Structural Basis for New Pattern of Conserved Amino Acid Residues Related to Chitin-binding in the Antifungal Peptide from the Coconut Rhinoceros Beetle *Oryctes rhinoceros*

Hikaru Hemmi†‡, Jun Ishibashi*,‡, Tetsuya Tomie, and Minoru Yamakawa†

From the †National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan, ‡National Institute of Agrobiological Sciences, 1-2 Ouwashi, Tsukuba, Ibaraki 305-8634, Japan, and ‡Biological Research Laboratories, Nissan Chemical Industries, Ltd., Shiraoka, Sattama 349-0294, Japan

Scarabaecin isolated from hemolymph of the coconut rhinoceros beetle *Oryctes rhinoceros* is a 36-residue polypeptide that has antifungal activity. The solution structure of scarabaecin has been determined from two-dimensional 1H NMR spectroscopic data and hybrid distance geometry-simulated annealing protocol calculation. Based on 492 interproton and 10 hydrogen-bonding distance restraints and 36 dihedral angle restraints, we obtained 20 structures. The average backbone root-mean-square deviation for residues 4–35 is 0.728 ± 0.217 Å from the mean structure. The solution structure consists of a two-stranded antiparallel β-sheet connected by a type-I β-turn after a short helical turn. All secondary structures and a conserved disulfide bond are located in the C-terminal half of the peptide, residues 18–36. Overall folding is stabilized by a combination of a disulfide bond, seven hydrogen bonds, and numerous hydrophobic interactions. The structural motif of the C-terminal half shares a significant tertiary structural similarity with chitin-binding domains of plant and invertebrate chitin-binding proteins, even though scarabaecin has no overall sequence similarity to other peptide/polypeptides including chitin-binding proteins. The length of its primary structure, the number of disulfide bonds, and the pattern of conserved functional residues binding to chitin in scarabaecin differ from those of chitin-binding proteins in other invertebrates and plants, suggesting that scarabaecin does not share a common ancestor with them. These results are thought to provide further strong experimental evidence to the hypothesis that chitin-binding proteins of invertebrates and plants are correlated by a convergent evolution process.

A 36-residue peptide named scarabaecin isolated from hemolymph of the coconut rhinoceros beetle *Oryctes rhinoceros* has been found to show strong antifungal activity against phytopathogenic fungi. Scarabaecin also has chitin-binding activity. Antimicrobial activity was initially identified for chitin-binding proteins isolated from plants (2, 3), which commonly consist of single or multiple copies of the chitin-binding domain. The plant chitin-binding domain consists mostly of 30–43 residues, including eight cysteines, three aromatic residues, and glycines, and is frequently referred to as a hevein domain (4). This domain is indispensable to antimicrobial activity and shows significant conservation in the primary sequence (>40%) and in the tertiary structure (5–8). Although chitin is an important structural component of invertebrates, fungi, and bacteria, chitin-binding proteins of invertebrates have only recently been characterized (9–13). The invertebrate chitin-binding domain is assumed to consist of about 65 residues (14) involving a high percentage of cysteine and aromatic residues similar to the plant chitin-binding domain. However, scarabaecin has only two cysteine residues in the sequence and shows no sequence similarity to other chitin-binding proteins in plants and invertebrates. Moreover, a computer-aided homology search of scarabaecin indicated that this antifungal peptide has no significant sequence similarity to peptides/polypeptides so far reported, suggesting that a novel immune peptide occurs in the coleopteran insect.

The solution structure of tachycitin, a 73-residue polypeptide from the horseshoe crab that has chitin-binding and antimicrobial activity (9), was determined by NMR spectroscopy (15), which provided the first 3D structural information on invertebrate chitin-binding protein. A comparison of the tertiary structure of tachycitin with that of a plant chitin-binding protein, hevein, showed that tachycitin shared a remarkable local structural similarity with hevein (15). The local structure consists of an antiparallel β-sheet and a helical turn, and the secondary structures are constructed in both proteins in highly similar manners. Because the conserved structural motif was identified as an essential chitin-binding domain for hevein, the structural domain of tachycitin was assumed to serve as an essential chitin-binding site. The residues of cysteine, proline, and glycine, all of which may significantly influence structural construction, were well conserved in the structural motif of both proteins. Conservation of polar and hydrophobic residues

*This study was supported in part by the Program for Promotion of Basic Research Activities for Innovative Biosciences of the Bio-Integrated Technology Research Advancement Institution, Japan, and by a grant (Insect Technology Program), MAFF, Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

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

The 1H and 13C α-carbon chemical shifts of scarabaecin reported in this article were deposited in BioMagResBank (BMRB; www.bmrbr.wisc.edu) under accession number 5491.

§ To whom correspondence should be addressed. Tel.: 81-29-838-9033; Fax: 81-29-838-7996; E-mail: hemmi@nfrl.affrc.go.jp.

1 T. Tomie, J. Ishibashi, S. Furukawa, S. Kobayashi, R. Sawahata, A. Asakura, M. Tagawa, and M. Yamakawa, submitted for publication.

2 The abbreviations used are: 3D, three-dimensional; 2D, two-dimensional; DQF, double-quantum-filtered; COSY, correlated spectroscopy; TOCST, total correlation spectroscopy; NOE, nuclear Overhauser effect; WGA, wheat germ agglutinin; Ac-AMP2, Amaranthus caudatus antimicrobial protein 2; UDA, Urtica dioica agglutinin; r.m.s.d., root-mean-square deviation.
was identified for putative chitin-binding residues (15). The tertiary structure of tachycitin and 3D structure-based sequence alignment between chitin-binding proteins in invertebrates and plants for the region corresponding to Cys-40–Gly-60 of tachycitin as the conserved structural motif are thought to provide faithful evidence for the recent hypothesis (14) that chitin-binding proteins of invertebrates and plants are evolutionarily related by a convergent evolutionary process (15). The residues of cysteine, proline, and glycine in the amino acid sequence of scarabaein were well conserved by comparing amino acid sequences focusing on putative chitin-binding domains among scarabaein, tachycitin, hevein, and other invertebrate and plant chitin-binding proteins.\(^1\) Conservation of polar and hydrophobic residues was further identified for putative chitin-binding residues.\(^1\) Thus, it is worth determining the tertiary structure of scarabaein and comparing it with those of tachycitin and hevein to clarify that scarabaein has a conserved structural motif as a chitin-binding site, even though this peptide has no overall sequence similarity to other peptides/polyptides, including chitin-binding proteins.

In this study, we determined the 3D structure of scarabaein in solution by 2D \(^1\)H NMR spectroscopy and distance geometry-simulated annealing calculation. The solution structure showed that the backbone conformation of the conserved structural motif as a putative chitin-binding site in scarabaein was very similar to those in hevein and tachycitin. The length of the primary sequence and the number of disulfide bonds in scarabaein, however, differ from those in hevein and tachycitin, indicating a new pattern of conserved amino acid residues at the chitin-binding site. These results provide additional evidence to support the convergent evolution proposed for horse-shoe crab and plant chitin-binding proteins (15). Finally, we discuss conserved functional residues crucial for binding to saccharides and the relationship between the conserved structural motif and antimicrobial specificity.

**EXPERIMENTAL PROCEDURES**

**Sample Preparation**—The peptide scarabaein was synthesized by solid-phase methodology with Fmoc (N-[(9-fluorenyl)methoxy carbonyl]-amino acids and formed a disulfide bond between Cys-18 and Cys-29 using 2,2′-bispyridyl disulfide as described elsewhere.\(^1\) The resulting peptide was purified by reversed-phase high pressure liquid chromatography, and the molecular mass and amino acid sequence were determined using a matrix-assist laser desorption ionization time-of-flight mass spectrometer and a protein sequencer.\(^1\)

**NMR Spectroscopy**—A NMR sample was prepared by dissolving the synthetic peptide in 500 \(\mu\)l of 90% H\(_2\)O/10% D\(_2\)O or 99.96% D\(_2\)O solution. The final peptide concentration was \(\sim 3\) mM and pH was 2.4. All NMR spectra were obtained on Bruker Avance500 and Avance800 spectrometers with quadrature detection in the phase-sensitive mode by time proportional phase incrementation (TPPI) (16) and States-TPPI (17). The following spectra were recorded at 20 °C, 25 °C, 30 °C, and 35 °C with 15 ppm spectral widths in the \(t_1\) and \(t_2\) dimensions; 2D double quantum-filtered correlated spectroscopy (DQF-COSY) (18), recorded with 512 and 2048 complex points in the \(t_1\) and \(t_2\) dimensions; 2D homonuclear total correlation spectroscopy (TOCSY) (19) with DIPSI-2 mixing sequence, recorded with mixing times of 35, 60, and 80 ms, 512 and 2048 complex points in the \(t_1\) and \(t_2\) dimensions; 2D nuclear Overhauser effect (NOE) spectroscopy (20), recorded with mixing times of 60, 100, 200, and 400 ms, 512 and 2048 complex points in the \(t_1\) and \(t_2\) dimensions; and 2D rotating frame NOE spectroscopy (21), recorded with high digital resolution 2D DQF-COSY spectra and intraresidue and sequential NOEs. We obtained 30 \(\phi\) angle restraints. Backbone \(\phi\) angles were restrained to \(-60^\circ \pm 30^\circ\) for \(\phi_{\text{HN-HN}}<6\) Hz and \(-120^\circ \pm 40^\circ\) for \(\phi_{\text{HN-HH}}>8\) Hz. The additional \(\phi\) angle restraint of \(100^\circ \pm 80^\circ\) was applied to residues for which the intraresidue HN-HA NOE was clearly weaker than the NOE between HN and HA of the preceding residue (29). An \(\psi\) angle restraint was used for residues in \(\beta\)-strand structures, as predicted from the chemical shift index (30) and from NOE patterns characteristic of the secondary structure. Five \(\psi\) angles in the \(\beta\)-sheet region were restrained to \(120^\circ \pm 60^\circ\). Side-chain \(\chi_1\) and \(\chi_2\) dihedral angles were determined by \(\chi_1\)-angle restraints from exclusive 2D scalar COSY and short-mixing TOCSY connectivities combined with NH-HB and HA-H\(_\beta\) NOE (31). We obtained seven \(\chi_1\)-angle restraints. The \(\chi_1\)-angle restraints were normally restricted to \(\pm 60^\circ\) from staggered conformations, \(g\ (\pm 60^\circ)\), t \((180^\circ)\), or \(g\ (\pm 60^\circ)\).

**Structure Calculations**—NOE-derived distance restraints were classified into three ranges, 1.8–2.7 Å, 1.8–3.5 Å, and 1.8–5.0 Å, according to the relative NOE intensities. Upper distance limits for NOEs involving methyl protons and nonstereochemically assigned methylene protons were corrected appropriately for center averaging (27). In addition, a distance of 0.5 Å was added to the upper distance limits only for NOEs involving methyl proton (28) after correction for center averaging. Torsion angle restraints on the backbone \(\phi\)-angle were derived by \(3\_\text{DIN-HH}\), coupling constants from the high digital resolution 2D DQF-COSY spectra and intraresidue and sequential NOEs. We obtained 30 \(\phi\) angle restraints. Backbone \(\phi\)-angles were restrained to \(-60^\circ \pm 30^\circ\) for \(\phi_{\text{HN-HN}}<6\) Hz and \(-120^\circ \pm 40^\circ\) for \(\phi_{\text{HN-HH}}>8\) Hz. The additional \(\phi\) angle restraint of \(100^\circ \pm 80^\circ\) was applied to residues for which the intraresidue HN-HA NOE was clearly weaker than the NOE between HN and HA of the preceding residue (29). An \(\psi\) angle restraint was used for residues in \(\beta\)-strand structures, as predicted from the chemical shift index (30) and from NOE patterns characteristic of the secondary structure. Five \(\psi\)-angles in the \(\beta\)-sheet region were restrained to \(120^\circ \pm 60^\circ\). Side-chain \(\chi_1\)-angles were determined by \(\chi_1\)-angle restraints from exclusive 2D scalar COSY and short-mixing TOCSY connectivities combined with NH-HB and HA-H\(_\beta\) NOE (31). We obtained seven \(\chi_1\)-angle restraints. The \(\chi_1\)-angle restraints were normally restricted to \(\pm 60^\circ\) from staggered conformations, \(g\ (\pm 60^\circ)\), t \((180^\circ)\), or \(g\ (\pm 60^\circ)\).

**Hydrogen-deuterium exchange experiments** identified seven hydrogen bond donors. Corresponding hydrogen bond acceptors were determined based on NOE patterns observed for regular secondary structural regions and preliminary calculated structures without restraints regarding hydrogen bonds. Hydrogen-bond restraints were applied to N–H and C=O groups: 1.7–2.4 Å for the H–O distance and 2.7–3.4 Å for the N–O distance.

Structure calculations were performed using the hybrid distance geometry-simulated annealing method using X-PLOR 3.851 (32). A total of 502 interproton distance restraints and 36 dihedral angle restraints were used to calculate an ensemble of structures. The structure calculation proceeded in two stages by using the standard X-PLOR protocol. In the first stage, a low-resolution structure was preliminarily determined using NOE-derived distance restraints and dihedral angle restraints, except for \(\psi\)-angle restraints. In the second stage, the protocol was applied by adding hydrogen bond restraints and \(\psi\)-angle restraints. The force constants for the distance restraints were set to 50 kcal mol\(^{-1}\) Å\(^{-2}\) throughout all the calculations, and dihedral angle restraints were initially set to 5 kcal mol\(^{-1}\) rad\(^{-2}\) during the high-resolution calculation and increased to 20 kcal mol\(^{-1}\) rad\(^{-2}\) during the annealing stage. The final round of calculations began with 200 initial structures, and the best 20 structures were selected and analyzed with MOLMOL (33), InsightII (MSI, San Diego, CA), and PROCHECK-NMR (34). The average coordinates of ensembles of the 20 lowest energy structures were subjected to 1000 cycles of Powell restrained minimization to improve stereochemistry and nonbonded contacts. Structure figures were generated using MOLMOL.

**RESULTS**

**Resonance Assignment and Secondary Structure**—Essentially complete \(^1\)H resonance assignments were obtained for the peptide using spin system identification and sequential assignment (35). Any ambiguous peaks caused by overlapping with water resonance or degeneration in the chemical shift were resolved by comparing NOE spectroscopy and TOCSY spectra at four different temperatures (20 °C, 25 °C, 30 °C, and 35 °C) and by combining with the 2D \(^1\)H,\(^1\)C heteronuclear single quantum correlation spectrum. The fingerprint region of the DQF-COSY spectrum collected at 25 °C is shown in Fig. 1 with sequence-specific resonance assignments. In these assignments, \(\text{H}a(i)-\text{H}b(i+1)\) (d(aO)) or \(\text{H}a(i)-\text{H}b(i+1)\) (d(aO)) NOEs instead of dNOE were used for Pro residues. All four proline residues (Pro-3, Pro-6, Pro-19, and Pro-32) showed strong d(aO) NOEs, indicating that all proline residues in the peptide have a trans configuration. Finally, the resonance as-

Downloaded from http://www.jbc.org/ by guest on July 24, 2018
The resonance assignment for the backbone and side-chain $^1H$ of Glu-1. The assignment was extended by determining stereospecific assignments of amide protons to obtain high-precision NMR structures. Stereospecific assignments of $\beta$-methylene protons were obtained for 7 of 28 residues of the peptide using information on $^3J_{\text{H}_{N}-\text{H}_{\alpha}}$ coupling constants from exclusive 2D scalar COSY and qualitatively estimated from short-mixing time TOCSY spectra combined with intraresidue NH-H NOEs.

A qualitative analysis of short- and medium-range NOEs, $^3J_{\text{H}_{N}-\text{H}_{\alpha}}$ coupling constants, and a slowly exchanging amide proton pattern was used to characterize the secondary structure of scarabaecin (Fig. 2A). The absence of consecutive $\text{H}_{\alpha}(i)-\text{H}_{\alpha}(i+3)$ NOEs and only one small $^3J_{\text{H}_{N}-\text{H}_{\alpha}}$ coupling constant (<6 Hz) indicated that scarabaecin did not have an $\alpha$- or $\beta$-helix. An antiparallel $\beta$-sheet was identified from large $^3J_{\text{H}_{N}-\text{H}_{\alpha}}$ coupling constant (>8 Hz) and its characteristic NOE pattern: strong consecutive sequential $\delta\alpha(i,j)$ and long-range interstrand $\delta\alpha(i,j)$ cross-peaks. From these data, a two-stranded antiparallel $\beta$-sheet consisting of the two regions of residues Val-23–Ans-25 and Asp-28–Lys-30 was identified in scarabaecin (Fig. 2B). The existence of the $\beta$-sheet was supported by four slowly exchanging NH protons at positions 23, 25, 28, and 30, and chemical shift index pattern (Fig. 2A). The region Ans-25–Asp-28, linking the two $\beta$-strands and residues 23–25 and 28–30, was identified as a $\beta$-turn based on characteristic NOE patterns, weak $\text{H}_{\alpha}(i)-\text{H}_{\alpha}(i+2)$ connectivity, and weak $\text{H}_{\alpha}(i)-\text{H}(i+2)$ connectivity between Gly-26 and Asp-28 (Fig. 2B). The slow exchanging NH proton at position Asp-28 supports the presence of the tight turn (Fig. 2A).

Tertiary Structure of Scarabaecin—The 3D structure of scarabaecin was determined based on the distance and dihedral angle restraints from the NMR data using the hybrid distance geometry-simulated annealing approach. 200 structures were calculated. Of these, 20 final structures showing the lowest energy values, no distance constraint violation of >0.5 Å, and no dihedral constraint violation of >5° were selected. The restraints used and the structural statistics for the final structures are summarized in Table I. The structures exhibited good covalent geometry and stereochemistry, as evidenced by the low r.m.s.d. values for bond, angle, and improper from idealized geometry. In a Ramachandran plot, 100% of backbone dihedral angles of the 20 structures fall in either core or allowed regions. Fig. 3A shows best-fit superposition of backbone atoms of the 20 structures of scarabaecin. The r.m.s.d. of these structures from the average structure were 0.72 ± 0.21 Å for backbone heavy atoms in the whole sequence (excluding three N-terminal residues and one C-terminal residue, Lys-4–Phe-35) and 1.30 ± 0.23 Å for all heavy atoms in the same region, whereas the corresponding values were 0.35 ± 0.14 Å for backbone heavy atoms and 0.84 ± 0.23 Å for all heavy atoms in the C-terminal half region of residues 18–35 involving all secondary structure elements and a conserved disulfide bond. This data indicates that the C-terminal half region of residues 18–35 converges very well in the calculated structures.

A ribbon representation of the restrained energy minimized average structure of scarabaecin is shown in Fig. 3. B and C. The solution structure of the peptide consists of an antiparallel $\beta$-sheet connected by a type-I $\beta$-turn (Val-23–Lys-30) after a short helical turn (Ser-31–Ala-34). The $\beta$-turn, residues 25–28, was classified as type I based on dihedral angles of Gly-26 and Phe-27, the hydrogen bonding between Asn-25 and Asp-28, and the spatially positions of carbonyl oxygen of Gly-26 and Cβ of Phe-27. All secondary structures of scarabaecin are located in the C-terminal half of the peptide (residues 18–36), not in the N-terminal half (residues 1–17).

Seven hydrogen bond donors were observed in hydrogen-deuterium exchange experiments. Four hydrogen bonds (Val-23–H–Lys-30–O, Ans-25–H–Asp-28–O, Asp-28–H–Asn-25–O, and Lys-30–H–Val-23–O) were found within the secondary structures, and one hydrogen bond (Ile-12–H–Asn-17–O(1)) was determined from the preliminary structures; five hydrogen bonds were confirmed from the final 20 structures. Evidence for two additional hydrogen bonds (Leu-5–H–Pro-3–C and Val-10–H–Asp-8–O) was observed in a subset of the family of structures. Within the ordered portion of the molecule, a subset of side chains forms a well-defined hydrophobic core centered on Pro-19 and the disulfide bond between Cys-18 and Cys-29. These residues have NOE contacts to side chains of Leu-5, Pro-6, Leu-11, and Trp-24. The chitin-binding domains of plants and invertebrates involve a high percentage of cysteine residues in their sequences, and cysteine residues have been elucidated in maintaining protein fold (14). However, scarabaecin has only two cysteine residues, Cys-18 and Cys-29, to form a disulfide bond. The N-terminal half of the peptide (residues 1–17) has no secondary structure or disulfide bond. Thus, the overall folding of scarabaecin is stabilized by a combination of one covalent disulfide linkage (Cys-18–Cys-29), seven hydrogen bonds, and hydrophobic side chain packing consisting of seven residues (Leu-5, Pro-6, Leu-11, Cys-18, Pro-19, Trp-24, and Cys-29) (Fig. 3C).

Comparison of the Tertiary Structure of the Conserved Structural Motif in Scarabaecin with Those in Tachycitin and Hevein—So far, the plant chitin-binding domains have been well characterized by information on their primary sequences, 3D structures, and functional experiments. Less is known about invertebrate chitin-binding proteins, however. Recently, the solution structure of tachycitin was determined as the first 3D structure of invertebrate chitin-binding protein by using 2D NMR methods (15). A comparison of the tertiary structure of tachycitin with that of hevein showed that the region Cys-40–Gly-60 of tachycitin had the same structural motif essential for chitin-binding as the corresponding region, Cys-12–Ser-32, of hevein (backbone r.m.s.d. of the structural motif between them was −1.5 Å) (15). The structural motif consists of an antiparallel $\beta$-sheet after a short helical turn. The backbone conformation of the corresponding region in scarabaecin, Cys-18–Ser-36, was compared with those of tachycitin and hevein. The backbone conformation of the corresponding region in scarabaecin is very similar to those of hevein and tachycitin, with backbone r.m.s.d. of 1.42 and 1.63 Å, excluding the C-terminal.
short helical turn in the structural motif (Fig. 4). Because the full length of the C-terminal amino acid sequence after the last \( \beta \)-strand in scarabaecin is much shorter than those in hevein and tachycitin, the conformation of the helical turn in the region of scarabaecin is very different from those in the corresponding regions of hevein and tachycitin. The aromatic side-chain groups of Trp-21, Trp-23, and Tyr-30 of hevein are known to bind specifically to a chitin-derived oligosaccharide through hydrophobic interactions (36, 37). The binding is further strengthened by two hydrogen bonds with Ser-19 and Tyr-30 of hevein (37). As shown in Fig. 4, the residues of Asn-25 and Phe-27 in the conserved secondary structural region of scarabaecin are located at perfectly corresponding positions to the residues Ser-19 and Trp-21 of hevein. The aromatic residue Phe-35 of scarabaecin is also located at a remarkably corresponding position to the aromatic residue Tyr-30 of hevein, despite their different conformations for the helical turn, including the aromatic residue. Scarabaecin has no hydrophobic...
residue at the position corresponding to the residue Trp-23 of hevein. In the corresponding region of tachycitin, the residues of Asn-47, Tyr-49, and Val-52 correspond to the residues of Ser-19, Trp-21, and Trp-23 of hevein (15). The aromatic residue at the position corresponding to the aromatic residue Tyr-30 of hevein is lacking in the region of tachycitin. However, it was assumed that the region of tachycitin served as an essential chitin-binding site because of similarities of the backbone conformation, including the secondary structural pattern (an antiparallel $\beta$-sheet and a helical turn) and local positions of the putative functional three residues (Asn-47, Tyr-49, and Val-52) with those in the hevein domain (15). The region Cys-18–Ser-36, consisting of an antiparallel $\beta$-sheet after a short helical turn in the C-terminal half of scarabaecin, thus may serve as an essential chitin-binding site that protrudes through the side chains of putative functional residues Asn-25, Phe-27, and Phe-35 in the same manner as hevein, even though scarabaecin has no overall sequence similarity to other peptides/polypeptides including chitin-binding proteins.

**DISCUSSION**

In this study, we determined the 3D structure of the antifungal peptide scarabaecin by means of $^1$H-NMR measurements and distance geometry-simulated annealing calculation. The structure has a two-stranded antiparallel $\beta$-sheet after a

---

**REFERENCES**

1. **FIG. 3.** A, superimposition of 20 selected structures of scarabaecin with lowest total energy, calculated by means of the hybrid distance geometry-simulated annealing procedure of X-PLOR 3.851 (31). B, schematic ribbon drawing of the restrained minimized average structure of scarabaecin. Scarabaecin is shown with the $\beta$-sheet in front. Disulfide bridges are shown as ball-and-stick models. C, scarabaecin is rotated by 180° about the vertical axis. Nonhydrogen side chain atoms of the residues forming hydrophobic interaction are shown.

**FIG. 4.** Comparison of the structure of the putative chitin-binding region in scarabaecin (Cys-18–Ser-36) with that in tachycitin (Cys-40–Gly-60) and the structure of the previously identified chitin-binding region in hevein (Cys-12–Ser-32). Backbone atoms of residues 18–30 (scarabaecin, blue); 40–49, 52–54 (tachycitin, red); and 12–21, 23–25 (hevein, green) were used for superposition. The side chains of the conserved cysteine and putative or really functional residues of all three structures are shown, with cysteine sulfur atoms colored yellow.

**FIG. 5.** 3D structure-based sequence alignment of the region, Cys-18–Ser-36, of scarabaecin with the corresponding region of several chitin-binding proteins in invertebrates and plants. Invertebrates are as follows: Scarabaecin, antifungal peptide from Oryctes rhinoceros (Footnote 1); Tachycitin, Tachypleus tridentatus tachycitin (15); Ag-chit, Anopheles gambiae chitinase (11); Pj-chit1, Penaeus japonica chitinase 1 (48); Ch-chit, Chelonus sp. chitinase (49); Peritrophin-44, 44-kDa glycoprotein from Lucilia cuprina (10); and Tn-IM, Trichoplusia ni intestinal mucin (50). Plants are as follows: Hevein, hevein from the rubber tree (5); Ac-AMP2, Amaranthus caudatus antimicrobial protein 2 (6); UDA, Urtica dioica agglutinin (1); and WGA-A, -B, -C, and -D, four homologous domains of wheat germ agglutinin (7). Residue numbers for each region are indicated in parentheses. Amino acids conserved between invertebrate and plant chitin-binding proteins are indicated with bold letters. The chitin-binding residues in plants and the corresponding residues in invertebrates are found to be aligned, for which polar and hydrophobic residues are represented by boxed letters. The hydrophobic residues are indicated by A1, A2, and A3 in order from N-terminal sequence, and polar residues are indicated by the asterisk at the bottom. The conserved disulfide bond and the secondary structures ($\beta$-strand represented by the arrow and helical turn by the box) are also shown at the bottom.
short helical turn and one disulfide bond. All secondary structures and the disulfide bond in scarabaecin are located in the C-terminal half of the peptide (residues 18–36) and the backbone conformation of the region from Cys-18 to Ser-36 is very similar to those of the corresponding regions identified as an essential chitin-binding domain in hevein and tachycitin. Conservation of the chitin-binding structural motif among chitin-binding proteins in invertebrates and plants based on 3D structure-based sequence alignment for the corresponding region, Cys-40–Gly-60, of tachycitin was reported by Suetake et al. (15). Based on this alignment, we compared amino acid sequences focusing on putative chitin-binding domains among scarabaecin and several chitin-binding proteins in invertebrates and plants. In this article, we realigned the corresponding region of scarabaecin, Cys-18–Ser-36, with the structural motif of other chitin-binding proteins by using the 3D structure of scarabaecin (Fig. 5). The residues of Cys, Pro, and Gly, all of which significantly influence structural constructions, were well conserved in chitin-binding proteins of both plant and invertebrates (15). The corresponding residues, Cys and Pro, but not Gly, are also conserved in the corresponding region of scarabaecin. Conservation of polar and hydrophobic residues based on the alignment is further identified for putative chitin-binding residues of invertebrate chitin-binding proteins (e.g. Asn-25, Phe-27, and Phe-35 for scarabaecin). The pattern of conserved functional residues in invertebrate chitin-binding proteins is two hydrophobic residues and one polar residue, whereas that in plant chitin-binding proteins is three hydrophobic (aromatic) residues and one polar residue. Scarabaecin also has two conserved hydrophobic (aromatic) residues, Phe-27 and Phe-35, and one polar residue, Asn-25, but the pattern of local positions for two conserved hydrophobic residues in scarabaecin is different from those in other invertebrates chitin-binding proteins. We denote the three conserved hydrophobic residues of plant chitin-binding protein in order of primary sequence from the N terminus as A1, A2, and A3 (see also Fig. 5). The pattern of local positions for the two conserved hydrophobic residues of the invertebrate chitin-binding proteins, including tachycitin, is A1 and A2, but that for the two conserved hydrophobic residues of scarabaecin is A1 and A3. Accordingly, scarabaecin showed a new pattern for conserved functional hydrophobic residues related to the chitin binding among invertebrate chitin-binding proteins. Shen and Jacobs-Lorena (14) hypothesize that a rare evolutionary process, convergent evolution, correlates chitin-binding proteins in invertebrates and plants. The tertiary structure of tachycitin and the 3D structure-based sequence alignment provided the faithful evidence for the proposed idea of the convergent evolution relationship between invertebrate and plant chitin-binding proteins (15). Tachycitin and scarabaecin share the structural motif, but the length of the primary sequence, the number of cysteine residues forming disulfide bridges, and the pattern of putative chitin-binding residues in scarabaecin are different from those in other invertebrate chitin-binding proteins, including tachycitin and plant chitin-binding proteins. Thus, these findings suggest that scarabaecin does not share a common ancestor with other invertebrate and plant chitin-binding proteins, which are thought to provide further strong experimental evidence for the idea of convergent evolution.

We found that scarabaecin has antifungal activity and very weak antibacterial activity against the Gram-negative bacteria, *Pseudomonas syringae*. In Table II, we summarize reported data on some chitin-binding proteins having the conserved structural motif versus their antimicrobial specificity. Tachycitin, an invertebrate chitin-binding protein, showed antifungal and antibacterial activity against Gram-positive and -negative bacteria (9). Only antifungal activity has been reported for hevein (38) and wheat germ agglutinin (WGA) (39),

| Peptide     | Source          | Antimicrobial specificity                                      | Ref.   |
|-------------|-----------------|---------------------------------------------------------------|--------|
| Scarabaecin | Beetle          | Fungi, Gram − (P. syringae, very weak)                        | Footnote 1 |
| Tachycitin  | Horseshoe crab  | Fungi, Gram −/+                                               | (9)    |
| Hevein      | Rubber tree     | Fungi                                                         | (38)   |
| Ac-AMP2     | Amaranth        | Fungi, Gram +                                                | (2)    |
| UDA         | Nettle          | Fungi, Gram + (B. megaterium, very weak)                      | (40)   |
| WGA         | Wheat germ      | Fungi                                                         | (39)   |

FIG. 6. Surface electrostatic potentials (calculated in MOLMOL (33)) of the conserved structural motif in scarabaecin (a), hevein (b), tachycitin (c), Ac-AMP2 (d), and UDA (e), colored by electrostatic potential with positive regions in blue and negative regions in red.
plant chitin-binding proteins. Other plant chitin-binding proteins, *Amaranthus caudatus* antimicrobial protein 2 (Ac-AMP2) (2) and *Urtica dioica* agglutinin (UDA) (40), showed antifungal activity and antibacterial activity against Gram-positive bacteria, but UDA has very weak antibacterial activity against Gram-negative bacteria. The chitin-binding proteins show different antimicrobial specificity, but all have antibacterial activity. Chitin is a component of the cell wall of fungi, and chitin-binding proteins identified in plants such as antibacterial substances, lectins, and chitinase are considered to be defense molecules to kill fungal pathogens when recognizing cell wall component(s), perturbing the plasma membrane, or hydrolyzing the polysaccharides on the cell wall (9). Further, it is believed that the capability for chitin binding may play a role in the antibacterial activities of hevein and hevein-like proteins (41). Thus, we assume that the conserved polar and hydrophobic residues identified as putative chitin-binding residues in the conserved structural motif of the antimicrobial proteins are critical to antibacterial activity. On the other hand, four of six antimicrobial proteins show antibacterial activity; two of them, scarabaeacin and UDA, have very weak antibacterial activity (Table II). These findings indicate that the mechanism for antibacterial activity of the chitin-binding proteins is different from that for antibacterial activity.

The 3D structures of antibacterial peptides from plants and animals have been determined by NMR and x-ray crystallography, showing them to be amphipathic helices, β-sheets stabilized by disulfide bridges, or to contain a helix–β-sheet motif (42). The tertiary structures of the two antibacterial peptides, tachyplesin I from the horseshoe crab (43, 44) and protegrin-1 from porcine leukocytes (42), consist of a two-stranded antiparallel β-sheet stabilized by disulfide bridges, and the β-sheet of the peptides is amphipathic with basic and hydrophobic faces. Matsuzaki et al. (45) suggested that the mechanism for antibacterial activity of peptides with amphipathic β-sheet structures was the same as that with an amphipathic α-helical structure forming pores and then increasing membrane permeability to kill the bacteria. The tertiary structures of the antimicrobial proteins listed in Table II consist predominantly of a β-sheet structure. The secondary structure of scarabaeacin and Ac-AMP2 having antibacterial activity consists of an antiparallel β-sheet and a short helical turn of the conserved structural motif in the molecule. These findings suggest that this antiparallel β-sheet is mainly responsible for the antibacterial activity of the proteins. Why hevein and WGA do not have antibacterial activity even though they have the conserved structural motif of the β-sheet structure is unknown. Scarabaeacin, tachycitin, Ac-AMP2, and UDA but not hevein and WGA have basic amino acid residues, Arg and Lys, in the sequence of the conserved structural motif. The surface of the bacterial membrane is negatively charged, indicating that the positively charged surface from basic amino acid residues is necessary for the antibacterial proteins to interact with the surface of the bacterial membrane. The surface electrostatic potential of the conserved structural motif in hevein shows no positively charged region because of the lack of basic residues; therefore, hevein and WGA probably do not have antibacterial activity. Other proteins, such as scarabaeacin, tachycitin, Ac-AMP2, and UDA, have basic amino acid residues in the sequence of the conserved structural motif, but the number and local positions of basic residues in the conserved structural motif of the proteins differ. The pattern of surface electrostatic potentials of the conserved structural motif in the proteins also differs (Fig. 6). The antimicrobial proteins have different antibacterial activity and antibacterial specificity. Ac-AMP2 has the strongest antibacterial activity among the antimicrobial proteins. The surface electrostatic potential of Ac-AMP2 is perfectly amphipathic with positively charged and hydrophobic faces. This amphipathic character is common to the helices (46) or the β-sheet structures (42–44) of many antibacterial peptides. Recently, it was reported that one of the variants of heliomicin, an antibacterial peptide from the lepidopteran *Heliosis virescens*, showed significant activity against Gram-positive bacteria and remained efficient against fungi, although the wild-type heliomicin was inactive against bacteria (46). In a comparison of the solution structure of the wild-type peptide with that of the variant, their global folding involving an α-helix and a three-stranded antiparallel β-sheet are very similar but the substitution of two basic residues at the end of the helix with two leucines gives rise to an amphipathic character in the α-helix of the variant (46). Furthermore, it was suggested that the amphipathic topology of the cationic charge and hydrophobic clusters may provide a useful guide to the design of selective peptide antibiotics against Gram-positive or -negative bacteria using the scaffolds consisting of β-sheet stabilized by disulfide bridges such as tachyplesin I and protegrin-1 (47). Thus, we assume that the pattern of the surface electrostatic potential of the conserved structural motif in the proteins may correlate with the antimicrobial specificity and the strength of antibacterial activity. To confirm this hypothesis, further study is needed of the antibacterial activity of the scarabaeacin analog giving rise to a perfectly amphipathic character in the β-sheet.

**Acknowledgment—**We thank Dr. M. Yoshioka, National Food Research Institute, for gracious help.

**REFERENCES**

1. Saul, F. A., Rovira, P., Boulot, G., Van Damme, E. J. M., Peumans, W. J., Truffa-Bachi, P., and Bentley, G. A. (2000) Structure 8, 595–603
2. Broekgaert, W. F., Marien, W., Terras, F. R., De Bolle, M. F., Prout, P., Van Damme, J., Dillen, L., Claey, M., Rees, S. B., Vanderleyden, J., and Cammue, B. P. (1992) Biochemistry 31, 4308–4314
3. Koo, J. C., Lee, S. Y., Chun, H. J., Cheong, Y. H., Choi, J. S., Kawabata, S., Miyagi, M., Tsunasawa, S., Ha, K. S., Bae, D. W., Han, C. D., Lee, B. L., and Cho, M. J. (1998) Biochim. Biophys. Acta 1382, 80–90
4. Bentema, J. J. (1994) FEBS Lett. 350, 159–163
5. Andersen, N. H., Can, B., Rodriguez-Romero, A., and Arreguin, B. (1993) Biochemistry 32, 1407–1422
6. Martina, J. C., Maes, D., Loris, R., Pepermans, H. A., Wyna, L., Willem, R., and Verheyden, P. (1998) J. Mol. Biol. 285, 2323–2333
7. Wright, C. S. (1990) J. Mol. Biol. 215, 635–651
8. Weaver, J. L., and Prestegard, J. H. (1998) Biochemistry 37, 116–128
9. Kawahata, S., Nagayama, B., Hirata, R., Inouye, T., Azagurula, K. L., Saito, T., Cho, J., Nakajima, H., Takagi, T., and Iwamama, S. (1998) J. Biochim. Biophys. Acta 1382, 181–185
10. Elvin, C. M., Vuocid, T., Pearson, R. D., East, I. J., Riding, G. A., Eisenmann, C. H., and Tellam, R. L. (1996) J. Biol. Chem. 271, 8925–8935
11. Shen, Z., and Jacobs-Lorena, M. (1997) J. Biol. Chem. 272, 28895–28900
12. Shen, Z., and Jacobs-Lorena, M. (1998) J. Biol. Chem. 273, 17665–17670
13. Watanabe, T., Kono, M., Aida, K., and Nagasawa, H. (1998) Biochim. Biophys. Acta 1382, 1–185
14. Shen, Z., and Jacobs-Lorena, M. (1999) J. Mol. Biol. 48, 341–347
15. Suetake, T., Tsuda, S., Kawabata, S., Miura, K., Iwanaga, S., Hikichi, K., Nitta, K., and Kawano, K. (2000) J. Biol. Chem. 275, 17929–17932
16. Marion, D., and Wurtz, K. (1983) Biochim. Biophys. Res. Commun. 113, 567–574
17. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) J. Magn. Reson. 85, 393–399
18. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 117, 479–485
19. Davis, D. G., and Bax, A. (1985) J. Am. Chem. Soc. 107, 2820–2821
20. Kumar, A., Ernst, R. R., and Wurtz, K. (1980) Biochem. Biophys. Res. Commun. 85, 1–6
21. Griesinger, C., and Ernst, R. (1987) J. Magn. Reson. 75, 261–271
22. Griesinger, C., Sorensen, O. W., and Ernst, R. R. (1987) J. Magn. Reson. 75, 474–492
23. Pietto, M., Saudek, V., and Sklenar, V. (1992) J. Biomol. NMR 2, 661–666
24. Sklenar, V., Pietto, M., Leppik, R., and Saudek, V. (1993) J. Magn. Reson. Ser. 102, 241–245
25. Bodenhausen, G., and Ruben, D. J. (1980) Chem. Phys. Lett. 69, 185
26. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, D., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1985) J. Magn. Reson. 6, 135–140
27. Wuthrich, K., Billette, M., and Braun, W. (1988) J. Mol. Biol. 199, 949–961
28. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) Biochim. Biophys. Acta 8612–8625
NMR Structure of Antifungal Peptide from Oryctes rhinoceros

29. Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry* 33, 2761–2772
30. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry* 31, 1647–1651
31. Wagner, G., Braun, W., Havel, T. F., Schaumman, T., Go, N., and Wuthrich, K. (1987) *J. Mol. Biol.* 196, 611–639
32. Brünger, A. T. (1993) *X-PLOR Version 3.1 Manual*, Yale University, New Haven, CT.
33. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graph.* 14, 51–55
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291
35. Wuthrich, K. (1990) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
36. Asesio, J. L., Canada, F. J., Bruix, M., Rodriguez-Romero, A., and Jimenez-Barbero, J. (1995) *Eur. J. Biochem.* 230, 621–633
37. Asesio, J. L., Canada, F. J., Bruix, M., Gonzalez, C., Khiar, N., Rodriguez-Romero, A., and Jimenez-Barbero, J. (1998) *Glycobiology* 8, 569–577
38. Parijo, J. V., Broekaert, W. F., Goldstein, I. J., and Peumans, W. J. (1991) *Planta* 183, 258–264
39. Ciopraga, J., Gozia, O., Tudor, R., Brezuica, L., and Doyle, R. J. (1999) *Biochim. Biophys. Acta* 1428, 424–432
40. Cammune, B. A., De Bolle, M. F. C., Terras, F. R. G., Proost, P., and Van Damme, J., Rees, S. B., Vanderleyden, J., and Broekaert, W. F. (1992) *J. Biol. Chem.* 267, 2228–2233
41. Ha, S. C., Min, K., Koo, J. C., Kim, Y., Yun, D. J., Cho, M. J., and Kim, K. K. (2001) *Acta Crystallogr. D Biol. Crystallogr.* 57, 263–265
42. Aumelas, A., Mangoni, M., Roumeostand, C., Chiche, L., Despaux, E., Grasse, G., Calas, B., and Chavanieu, A. (1998) *Eur. J. Biochem.* 237, 575–583
43. Kawano, K., Yoneya, T., Miyata, T., Yoshikawa, K., Tokunaga, F., Terada, Y., and Iwanaga, S. (1990) *J. Biol. Chem.* 265, 15365–15367
44. Tamura, H., Kuroda, M., Masuda, M., Otaka, A., Funakoshi, S., Nakashima, H., Yamamoto, N., Waki, M., Matsumoto, A., Lancelin, J. M., Kohda, D., Tate, S., Inagaki, F., and Fujii, N. (1993) *Biochim. Biophys. Acta* 1163, 209–216
45. Matsuzaki, K., Yoneyama, S., Fujii, N., Miyajima, K., Yamada, K., Kirino, Y., and Anzai, K. (1997) *Biochemistry* 36, 9799–9806
46. Lamberty, M., Caille, A., Landon, C., Tassin-Moindrot, S., Hetru, C., Bulet, P., and Vovelle, F. (2001) *Biochemistry* 40, 11995–12003
47. Muhle, S. A., and Tam, J. P. (2001) *Biochemistry* 40, 5777–5785
48. Watanabe, T., Kono, M., Aida, K., and Nagasawa, H. (1996) *Mol. Mar. Biol. Biotechnol.* 5, 299–303
49. Krishnan, A., Nair, P. N., and Jones, D. (1994) *J. Biol. Chem.* 269, 20971–20976
50. Wang, P., and Granados, R. R. (1997) *J. Biol. Chem.* 272, 16663–16669
Structural Basis for New Pattern of Conserved Amino Acid Residues Related to Chitin-binding in the Antifungal Peptide from the Coconut Rhinoceros Beetle

Oryctes rhinoceros

Hikaru Hemmi, Jun Ishibashi, Tetsuya Tomie and Minoru Yamakawa

J. Biol. Chem. 2003, 278:22820-22827.
doi: 10.1074/jbc.M301025200 originally published online April 3, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M301025200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 48 references, 8 of which can be accessed free at http://www.jbc.org/content/278/25/22820.full.html#ref-list-1