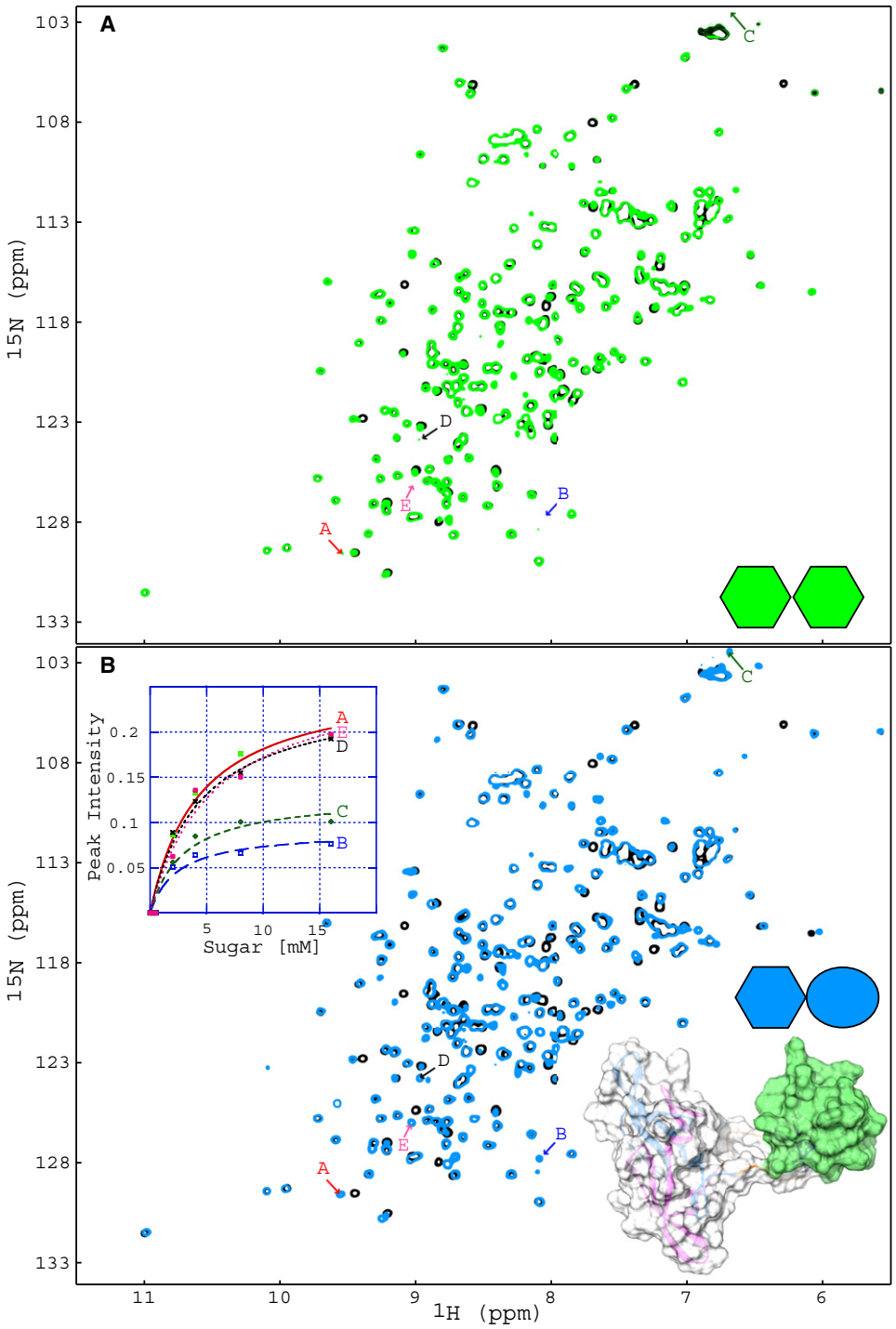

Figure 7. Interaction of MoCVNH-LysM with N-GlcNAc-β-(1,4)-N-GlcNAc and N-GlcNAc-β-(1,4)-N-MurNAc Disaccharides

(A) and (B) display superpositions of ^{1}H-^{15}N HSQC spectra in the absence (black) and presence of ~80 molar excess of disaccharides of N-GlcNAc-β-(1,4)-N-GlcNAc (green; A) and N-GlcNAc-β-(1,4)-N-MurNAc (light blue; B). Five new bound amide resonances (binding is in slow exchange on the chemical shift scale) were chosen for extracting the binding constant for the interaction with the N-GlcNAc-β-(1,4)-N-MurNAc disaccharide, yielding an apparent K_d value of ~3.93 ± 1.45 mM. No binding constant could be extracted from the titration with N-GlcNAc-β-(1,4)-N-GlcNAc since binding is very weak. Affected residues by the N-GlcNAc-β-(1,4)-N-MurNAc disaccharide were all from the LysM domain and highlighted in light green surface representation. See also Figure S2. As indicated in the inset, the standard deviation was calculated based on five data points.

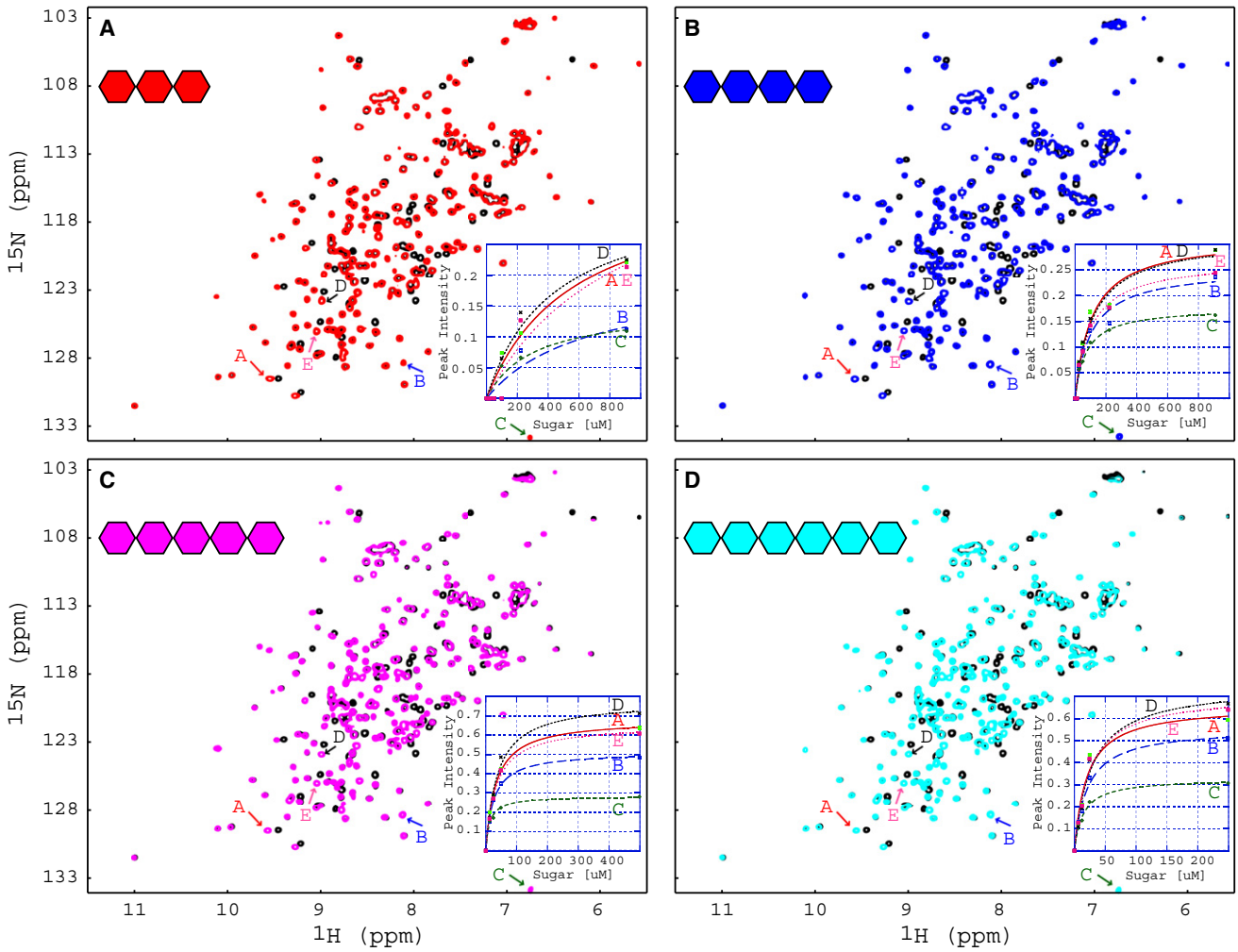

Figure 8. Interaction of MoCVNH-LysM with Oligomers of β-(1,4) Linked N-Acetylglucosamine

(A–D) Superpositions of the ^{1}H-^{15}N HSQC spectra without (black) and with excess carbohydrates; addition of $(N\text{-GlcNAc})_3$ trisaccharide (red; A), $(N\text{-GlcNAc})_4$ tetrasaccharide (green; B), $(N\text{-GlcNAc})_4$ pentasaccharide (magenta; C), and $(N\text{-GlcNAc})_6$ hexasaccharide (cyan; D). Similar to Figure 7B, five new bound amide resonances were chosen to plot binding isotherms and to extract apparent K_d values, yielding values of ~560 ± 399, ~100 ± 27, ~28 ± 3, and ~21 ± 4 μM for the $(N\text{-GlcNAc})_3$ trisaccharide, the $(N\text{-GlcNAc})_4$ tetrasaccharide, the $(N\text{-GlcNAc})_5$ pentasaccharide, and the $(N\text{-GlcNAc})_6$ hexasaccharide, respectively. See also Figure S3. As indicated in the inset, Kd and standard deviations were calculated based on five data points.

Since our localization experiments revealed that the full-length protein bearing the dual-domain CVNH-LysM module is preferentially expressed in and localized to cells specialized in plant infection, it is tempting to speculate that MGG_03307 may play a role in appressorium development. Indeed, the observation that septum formation and spatial uncoupling of nuclear division from cytokinesis is a key event in the pathogenic differentiation of *M. oryzae* asexual spores (Saunders et al., 2010), supports such a link. Specifically, the role of MGG_03307 in appressorium development and plant infection by *M. oryzae* may be reminiscent of various roles played by prokaryotic LysM proteins in the development of bacterial spores (Buist et al., 2008), for example, as morphogenetic signals, involved in the assembly of the bacterial endospore coat. In this case, various protein structural components are guided to the surface of the developing spore and act as mediators of protein-protein interactions at the spore coat-cortex interface (Costa et al., 2006). Experimental verification of this and related hypotheses regarding the physiological function of MGG_03307 will require further in-depth analyses of the full-length protein, including the functionally uncharacterized Rick 17 kDa domain.

EXPERIMENTAL PROCEDURES

Recombinant Protein Expression

The region encoding amino acids G174 to C340 of MGG_03307 was amplified by PCR using *M. oryzae* genomic DNA as template and the MoCVNHIIIc(+) and MoCVNHIIIc (–) oligonucleotides as primers (see Table S1). Following sequence verification, the resulting amplicon was directly inserted into a modified pET28b(+)-*SnaB*I expression vector (Bolchi et al., 2005), that encodes a N-terminal 6xHis tag for metal-affinity purification, and transformed into *E. coli* codon plus BL21 (DE3) RIPL cells (Stratagene). Cultures were grown at 37°C, induced with 1 mM IPTG, and further grown at 16°C for 18 hr for protein production. Cells were harvested by centrifugation, resuspended in TBS buffer, and lyzed by sonication. Protein was purified by metal-affinity

chromatography on Ni^{2+}-derivatized HisTrap columns (GE Healthcare) using a linear (20–500 mM) imidazole gradient for elution. Aggregation and precipitation behavior was evaluated using a set of 96 different buffers of varying pH and composition, with 50 mM sodium acetate (pH 5.0) containing 20 mM NaCl plus 0.02% NaN_3 selected as the most favorable one for stability. Purified protein was digested with thrombin in the above buffer for removal of the 6xHis tag (13 out of the 20 tag residues removed), followed by gel filtration on Superdex75 (GE Healthcare). Isotopic labeling was carried out by growing bacterial cell cultures in modified minimal media containing $^{15}NH_4Cl$ and/or ^{13}C-glucose as sole nitrogen and/or carbon sources, respectively.

Construction and In Vivo Imaging of MGG_03307-GFP Fusions

DNA constructs containing the MGG_03307 coding sequence fused to eGFP at either the N- or the C terminus, under the control of the endogenous *MGG_03307* promoter, were assembled in the pCB1004 vector (kindly provided by the Fungal Genetics Stock Center). Promoter (−1500) and coding sequence regions were amplified from *M. oryzae* genomic DNA. For the C-terminal fusion construct, the promoter region (prm-MGG_03307: −1500; −1) was amplified with the CVNHprm.fw SpeI (forward) and CVNHprm.rv NcoI (reverse) primers (see Table S1), and inserted into the SpeI/NcoI restriction sites of the pMJK-80 plasmid (Kershaw et al., 1998) upstream of the SGFP coding sequence. Prm-MGG_03307/SGFP was then amplified with the CVNHprm.fw SpeI and GFP.rv EcoRI primers and inserted into pBluescript-SK as a SpeI/EcoRI fragment. The MGG_03307 coding sequence plus a portion of the 3′-flanking region (cds-MGG_03307: +1; +1500) was amplified with the CVNHcds.fw EcoRI and CVNHcds.rv SalI primers and inserted into EcoRI/SalI cut prm-MGG_03307/SGFP pBluescript-SK. Finally, the entire prm-MGG_03307/SGFP/cds-MoCVNHIII sequence was inserted into pCB1004 as a SpeI/SalI fragment. To construct the N-terminally fused MGG_03307-GFP derivative, the promoter region plus the coding sequence (prm-cds-MGG_03307: −1500; +1021) was first amplified with the CVNHprm.fw SpeI (forward) and CVNHcds.rv NcoI (reverse) primers and inserted into SpeI/NcoI cut prm-MGG_03307/SGFP pBS-SK. The entire prm-cds-MGG_03307/SGFP sequence was then transferred to pCB1004 as a SpeI/SalI fragment. All constructs were sequence-verified prior to transformation of *M. oryzae*.

Protoplast preparation and transformation were performed as described previously (Talbot et al., 1993). MGG_03307-SGFP constructs were transformed into the *M. oryzae* wild-type strain Guy_11 and transformants were selected for hygromycin resistance (200 μg/ml Hygromycin B; Boehringer Mannheim). Putative transformants were confirmed by DNA gel blot hybridization, and those carrying single plasmid integrations were selected. For epifluorescence examination of MGG_03307-GFP transformants, conidia were deposited onto glass coverslips and placed onto a 2% agar cushion. They were observed using a IX81 motorized inverted microscope (Olympus) equipped with a UPlanSApo 100X/1.40 Oil objective (Olympus). Excitation of fluorescently labeled proteins was carried out using a VS-LMS4 Laser-Merge-System with solid state lasers (488 nm/50 mW). The laser intensity was controlled by a VS-AOTF100 System and coupled into the light path using a VS-20 Laser-Lens-System (Visitron Systems). Images were captured using a Charged-Coupled Device camera (Photometric CoolSNAP HQ2, Roper Scientific). All parts of the system were under the control of the software package MetaMorph (Molecular Devices).

NMR Spectroscopy and Structure Calculations

All NMR spectra for resonance assignments and NOE identification were recorded on Bruker AVANCE700 and AVANCE800 spectrometers, equipped with 5-mm triple-resonance, three-axis gradient probes or z axis gradient cryoprobes at 25°C, using ~0.5 mM $^{13}C/^{15}N$-labeled protein samples. Experiments included 3D HNCACB, CBCA(CO)NH, H(CCO)-NH, (H)C(CO)-NH, HCCH-TOCSY (Bax and Grzesiek, 1993; Clore and Gronenborn, 1998) and distance constraints were derived from 3D simultaneous ^{13}C- and ^{15}N-NOESY experiments (mixing time 120 ms) (Sattler et al., 1995). Spectra were processed with NMRPipe (Delaglio et al., 1995) and analyzed using NMRView (Johnson and Blevins, 1994). Assignments are 100% complete for backbone and >80% for side-chain resonances. Using the automatic NOE assignment and intensity-to-distance conversion modules incorporated in CYANA (Güntert et al., 1997) yielded a total of 2236 proton distance constraints. 130 hydrogen bond constraints were derived from observing slowly exchanging amide protons in hydrogen-deuterium exchange experiments and from analyzing NOESY patterns in the ^{15}N- and ^{13}C-NOESY spectra. Distance constraints of 1.8–2.8 Å (H-O) and 1.8–3.2 Å (N-O) were used for H-bonded amides. Backbone torsion angles (ϕ and ψ) were predicted using TALOS (Cornilescu et al., 1999) and a total of 161 ψ and ϕ angles were well predicted. All restraint information was applied in a simulated annealing protocol using CNS. A summary of the experimental constraints as well as pertinent structural statistics is provided in Table 1. 1000 structures were calculated, from which 25 structures with the lowest restraint energy values were further refined with implicit water. The quality of the structures was analyzed with MOLPROBITY (Davis et al., 2007). For the 25 final conformers of MoCVNH-LysM, 91.1% and 98.5% of all residues were found in the favored and allowed regions of the Ramachandran plot, respectively. All structural figures were generated with Chimera (Pettersen et al., 2004).

Determination of Dissociation Constants for Sugar Binding

Sugar binding was determined by NMR titration experiments using ^{15}N-labeled protein (0.1mM) in 50 mM NaAcetate, 0.02% NaN_3, 90/10% H_2O/D_2O, pH 5.0, at 25°C. For Manα(1-2)Man titrations, increasing amounts of sugar were added to the protein and ^{1}H-^{15}N HSQC spectra were recorded for each addition up to a final molar ratio of sugar to protein of ~51:1. Titration curves were plotted for eight resonances that exhibited sizeable, saturable shifts with no peak overlap, yielding reliable apparent K_d values. The eight amide resonances used were G5, F7, R27, and N161 in domain A and N45, N47, F120, and D122 in domain B of the CVNH domain of MoCVNH-LysM. Titrations with monosaccharides of N-GlcNAc or N-MurNAc, were carried out up to a final molar ratio of sugar to protein of ~250:1 and ~500:1, respectively, and for the disaccharide titrations with N-GlcNAc-β-(1,4)-N-GlcNAc and N-GlcNAc-β-(1,4)-N-MurNAc, final molar ratios of sugar to protein were ~160:1. Binding curves were based on signal intensities of five bound resonances (peaks A to E) and apparent K_d values were obtained by non-linear best fitting of the titration curves using KaleidaGraph (Synergy Software, Reading, PA), averaging over four or five curves.

Other Procedures

The full-length MGG_03307 amino acid sequence was used as query in BlastP searches that were carried out at the Broad Institute, GenBank, and SwissProt databases. The pfam database was used to search for conserved domains. Subcellular localization signals and other sequence features were searched with PsortII. Multiple alignments were constructed with the MEGA4 software (Tamura et al., 2007) and were visualized with GeneDoc (http://www.psc.edu/biomed/genedoc). Aligned N-terminal regions (up to amino acid 160) were used to produce a consensus sequence logo with the WebLogo application tool (http://weblogo.threeplusone.com/).

For nuclear localization trap assays (Marshall et al., 2007), full-length MGG_03307 was expressed in the yeast strain L40 (MATa, trp1, leu2, his3, LYS2::lex A-HIS3, URA3::lex A-lacZ) as a fusion with the mLexA-MBP-Gal4AD polypeptide. To this end, the coding sequence plus a portion of the 3′-flaking region of MGG_03307 (cds-MGG_03307: +1; +1500) was amplified with the CVNHcds.fw EcoRI and CVNHcds.rv SalI primers (Table S1) and inserted into the EcoRI/SalI digsted pNIA-CEN-MBP vector (Marshall et al., 2007). The expression of two reporter genes (HIS3 and LacZ) was analyzed using previously described methods (Walhout et al., 2001) with the empty vector as a negative control.

Polyclonal rabbit anti-MoCVNH-LysM serum (SeqLab, Göttingen) was used for detection in glycan array screening experiments. The latter were conducted at the Consortium for Functional Glycomics (Core H, Emory University; http://www.functionalglycomics.org/static/index.shtml) by incubating a set of over 250 different glycans printed on glass slides with the purified MoCVNH-LysM protein, followed by detection of the glycan-bound protein with anti-MoCVNH-LysM and an Alexa 488-labeled anti-rabbit secondary antibody (Paulson et al., 2006).

Chitin binding assays were performed by incubating MoCVNH-LysM (0.2 mg/ml) with 10 mg of chitin beads (New England Biolabs) for 90 min at room temperature on a roller wheel in a final volume of 0.1 ml, followed by centrifugation and SDS-polyacrylamide gel electrophoresis analysis as described (van den Burg et al., 2006).

ACCESSION NUMBERS

The atomic coordinates and NMR constraints have been deposited in the RCSB Protein Data Bank under accession code 2I9Y and chemical shift assignments has been deposited in BMRB with accession number 17493.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures and one table can be found with this article online at doi:10.1016/j.str.2011.03.004.

ACKNOWLEDGMENTS

We thank M. Delk for NMR technical support and the Consortium for Functional Glycomics (CFG), particularly David Smith (Emory University; Core H), for glycan array screening. This work was supported, in part, by start-up funds from the University of Pittsburgh School of Medicine and a National Institutes of Health Grant GM080642 (to A.M.G.), and by grants from the Fondazione Cariparma and the Ministry of Education, University and Research (MIUR) "Progetti di Rilevante Interesse Nazionale" (to S.O.). We declare no conflict of interest.

Received: February 3, 2011
Revised: March 7, 2011
Accepted: March 8, 2011
Published: May 10, 2011

REFERENCES

Barrientos, L.G., and Gronenborn, A.M. (2005). The highly specific carbohydrate-binding protein cyanovirin-N: structure, anti-HIV/Ebola activity and possibilities for therapy. Mini Rev. Med. Chem. *5*, 21–31.

Barrientos, L.G., Louis, J.M., Ratner, D.M., Seeberger, P.H., and Gronenborn, A.M. (2003). Solution structure of a circular-permuted variant of the potent HIV-inactivating protein cyanovirin-N: structural basis for protein stability and oligosaccharide interaction. J. Mol. Biol. *325*, 211–223.

Bax, A., and Grzesiek, S. (1993). Methodological advances in protein NMR. Acc. Chem. Res. *26*, 131–138.

Bewley, C.A. (2001). Solution structure of a cyanovirin-N:Man alpha 1-2Man alpha complex: structural basis for high-affinity carbohydrate-mediated binding to gp120. Structure *9*, 931–940.

Bewley, C.A., Gustafson, K.R., Boyd, M.R., Covell, D.G., Bax, A., Clore, G.M., and Gronenborn, A.M. (1998). Solution structure of cyanovirin-N, a potent HIV-inactivating protein. Nat. Struct. Biol. *5*, 571–578.

Bolchi, A., Ottonello, S., and Petrucco, S. (2005). A general one-step method for the cloning of PCR products. Biotechnol. Appl. Biochem. *42*, 205–209.

Botos, I., O'Keefe, B.R., Shenoy, S.R., Cartner, L.K., Ratner, D.M., Seeberger, P.H., Boyd, M.R., and Wlodawer, A. (2002). Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. J. Biol. Chem. *277*, 34336–34342.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., et al. (1997). Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. Antimicrob. Agents Chemother. *41*, 1521–1530.

Buist, G., Steen, A., Kok, J., and Kuipers, O.P. (2008). LysM, a widely distributed protein motif for binding to (peptido)glycans. Mol. Microbiol. *68*, 838–847.

Clore, G.M., and Gronenborn, A.M. (1998). Determining the structures of large proteins and protein complexes by NMR. Trends Biotechnol. *16*, 22–34.

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR *13*, 289–302.

Costa, T., Isidro, A.L., Moran, C.P., Jr., and Henriques, A.O. (2006). Interaction between coat morphogenetic proteins SafA and SpoVID. J. Bacteriol. *188*, 7731–7741.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S., and Richardson, D.C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. *35*, W375-W383.

de Jonge, R., and Thomma, B.P. (2009). Fungal LysM effectors: extinguishers of host immunity? Trends Microbiol. *17*, 151–157.

de Jonge, R., van Esse, H.P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M.H., and Thomma, B.P. (2010). Conserved fungal LysM effector Ecp6 prevents chitin-triggered immunity in plants. Science *329*, 953–955.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). Nmrpipe: a multidimensional spectral processing system based on Unix pipes. J. Biomol. NMR *6*, 277–293.

Donovan, D.M. (2007). Bacteriophage and peptidoglycan degrading enzymes with antimicrobial applications. Recent Pat. Biotechnol. *1*, 113–122.

Fromme, R., Katiliene, Z., Giomarelli, B., Bogani, F., Mc Mahon, J., Mori, T., Fromme, P., and Ghirlanda, G. (2007). A monovalent mutant of cyanovirin-N provides insight into the role of multiple interactions with gp120 for antiviral activity. Biochemistry *46*, 9199–9207.

Garvey, K.J., Saedi, M.S., and 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*, 10001–10008.

Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. J. Mol. Biol. *273*, 283–298.

Johnson, B.A., and Blevins, R.A. (1994). NMRView: a computer-program for the visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

Kershaw, M.J., Wakley, G., and Talbot, N.J. (1998). Complementation of the mpg1 mutant phenotype in Magnaporthe grisea reveals functional relationships between fungal hydrophobins. EMBO J. *17*, 3838–3849.

Knogge, W., and Scheel, D. (2006). LysM receptors recognize friend and foe. Proc. Natl. Acad. Sci. USA *103*, 10829–10830.

Koharudin, L.M., Furey, W., and Gronenborn, A.M. (2009). A designed chimeric cyanovirin-N homolog lectin: structure and molecular basis of sucrose binding. Proteins *77*, 904–915.

Koharudin, L.M.I., Viscomi, A.R., Jee, J.G., Ottonello, S., and Gronenborn, A.M. (2008). The evolutionarily conserved family of cyanovirin-N homologs: structures and carbohydrate specificity. Structure *16*, 570–584.

Layec, S., Gerard, J., Legue, V., Chapot-Chartier, M.P., Courtin, P., Borges, F., Decaris, B., and Leblond-Bourget, N. (2009). The CHAP domain of Cse functions as an endopeptidase that acts at mature septa to promote Streptococcus thermophilus cell separation. Mol. Microbiol. *71*, 1205–1217.

Marshall, K.S., Zhang, Z., Curran, J., Derbyshire, S., and Mymryk, J.S. (2007). An improved genetic system for detection and analysis of protein nuclear import signals. BMC Mol. Biol. *8*, 6.

Matei, E., Furey, W., and Gronenborn, A.M. (2008). Solution and crystal structures of a sugar binding site mutant of cyanovirin-N: no evidence of domain swapping. Structure *16*, 1183–1194.

Miya, A., Albert, P., Shinya, T., Desaki, Y., Ichimura, K., Shirasu, K., Narusaka, Y., Kawakami, N., Kaku, H., and Shibuya, N. (2007). CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. Proc. Natl. Acad. Sci. USA *104*, 19613–19618.

Moll, A., Schlimpert, S., Briegel, A., Jensen, G.J., and Thanbichler, M. (2010). DipM, a new factor required for peptidoglycan remodelling during cell division in Caulobacter crescentus. Mol. Microbiol. *77*, 90–107.

Oh, Y., Donofrio, N., Pan, H., Coughlan, S., Brown, D.E., Meng, S., Mitchell, T., and Dean, R.A. (2008). Transcriptome analysis reveals new insight into appressorium formation and function in the rice blast fungus Magnaporthe oryzae. Genome Biol. *9*, R85.

Ohnuma, T., Onaga, S., Murata, K., Taira, T., and 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.

Paulson, J.C., Blixt, O., and Collins, B.E. (2006). Sweet spots in functional glycomics. Nat. Chem. Biol. *2*, 238–248.

Pennisi, E. (2010). Sowing the seeds for the ideal crop. Science *327*, 802–803.

Percudani, R., Montanini, B., and Ottonello, S. (2005). The anti-HIV cyanovirin-N domain is evolutionarily conserved and occurs as a protein module in eukaryotes. Proteins *60*, 670–678.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Ponting, C.P., Aravind, L., Schultz, J., Bork, P., and Koonin, E.V. (1999). Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. J. Mol. Biol. *289*, 729–745.

Saedi, M.S., Garvey, K.J., and Ito, J. (1987). Cloning and purification of a unique lysozyme produced by Bacillus phage phi 29. Proc. Natl. Acad. Sci. USA *84*, 955–958.

Sattler, M., Maurer, M., Schleucher, J., and Griesinger, C. (1995). A Simultaneous ^{15}N, ^{1}H-HSQC and ^{13}C, ^{1}H-HSQC with sensitivity enhancement and a heteronuclear gradient-echo. J. Biomol. NMR *5*, 97–102.

Saunders, D.G., Dagdas, Y.F., and Talbot, N.J. (2010). Spatial uncoupling of mitosis and cytokinesis during appressorium-mediated plant infection by the rice blast fungus Magnaporthe oryzae. Plant Cell *22*, 2417–2428.

Shenoy, S.R., Barrientos, L.G., Ratner, D.M., O'Keefe, B.R., Seeberger, P.H., Gronenborn, A.M., and Boyd, M.R. (2002). Multisite and multivalent binding between cyanovirin-N and branched oligomannosides: calorimetric and NMR characterization. Chem. Biol. *9*, 1109–1118.

Talbot, N.J., Ebbole, D.J., and Hamer, J.E. (1993). Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus Magnaporthe grisea. Plant Cell *5*, 1575–1590.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. *24*, 1596–1599.

van den Burg, H.A., Harrison, S.J., Joosten, M.H., Vervoort, J., and de Wit, P.J. (2006). Cladosporium fulvum Avr4 protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. Mol. Plant Microbe Interact. *19*, 1420–1430.

Walhout, A.J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods *24*, 297–306.

Wan, J., Zhang, X.C., Neece, D., Ramonell, K.M., Clough, S., Kim, S.Y., Stacey, M.G., and Stacey, G. (2008). A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. Plant Cell *20*, 471–481.

Williams, D.C., Jr., Lee, J.Y., Cai, M., Bewley, C.A., and Clore, G.M. (2005). Crystal structures of the HIV-1 inhibitory cyanobacterial protein MVL free and bound to Man3GlcNAc2: structural basis for specificity and high-affinity binding to the core pentasaccharide from n-linked oligomannoside. J. Biol. Chem. *280*, 29269–29276.

Wilson, R.A., and Talbot, N.J. (2009). Under pressure: investigating the biology of plant infection by Magnaporthe oryzae. Nat. Rev. Microbiol. *7*, 185–195.

Zhang, X.C., Wu, X., Findley, S., Wan, J., Libault, M., Nguyen, H.T., Cannon, S.B., and Stacey, G. (2007). Molecular evolution of lysin motif-type receptor-like kinases in plants. Plant Physiol. *144*, 623–636.

