Solution Structure of the Recombinant Penaeidin-3, a Shrimp Antimicrobial Peptide

Yinshan Yang, Joël Poncet, Julien Garnier, Céline Zatylny-Gaudin, Evelyne Bachère, André Aumelas

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 38, Issue of September 19, pp. 36859–36867, 2003
Printed in U.S.A.

Solution Structure of the Recombinant Penaeidin-3, a Shrimp Antimicrobial Peptide*

Received for publication, May 23, 2003, and in revised form, June 23, 2003
Published, JBC Papers in Press, July 3, 2003, DOI 10.1074/jbc.M305450200

Yinshan Yang‡, Joël Poncet§, Julien Garnier¶, Céline Zatylny¶, Evelyne Bachère¶, and André Aumelas‡‖

From the ‡Centre de Biochimie Structurale, CNRS UMR 5048, INSERM U414, Université Montpellier 1, Faculté de Pharmacie, 15 avenue Charles Flahault, 34093 Montpellier Cedex 5, France, the §Laboratoire de Génomique Fonctionnelle, CNRS UPR 2580,141, rue de la Cardonille, 34094 Montpellier Cedex 5, France, and the ¶IFREMER-CNRS-Université Montpellier 2, Défense et Résistance chez les Invertébrés Marins (DRIM), FRE 2626, Université Montpellier 2, CC80 Place Eugène Bataillon 34095 Montpellier Cedex 5, France

Penaeidins are a family of antimicrobial peptides of 47–63 residues isolated from several species of shrimp. These peptides display a proline-rich domain (N-terminal part) and a cysteine-rich domain (C-terminal part) stabilized by three conserved disulfide bonds whose arrangement has not yet been characterized. The recombinant penaeidin-3a of *Litopenaeus vannamei* (63 residues) and its [T8A]-Pen-3a analogue were produced in *Saccharomyces cerevisiae* and showed similar antimicrobial activity. The solution structure of the [T8A]-Pen-3a analogue was determined by using two-dimensional ^{1}H NMR and simulated annealing calculations. The proline-rich domain, spanning residues 1–28 was found to be unconstrained. In contrast, the cysteine-rich domain, spanning residues 29–58, displays a well defined structure, which consists of an amphipathic helix (41–50) linked to the upstream and the downstream coils by two disulfide bonds (Cys^{32}–Cys^{47} and Cys^{48}–Cys^{55}). These two coils are in turn linked together by the third disulfide bond (Cys^{36}–Cys^{54}). Such a disulfide bond packing, which is in agreement with the analysis of trypsin digests by ESI-MS, contributes to the highly hydrophobic core. Side chains of Arg^{45} and Arg^{50}, which belong to the helix, and side chains of Arg^{37} and Arg^{53}, which belong to the upstream and the downstream coils, are located in two opposite parts of this globular and compact structure. The environment of these positively charged residues, either by hydrophobic clusters at the surface of the cysteine-rich domain or by sequential hydrophobic residues in the unconstrained proline-rich domain, gives rise to the amphipathic character required for antimicrobial peptides. We hypothesize that the antimicrobial activity of penaeidins can be explained by a cooperative effect between the proline-rich and cysteine-rich features simultaneously present in their sequences.

Antimicrobial peptides are major elements of host-defense systems represented in all species from plants to vertebrate and invertebrate animals. Among these antimicrobial molecules, cysteine-rich peptides are the most widespread. They are structurally classified into (i) peptides with a β-sheet structure such as the mammalian defensins (1), (ii) peptides with a β-hairpin-like fold such as tachyplesins from horseshoe crabs (2), thanatin (3), porcine protegrins (4, 5), androctonin (6), or gomesin (7), and (iii) peptides adopting the cystine-stabilized αβ-motif, including invertebrate and plant defensins (8–10). Only recently, such effectors of innate immunity were isolated from crustaceans, whereas numerous peptides have been characterized from other arthropods, both insects and chelicerates. Three peptides, named penaeidins (Pen),[1] were initially purified in their active form (5.48–6.62 kDa) from the hemocytes of the shrimp *Litopenaeus (Penaeus) vannamei*, and they were fully characterized at the amino acid level (Pen-1, Pen-2, Pen-3a) and by cDNA cloning from a hemocyte library (pen-2, -3a, -3b, and-3c) (11). Penaeidins are an original peptide family composed of an N-terminal proline-rich sequence and a C-terminal region containing six cysteines engaged in three intramolecular disulfide bridges. Recently, using the molecular EST approach (Expression Sequence Tag) and sequence homology search, members of the penaeidin family have been identified in several shrimp species, including *Litopenaeus setiferus* (12, 13), *Penaeus monodon*,[2] *Fenneropenaeus chinensis*,[3] *Litopenaeus stylirostris*,[4] and *Penaeus japonicus* (14). These data confirm that the penaeidin family is characterized by the presence of two distinct domains and by the conserved position of the six cysteines in the C-terminal domain. Interestingly, these studies have also revealed a new class of penaeidins, Pen-4 (Fig. 1), and as evidenced for the Pen-3 class, an abundance and diversity of isoforms can be produced in a single species (13). Penaeidins are synthesized as precursor molecules consisting of a peptide signal (19–21 residues) preceding the mature bioactive peptide. Pen-1, -2, and Pen-4 have between 48

* 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 1ueo) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

The chemical shifts have been submitted to the BMRB under accession no. 5806.

‖ To whom correspondence should be addressed: Centre de Biochimie Structurale, CNRS UMR 5048, INSERM U414, Faculté de Pharmacie, 15 avenue Charles Flahault, 34093 Montpellier Cedex 5, France. Tel.: 33-0-4-67-04-34-32; Fax: 33-0-4-67-52-96-23; E-mail: aumelas@cbs.cnrs.fr.

[1] The abbreviations used are: Pen, penaeidins; CSI, chemical shift index; DG, distance geometry; DQF-COSY, two-dimensional double-quantum filter correlation spectroscopy; ESI-MS, electrospray ionization-mass spectrometry; MIC, minimal inhibitory concentration; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; Pen-3a, penaeidin-3a isoform of *L. vannamei*; RP-HPLC, reverse-phase high performance liquid chromatography; rmsd, root mean square deviation; SA, simulated annealing; TOCSY, total correlation spectroscopy; TPPI, time proportional phase incrementation; PDB, protein data bank.

[2] A. Tassanakajon, personal communication.

[3] J. Xiang, personal communication.

[4] M. Muñoz, personal communication.

and 50 amino acids, whereas Pen-3 class peptides are the longest peptides (63 amino acids), characterized by the insertion of a medial sequence (3–5 amino acids), linking together the proline-rich and the cysteine-rich domains, and by additional residues at the C-terminal sequence. Pen-2 and Pen-3a were produced in the recombinant system *Saccharomyces cerevisiae* to study their biological properties and their spectra of antimicrobial activities (15). These peptides exhibit broad activity at rather low concentration (below 10 μM) against mainly Gram-positive bacteria and filamentous fungi, and comparatively weak activity against the Gram-negative bacteria (ranging from 10 to 50 μM). In addition, for Pen-3a, a chitin-binding property has been evidenced, which appeared to be displayed by the C-terminal region (16).

Penaeidins are original peptides, which possess six cysteine residues organized in two doublets in their C-terminal domain. This is unusual in antimicrobial peptides since only vertebrate defensins, namely β-defensins, and hepcidin are known to present one and three cysteine doublets, respectively. Hepcidin (initially called liver-expressed antimicrobial peptide, LEAP-1), which displays antimicrobial activity, is presently known as an iron-regulatory hormone (17, 18). Neither the disulfide bond arrangement nor the three-dimensional structure of the penaeidins has yet been reported. For this reason, the Pen-3 family, which appears to be most abundantly produced in shrimp, was chosen for such a structural study. Given the fact that native peptides are difficult to purify from natural samples, the Pen-3a isoform was expressed in a heterologous expression system (15). In this study, we report the solution structure of the recombinant Pen-3a, as determined by ^{1}H NMR derived-constraints, and its disulfide arrangement. Finally, from the determined structure, the distribution of positively charged and hydrophobic side chains, both in the unconstrained proline-rich and in the highly constrained cysteine-rich domains is discussed in terms of the amphipathic character required for antimicrobial activity of peptides.

EXPERIMENTAL PROCEDURES

Expression and Purification of Pen-3a and of Its [T8A] Analogue—Pen-3a used in this study was obtained as previously described (15) by recombinant expression in the *S. cerevisiae* strain TGY 48-1 (20). Briefly, the Pen-3a coding sequences fused to the yeast β-glucanase/MFα1 preprosequence, were subcloned into a shuttle vector pTG4812. This construct was used to transform the yeast. An analogue of Pen-3a having the Thr to Ala substitution at residue 8, and termed [T8A]-Pen-3a analogue, was also expressed in the same yeast strain. The substitution was performed using the QuickChange site-directed mutagenesis kit (Stratagene) by changing only one nucleotide from the Pen-3a coding sequence.[5] *S. cerevisiae* TGY-transformed cells, expressing respectively Pen-3a and its [T8A]-Pen-3a analogue, were inoculated into selective medium YNBG (yeast nitrogen base 0.67%, glucose 1%) supplemented with 0.5% casamino acids lacking uracil. After 72 h of incubation at 30 °C under vigorous shaking, the cultures were centrifuged (1000 × *g*, 20 min, 4 °C) and the supernatants acidified to pH 2.9 by addition of trifluoroacetic acid. Large scale purifications of recombinant penaeidins were essentially performed as previously described (15). To increase peptide purity, the molecules were finally purified on a RP18 reversed-phase column (Symetry Shield TM RP18; 4.6 × 250; 5 mm; Waters) by a linear biphasic gradient of 2–18% acetonitrile in acidified water over 10 min, and of 18–23% over 40 min at a flow rate of 0.8 ml/min. Finally, the fractions containing the penaeidins were pooled, lyophilized, and kept as dry powder.

Mass Spectrometry—Compounds were recorded offline in a QSTAR-Pulsar-*i* mass spectrometer (Applied Biosystems, Foster City, CA) fitted with a Protana nanospray inlet system (Protana, Odense, Denmark) and analyzed using the Analyst QS software (Applied Biosystems). Parameters were adjusted as follows: ion spray voltage (IS), 900 V; curtain gas (CUR), 25; declustering potential (DP), 45–75 V; focusing potential (FP), 265 V; declustering potential 2 (DP2), 15 V. Capillaries were loaded with ~2 μl of sample and placed in the source tip holder. Prior to analysis by nanospray, each aliquot from trypsin cleavage was solubilized in 50 μl of deionized water. 2 μl were desalted on Poros 20 R2 packed in a gel-loader pipette tip and eluted with 50:50:1 methanol/water/formic acid following an already described procedure (21). Analysis of fragments was carried out by using the Protein Prospector (V4.0.4) program (prospector.ucsf.edu/mshome4.0.htm).

Antimicrobial Assays—Antimicrobial activity of the [T8A]-Pen-3a analogue was assayed against different micro-organisms already used in a previous study relative to the activity spectrum study of the recombinant Pen-3a (15). The test microorganisms were the Gram-positive bacteria *Aerococcus viridans*, *Micrococcus luteus*, and *Bacillus megaterium*; the Gram-negative strains, *Escherichia coli 363*, *Salmonella thyphimurium*, *Klebsiella pneumoniae*, and marine bacteria *Vibrio harveyi* pathogenic for shrimp. Finally, the filamentous fungus *Fusarium oxysporum* isolated from shrimp was also used.

MICs were determined in triplicate by the liquid growth inhibition assay as described (11). Poor broth nutrient medium (PB: 1% bactotryptone, 0.5% NaCl, w/v, pH 7.5) or saline peptone water (1.5% peptone, 1.5% NaCl w/v, pH 7.2) was used for the standard bacterial strains and the marine bacteria, respectively. Briefly, in a sterile microtitration plate, 10 μl of peptide, or deionized water as a control was added to 90 μl of a mid-logarithmic growth phase culture of bacteria diluted in culture medium to OD_{600} = 0.001. Plates were incubated for 24 h at 30 °C with vigorous shaking and monitored spectrophotometrically at 620 nm. An identical liquid growth inhibition assay as that used to measure anti-bacterial activity was used to measure anti-fungal activity on Potato Dextrose Broth (Difco) medium. After incubation for 48 h at 25 °C in the dark without shaking and in a moist chamber, growth inhibition of the fungus was measured at 600 nm.

Trypsin Cleavage—Pen-3a and its [T8A]-Pen-3a analogue (100 μg) were subjected to bovine trypsin (Promega Corporation) treatment with a trypsin/substrate ratio of 1/20 (w/w) in 500 μl of a 200 mM Tris-HCl buffer (pH 8.0) at 37 °C. After 2, 5, 10, and 17 h of incubation, 100-μl aliquots of the solution were frozen to stop the reaction and freeze-dried. ESI-MS spectra were recorded and analyzed as described above.

NMR Spectroscopy—D_2O (99.95%) was purchased from the CEA (Saclay, France). The pH values were measured at room temperature with a 3-mm electrode and are given uncorrected for the deuterium isotopic effect. ^{1}H chemical shifts were referenced with respect to sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) according to the IUPAC recommendations (22).

Two 300-μl samples containing either Pen-3a or the [T8A]-Pen-3a analogue (0.7 mM) were prepared in a Shigemi tube (Shigemi Co., Ltd., Tokyo, Japan) in 95:5 H_2O:D_2O and in 99.98% D_2O, respectively. The pH was adjusted to 3.9 by addition of DCl or NaOD.

Proton NMR experiments were performed both on a Bruker Avance 600 and Avance 500 spectrometers equipped with a triple resonance probe and with a cryoprobe, respectively. In all experiments, the carrier frequency was set at the center of the spectrum at the water frequency. To identify several overlapping spin systems, three sets of spectra were recorded at 12, 20, and 27 °C. DQF-COSY (23), z-TOCSY (24), and NOESY (25) spectra were acquired in the phase-sensitive mode using the States-TPPI method (26). For spectra recorded in H_2O, and except for the DQF-COSY spectra (where low-power irradiation was used), the water resonance was suppressed by the WATERGATE method (27). z-TOCSY spectra were obtained with a mixing time of 50 and 70 ms and NOESY spectra with a mixing time of 100, 150, and 200 ms.

Slowly exchanging amide protons were identified from successive TOCSY and NOESY spectra recorded at 27 °C following the solubilization of Pen-3a and of [T8A]-Pen-3a in D_2O from 15 min to 24 h.

Data were processed by using either the XWINNMR or the GIFA (28) software. The data were zero-filled before processing and shifted sine-bell functions were used for apodization. The processed data were baseline corrected using a five-order polynomial function. The full sequential assignment was achieved using the general strategy described by Wüthrich (29).

Structure Calculation—The NOESY cross-peaks were measured from the NOESY spectrum acquired at 27 °C with a mixing time of 150 ms and were 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-stereospecific assigned protons, pseudo-atoms were introduced. The ϕ angle restraints were derived from the $^3J_{HN\text{-}C\alpha H}$ coupling constants, and the χ1 angle restraints were derived from the combined analysis of the $^3J_{H\alpha\text{-}H\beta}$ coupling constants and intra-residues NOEs, respectively.

To calculate three-dimensional structures, these distance and dihe-

[5] C. Zatylny, unpublished results.

dral angle restraints were used as input in the standard distance geometry (DG)/simulated annealing (SA) refinement and energy-minimization protocol using X-PLOR 3.8 (30). In the first stage of the calculation, an initial ensemble of 60 structures was generated from a template structure with randomized ϕ, φ dihedral angles and extended side chains, using a DG protocol followed by restrained SA and refinement (31). No hydrogen bond restraint was used. Since the arrangement of the disulfide bonds had to be determined, structures were generated without using any disulfide bond constraints. Analyzing the obtained structures and comparing them with the NMR data allowed us to identify more additional NOE restraints, which were introduced into the subsequent calculation. After a number of these processes, 552 NOE-derived distance restraints (73 medium range and 93 long range) and 28 dihedral angles (15 ϕ and 13 χ1) were used as final input data. Analysis of the average Sγ–Sγ distances in the lowest energy structures allowed us to unambiguously determine the arrangement of the three disulfide bonds. Finally, a calculation of 60 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 rmsd were performed with INSIGHT 97 (Molecular Simulation Inc., San Diego). 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 (32), and the limits of the secondary structure elements and the van der Waals surfaces were determined with the STRIDE program (33). The chemical shifts and coordinates of the energy-minimized conformers of [T8A]-Pen-3a are deposited in the BMRB (accession number 5806) and in the Protein Data Bank (PDB entry: 1ueo), respectively.

RESULTS

Production of Pen-3a and of Its [T8A] Analogue—S. cerevisiae-expressed Pen-3a was purified by reversed-phase HPLC. As already shown in a previous study, in addition to molecules similar to the native one, several forms of the recombinant penaeidins were found to be glycosylated (15). Despite additional successive purification steps performed to obtain highly pure Pen-3a suitable for the structural study, residual contaminating glycosylated molecules were observed (see below). Nevertheless, a small amount of the non-glycosylated peptide was isolated and characterized by ESI-MS (6686.62 Da). It displayed a glutamine residue in its N terminus and therefore corresponded to the non-post-translationally modified Pen-3a [the N terminus of native Pen-3a was shown to be a pyroglutamic amino acid (11)]. To obtain enough non-glycosylated material to carry out the structural study, a [T8A]-Pen-3a analogue was further produced. Since in a previous study (15), the threonine in position 8 was identified as a site of *O*-glycosylation, this residue was substituted for alanine by directed mutagenesis to avoid glycosylation events in the yeast. The [T8A]-Pen-3a analogue was expressed in *S. cerevisiae*, purified by RP-HPLC, and characterized by ESI-MS and by NMR. Its monoisotopic mass (6656.30 Da) was in agreement with the theoretical one of 6656.26 Da (Table I). Its ^{1}H NMR spectrum was similar to that of the parent peptide suggesting that they share the similar fold.

Antimicrobial Activity—The antimicrobial activity of the recombinant Pen-3a and of its [T8A]-Pen-3a analogue was measured using a panel of microorganisms including Gram-positive and Gram-negative bacteria, and the filamentous fungus *F. oxysporum.* The MIC values obtained for the recombinant Pen-3a and the [T8A]-Pen-3a analogue were found to be almost identical and revealed similar ranges of activity (Table II). The antimicrobial activity appeared to be target-specific with great inhibition of some of the Gram-positive bacteria strains such as *A. viridans* (with MIC less than 1.6 μM). As for Pen-3a, an anti-Gram-negative activity was recorded for [T8A]-Pen-3a against *E. coli*, whereas no inhibition of the other bacterial strains tested could be observed for either peptides up to 50 μM.

TABLE I
ESI-MS data for the recombinant Pen-3a and its [T8A]-Pen-3a analogue and for their fragments resulting of the trypsin hydrolysis

	Monoisotopic experimental mass	Monoisotopic theoretical mass
	Da	*Da*
Peptide		
Pen-3a	6686.62[a,b]	6686.26
[T8A]-Pen-3a	6656.30[c]	6656.26
Fragment		
5-59	5773.96[c]	5773.83
10-59	5269.72[d]	5269.58
20-59	4113.00[e]	4112.89
20-37/46-50/54-59	2976.33[f]	2976.27
10-37/46-50/54-59	4133.07[g]	4132.99

[a] Means values obtained from deconvoluted charged species are shown in Footnotes *b–g*.
[b] 6+ to 9+.
[c] 5+ to 9+.
[d] 5+ to 8+.
[e] 4+ to 6+.
[f] 3+ to 5+.
[g] 3+ to 6+.

TABLE II
Antimicrobial spectrum activity comparison of the recombinant Pen-3a and its [T8A]-Pen-3a analogue

Microorganisms	Pen-3a	[T8A]-Pen-3a
	μM	μM
Gram (+) bacterial		
Aerococcus viridans	0.8–1.6[a]	0.4–0.8
Micrococcus luteus	0.8–1.6	0.8–1.6
Bacillus megaterium	3.12–6.25	3.12–6.25
Gram (−) bacteria		
Escherichia coli 363	6.25–12.5	3.12–6.25
Vibrio harveyi	>50	>50
Salmonella thyphimurium	>50	>50
Klebsiella pneumoniae	>50	>50
Filamentous fungi		
Fusarium oxysporum	3.12–6.25	3.12–6.25

[a] MIC are expressed as the interval *a-b,* where *a* is the highest concentration tested at which the growth of the microorganisms is not inhibited and *b* the lowest concentration that causes 100% growth inhibition.

Thus, we concluded that the [T8A] mutation did not affect the antimicrobial activity against the selected microorganisms.

Attempts to Identify the Disulfide Pattern by Trypsin Cleavage and ESI-MS—The Pen-3a sequence of *L. vannamei* is characterized by a high content in glycine (17.5%), proline (15.9%), arginine (11.1%), serine (9.5%), and cysteine (9.5%). Together, these five amino acids account for 63.5% of the sequence. In contrast, Pen-3a does not contain either aspartic acid, or tryptophan, nor methionine. Class 3 sequences, which appear to be highly conserved within the shrimp species, significantly differ between species (ex: *L. vannamei* and *L. setiferus*) (13).

Concerning the location of the six cysteines, it should be noted that for all classes four cysteines belong to 2 doublets (Cys^{47}, Cys^{48}, and Cys^{54}, Cys^{55} for Pen-3a) and that the first two cysteines are separated by three residues (Pro^{33}§NVal^{34}–Ser^{35}) in class 3, whereas they are separated by only one for class 1 (Ala^{26}) or by two for class 2 (Asn^{27}–Ala^{28}) and class 4 (Asp^{24}–Val^{25}) (Fig. 1).

The arrangement of the three disulfide bonds in penaeidins was as yet unknown. The ESI-MS analysis of trypsin digests allowed us to identify fragments essentially resulting from cleavages after Lys^4, Arg^9, Arg^{19}, Arg^{37}, Arg^{45}, Arg^{50}, Arg^{53}, and Lys^{59} residues (Fig. 1 and Table 1B). The molecular mass of peptides observed suggested that there is not a disulfide bond between Cys^{32} and Cys^{36} cysteines. In addition, neither the fragment S^{46}-R^{50} (552.64 Da) nor the fragment Leu^{51}–Lys^{59} (970.19 Da) and/or Cys^{54}–Lys^{59} (643.27 Da), each of them

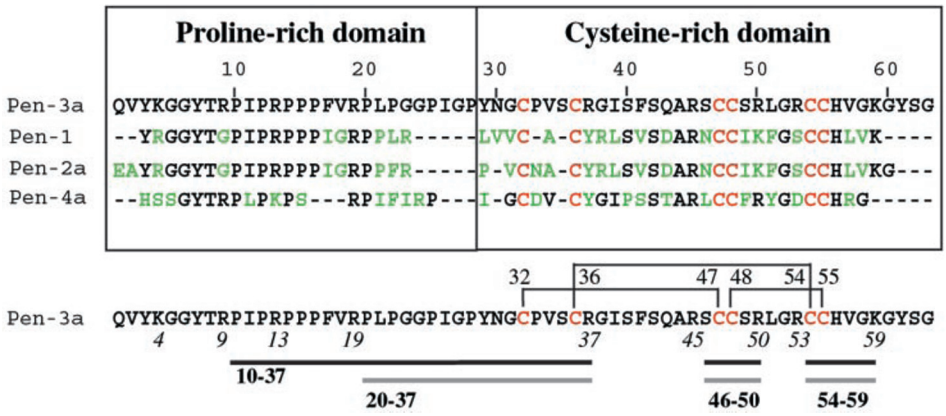

FIG. 1. *Upper part*, selected penaeidins from *L. vannamei* belonging to the four classes and their sequence alignment, with regard to the Pen-3a sequence, as proposed by ClustalW (19). Notice that penaeidins of class 3 significantly differ from the others, mainly by an insertion of 3–5 residues in the middle of the sequence and by its C-terminal extension. The numbering on the *top* is that of Pen-3a. The N-terminal proline-rich and the C-terminal cysteine-rich domains are *boxed*. Mutations are displayed in *green* and conserved cysteines in *red*. *Lower part*, to determine the disulfide pattern, a trypsin cleavage was carried out. The trypsin cleavage sites of Pen-3a are labeled by the numbers in *italics*. The *three black lines* represent the Pro^{10}–Arg^{37}/Ser^{46}–Arg^{50}/Cys^{54}–Lys^{59} heterotrimer cross-linked by three disulfide bonds (4133.07 Da) that was observed in the trypsin digests. The three gray lines are for a shorter peptide corresponding to the Arg^{19}-Pro^{20} cleavage (2976.33 Da). The disulfide pattern determined in the present study is displayed above the sequence.

potentially including a disulfide bond between two adjacent cysteines, could be identified. This suggests that neither the Cys^{47}–Cys^{48} nor the Cys^{54}–Cys^{55} disulfide bonds are present in Pen-3a. In contrast, two peptides corresponding to three fragments cross-linked together by the three disulfide bonds were characterized from trypsin digests of Pen-3a and of its [T8A]-Pen-3a analogue. A first one of 4133.07 Da, resulting from the Arg^{9}, Arg^{37}, Arg^{45}, Arg^{50}, and Lys^{59} cleavages, which corresponds to the Pro^{10}–Arg^{37}/Ser^{46}–Arg^{50}/Cys^{54}–Lys^{59} fragments (theoretical mass 4132.99 Da), and another shorter one of 2976.33 Da resulting from the Arg^{19}, Arg^{37}, Arg^{45}, Arg^{50}, Arg^{53}, and Lys^{59} cleavages, which corresponds to the Pro^{20}–Arg^{37}/Ser^{46}–Arg^{50}/Cys^{54}–Lys^{59} fragments (theoretical mass 2976.27 Da) (Fig. 1). These three-fragment peptides only differ by the presence of the Pro^{10}–Arg^{19} sequence. On the basis of the trypsin cleavage data, the Cys^{32}–Cys^{36}, the Cys^{47}–Cys^{48}, and the Cys^{54}–Cys^{55} disulfide bonds were excluded discarding seven arrangements among the 15 theoretically possible. Even using chymotrypsin, no more information could have been obtained from such an unusual disulfide pattern. In conclusion, the trypsin cleavage alone did not allow us to fully determine the disulfide bond pattern of Pen-3. Its determination has to be completed by modeling.

NMR Study—Preliminary NMR studies showed that Pen-3a gave a few broad resonances. The ESI-MS data showed that Pen-3a was contaminated by glycosylation on Thr^{8} with polysaccharide chains from 3 to 15 sugar units. These polysaccharide units were also observed on the DQF-COSY experiment recorded in D_2O, showing at least 12 typical signals of anomeric protons in the 4.8–5.4 ppm range. Moreover, two spin systems of practically equal intensity were observed for the Tyr^{7}, Thr^{8}, and Arg^{9} residues. For one spin system of Thr^{8}, the NOE between the methyl group and the sugar protons (3.64 and 5.10 ppm) clearly indicated its glycosylation. Therefore, to avoid the glycosylation process, the [T8A]-Pen-3a analogue was expressed in yeast, purified to homogeneity, and used for the structural study.

Assignment—Two-dimensional spectra (TOCSY, DQF-COSY, and NOESY) of both Pen-3a and the [T8A]-Pen-3a analogue were recorded at several temperatures, ranging from 12 to 27 °C. The similarity of their spectra indicated that the two peptides were similarly folded. The identification of all the spin systems of [T8A]-Pen-3a was obtained by analysis and comparison of DQF-COSY, TOCSY, and NOESY spectra according to the strategy described by Wüthrich (29). Two parts of the NOESY experiment, showing the dNN and the $d\alpha(i)-\delta\delta'Pro(i+1)$ dipolar correlations, are displayed in Fig. 2.

Due to the spreading of spin systems of the cysteine-rich domain, their assignment was straightforward (Fig. 2). In contrast, due to the unusual number of prolines (three are successive, Pro^{14}–Pro^{15}–Pro^{16}) and glycine residues along with several repetitive sequences (three Arg–Pro, (Arg^{9}–Pro^{10}, Arg^{13}–Pro^{14}, Arg^{19}–Pro^{20}) and two Gly–Gly (Gly^{5}–Gly^{6}, Gly^{23}–Gly^{24})), the assignment of the proline-rich domain was more difficult. However, the unambiguous assignment of the 10 prolines was obtained from the $d\alpha(i)$-$\delta\delta'Pro(i+1)$ dipolar correlations which were well resolved, due to the dispersion of the alpha proton resonances of the preceding residues (4.06–5.08 ppm) (Fig. 2). When the TOCSY, DQF-COSY, and NOESY data recorded in H_2O and D_2O were combined, the full assignment of the peptide was obtained. Chemical shifts were deposited in the BMRB (accession number 5806).

To locate the elements of secondary structure, chemical shifts of the alpha protons were compared with statistical chemical shift values determined for random coil conformation to calculate the chemical shift indexes (CSI) (34, 35). Indeed, successive positive and negative CSI are indicative of a β-stranded or a helical structure. The CSI are reported in Fig. 3. They indicate a helical structure for the Phe^{41}–Arg^{50} sequence and β-stranded structures for the Arg^{13}–Pro^{15} and Leu^{51}–Arg^{53} short sequences. The Val^{34}, Ser^{35}, Cys^{36}, and Cys^{55} negative indexes suggest that these residues belong to either a short helix or a turn, respectively. The largest positive chemical shift deviations measured for Ile^{39} (0.55 ppm), Leu^{51} (0.40 ppm), Cys^{32} (0.37 ppm), and Arg^{53} (0.31 ppm) and the largest negative chemical shift deviations measured for Cys^{36} (−0.79 ppm), Ala^{44} (−0.68 ppm), Cys^{47} (−0.58 ppm), Arg^{45} (−0.47 ppm), Cys^{48} (−0.46 ppm), and Cys^{55} (−0.45 ppm) indicate that all these residues belong to constrained parts of the molecule. In contrast, on the basis of CSI, most of the residues of the proline-rich domain and of the His^{56}–Gly^{63} C-terminal part are unconstrained. The distribution and the summary of NOEs as well as the amide protons in slow exchange are displayed in Fig. 3.

Calculation of Structures—To calculate the [T8A]-Pen-3a structures with XPLOR, 552 NOE-derived distances and 28 dihedral constraints were used as input. Since the disulfide bond arrangement was unknown, disulfide bond constraints

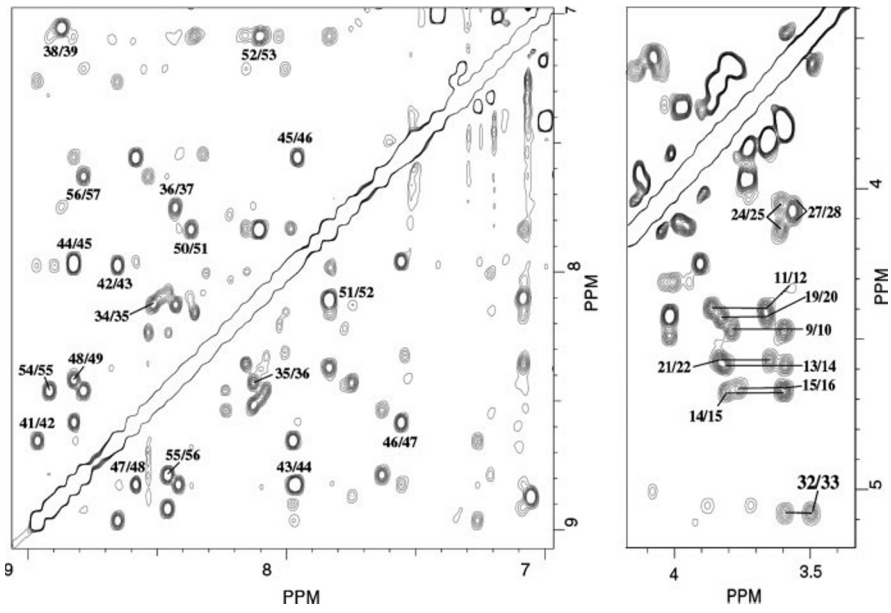

FIG. 2. **Two parts of the NOESY spectrum of [T8A]-Pen-3a showing the dNN and the dα(i)-δδ′Pro(i+1) NOEs recorded in H_2O (*left*) and D_2O (*right*), respectively (20 °C, pH 3.9, 150 ms of mixing time).**

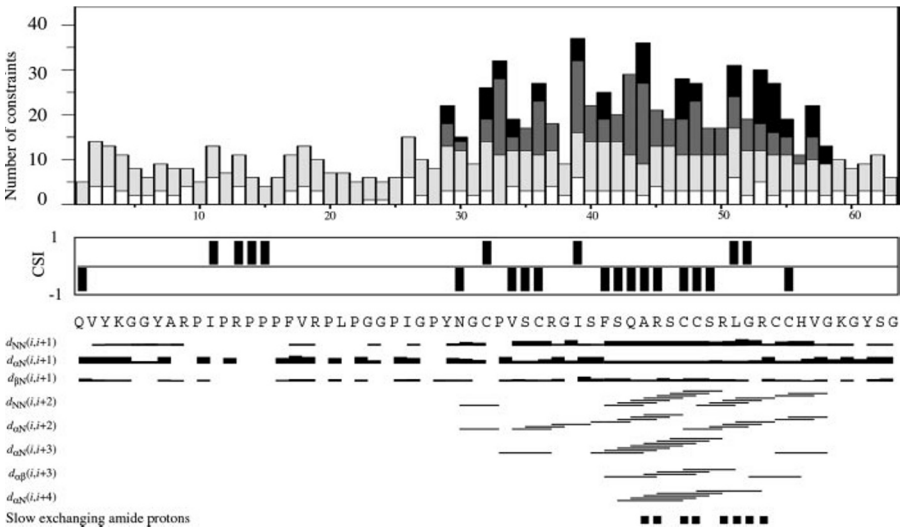

FIG. 3. **Distribution and summary of NOEs and CSI for [T8A]-Pen-3a.** *Upper part*, distribution of intraresidual (*white*), sequential (*light gray*), medium range (*dark gray*), and long range (*black*) constraints along the sequence. This plot clearly indicates that only the second half of the sequence, corresponding to the cysteine-rich domain, adopts a constrained structure. *Middle part*, only chemical shift deviations greater than 0.2 ppm were retained to plot the CSI of Hα resonances as a function of the sequence. Random coil chemical shifts are those from Wishart & Case (35). *Lower part*, summary of the sequential, medium-range, and long range NOEs. The relative intensity of NOEs is represented by the thickness of the *bars*. Slowly exchanging amide protons are identified by *filled squares*.

were not taken into account in the preliminary calculations. These calculations revealed that the [T8A]-Pen-3a structure consists of an unconstrained and a constrained part, which roughly match with the proline-rich (1–28 residues) and the cysteine-rich (29–58 residues) domains, respectively.

The global fold of the cysteine-rich domain is stabilized by three disulfide bonds whose arrangement has yet to be identified. As a result of trypsin cleavage experiment seven arrangements out of the 15 theoretically possible were withdrawn from consideration. From the global fold of the cysteine-rich domain, the eight arrangements still possible had to be taken into account by a statistical analysis of the mean distances between the beta carbons of cysteines. Therefore, an initial calculation of structures was carried out without any disulfide bond.

Such a distance analysis allowed us to unambiguously identify a first disulfide bond between Cys^{48} and Cys^{55}. However, the inter-cysteine dαβ NOEs, that usually characterize a disulfide bond, could not be observed due to the superimposition of the resonances of these protons. The mean distance between their Cβ was about 3.6 Å, whereas the distances with the other cysteine Cβ were in the range of 7.5–10.5 Å. Consequently, the two other disulfide bonds involve the four remaining cysteines (Cys^{32}, Cys^{36}, Cys^{47}, and Cys^{54}) in three possible arrangements. In fact, in the preliminary calculations these four cysteine Cβ occupied the corners of a square whose side is 4–5 Å in length. Their location agrees well with NOEs observed between Hα of Cys^{32} and Hβ of Cys^{47}, between Hα of Cys^{54} and Hβ of Cys^{36}, and between Hβ of Cys^{54} and Hβ of Cys^{47}. Nevertheless, due to the packing, some NOEs must involve protons belonging to two disulfide bonds, thus making ambiguous the identification of the correct disulfide arrangement, which in this case remains a challenge.

To solve the disulfide arrangement, the three possible disulfide arrangements -Cys^{32}–Cys^{36}/Cys^{47}–Cys^{54}, Cys^{32}–Cys^{54}/Cys^{36}–Cys^{47}, and Cys^{32}–Cys^{47}/Cys^{36}–Cys^{54}- were taken into account in modeling and three calculations were carried out to evaluate their agreement with experimental data. As shown in Table III, for the Cys^{32}–Cys^{36}/Cys^{47}–Cys^{54} pattern, inconsistent with the trypsin cleavage; the Cys^{54} side chain adopted an unfavorable rotamer ($\chi 1$ around 0°). In the structures obtained with the Cys^{32}–Cys^{54}/Cys^{36}–Cys^{47} forced pattern, the distance between the Cys^{36} Hα and Cys^{32} Hβ protons and between the Cys^{36} amide proton and the Cys^{47} Hβ protons are too short ($d\alpha\beta \leq 1.8$ Å) and give rise to severe van der Waals contacts. These two strong contacts are inconsistent with the NOESY spectrum since no strong NOE between these protons was observed. Moreover, the resulting disulfide bonds interfere each other. All these inconsistencies as well as the NOEs vio-

TABLE III
Summary of unsatisfied constraints for the cysteine-rich domain structures of Pen-3a

These were calculated with the three possible disulfide arrangements involving the Cys^{32}, Cys^{36}, Cys^{47}, and Cys^{54}. Their consistency with trypsin cleavage results are indicated.

Disulfide bonds	Unsatisfied NOEs	Trypsin cleavage
32–36/47–54	C^{54} $\chi1=0°$	Inconsistent
32–54/36–47	$d\alpha\beta C^{36}C^{32}$	Agreement
	$d\alpha\beta C^{36}C^{47}$ [a]	
32–47/36–54	None	Agreement

[a] The two disulfide bonds interfere each other.

lations, allowed us to conclude that the likely disulfide bond pattern of [T8A]-Pen-3a is the third one (Cys^{32}–Cys^{47} and Cys^{36}–Cys^{54}). Indeed, this arrangement, which is in agreement with the trypsin cleavage satisfies all the experimental NOEs since neither violation nor strong van der Waals contact was observed.

A final round of calculations including the Cys^{32}–Cys^{47}, Cys^{36}–Cys^{54}, and Cys^{48}–Cys^{55} disulfide bonds was carried out and the 20 lowest energy conformers out of the 60 refined structures were minimized and analyzed. The statistics for this ensemble of structures are summarized in Table IV. Since the proline-rich domain of [T8A]-Pen-3a was found to be unstructured, only the superimposition of conformers of the cysteine-rich domain is shown in Fig. 4. This globular part is well defined and the pairwise mean rmsd value for the superimposition of the 30–58 backbone atoms was 0.636 ± 0.166 Å (0.533 ± 0.154 Å for 31–57 backbone atoms). The distribution in the Ramachandran plot of all residues (except for the glycines and prolines) of the 20 best structures indicates their quality: 73.0% are located in the most favored regions, 24.1% in the additional allowed regions, and 1.5% in the generously allowed regions. 1.3% of the residues, mainly located in the unconstrained proline-rich domain, are located in the disallowed regions (32). Limits of secondary structure elements were determined by using STRIDE.

Finally, the main feature of the [T8A]-Pen-3a structure is that the proline-rich domain does not adopt a constrained structure, whereas the cysteine-rich domain adopts a well-defined structure that consists of a helix (41–50 residues) tightly bound to the upstream (27–40 residues) and the downstream (51–63 residues) coils by Cys^{32}–Cys^{47} and Cys^{48}–Cys^{55} disulfide bonds. In turn, these two coils are linked together by the third, Cys^{36}–Cys^{54}, disulfide bond (Fig. 4). The structure of the Pro^{33}–Cys^{36} and Cys^{54}–Val^{57} sequences consists of beta turn structures of type I and IV, respectively. Such a disulfide bond arrangement, which tightly links the three segments gives rise to a globular shape for the cysteine-rich domain (20 and 23 Å for its diameter and height, respectively). This is supported by the slow H-D exchange measured for eight amide protons involved in hydrogen bonds. Four hydrogen bonds between Ala^{44}, Arg^{45}, Cys^{47}, and Cys^{48} amide protons and Ser^{40}, Phe^{41}, Gln^{43}, and Ala^{44} carbonyl groups, respectively, are located in the helical part. Four other hydrogen bonds are located in the loop following the helix and involve Arg^{50}, Leu^{51}, Gly^{52}, and Arg^{53} amide protons and Ser^{46}, Cys^{47}, Cys^{48}, and Cys^{47} carbonyl groups, respectively.

Conformation of the Disulfide Bridges and of the X-Pro Amide Bonds—It is worthy to note that two of the three disulfide bridges are buried in the core of the molecule and significantly contribute to its hydrophobicity. In contrast, the third one (Cys^{48}–Cys^{55}) is exposed to the solvent. Concerning their geometry, the statistical analysis shows that the Cys^{32}–Cys^{47}, Cys^{36}–Cys^{54}, and Cys^{48}–Cys^{55} disulfide bridges have $\chi3$ angle values of −105.6 ± 15.0°, −110.31 ± 12.5°, and 109.3 ± 3.8°, respectively. These values are those measured for classical conformations of disulfide bridges ranging in the range of ± 90 ± 45° (36). Concerning the geometry of the ten *X*-Pro amide bonds (*X* for any residue), the observation of the typical $d\alpha$(i-1)-$\delta\delta'$Pro(i) NOEs was indicative of a *trans* conformation for all of them (Fig. 2).

TABLE IV
Experimental and structural statistics for the family of 15 structures of [T8A]-Pen-3a

For these calculations, the XPLOR all-hydrogen force fields "topoallhdg" and "parallhdg all" were used. The final minimization of the 21 structures was made with force constants of 15 kcal mol^{-1} $Å^{-2}$ and 50 kcal mol^{-1} rad^{-2} for the NOE and dihedral-angle potentials, respectively.

Distance restraints	
Intraresidue (i − j = 0)	159
Sequential (i − j = 1)	227
Medium-range (i − j ≤ 5)	73
Long-range (i − j > 5)	93
Total	552
Dihedral-angle restraints	
ϕ	15
χÅ1	13
total	28
NOE violations	
Number >0.2 and <0.5 (Å)	5.70 ± 1.9
NOE	0.0594 ± 0.0026
Dihedral violations	
Number >5°	none
cdih	0.54 ± 0.24
Mean rmsd from idealized covalent geometry	
Bonds (Å)	0.00338 ± 0.00007
Angles (°)	0.6182 ± 0.0046
Impropers (°)	0.3839 ± 0.0042
Mean energies (kcal·mol^{-1})	
E noe	29.29 ± 2.59
E cdih	0.206 ± 0.153
E vdW	−43.80 ± 3.59
E bond	10.80 ± 0.47
E improper	11.49 ± 0.25
E angle	99.65 ± 1.47
E total	−57.79 ± 6.14
Ramachandran (%)[a]	
Most favored	73.0
Additionally allowed	24.2
Generously allowed	1.5
Disallowed	1.3
Pairwise atomic rmsd (Å)	
Backbone atoms (31–57)	0.533 ± 0.154
Heavy atoms (31–57)	1.265 ± 0.188
Backbone atoms (30–58)	0.636 ± 0.166
Heavy atoms (30–58)	1.313 ± 0.195

[a] Calculated with PROCHECK.

DISCUSSION

As a general feature, antimicrobial peptides are positively charged and most of them display an amphipathic character. According to their sequence or their three-dimensional structures, antimicrobial peptides are divided into four classes: the proline-rich, the glycine-rich, those which adopt a helical structure, and those which have disulfide bonds. Penaeidin sequences contain all these features. A sequence identity search indicated that neither the proline-rich nor the cysteine-rich domains of Pen-3a share significant identity with known antimicrobial proteins. Nevertheless, some low and limited similarities of the proline-rich sequence (Pro^{12}–Arg^{13}–Pro^{14}–Pro^{15}) with apidaecin, drosocin, pyrrhocoricin, metalkowin and (^{9}RPIPRP^{14}) with the bactenecin-7 antimicrobial peptides have been reported (10). Concerning the cysteine-rich domain, some limited sequence similarities were observed with the chitin binding motif (^{47}CCS^{49} sequence) and tachycitin (10, 37).

The presence of these features in the same molecule might explain the multifonctional properties of the penaeidins and their broad spectrum of activity. Indeed, penaeidins display antimicrobial activity against Gram-positive bacteria, either

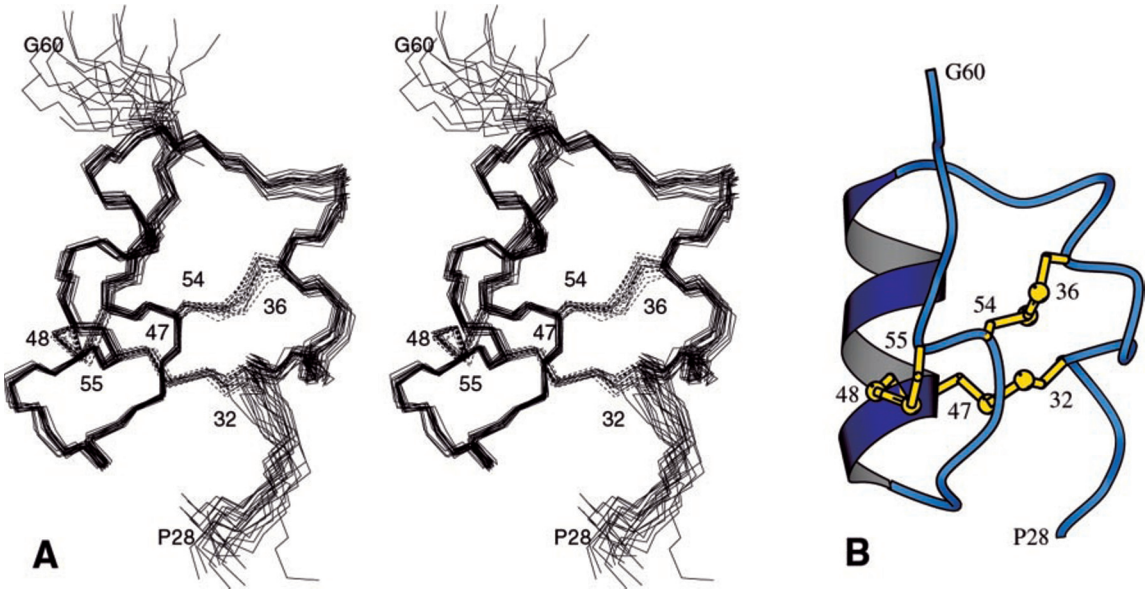

FIG. 4. **Structure of the cysteine-rich domain of [T8A]-Pen-3a.** *A*, stereo view of the 20 best energy minimized conformers. The heavy atoms of the backbone were used for the superimposition. The three disulfide bridges are labeled and displayed as *dashed lines*. *B*, the structure prepared using MOLSCRIPT shows the arrangement of the three disulfide bonds (in *yellow*) tightly linking together the helix and the upstream and downstream coils.

bactericidal or bacteriostatic depending on the strains, and a fungicidal effect activity against filamentous fungi (15). Additionally, a chitin-binding property has been experimentally evidenced for the peptide (16) as well as a capability to bind to Gram-negative bacteria of the genus *Vibrio* (38). Thus, to understand the function of penaeidins, the elucidation of their three-dimensional structure and organization of the two domains appeared to be of prime importance.

Since the amount of the native Pen-3a needed for a structural study was difficult to purify from shrimp, we expressed it in the yeast *S. cerevisiae*.

Whereas the native peptides purified from shrimp hemocytes are not post-translationally modified, the yeast-expressed Pen-3a appeared to be partially *O*-glycosylated on Thr^{8}. Despite an intensive purification procedure to eliminate the glycosylated forms, traces of residual modified penaeidins rendered difficult the NMR study. To circumvent this *O*-glycosylation the [T8A]-Pen-3a analogue was expressed in yeast, purified, and used instead. No differences were seen between the antimicrobial activity of the analogue and its parent peptide. Their ^{1}H NMR spectra also appeared to be similar, thus suggesting an identical three-dimensional structure. Consequently, the structural study of the [T8A]-Pen-3a analogue was carried out by NMR to determine both its three-dimensional structure and the pattern of the three disulfide bonds. This structural study allowed us to gain insight into the amphipathic character of penaeidins and to initiate a structure-activity relationship analysis for this family of antimicrobial peptides.

Disulfide Bonds and Three-dimensional Structure—Because of the two cysteine doublets, the determination of the full disulfide bond arrangement by the enzymatic cleavage alone was destined to fail. Nevertheless, given the numerous trypsin cleavage sites in the molecule, the absence of fragments including the potential Cys^{32}–Cys^{36}, Cys^{47}–Cys^{48}, and Cys^{54}–Cys^{55} disulfide bonds suggested that these three disulfide bonds were not present in Pen-3a. The absence of disulfide bonds between two adjacent cysteines (Cys^{47}–Cys^{48} and Cys^{54}–Cys^{55}) is not surprising since such a disulfide bond is known to be energetically unfavorable, although such a bond formation has been reported for a protein, the carboxypeptidase T (PDB code: 1obr) (39), and recently for the hepcidin peptide (PDB code: 1m4e.pdb and 1m4f.pdb), an iron-regulatory hormone which also displays antimicrobial activity (40). In contrast, the mass of two fragments from Pen-3a corresponding to a three fragment-peptide cross-linked by the three disulfide bonds was measured. Although these trypsin cleavage data did not solve the arrangement of the three disulfide bonds, they allowed us to reject seven out of the 15 disulfide bond patterns theoretically possible. For these three packed disulfide bonds, the arrangement proposed by the modeling step is Cys^{32}–Cys^{47}, Cys^{36}–Cys^{54}, and Cys^{48}–Cys^{55}. One disulfide bond (Cys^{48}–Cys^{55}) is solvent-exposed, whereas the two others belong to the hydrophobic core of the cysteine-rich domain.

The established three-dimensional structure together with the sequence allowed us to more rationally define the limits of the proline-rich and cysteine-rich domains. The absence of NOEs between the Pro^{28}, Gly^{58}, and Lys^{59} side chains with the core of the cysteine-rich domain and the presence of NOEs between side chains of Tyr^{29} and Leu^{51} and between Val^{57} and Cys^{36} were indicative of the fact that Tyr^{29} and Val^{57} belong to the cysteine-rich domain whereas Pro^{28} and Gly^{58} do not. Therefore the cysteine-rich domain spans residues Tyr^{29}–Val^{57}. This is clearly shown in Fig. 4. Therefore, concerning the proline-rich domain, which is mainly unconstrained, it spans the N-terminal part up to Pro^{28}. Nevertheless, it should be noted that no proline is located in the first nine residues and that the two last prolines (Pro^{25} and Pro^{28}) belong to an additional sequence that is absent in the three other classes of penaeidins (Pro^{25} is present in Pen-4a). Moreover, this additional sequence contains three glycine residues (Gly^{23}, Gly^{24}, and Gly^{27}), which probably provide this sequence with high flexibility, which therefore could act as a hinge between the two domains.

The [T8A]-Pen-3 structure, as determined from the NMR-derived constraints, is characterized by a typically unstructured part corresponding to the proline-rich domain (1–28 residues), which contrasts with the well-defined structure adopted by the cysteine-rich domain spanning 29–57 residues. The global fold of this latter domain consists of a helix inserted between an upstream and a downstream coil antiparallel to the helix, all tightly bound together by the disulfide bonds.

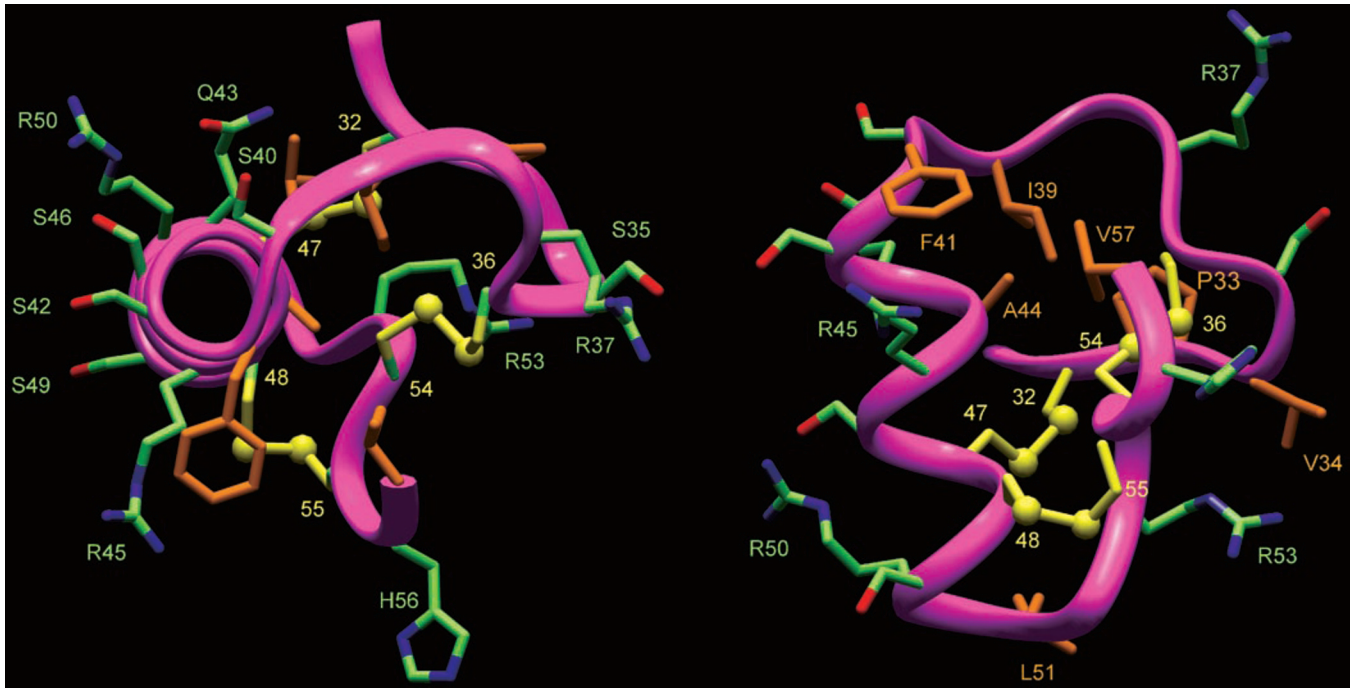

FIG. 5. **Two views 90° apart, of the structure of the 31–57 part of the cysteine-rich domain, showing the location of the hydrophilic and hydrophobic side chains.** *Left*, the solvent-exposed face of the helix exclusively contains hydrophilic residues, whereas hydrophobic residues belong to the core (*right*). Hydrophobic side chains are colored in *orange* and in *yellow* (disulfide bonds). Notice that the Arg^{45}, Arg^{50} side chains belong to the helix whereas the Arg^{37}, Arg^{53} side chains belong to the upstream and downstream coils. As a result, the arginine side chains are located in two opposite parts of the globular structure. The Lys^{59} side chain is not displayed.

The three-dimensional structure of the cysteine-rich domain was submitted to DALI (41) for a search of similar structures (Z score: >2). On the basis of the values of the Z score and of the rmsd, no similar fold was proposed. Indeed, with a Z score of 0.3 and an rmsd of 2.4 Å, the frur protein (PDB code: 1uxc) was the best one (42). Their superimposition confirmed that the two structures share poor similarity. This research suggests that the Pen-3a elementary fold is not common, possibly due to the constraints imposed by three disulfide bond linkages.

The Amphipathic Character of Pen-3a—As a general feature, the proline-rich and cysteine-rich domains contain four (Lys^4, Arg^9, Arg^{13}, Arg^{19}) and five (Arg^{37}, Arg^{45}, Arg^{50}, Arg^{53}, Lys^{59}) positively charged residues, respectively (Fig. 1).

Interestingly, the helix spanning residues 41–50 exhibits marked amphipathic character (Fig. 5). The Ser^{40}, Ser^{42}, Gln^{43}, Arg^{45}, Ser^{46}, Ser^{49}, Arg^{50} hydrophilic side chains are gathered on the solvent-exposed side, whereas the Phe^{41}, Ala^{44}, and Cys^{47} side chains belong to the hydrophobic core, which in addition includes the Ile^{39}, Val^{57}, Pro^{33}, Leu^{51} side chains and the Cys^{32}–Cys^{47} and Cys^{36}–Cys^{54} disulfide bonds.

Moreover, from the well-defined structure of the cysteine-rich domain, it appears that positively charged side chains are located in two opposite regions of the globular surface. Arg^{45} and Arg^{50} belong to the solvent-exposed face of the helix, whereas Arg^{37} and Arg^{53} belong to the upstream and downstream coils. These two patches of positively charged side chains are more or less separated by hydrophobic clusters. The Arg^{53} side chain is surrounded by Leu^{51}, Cys^{48}, Cys^{55}, Cys^{36}, Val^{34}, and Tyr^{29} hydrophobic side chains. The chemical shift non-equivalency observed for the R^{53} delta protons (2.98–3.08 ppm) suggests that the extremity of this side chain is constrained in this hydrophobic environment. A similar location is also observed for the Arg^{45} side chain, which is surrounded by the Phe^{41}, Val^{57}, Ala^{44}, Cys^{55}, and Cys^{48} hydrophobic side chains. In contrast, the environment of the Arg^{37}, Arg^{50}, and Lys^{59} side chains is essentially due to side chains of polar residues. As a result, the surface of the cysteine-rich domain exhibits an amphipathic character, a feature required for antimicrobial peptides.

Interestingly, the five positively charged residues (Arg^{37}, Arg^{45}, Arg^{50}, Arg^{53}, and Lys^{59}) are conserved for the penaeidins of class 3. The alignment of all classes of penaeidins revealed that only Arg^{45} and Arg^{50} (except for the *L. setiferus* Pen-2d, where there is a phenylalanine in this position) are conserved (13). In contrast, the Arg^{37} position is replaced by a tyrosine, whereas Arg^{53} is replaced either by a serine or an aspartic acid as in Pen-2d from *L. setiferus*. This suggests that the Arg^{45} and Arg^{50} conserved cationic side chains of the helical structure are essential for antimicrobial activity.

Regarding the proline-rich domain (1–28 residues), since it is unconstrained, such a precise information about its three-dimensional structure cannot be obtained. Nevertheless, it is interesting to mention that three out of the four positively charged residues are sequentially surrounded by hydrophobic residues. This is the case for R^9 with the Tyr^7, Ala^8, and Pro^{10} side chains and for Arg^{13} with the Ile^{11}, Pro^{12}, Pro^{14}, Pro^{15}, and Pro^{16} and Arg^{19} with the Phe^{17}, Val^{18}, Pro^{20}, Leu^{21}, and Pro^{22} side chains. The alignment of all classes of penaeidins shows that two RP sequences (Arg^{13}-Pro^{14} and Arg^{19}-Pro^{20}) are well conserved. In contrast to the highly constrained structure of the cysteine-rich domain, which probably remains unchanged upon interacting with the membrane, we hypothesize that such an alternate distribution of positively and hydrophobic side chains in the proline-rich domain can be substantially modified upon interaction. In this case the positively charged side chains would initiate weak anchoring to the membrane and then the hydrophobic interaction would drive to a new conformation able to improve the anchoring. Then, the peptide would stick on the surface of the membrane or incorporate into it to form pores in accordance with its mechanism of action. The carpet-like or barrel stave mechanisms are currently proposed (43–45).

In a previous study it has been shown that the synthetic peptide corresponding to the proline-rich domain was devoid of antimicrobial activity and a possible role of membrane target-

ing was proposed (15). The presence of the proline- and glycine-rich features and the cysteine-rich feature gathered in penaeidins in unconstrained and highly constrained domains, respectively, suggests that each domain alone does not exhibit an efficient antimicrobial effect to be selected by evolution. Therefore, we hypothesize a complementary effect of the two domains, the proline-rich domain bearing the membrane-anchoring feature and the cysteine-rich domain responsible for the antimicrobial activity. Further studies are necessary to address the question of the respective involvement of the two domains in the delivery, targeting, or binding to membranes, and in the membrane permeabilization or peptide internalization character, leading to the death of microorganisms.

Acknowledgments—We thank Dr. Delphine Destoumieux-Garzon for stimulating and helpful discussions, Dr. Jérome Gracy for his help in screening the PDB and are indebted to Dr. S. L. Salhi for the editorial revision of the manuscript. Mass spectroscopy was realized using the facilities of the Proteomic Platform of the Génopole de Montpellier-Languedoc-Roussillon.

REFERENCES

1. Ganz, T. (1999) *Science* **286,** 420–421
2. Kawano, K., Yoneya, T., Miyata, T., Yoshikawa, K., Tokunaga, F., Terada, Y., and Iwanaga, S. (1990) *J. Biol. Chem.* **265,** 15365–15367
3. Mandard, N., Sodano, P., Labbe, H., Bonmatin, J. M., Bulet, P., Hetru, C., Ptak, M., and Vovelle, F. (1998) *Eur. J. Biochem.* **256,** 404–410
4. Aumelas, A., Mangoni, M., Roumestand, C., Chiche, L., Despaux, E., Grassy, G., Calas, B., and Chavanieu, A. (1996) *Eur. J. Biochem.* **237,** 575–583
5. Fahrner, R. L., Dieckmann, T., Harwig, S. S., Lehrer, R. I., Eisenberg, D., and Feigon, J. (1996) *Chem. Biol.* **3,** 543–550
6. Mandard, N., Sy, D., Maufrais, C., Bonmatin, J. M., Bulet, P., Hetru, C., and Vovelle, F. (1999) *J. Biomol. Struct. Dyn.* **17,** 367–380
7. Mandard, N., Bulet, P., Caille, A., Daffre, S., and Vovelle, F. (2002) *Eur. J. Biochem.* **269,** 1190–1198
8. Dimarcq, J. L., Bulet, P., Hetru, C., and Hoffmann, J. (1998) *Biopolymers* **47,** 465–477
9. Cornet, B., Bonmatin, J. M., Hetru, C., Hoffmann, J. A., Ptak, M., and Vovelle, F. (1995) *Structure* **3,** 435–448
10. Bachère, E., Destoumieux, D., and Bulet, P. (2000) *Aquaculture* **191,** 71–88
11. Destoumieux, D., Bulet, P., Loew, D., Van Dorsselaer, A., Rodriguez, J., and Bachère, E. (1997) *J. Biol. Chem.* **272,** 28398–28406
12. Gross, P. S., Bartlett, T. C., Browdy, C. L., Chapman, R. W., and Warr, G. W. (2001) *Dev. Comp. Immunol.* **25,** 565–577
13. Cuthbertson, B. J., Shepard, E. F., Chapman, R. W., and Gross, P. S. (2002) *Immunogenetics* **54,** 442–445
14. Rojtinnakorn, J., Hirono, I., Itami, T., Takahashi, Y., and Aoki, T. (2002) *Fish Shellfish Immunol.* **13,** 69–83
15. Destoumieux, D., Bulet, P., Strub, J. M., Van Dorsselaer, A., and Bachère, E. (1999) *Eur. J. Biochem.* **266,** 335–346
16. Destoumieux, D., Muñoz, M., Cosseau, C., Rodriguez, J., Bulet, P., Comps, M., and Bachère, E. (2000) *J. Cell Sci.* **113,** 461–469
17. Krause, A., Neitz, S., Magert, H. J., Schulz, A., Forssmann, W. G., Schulz-Knappe, P., and Adermann, K. (2000) *FEBS Lett.* **480,** 147–150
18. Nicolas, G., Bennoun, M., Devaux, I., Beaumont, C., Grandchamp, B., Kahn, A., and Vaulont, S. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 8780–8785
19. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
20. Reichhart, J. M., Petit, I., Legrain, M., Dimarcq, J. L., Keppi, E., Lecocq, J. P., Hoffmann, J. A., and Achstetter, T. (1992) *Invert. Reprod. Dev.* **21,** 15–24
21. Wilm, M., and Mann, M. (1996) *Anal. Chem.* **68,** 1–8
22. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wuthrich, K. (1998) *Eur. J. Biochem.* **256,** 1–15
23. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117,** 479–485
24. Rance, M. (1987) *J. Magn. Reson.* **74,** 557–564
25. Macura, S., Huang, Y., Sutter, D., and Ernst, R. R. (1981) *J. Magn. Reson.* **43,** 259–281
26. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson.* **85,** 393–399
27. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* **2,** 661–665
28. Pons, J. L., Malliavin, T. E., and Delsuc, M. A. (1996) *J. Biolmol. NMR* **8,** 445–452
29. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York
30. Rice, L. M., and Brunger, A. T. (1994) *Proteins* **19,** 277–290
31. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett.* **229,** 317–324
32. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
33. Frishman, D., and Argos, P. (1995) *Proteins* **23,** 566–579
34. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR* **5,** 67–81
35. Wishart, D. S., and Case, D. A. (2001) *Methods Enzymol.* **338,** 3–34
36. Srinivasan, N., Sowdhamini, R., Ramakrishnan, C., and Balaram, P. (1990) *Int. J. Pept. Protein Res.* **36,** 147–155
37. Kawabata, S., Nagayama, R., Hirata, M., Shigenaga, T., Agarwala, K. L., Saito, T., Cho, J., Nakajima, H., Takagi, T., and Iwanaga, S. (1996) *J. Biochem.* (*Tokyo*) **120,** 1253–1260
38. Muñoz, M., Vandenbulcke, F., Saulnier, D., and Bachère, E. (2002) *Eur. J. Biochem.* **269,** 2678–2689
39. Teplyakov, A., Polyakov, K., Obmolova, G., Strokopytov, B., Kuranova, I., Osterman, A., Grishin, N., Smulevitch, S., Zagnitko, O., Galperina, O. *et al.* (1992) *Eur. J. Biochem.* **208,** 281–288
40. Hunter, H. N., Fulton, D. B., Ganz, T., and Vogel, H. J. (2002) *J. Biol. Chem.* **277,** 37597–37603
41. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123–138
42. Penin, F., Geourjon, C., Montserret, R., Bockmann, A., Lesage, A., Yang, Y. S., Bonod-Bidaud, C., Cortay, J. C., Negre, D., Cozzone, A. J., and Deleage, G. (1997) *J. Mol. Biol.* **270,** 496–510
43. Shai, Y. (2002) *Biopolymers* **66,** 236–248
44. Shai, Y. (1999) *Biochim. Biophys. Acta* **1462,** 55–70
45. Oren, Z., and Shai, Y. (1998) *Biopolymers* **47,** 451–463