Solution Structure of Synthetic Penaeidin-4 with Structural and Functional Comparisons with Penaeidin-3*

Brandon J. Cuthbertson*, Yinshan Yang†, Evelyne Bachère‡, Erika E. Büellesbach**, Paul S. Gross***, and André Aumelas††‡‡

From the *Centre de Biochimie Structurale, CNRS UMR 5048, INSERM U414, Université Montpellier 1, Faculté de Pharmacie, 15 Avenue Charles Flochault, 34093 Montpellier Cedex 5, France, ‡IFREMER-CNRS-Université Montpellier 2, UMR 5171, Université Montpellier 2, CC80 Place Eugène Bataillon 34095 Montpellier Cedex 5, France, §Marine Biomedicine and Environmental Sciences, Medical University of South Carolina, Charleston, South Carolina 29412, the ¶Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, South Carolina 29425, and the **Proteomics Facility, Clemson University Genomics Institute, 317C Biosystems Research Complex, Clemson, South Carolina 29634.

Antimicrobial peptide structure has direct implications for the complexity of functions and mechanisms of action. The penaeidin antimicrobial peptide family from shrimp is divided into multiple class designations based on primary structure. The penaeidin classes are not only characterized by variability in primary sequence but also by variation in target specificity and effectiveness. Whereas class 4 exhibits low isoform diversity within species and is highly conserved between species, the primary sequence of penaeidin class 3 is less conserved between species and exhibits considerable isoform diversity within species. All penaeidins, regardless of class or species, are composed of two dramatically different domains: an unconstrained proline-rich domain and a disulfide bond-stabilized cysteine-rich domain. The proline-rich domain varies in length and is generally less conserved, whereas the spacing and specific residue content of the cysteine-rich domain is more conserved. The structure of the synthetic penaeidin class 4 (PEN4-1) from Litopenaeus setiferus was analyzed using several approaches, including chemical mapping of disulfide bonds, circular dichroism analysis of secondary structural characteristics, and complete characterization of the solution structure of the peptide by proton NMR. L. setiferus PEN4-1 was then compared with the previously characterized structure of penaeidin class 3 from Litopenaeus vannamei. Moreover, the specificity of these antimicrobial peptides was examined through direct comparison of activity against a panel of microbes. The penaeidin classes differ in microbial target specificity, which correlates to variability in specific domain sequence. However, the tertiary structure of the cysteine-rich domain and indeed the overall structure of penaeidins is conserved across classes.

Antimicrobial peptides (AMPs) can form a physicochemical barrier that deters infection by potential pathogens in an individual organism. This is true for many multicellular organisms including both vertebrate and invertebrate animals (1–4). Many families of AMPs have been identified, and some are of particular interest for direct application to human health (5–7). In addition to direct therapeutic application of native AMPs, the potential exists for the enhancement of activity through the use of synthetic peptide derivatives that incorporate atypical or nonnative amino acids, so as to resist natural metabolic proteases, enhance pathogen specificity, and increase effectiveness (8–10).

Penaeidins are known to be a diverse family of cationic peptides that, despite their extensive variability in primary structure, target in an almost exclusive manner Gram(-) bacteria and fungi (11, 12). Variability in primary structure has been noted in the form of multiple classes (designated 1/2, 3, and 4) and multiple isoforms of some classes (13).

The typical penaeidin molecule is composed of two domains, an unconstrained proline-rich N-terminal domain (PRD) and a cysteine-rich domain (CRD) with a stable α-helical structure (14). Classes 3 and 4 are the most dissimilar representatives from this peptide family, being distinguished by differences in length and amino acid composition, with class 4 being the shortest and class 3 being the longest member of the family (15). Class 4 does not share the conserved Pro-Arg-Pro motif found in many members of classes 2 and 3 across species, and the PRD of class 4 exhibits an independent antimicrobial function not seen in the PRD of class 3 (15, 16). In contrast, penaeidin class 3 is characterized by the addition of a 9-amino acid linker region between the PRD and the CRD that is not found in class 4 (Fig. 1).

Pairwise comparisons and alignments of many penaeidin family members (PEN4, -3, and -2) from multiple species were the subjects of several descriptions of penaeidin sequences (15, 16). These alignments revealed that PEN4 had a significantly shorter PRD and CRD, mainly due to two deletions of 9 and 6...
residues at their C-terminal parts, respectively, resulting in a 47-residue peptide rather than a 63-residue peptide as in PEN3. The common PRD segment of 21 residues displayed only a 47.6% identity with conservation of 4 of 5 proline and 2 of 3 arginine residues. Divergence in primary structure between penaeidin classes is most pronounced in the PRD, but enough sequence dissimilarity exists between the CRD of each penaeidin class to permit phylogenetic classification into the three distinct and accepted classes across the taxonomic family, Penaeidae (17).

All penaeidins characterized thus far are generally inactive against Gram(−) bacteria in vitro, including known representatives of marine bacteria that are either normal components of shrimp gut flora or specific aquaculture-related shrimp pathogens (18). As a group, penaeidins exhibit class-specific activity against specific microbes, with each class having a unique level of effectiveness against certain Gram(+) bacteria and fungi (15), suggesting a partitioning of function in the shrimp immune system that relies on diversification of penaeidin sequence. For example, penaeidin class 4 (PEN4), a recently discovered family member, has a high relative effectiveness against fungal species and certain Gram(+) bacterial species (15), whereas PEN3 exhibits greater effectiveness against other Gram(+) bacterial species. Microbial target specificity may relate to structural characteristics that cannot be deduced directly from primary amino acid sequence comparisons between penaeidin isoforms. In light of this, complete structural characterization of PEN4 was undertaken for comparison with the recently characterized PEN3 (14).

Reported here is the chemical determination of the disulfide pattern and complete solution structure as determined by 1H NMR of synthetic PEN4. The PEN4 structure is also compared with that of PEN3, the shortest and longest sequences in this family, respectively. Overall distribution of positively charged and hydrophobic side chains both in the PRD and in the CRD are discussed in terms of antimicrobial activity.

MATERIALS AND METHODS

**Chemical Synthesis**—PEN4 was synthesized and initially characterized as described by Cuthbertson et al. (15). Briefly, the proline-rich segment (1–22 residues) was constructed through tert-butylxycarbonyl chemistry on 3-thiopropanic acid-4-methylbenzhydrolamine resin by solid phase peptide synthesis (19, 20). The crude peptide was extracted with 20% acetic acid, desalted, and lyophilized. The cysteine-rich segment (23–47 residues) was synthesized via conventional fluorenyl-methyloxycarbonyl chemistry and HPLC-purified. The two segments were resuspended in 2 M guanidine HCl, 1M NH4OH for 1 h resulting in cleavage of the peptide backbone. The solvent was removed by lyophilization, and fragments were resuspended in 2 μl of 0.2 M triis(2-carboxyethyl)phosphine (1 h at 37 °C) for complete reduction of remaining disulfide bonds and diluted with 18 μl of solution A in preparation for analysis by MALDI-TOF MS.

**Mass Spectrometry**—MALDI-TOF MS analysis was performed at the Medical University of South Carolina mass spectrometry facility, and spectra were recorded on either a PerSeptive Biosystems Voyager DR instrument or an Applied Biosystems (ABI) Voyager DR-STE instrument using α-cyano-4-hydroxycinnamic acid as the ionization matrix.

For peptide sequencing of the native PEN4, the ABI 4700 proteins analyzer was used. For each experiment, peptide was dissolved in 0.1% trifluoroacetic acid in water (1 μg/μl) and diluted with 3 volumes of 50 mM α-cyano-4-hydroxycinnamic acid in 80% acetonitrile. The m/z values are presented directly from the observed spectra in comparison with predicted values for each peptide examined.

**Circular Dichroism**—Molecular ellipticity was measured in the far UV (190–250 nm) using a Jasco J-710 spectropolarimeter. The protein was dissolved in water, and the concentration was determined by UV spectroscopy (the calculated molar absorption coefficient for PEN4 based on the presence of 3 tyrosine residues is 4020 cm⁻¹ M⁻¹). The peptide was diluted (1:1, v/v) to a final concentration of 5–50 μM with 50 mM phosphate buffer to assess the effect of concentration on secondary structural characteristics. The structural stabilizing reagent trifluoroethanol (TFE) was added to the 5 μM concentration up to 70% TFE. Measurements were performed using a cell of 0.1-cm path length, at a resolution of 0.2 nm and a bandwidth of 2 nm, and 10 spectra were averaged.

**NMR Spectroscopy**—Solution structure of PEN4 was obtained through proton NMR using the exact method that was used previously for PEN3 (14). D2O (99.95%) was purchased from the Commissariat à l’Énergie Atomique (Saclay, France). The pH values were measured at room temperature with a 3-nm electrode and are given uncorrected for the deuterium isotope effect. Chemical shifts were referenced with respect to sodium 4,4-dimethyl-4-silapentane-1-sulfonate according to the IUPAC recommendations. Two samples containing PEN4 (1.0 mM) were prepared in 95.5% H2O/D2O and in 99.98% D2O, respectively. The pH was adjusted to 3.6 by the addition of DCl or NaOD.

Proton NMR experiments were performed on both Bruker Avance 600 and Avance 500 spectrometers equipped with a triple resonance probe equipped with a cryoprobe and a triple resonance probe equipped with a cryoprobe, respectively. In all experiments, room temperature was set at the center of the spectrum at the water frequency. To identify overlapping spin systems, three sets of spectra were recorded at 12, 20, and 25 °C. Double quantum-filtered correlated spectroscopy (DQF-COSY) (22), z-filtered total correlated spectroscopy (z-TOCSY) (23), and nuclear Overhauser effect spectroscopy (NOESY) (24) spectra were acquired in the phase-sensitive mode using the States-organized spectro-temporal phase incrementation method (26). For spectra recorded in H2O, and except for the DQF-COSY spectra (where low power irradiation was used), the water resonance was suppressed by the WATERGATE method (26). The z-TOCSY spectra were obtained with a mixing time of 90 ms, and NOESY spectra were obtained with mixing times of 100, 150, and 300 ms. Slowly exchanging amide protons were identified in the NOE spectra and were observed at 25 °C following the solubilization of PEN4 in D2O from 15 min to 24 h.

Data were processed by using the same software packages utilized for PEN3 structural analysis (14). The full sequential assignment was achieved using the general strategy described by Wüthrich (27).

**Structure Calculation**—The NOESY cross-peaks were measured from the NOESY spectrum acquired at 25 °C with a mixing time of 150 ms and were subsequently divided into five classes, according to their intensities. Very strong, strong, medium, weak, and very weak NOEs were then converted into 1.8–2.4, 1.8–2.8, 1.8–3.6, 1.8–4.4, and 1.8–5.0 Å distance constraints, respectively. For equivalent protons or non stereospecifically assigned protons, pseudoatoms were introduced. The φ angle restraints were derived from the 'JHNH-CαH coupling constants, and the ψ angle restraints were derived from the combined analysis of the 'JHH-Hα,Hβ coupling constants and intraresidue NOEs, respectively.

To calculate three-dimensional structures, distance and dihedral angle restraints were used as input in the standard distance geometry/simulated annealing refinement and energy minimization protocol using XP-PLOR 3.8 (28). In the first stage of the calculation, an initial structure was generated from a template structure with randomized φ, ψ dihedral angles and extended side chains. In preliminary calculations, neither hydrogen bond nor disulfide bond was used as restraint. Analyzing the obtained structures and comparing them with the NMR data permitted identification of more additional NOE restraints, which were introduced into the subsequent calculation. Analysis of the average Cβ–Cβ distances between cysteines in the lowest energy structures permitted unambiguous determination of the...
of conditions in the presence of 10% Me₂SO and monitored by length sequence shown in Fig. 1) was assessed under a variety mediated methods in solution (15). Reduced full-length peptide was oxidized by chemically reduced cysteine-rich segment after previously thioester segment to the reduced cysteine-rich segment. Disulfide isomerization of native ligation of PEN4 proline-rich segment after previously evaluated by MALDI-TOF MS and 1H NMR.

Arrangement of the three disulfide bonds. After a number of these processes, 511 NOE-derived distance restraints (103 medium range and 53 long range) and 25 dihedral angles (8 φ, 14 χ₁, and 3 χ₂) were used as final input data. Finally, a calculation of 40 conformers including the disulfide bonds was carried out, and the resulting 20 structures with a minimum of restrained violations were submitted to 5000 cycles of restrained Powell energy minimization.

Structure Analysis—The visual display and the calculation of r.m.s. deviation were performed with INSIGHT 97 (Molecular Simulation Inc., San Diego, CA). Hydrogen bonds were considered as present if the distance between heavy atoms was less than 3.5 Å and the donor-hydrogen-acceptor angle was greater than 120°. The Ramachandran analysis was performed with PROCHECK (29), and the limits of the secondary structure elements and the van der Waals surfaces were determined with the STRIDE (30) and PROMOTIF (31) programs. The chemical shifts and coordinates of the energy-minimized conformers of PEN4 are deposited in the BioMagResBank (BMRB entry 6388) and the chemical shifts and coordinates of the energy-minimized conformers of PEN3 (14).

Antimicrobial Assays—A variety of microorganisms were used to measure antimicrobial activity of the synthetic PEN4 (see Table II). Minimum inhibitory concentrations were determined in triplicate by measuring antimicrobial activity of the synthetic PEN4 and the identical chemical shifts and coordinates of the energy-minimized conformers of PEN3 (14).

Chemical Synthesis—The oxidative folding of PEN4 (full-length sequence shown in Fig. 1) was assessed under a variety of conditions in the presence of 10% Me₂SO and monitored by analytical HPLC. Oxidation at pH 7.0 yielded a robust peak at 17.8 min (Fig. 2), indicating a single predominant isomer, and a highly pure disulfide isomer. Amino acid analysis and N-terminal peptide sequencing confirmed the predicted sequence of the full-length oxidized PEN4 (15).

Chromatographic and analytical methods in solution (15). B, oxidized PEN4 appeared as a mixture of disulfide isomers with retention times ranging from 17.8 to 21.5 min. At pH 7.0 in the absence of 0.15 NaCl, a most intense peak was observed for the most polar disulfide isomer (17.8-min retention time). Bulk oxidations under these conditions were done, and the predominant isomer was purified by HPLC. Analytical HPLC of highly pure PEN4 is shown at the bottom. The quality of the peptide was evaluated by MALDI-TOF MS and 1H NMR.

FIG. 1. Alignment of PEN4 and PEN3 sequences. The analogous PRD and CRD domains are boxed, and the percentage of sequence identity is displayed. The conservation of the disulfide arrangement for the two classes coincides with a more conserved CRD amino acid sequence relative to that of the PRD.

FIG. 2. Synthesis of PEN4. A, illustration of native ligation of PEN4 proline-rich thioester segment to the reduced cysteine-rich segment after previously described methods (19, 20). Reduced full-length peptide was oxidized by chemically mediated methods in solution (15). B, oxidized PEN4 appeared as a mixture of disulfide isomers with retention times ranging from 17.8 to 21.5 min. At pH 7.0 in the absence of 0.15 NaCl, a most intense peak was observed for the most polar disulfide isomer (17.8-min retention time). Bulk oxidations under these conditions were done, and the predominant isomer was purified by HPLC. Analytical HPLC of highly pure PEN4 is shown at the bottom. The quality of the peptide was evaluated by MALDI-TOF MS and 1H NMR.
Gram(+) bacteria, Gram(−) bacteria, and filamentous fungi. Dramatic differences in effectiveness are observed for PEN4 against filamentous fungi with respect to PEN3. Furthermore, PRD PEN4 alone was more active against these fungi than either PEN3 or PRD PEN3 and retained the same functional range of targets as PEN4 (Table II). In addition and as previously reported, target specificity for PEN3 against certain Gram(+) bacteria is also noted (15). Despite extensive testing, no penaeidin isoform described thus far exhibits activity against marine Gram(−) bacteria at peptide concentrations lower than 50 µM (15).

Class 4 Structural Stability—To further investigate the stability of the PEN4 structure, an effective concentration for antimicrobial activity (5 µM) was chosen for analysis (Fig. 4). Examination of the effect of TFE, an organic solvent that stabilizes intramolecular hydrogen bonds and promotes secondary structures such as α-helices and β-turns and destabilizes intermolecular hydrogen bonds as in β-sheets, further emphasized the stability of the PEN4 structure in solution. The minimum signal at 210 nm and a shoulder at 220 nm are indicative of an α-helix, but the low maximum at 195 nm implies a strong contribution of poorly ordered structures, possibly random coil or loops constrained by disulfides. Based on the assumption that the ellipticity at 220 nm for a 100% helical polypeptide results in an ellipticity of −40,000 (degrees cm² dmol⁻¹), the helix content of PEN4 in water at pH 7.5 is −23%, whereas in TFE the helix content increased to −30% as calculated using interpretative algorithms (32). This result implies that TFE had only a small effect on the stabilization of the secondary structure of PEN4.

NMR Structural Study and Assignment—The two-dimensional NMR spectra (TOCSY, DQF-COSY, and NOESY) of PEN4 were recorded at several temperatures, ranging from 12 to 25 °C. The identification of all of the spin systems of PEN4 was obtained by analysis and comparison of DQF-COSY, TOCSY, and NOESY spectra according to the general strategy previously described by Wu¨ thrich (27). Two selected parts of the TOCSY and of the NOESY experiments are displayed in Fig. 5, with the assignment of the spin systems and the dNN dipolar correlations, respectively. Due to the spreading of the amide protons of residues belonging to the cysteine-rich domain, their assignment was straightforward (Fig. 6). The overlapped amide signals and the three repetitive sequences (R7P8, R14P15, and R19P20) were assigned by using data recorded at different temperatures. (see Supplementary Materials for the list of chemical shifts).

**FIG. 3.** Chemically determined disulfide arrangement of PEN4. Following the protocol described by Wu and Watson (21), the purified PEN4 was partially reduced, cyanlated, and cleaved. After complete reduction, the cleavage products were analyzed by MALDI-TOF MS, revealing fragments that confirm the disulfide arrangement shown in Fig. 1. Included in this spectrum are three peaks that confirm a single disulfide bond (m/z 4103.71, 888.95, and 396.89) linking Cys38 to Cys45, labeled C. Also resolved are the two other largest fragments that would be generated from the A and B disulfide bonds that exist between Cys25 and Cys37 (m/z 2535.68) and between Cys26 and Cys44 (m/z 2582.17), respectively. Table I summarizes all fragments that were observed in a collection of several fractions obtained in the experiment, further confirming the arrangement.

| Table I | Summary of fragments detected from chemical approach to determination of PEN4 disulfide bond arrangement following the methods of Wu and Watson (21) |
|---|---|
| PEN4 intact sequence | HSSGYTRPLRKPSPRIFIRPIFGCDVCYQIPSSTARLCFCRYGDCCCHL* |
| Fragmentation at reduced and modified cysteine residues | / |
| Disulfide bond A (C23-C37) | HSSGYTRPLRKPSPRIFIRPIFGCDVCYQIPSSTARLCCFCRYGDCCCHL* |
| | +CDVCYQIPSSTARL* |
| | +CCFCRYGDCCCHL* |
| Disulfide bond B (C26-C44) | HSSGYTRPLRKPSPRIFIRPIFGCDV* |
| | +CYQIPSSTARLCCFCRYGD* |
| | +CCCHL* |
| Disulfide bond C (C38-C45) | HSSGYTRPLRKPSPRIFIRPIFGCDVCYQIPSSTARLC* |
| | +CCRYGC* |
| | +CCHL* |

*Partial reduction of PEN4 followed by cyanlation and cleavage resulted in fragments that could be analyzed by direct MALDI-TOF MS (Fig. 3). Analysis of the cleavage fragments proved that the disulfide arrangement of the chemically synthesized PEN4 is identical to that of the recombinant PEN3 determined by NMR (14). Individual fragments of a single bond are grouped according to bond letter (A–C), as shown in Fig. 3. All possible cleavage points are shown below the full-length sequence by placement of slashes. Plus signs indicate additional CN mass added as a result of possible bond cleavage at the amino terminus of cyanlated cysteine residues, and asterisks symbolize C-terminally amidated fragments. The calculated values were obtained by using the www.expasy.org/tools/peptide-mass.html Web site. Calculated and experimental (m/z) values correspond to the MH⁺ ions.
Several resonances display unusual chemical shift values. This was the case for those of the Arg$^{49}$ side chain, which are upfield-shifted in the 1.44 to 0.86 ppm range. The resonance of the $\text{H}_\alpha$ proton of Asp$^{33}$ is downfield-shifted (5.18 ppm), whereas those corresponding to the $\beta\beta'$ protons were strongly nonequivalent (0.90 ppm). Two other remarkable $\beta\beta'$ non-equivalencies were also observed for C23 (0.67 ppm) and for Y41 (0.51 ppm).

The chemical shift deviations measured for $\alpha$ protons with regard to random coil statistical values are displayed in Fig. 6. They clearly show two parts in the sequence. The first part (residues 1–28) suggested a poorly constrained structure, although several isolated deviations in the $\pm 0.2$–0.3 ppm range were observed. In contrast, for the second part (residues 29–45), large and successive deviations were measured. The negative and positive deviations suggested a helical structure for residues 31–40 and a short $\beta$ strand spanning residues 41–43. The largest deviation was measured for Ile$^{29}$ (+0.67 ppm). Although isolated, such a deviation is probably due to the particular environment for this residue. The distribution of NOEs (Fig. 6) also supported a poorly constrained segment (1–18 residues, which is shorter than that inferred above from the deviation of chemical shifts) and a well defined region (19–47 residues), which roughly match with the PRD and the CRD, respectively. The presence of successive dNN, $\text{da}\beta\beta, i + 3$ and $\text{da}\beta\beta, i + 4$ suggested a helical structure for residues 31–40. Most of the 10 successive amide protons found in slow exchange (from residues Ala$^{34}$ to Asp$^{43}$) belong to this helical structure and are therefore involved in hydrogen bonds.

**Calculation of Structures**—To calculate the PEN4 three-dimensional structure with XPLOR, 511 NOE-derived distances and 25 dihedral constraints were used as input. Although the disulfide arrangement was chemically determined, the disulfide bond constraints were not taken into account in the preliminary calculations. Then, since the $C\beta$-$C\beta$ distances between the cysteine pairs of the chemically determined disulfide bonds were consistent with disulfide bonds, the disulfide constraints were added for the final calculation. The 20 best conformers were energy-minimized, and the statistics for this ensemble are summarized in Table III. These calculations revealed that the PEN4 structure consisted of an unconstrained and a constrained part, which roughly match with the PRD (residues 1–17) and the CRD (residues 18–47), respectively.

The structure of the CRD was well defined. The final set of 20 conformers had a pairwise average backbone r.m.s. deviation of 0.48 ± 0.16 Å measured for the superimposition of the backbone atoms spanning 18–47 residues (Fig. 7). The Ramachandran plot of all residues (except for the glycines and prolines) of the 20 best conformers indicated that 71.5% were located in the most favored regions, 23.6% in the additional allowed regions, and 1.1% in the generously allowed regions. 1.8% of the residues, mainly belonging to the unconstrained PRD, were found in the disallowed regions. Limits of secondary structure elements were determined by using STRIDE and PROMOTIF.

The global fold of the CRD mainly consisted of a helix (residues 31–41), two turns of type IV (Cys$^{23}$–Cys$^{26}$ and Gly$^{42}$–Cys$^{45}$ residues), and one loop spanning 27–30 residues. All of these secondary structure elements were tightly linked together by the Cys$^{23}$–Cys$^{37}$, Cys$^{26}$–Cys$^{44}$, and Cys$^{38}$–Cys$^{45}$ disulfide bonds. Concerning their geometry, the statistical analysis showed that the Cys$^{23}$–Cys$^{37}$, Cys$^{26}$–Cys$^{44}$, and Cys$^{38}$–Cys$^{45}$ disulfide bridges have $\chi^3$ angle values of $-115.2 \pm 1.8^\circ$.

**Fig. 4. Structural stability of PEN4.** Circular dichroism analysis of PEN4 in phosphate buffer was done with different TFE concentrations. The lowest percentage of TFE that exhibited the maximal shift that is reflective of increased $\alpha$-helix content is around 30%. The molecular ellipticity (shown in units of $\theta$ [10$^{-3}$ degrees cm$^2$ dmol$^{-1}$] as the x axis) is an average of 10 spectra recorded over a range of wavelengths (190–250 nm, shown as the y axis).

---

**TABLE II**

Antimicrobial spectrum of activity of the chemically synthesized PEN4 and PRD PEN4 compared with PEN3 and PRD PEN3 expressed as minimum inhibitory concentration (MIC) range ($\mu$g) (updated from Ref. 15)

| Microorganisms | PEN4 | PRD PEN4 | PEN3 | PRD PEN3 |
|---------------|------|---------|------|---------|
| Filamentous fungi |       |         |      |         |
| *Fusarium oxysporum* | 0.84–1.26 | 2.94–4.38 | 3.12–6.25 | >50 |
| *Botrytis cinerea* | 4.38–6.57 | 14.8–22.2 | >50$^b$ | >50 |
| *Penicillium crustosum* | 1.26–1.9 | 4.38–6.57 | >50 | >50 |
| Gram(+) bacteria |       |         |      |         |
| *Aerococcus viridans* | 1.9–2.92 | 9.86–14.8 | 0.8–1.6 | >50 |
| *Bacillus megaterium* | >50 | >50 | 3.12–6.25 | >50 |

$^a$ PEN4 was compared directly in antimicrobial assays to PRD PEN4, PEN3, and PRD PEN3. MIC values are expressed in this table as the interval [X] – [Y], where [X] is the highest concentration tested at which the growth of the microbe is not affected and [Y] the lowest concentration that causes 100% growth inhibition.

$^b$ Previously unreported findings.
residues 31–41. Among the solvent-exposed residues, most of them were hydrophilic (Ser31, Ser32, Thr35, Arg36, and Arg40), but two were hydrophobic (Leu36 and Phe39). The Cys37, Cys38, and Ala44 residues belonged to the hydrophobic core of the CRD, which also included the Pro20, Ile21, Cys23, Cys26, Ile29, and Tyr41 residues. Several NOEs between aromatic protons of Tyr41 and protons of the Pro20 and Arg19 supported the spatial proximity of these three side chains. As a result, the Tyr41 and the Pro20 side chains that are relatively buried contributed to the hydrophobic core.

Interestingly, the Tyr41 side chain was surrounded by the Asp43 and Cys23 side chains, for which unusual $\beta^\prime$ nonequivalencies were observed. Both the Tyr41 steric hindrance and its ring current effect explained the unusual chemical shifts measured for these residues belonging to this constrained part of the molecule. The structure also revealed that the Arg40 side chain was located between the Phe39 and the Tyr41 aromatic rings, thus explaining the upfield shifts observed for the $\gamma\beta$ resonances.

The PEN4 sequence contains two aspartic residues (Asp24 and Asp43), which were located at the surface of the molecule. As discussed above, the Asp43 environment and the large $\beta^\prime$ proton nonequivalence (0.90 ppm) suggested that this side chain is particularly constrained. In contrast, the equivalency observed for the Asp24 $\beta^\prime$ protons suggested a great mobility for this residue. Both aspartic acid side chains were in the neighborhood of the Arg19 guanidine group and could interact electrostatically. However, neither the chemical shifts nor the nonequivalency measured for the protons of the Arg19 side chain support such interactions.

Comparison of the PEN4 and PEN3 CRDs—The solution structure of PEN3 was previously determined (Protein Data Bank entry 1ueo) (14). The two molecules shared similar features with an unconstrained PRD, which contrasts with the well defined structure of the CRD. The alignment of their cysteine-rich domains showed 61.5% identity with the conservation of the 6 cysteines and 2 of 3 arginine residues (Arg36 and Arg40), as well as Val24, Ile28, and Ala33 hydrophobic residues (Fig. 1). Therefore, only the structures of the two CRDs, which display a similar fold, were compared (Fig. 8). Their superimposition clearly showed that they share the same disulfide pattern that was chemically determined for the synthetic PEN4. For the 27–44 (PEN4) and 37–54 (PEN3) backbone superimposition, an r.m.s. deviation of 0.74 Å was measured, suggesting a similar structure for these two parts.

The PEN4/PEN3 conserved residues were similarly located in the globular structures including Ile29/Ile39, Ala34/Ala44, Ser32/Ser42, Arg35/Arg45, Arg40/Arg50, and Gly42/Gly52. Moreover, in the loop upstream from the helix, the Pro30–Ser31 sequence (for PEN4), equivalent to the Ser40–Phe41 sequence (for PEN3), had the same backbone structure. Downstream from the helix, the PEN4 (Y41GDC) and Pen3 (L51GRC) sequences also displayed a very similar backbone structure in which the Tyr41 and Leu51 hydrophobic side chains were similarly oriented.

The primary differences between the three-dimensional structures of two CRDs were localized in the surrounding of the two N-terminal cysteines (GC23DVC26Y sequence of PEN4 and GC32PVSC36R sequence of PEN3). The two fragments differed in their sequences (presence of Pro33 in PEN3) and in their cysteine spacing (see below). Although the disulfide pattern was conserved, the local structure of these fragments was significantly different. The C23DVC26 sequence adopted a type IV $\beta$-turn, whereas the structure of the C32PVSC36 sequence mainly consisted of a type I $\beta$-turn for the P33VSC36. The $\gamma$ angles of the two disulfide bonds involved in the structure of this sequence (Cys23–Cys37 (−115.2 ± 1.8°) and Cys26–Cys44

**FIG. 5.** Two selected parts of PEN4 $^1$H-NMR spectra recorded in H$_2$O (25 °C, pH 3.6). Upper part, the TOCSY spectrum shows the assignment of all spin systems. Lower part, the NOESY spectrum displays the dNN NOEs.
The penaeidin antimicrobial peptide family from shrimp is characterized by a great diversity of primary sequences (PEN2, -3, and -4), which has been the subject of several descriptions (15, 16). Penaeidin classes 3 and 4 were shown to differ in length and primary sequence, and they are the most diverged representatives of this peptide family in Litopenaeid species (13). Alignments revealed that PEN4 had a significantly shorter PRD and CRD (see above), resulting in a 47-residue peptide instead of a 63-residue peptide for PEN3 (the longest sequence in the family).

The native PEN4 from L. setiferus was affinity purified using a polyclonal rabbit antibody generated by inoculation with the synthetic peptide. Subsequent sequencing confirmed that the mature peptide (as deduced from expressed gene sequence analysis) is expressed in hemocytes (13, 15). However, for full characterization of this penaeidin member, PEN4 was synthesized using an original strategy based on the potential for...
reproducing a native peptide through efficient chemical methods employing native ligation and chemically mediated folding.

The major oxidized isomer was purified by HPLC and characterized by MALDI-TOF MS, N-terminal peptide sequencing, and amino acid analysis. The disulfide pattern of the peptide was chemically determined and was further confirmed by the NMR structural study carried out in water. The Cys23–Cys37, Cys26–Cys44, and Cys38–Cys45 disulfide pattern of PEN4 was identical to the corresponding arrangement previously determined for PEN3, resulting in a similarly folded C-terminal CRD. This similar disulfide pattern is notable, because, despite the 61.5% sequence identity between the two CRDs, they differ by the spacer length and sequence between the two first cysteines. PEN3 displays a 3-residue spacer, including a proline (C32PVSC36), whereas a two-residue spacer without proline is observed for PEN4 (C23DVC26). Although not clearly understood, it is known that a minor change in the sequence or spacing can induce a quite different disulfide arrangement (34). This is not observed for this family of peptides, since neither the spacer shortening nor the sequence changes altered the disulfide pattern.

PEN4 and PEN3 CRDs shared a similar fold and a well-defined and strongly conserved hydrophobic core (Cys23/Cys32, Cys26/Cys36, Cys37/Cys47, Cys44/Cys54, Ala34/Ala44, Ile37/Ile39, and Tyr50/Leu51). In contrast, the differences between their sequences gave rise to a significant alteration in the location of the hydrophobic and hydrophilic residues at the surface of the molecule. In addition, the two aspartate residues of PEN4 constituted a negatively charged cluster that does not exist in the PEN3 structure. Each of these qualities may influence the interaction with a microbial target component, particularly membrane lipids. Variations in hydrophobic and ionic regions of the peptides may account for the microbial target specificity observed in direct comparison of classes 3 and 4.

When comparing the two PRDs, the proline/glycine-rich region present in PEN3, which links the N-terminal and C-
terminal regions, is missing in PEN4 (Fig. 1). The common PRD fragment of 21 residues displayed 47.6% identity with conservation of 4 of 5 proline and 2 of 3 arginine residues. One potential source of acquired function of the PEN4 PRD is the regular spacing of the arginine-proline-hydrophobic residue 3-residue clusters. This is typically the case for R7P8L9, R14P15I16, and R19P20I21 sequences and to a lesser extent for the K11P12S13 sequence. Both these cationic and hydrophobic clusters are strengthened by Arg10 and Phe17 and by Ile18 residues, respectively. This pattern mimics many functional AMPs that possess both cationic and hydrophobic patches (35, 36). The comparison of the PEN4 and PEN3 PRDs shows that two of these four 3-residue clusters are conserved in PEN3 (R9P10I11 and R19P20I21), whereas the two others were merged in a larger one (R13P14P15P16V17F18), with a larger hydrophobic sequence including the three adjacent prolines. This analysis revealed that the two PRD sequences share some similarities but display significant sequence divergences that could be responsible for differences in antimicrobial activity and potency. The complex differences between the PRD of these two classes preclude prediction of functional differences based solely on structural characteristics.

The fact that the PRD of each class was neither constrained nor in contact with the CRD does not explain why the PEN4 CRD did not autonomously and correctly fold under conditions tested for the full-length molecule. This would indicate the potential that subtle and as yet undefined interactions between residues of the PRD and the CRD of PEN4 are essential to initiate and facilitate the correct folding process, much like a molecular chaperone. In contrast, an autonomous fold of the CRD of PEN3 was obtained with a 30–35% yield, suggesting that in this case the PRD sequence is not essential for the oxidative folding process.2

Results from CD analysis of PEN4 structural stability are in agreement with the overall three-dimensional structure derived by NMR spectroscopy. The α-helix content as calculated from the CD spectra is 23–30% and is in good agreement with the NMR spectroscopy data, where 11 of 47 residues (23.4%) form an α-helical backbone in the CRD. The resistance of the structure to changes upon the addition of TFE is explained both by the highly disulfide bond-constrained CRD and by the absence of helix propensity of the PRD mainly due to the presence of four proline residues in positions 8, 12, 15, and 20 (Figs. 1 and 7).

Similar to PEN3, PEN4 exhibited antimicrobial activity directed against Gram(+) bacteria. In contrast to PEN3, the synthetic PEN4 is highly effective against filamentous fungi. Target specificity inherent in these two different classes is particularly obvious when examined in the presence of two of the fungal species examined, B. cinerea and P. crustosum, and the Gram(+) bacterium B. megaterium (Table II) (15). In addition, the PEN4 PRD was an effective antimicrobial peptide particularly against fungi, exhibiting the same effect as PEN4

2 B. J. Cuthbertson, Y. Yang, E. Bachère, E. E. Bülliesbach, P. S. Gross, and A. Aumelas, unpublished data.
of suppressing spore production when in the presence of *F. oxysporum* (15). This single domain exhibits an activity against the same panel of microbes as the full-length PEN4, whereas the PEN3 PRD is inactive up to 50 μM in all tests. It appears that the CRDs in these two AMPs play a similar role (i.e. potentiation of activity) but to a much different degree, since both the CRD and PRD are essential to confer the maximal antimicrobial activity *in vitro* with respect to the particular microbial species tested. Replacement of a positively charged residue (Arg<sup>23</sup> of PEN3) by a negatively charged one (Asp<sup>43</sup> of PEN4) modifies the overall charge of the molecule but does not alter the antimicrobial activity, indicating that the substitution of this particular residue is not critical for activity.

Directed mutagenesis of the PEN4 PRD focused toward the identification of key residues and synthesis of truncated versions of this domain might be a useful approach for identification of a putative functional core, and further structural analyses of the interaction of these peptides with specific components of microbial membranes, such as membrane lipids, may shed light on the specific elements present in these peptides that promote activity. Therefore, the analysis of these peptides in the presence of synthetic or simulated membrane structures by solid state NMR (37, 38) or other types of structural analysis particularly suited to examination of membrane biology, such as x-ray lamellar diffraction and wide angle in-plane scattering (39, 40), may be necessary to examine and evaluate their potential for membrane interaction and penetration.

Beyond activity at the membrane, the seemingly complex function of PEN4 and PEN4 PRD in suppressing spore formation in *F. oxysporum* implies another level of function associated with the particulate microbial species tested. Replacement of a positively charged residue (Arg<sup>23</sup> of PEN3) by a negatively charged one (Asp<sup>43</sup> of PEN4) modifies the overall charge of the molecule but does not alter the antimicrobial activity, indicating that the substitution of this particular residue is not critical for activity.

Directed mutagenesis of the PEN4 PRD focused toward the identification of key residues and synthesis of truncated versions of this domain might be a useful approach for identification of a putative functional core, and further structural analyses of the interaction of these peptides with specific components of microbial membranes, such as membrane lipids, may shed light on the specific elements present in these peptides that promote activity. Therefore, the analysis of these peptides in the presence of synthetic or simulated membrane structures by solid state NMR (37, 38) or other types of structural analysis particularly suited to examination of membrane biology, such as x-ray lamellar diffraction and wide angle in-plane scattering (39, 40), may be necessary to examine and evaluate their potential for membrane interaction and penetration.

Beyond activity at the membrane, the seemingly complex function of PEN4 and PEN4 PRD in suppressing spore formation in *F. oxysporum* implies another level of function that is more complex than simple membrane disruption. At this point, a specific internal microbial component that may be targeted by penaeidin peptides has yet to be determined. The multiple distinct structural characteristics found in members of the penaeidin family that have been resolved through this comparative analysis may indicate multiple levels of function that include, but are not limited to, interaction with the microbial membrane.

REFERENCES

1. Hoffmann, J. A. (2003) *Nature* **426**, 33–38
2. Hoffmann, J. A., and Reichhart, J. M. (2002) *Nat. Immunol.* **3**, 121–126
3. Iwama, S. (2002) *Curr. Opin. Immunol.* **14**, 87–95
4. Lehrer, R. I., and Ganz, T. (1999) *Curr. Opin. Immunol.* **11**, 23–27
5. Jacob, L., and Zasloff, M. (1994) *Ciba Found. Symp.* **186**, 197–216; discussion 216–213

6. Ge, Y., MacDonald, D., Henry, M. M., Hait, H. I., Nelson, K. A., Lipsky, B. A., Zasloff, M. A., and Holroyd, K. J. (1999) *Diagn. Microbiol. Infect. Dis.* **35**, 45–53
7. Ge, Y., MacDonald, D. L., Holroyd, K. J., Thornsherry, C., Wexler, H., and Zasloff, M. (1999) *Antimicrob. Agents Chemother.* **43**, 782–788
8. Hananeto, K., Kida, Y., Zhang, Y., Shimizu, T., and Kowano, K. (2002) *Microbiol. Immunol.* **46**, 741–749
9. Patch, J. A., and Barron, A. E. (2002) *Curr. Opin. Chem. Biol.* **6**, 872–877
10. Fajgh, J. A., and Barron, A. E. (2003) *J. Am. Chem. Soc.* **125**, 12092–12093
11. Bache`re, E., Gueguen, Y., Gonzalez, M., de Lorgeril, J., Garnier, J., and Romestand, B. (2004) *Immunol. Rev.* **189**, 149–168
12. Destoumieux, D., Bulet, P., Loeser, D., Van Dorselaer, A., Rodriguez, J., and Bache`re, E. (1997) *J. Biol. Chem.* **272**, 28398–28406
13. Cuthbertson, B. J., Shepard, E. F., Chapman, R. W., and Gross, P. S. (2002) *Immunogenetics* **54**, 442–445
14. Yang, Y., Poncet, J., Garnier, J., Zatylny, C., Bachere, E., and Aumelas, A. (2003) *J. Biol. Chem.* **278**, 36859–36867
15. Cuthbertson, B. J., Bullesbach, E. F., Fievet, J., Bachere, E., and Gross, P. S. (2004) *Biochem. J.* **381**, 79–86
16. Bache`re, E., Destoumieux, D., and Bulet, P. (2000) *Aquaculture* **191**, 71–88
17. Cuthbertson, B. J. (2004) *Multiple Classes of Penaeidin Antimicrobial Peptides Expressed in Individual Shrimp Exhibit Microbial Target Specificity and a Class-Specific Gain of Function in a Single Domain*. Ph.D. dissertation, Medical University of South Carolina, Charleston, SC
18. Destoumieux, D., Bulet, P., Strub, J. M., Van Dorselaer, A., and Bache`re, E. (1999) *Eur. J. Biochem.* **266**, 335–346
19. Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) *Science* **266**, 776–779
20. Tam, J. P., Lu, Y. A., Liu, C. F., and Shao, J. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 12485–12489
21. Wu, J., and Watson, J. T. (1997) *Protein Sci.* **6**, 391–398
22. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) *Eur. J. Biochem.* **119**, 479–485
23. Rance, M. (1987) *J. Magn. Reson.* **47**, 557–564
24. Macura, S., Huang, Y., Suter, D., and Ernst, R. R. (1981) *J. Magn. Reson.* **43**, 259–281
25. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson.** **85**, 393–399
26. Piotto, M., Saudek, V., and Skenar, V. (1992) *J. Biol. Chem.* **266**, 661–665
27. Wuhrman, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York
28. Rice, L. M., and Brunger, A. T. (1994) *Proteins* **19**, 277–290
29. Lasowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–289
30. Frishman, D., and Argos, P. (1995) *Proteins* **23**, 566–579
31. Chan, A. W., Hutchinson, E. G., Harris, D., and Thornton, J. M. (1993) *Protein Sci.* **2**, 1574–1590
32. Delage, G., and Geourjon, C. (1993) *Comput. Appl. Biosci.** *9*, 197–199
33. Srinivasan, N., Sowdhamini, R., Ramakrishnan, C., and Balaram, P. (1990) *Int. J. Pept. Protein Res.* **36**, 147–155
34. Volkman, B. F., and Wemmer, D. E. (1997) *Biopolymers* **41**, 451–460
35. Zasloff, M. (2002) *Nature* **415**, 389–395
36. Shai, Y. (2002) *Biopolymers* **66**, 236–248
37. Bache`re, E. (1997) *Immunol. Rev.* **149–168
38. Bache`re, E. (1997) *Immunol. Rev.* **149–168
39. Bache`re, E. (1997) *Immunol. Rev.* **149–168
40. Bache`re, E. (1997) *Immunol. Rev.* **149–168
41. Wishart, D. S., and Case, D. A. (2001) *J. Appl. Crystallogr.* **34**, 451–460
42. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950
43. Kraulis, P. J. (1991) *Methods Enzymol.* **338**, 3–34
44. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950