Characterization of a defensin from the oyster Crassostrea gigas. Recombinant production, folding, solution structure, antimicrobial activities, and gene expression.

Yannick Gueguen, Amaury Herpin, André Aumelas, Julien Garnier, Julie Fievet, Jean-Michel Escoubas, Philippe Bulet, Marcelo Gonzalez, Christophe Lelong, Pascal Favrel, et al.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 1, pp. 313–323, January 6, 2006

Characterization of a Defensin from the Oyster *Crassostrea gigas*

RECOMBINANT PRODUCTION, FOLDING, SOLUTION STRUCTURE, ANTIMICROBIAL ACTIVITIES, AND GENE EXPRESSION[*] [S]

Received for publication, October 5, 2005, and in revised form, October 20, 2005 Published, JBC Papers in Press, October 24, 2005, DOI 10.1074/jbc.M510850200

Yannick Gueguen[‡1,2], Amaury Herpin[§¶1,3], André Aumelas[‖], Julien Garnier[‡], Julie Fievet[‡], Jean-Michel Escoubas[‡4], Philippe Bulet[**], Marcelo Gonzalez[‡], Christophe Lelong[§3], Pascal Favrel[§3], and Evelyne Bachère[‡]

*From the [‡]Ifremer, CNRS, Université de Montpellier II, UMR 5171, Génome Populations Interactions Adaptation, 2 Place E. Bataillon, CC80, F-34095 Montpellier cedex 5, France, the [§]Laboratoire de Biologie et Biotechnologies Marines, Institut de Biologie Fondamentale et Appliquée, UMR 100 Ifremer, Université de Caen, Physiologie et Ecophysiologie des Mollusques Marins, 14032 Caen cedex, France, the [‖]Centre de Biochimie Structurale, CNRS UMR 5048, INSERM U554, CNRS UMR5048, Université Montpellier I, 29 rue de Navacelles, F-34090 Montpellier Cedex, France, **Atheris Laboratories, Case Postale 314, CH-1233 Bernex-Geneva, Switzerland, and the [¶]SARS International Centre for Marine Molecular Biology, High Technology Centre, 5008 Bergen, Norway*

In invertebrates, defensins were found in arthropods and in the mussels. Here, we report for the first time the identification and characterization of a defensin (*Cg*-Def) from an oyster. *Cg-def* mRNA was isolated from *Crassostrea gigas* mantle using an expressed sequence tag approach. To gain insight into potential roles of *Cg*-Def in oyster immunity, we produced the recombinant peptide in *Escherichia coli*, characterized its antimicrobial activities, determined its solution structure by NMR spectroscopy, and quantified its gene expression *in vivo* following bacterial challenge of oysters. Recombinant *Cg*-Def was active *in vitro* against Gram-positive bacteria but showed no or limited activities against Gram-negative bacteria and fungi. The activity of *Cg*-Def was retained *in vitro* at a salt concentration similar to that of seawater. The *Cg*-Def structure shares the so-called cystine-stabilized α-β motif (CS-$\alpha\beta$) with arthropod defensins but is characterized by the presence of an additional disulfide bond, as previously observed in the mussel defensin (MGD-1). Nevertheless, despite a similar global fold, the *Cg*-Def and MGD-1 structures mainly differ by the size of their loops and by the presence of two aspartic residues in *Cg*-Def. Distribution of *Cg-def* mRNA in various oyster tissues revealed that *Cg-def* is mainly expressed in mantle edge where it was detected by mass spectrometry analyses. Furthermore, we observed that the *Cg-def* messenger concentration was unchanged after bacterial challenge. Our results suggest that *Cg-def* gene is continuously expressed in the mantle and would play a key role in oyster by providing a first line of defense against pathogen colonization.

Antimicrobial peptides (AMPs)[5] are important components of the innate immune system that have been conserved during evolution (1). They constitute a first line of host defense against pathogens in plants and animals (2, 3). We estimate that more than 1000 antimicrobial peptides have been described at the level of their primary structure (2). They are gathered in the Antimicrobial Sequence Database (www.bbcm.units.it/~tossi/amsdb.html). AMPs can be classified into three major groups: (i) linear peptides that form amphipathic α-helices, (ii) cyclic peptides containing cysteine-residue engaged in disulfide bonds, and (iii) peptides with an overrepresentation in certain amino acids (proline, arginine, glycine, or histidine). Despite their great diversity in terms of size, primary structure, amino acid composition, and mode of action, most AMPs are characterized by the preponderance in cationic and hydrophobic amino acids (2). In most of the cases, this amphipathic character is considered as crucial for the interaction of the effective peptide with the membrane of sensitive microorganisms. This first interaction seems to be essential whatever the exact mode of action: (i) through disruption of their negatively charged cytoplasmic membranes or (ii) through killing following translocation into the bacteria without membrane lyses and binding to a specific target protein (4). Depending on their tissue distribution, AMPs ensure either a systemic or a local protection of the organism against pathogens.

Among the AMPs, defensins represent an important peptide family. They are abundant and widely distributed in human and animal tissues that are involved in host defense against microbial infection. Defensins are compact cationic peptides, ~3–5 kDa in size, containing three or four disulfide bridges, and are active against a wide range of bacteria and

[*] This study was part of a collaborative project supported by the European Commission, DG XII, in the program International Co-operation with Developing Countries, INCO-DC, Contract ICA4-CT-2001-10023 (IMMUNAQUA). Work in the authors' laboratories was supported in part by the Ifremer, the CNRS, the INSERM, and the University of Montpellier II and Caen. This study was also part of a French program MOREST funded by Ifremer, the Regions Basse-Normandie, Bretagne, Pays de la Loire and Poitou-Charentes and by the Conseil Général du Calvados. Atheris Laboratories was supported by the Swiss Federal Office for Education and Science contract 02.0360. 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 nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AJ565499 and CAJ19280.

The atomic coordinates and structure factors (code 2B68) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains Table S1.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: Tel.: 33-46-714-4707; Fax: 33-46-714-4622; E-mail: ygueguen@ifremer.fr.

[3] Financially supported by the Conseil Régional de Basse-Normandie, Agence de l'Eau Seine Normandie and FEDER 4474 (program PROMESSE).

[4] Present address: Ecologie microbienne des insectes et interactions hôte-pathogène, UMR INRA Université de Montpellier II, 2 Place E. Bataillon, CC80, F-34095 Montpellier cedex 5, France.

[5] The abbreviations used are: AMP, antimicrobial peptide; *Cg*-Def, *C. gigas* defensin; DG, distance geometry; DQF-COSY, two-dimensional double-quantum filter correlation spectroscopy; MALDI-TOF MS; matrix-assisted laser desorption/ionization time of flight mass spectrometry; ESI-MS, electrospray ionization mass spectrometry; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; RP-HPLC, reversed-phase high performance liquid chromatography; r.m.s.d., root mean square deviation; TOCSY, total correlation spectroscopy; RT, reverse transcriptase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; MIC, minimum inhibitory concentration; EF, elongation factor; Ifremer, Institut Francais de Recherche pour l'Exploitation de la Mer; CFU, colony forming units.

fungi (5). The vertebrate defensins can be grouped into three subfamilies, the α-defensins and β-defensins, which are distinguished on the basis of the connectivity of their six cysteine residues, and more recently the cyclic θ-defensins, (6). The α-defensins are produced constitutively and stored in neutrophils of many animals and in human Paneth cells. β-Defensins, which have been identified in many cell types, including epithelial cells and neutrophils, were reported to be inducible or constitutively expressed (3, 7, 8). In mammals, apart from their antimicrobial activities, defensins play an important role in inflammation, wound healing, and regulation of specific immunity reactions (9). In contrast with the classification of the vertebrate defensins, based on their secondary structure, the grouping in clear distinct subfamilies of the invertebrate defensins is based on their biological properties, antibacterial *versus* antifungal (2). The invertebrate defensins differ from the vertebrate defensins by their disulfide bridging, (10). Defensins are the most widespread family of invertebrate AMPs, and >70 different defensins have been isolated in arthropods (insects, ticks, spiders, and scorpions) and mollusks (2). Most of the insect defensins were isolated from the hemolymph of experimentally infected animals, whereas in scorpions, termites, and mollusks, defensins are present in granular hemocytes of non-infected animals (2). In mollusks, AMPs have only been reported in bivalves, such as in the mussels *Mytilus edulis* (11) and *Mytilus galloprovincialis* (12). Interestingly, defensins from the Mediterranean mussels have been found to display sequence homology to defensins from arthropods, even though they belong to distant phylogenetic groups. Until now, in oyster, antimicrobial activities have been detected in the hemolymph of some species, however, AMPs have never been fully characterized despite many attempts to purify, from hemolymph and other tissues, such molecules by using RP-HPLC approaches (13).

In the present study, we report the characterization of the first AMP in oyster. The *Crassostrea gigas* defensin (*Cg*-Def) mRNA has been isolated from mantle edge using the expressed sequence tag approach. *Cg*-Def displays sequence homology with members of the arthropod defensin family and shares with the defensins (MGD-1 and -2) from the mussel *M. galloprovincialis*, a fourth pair of cysteine residues. To shed light on the evolutionary conservation of defensins in invertebrates and the function of this peptide in the defense reactions of the oyster, we expressed *Cg*-Def in *Escherichia coli* and determined, using the refolded peptide, its spectrum of activity against a panel of bacteria and fungi. To draw some structure-function features, the three-dimensional structure of the *Cg*-Def was determined in aqueous solution by ^{1}H NMR spectroscopy and molecular modeling, and then compared with MGD-1, its counterpart from mussels. Finally, the *Cg*-Def gene expression was analyzed in response to experimental bacterial challenge of oysters.

EXPERIMENTAL PROCEDURES

Animals and RT-PCR Amplification of Cg-def—Adult oysters, *C. gigas*, were purchased from a local oyster farm in Normandie (France) or Palavas (Gulf of Lion, France) and kept in seawater at 15 °C. A cDNA encoding part of a putative defensin was identified by randomly sequencing clones from a *C. gigas* mantle edge cDNA library constructed in λ-ZAP II (University of Caen). Specific primers were then used to isolate the full-length *Cg-def* cDNA from the same library. Briefly, the 5′ missing end of the transcript was identified using the orientated *C. gigas* mantle edge cDNA library as template and using a gene-specific oligonucleotide (*Cg*-DefR1: 5′-ACCAGAGCGTGGCTGCATCACAG-3′) and a vector-specific oligonucleotide (T3: 5′-AATTAACCCTCACTAAAGG-3′) as primers. PCR products were analyzed by electrophoresis on 1.5% (w/v) agarose gel. Fragment of the expected size was extracted from the gel using a Qiagen kit, cloned into the pGEM-T easy vector using a TA cloning kit (Promega, Madison, WI), and sequenced.

Screening of C. gigas Genomic Library—A genomic library of *C. gigas* was constructed in λ-DASH$_{II}$ (Stratagene). A total of 1.8×10^6 independent clones was recovered. After amplification to 4.5×10^{10} plaque-forming units/ml, 2.5×10^5 recombinant λ-DASH phages were plated at 5×10^4 per dish, adsorbed to nitrocellulose membranes, and screened at high stringency with digoxigenin-11-dUTP-labeled 322-bp encoding the full-length *Cg-def* cDNA. Positive clones were purified (Qiagen λ kit, Courtaboeuf, France) and subjected to restriction analysis and Southern blot hybridization using the original probe to confirm that the λ-DASH clones contained *Cg-def*-specific sequence. Genomic organization of positive clones was subsequently determined by direct sequencing. Exon/intron boundaries were determined by comparing the genomic sequence to the corresponding cDNA one.

Determination of Cg-def Expression in Different Tissues—Total RNA was isolated from adult tissues using Tri reagent (Sigma-Aldrich) according to the manufacturer's instructions. After treatment during 20 min at 37 °C with 1 unit of DNase I (Sigma-Aldrich) to prevent genomic DNA contamination, 1 μg of total RNA was reverse transcribed using 1 μg of random hexanucleotidic primers (Promega), 0.5 mM dNTPs, and 200 units of Moloney murine leukemia virus reverse transcriptase (Promega) at 37 °C for 1 h in the appropriate buffer. The reaction was stopped by incubation at 70 °C for 10 min. Then, quantitative RT-PCR analysis was performed using the iCycler apparatus (Bio-Rad). iQ™ SYBR Green supermix PCR kit (Bio-Rad) was used for real-time monitoring of amplification (5 ng of template cDNA, 40 cycles: 95 °C/15 s, 60 °C/15 s) with the following primers: QsDef, 5′-CCACAATCACTGCAAGTCCATTA-3′ and QaDef, 5′-CTTTCCATTACAATCGGTACATG-3′ as sense and antisense primers, respectively. Accurate amplification of the target amplicon was checked by performing a melting curve. Using QsGAPDH (5′-TTCTCTTGCCCCTCTTGC-3′) and QaGAPDH (5′-CGCCCAATCCTTGTTGCTT-3′) primers, a parallel amplification of oyster GAPDH transcript (EMBL CGI548886) was carried out to normalize the expression data of the *Cg-def* transcript. The relative level of *Cg-def* expression is calculated for 100 copies of the GAPDH housekeeping gene following the formula: $n = 100 \times 2^{(\text{Ct GAPDH} - \text{Ct Cg-Def})}$.

Recombinant Expression of Cg-Def—Recombinant *Cg*-Def was expressed in *E. coli* as an N-terminal His$_6$-tagged fusion protein using the pET-28a system (Novagen, Madison, WI). By PCR amplification, a Met-coding trideoxynucleotide was incorporated 5′ of the *Cg-def* cDNA and cloned in-frame with the N-terminal His$_6$ in the EcoRI/SalI sites of pET-28a. The *Cg-def* coding cDNA sequence was amplified by PCR using forward primer DefFw1 (5′-GCGCGAATTCATGGGATTTGGGTGTCCGGG) paired with reverse primer DefRv1 (5′-ATATATGTCGACTTACTTCTTTCCATTACAATCGG). The reaction was performed by incubating the reaction mixtures at 94 °C for 4 min, followed by successive cycles at 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min for 26 cycles using Isis™ DNA polymerase (Qbiogene). The underlined codon in the forward primer denotes a Met codon to incorporate a CNBr cleavage site immediately upstream of the N terminus of the designed peptide according to the method described for recombinant mouse α-defensin production (14).

Recombinant *Cg*-Def was expressed in *E. coli* Rosetta (DE3) pLysS cells (Novagen) transformed with the pET-28a/*Cg*-Def construct. The cells were grown at 37 °C to A_{600} 0.9 in Luria-Bertani (LB) medium (10 g of bacto-Tryptone, 5 g of bacto yeast extract, and 10 g of NaCl) supplemented with 50 μg/ml kanamycin. Expression of fusion proteins was induced with 0.5 mM isopropyl-β-D-1-thiogalactopyranoside. After

growth at 37 °C for 3 h, bacterial cells were harvested by centrifugation and stored at −20 °C. The cells were lysed by resuspending bacteria pellets in 6 M guanidine HCl in 100 mM Tris-HCl, pH 8.1, followed by sonication at 40% amplitude for 2 min using a Vibra cell Sonifier 450 (Branson Ultrasons, Annemasse, France). The lysate was clarified by centrifugation in a Sorvall SA-600 rotor at 10,000 × *g* for 30 min at 4 °C prior to protein purification.

Purification and Folding of Recombinant Cg-Def—His-tagged *Cg*-Def fusion protein was purified by affinity chromatography by incubating cell lysates with nickel-nitrilotriacetic acid resin (Novagen) at a ratio of 25:1 (v/v) in 6 M guanidine HCl, 20 mM Tris-HCl (pH 8.1) for 4 h at 4 °C. Fusion proteins were eluted with two column volumes of 6 M guanidine HCl, 1 M imidazole, 20 mM Tris-HCl (pH 6.4), dialyzed against 5% acetic acid (HOAc) in SpectraPor dialysis membranes (Spectrum Laboratories Inc., Rancho Dominguez), and lyophilized. The methionine residue introduced at the *Cg*-Def N terminus was subjected to CNBr cleavage by dissolving the lyophilized His_6 fusion proteins in 50% formic acid, addition of solid CNBr to 10 mg/ml (final concentration), and incubation of the mixtures for 8 h in the darkness at 25 °C. The cleavage reaction was terminated by adding 10 volumes of water, followed by freezing and lyophilization. Then, the cleaved fusion peptide mixture was reduced in 100 mM Tris-HCl buffer at pH 8 in the presence of 100 mM dithiothreitol. The reaction mixture was purified to homogeneity using a C_8 reverse-phase high performance liquid chromatography (RP-HPLC), and the fractions of interest were lyophilized. The peptide was resolved with a 0–55% acetonitrile gradient developed over 30 min at a flow rate of 5 ml/min on a UP10C8 column (Interchrom modulocart uptisphere 10 C_8, 250 × 10 mm). The refolding of the *Cg*-Def was performed at room temperature in 100 mM Tris-HCl buffer at pH 8 during 48 h. Then, the refolded *Cg*-Def was purified to homogeneity using an additional RP-HPLC step using the same column and eluting conditions as mentioned above. Alternatively, the cleaved fusion peptide mixture was directly folded at pH 8.1 in a buffer solution containing 0.1 M $NaHCO_3$ and 3 mM reduced and 0.3 mM oxidized glutathione in the presence of 2 M urea and 25% *N*,*N*-dimethylformamide (15). Then, the folded *Cg*-Def was purified to homogeneity by RP-HPLC using a C_8 column, as described above. Peptide purity was controlled by acid-urea PAGE, and the peptide concentration was estimated by amino acid analysis and UV absorption at 280 nm based on the extinction coefficients of the molecule. Molecular masses of the purified peptides were determined using matrix-assisted laser desorption ionization mode mass spectrometry (MALDI-TOF MS) and electrospray ionization mass spectrometry (ESI-MS).

Antimicrobial Assays—Antimicrobial activity of recombinant *Cg*-Def was assayed against several bacteria, including the Gram-positive *Micrococcus lysodeikticus, Staphylococcus aureus*, *S. hemeolyticus*, *Bacillus megaterium*, *Brevibacterium stationi*, *Microbacterium maritypicum*, the Gram-negative *E. coli* 363, *Vibrio alginolyticus*, and *Salmonella typhimurium.* The activity of the peptide was also investigated against the following filamentous fungi *Fusarium oxysporum*, *Botrytis cinerea*, and *Penicillium crustosum.* Minimum inhibitory concentrations (MICs) were determined in duplicate by the liquid growth inhibition assay based on the procedure described by Hetru and Bulet (16). Poor broth (PB: 1% bacto-Tryptone, 0.5% NaCl w/v, pH 7.5) nutrient medium was used for standard bacteria, and saline peptone water (1.5% peptone, 1.5% NaCl, pH 7.2) was used for marine bacteria. Antifungal assay was performed in potato dextrose broth (Difco, Sparks, MD) at half strength supplemented with tetracycline (10 μg/ml final concentration). Growth was monitored spectrophotometrically at 620 nm on a Multiscan microplate reader colorimeter (Dynatech).

Bactericidal or bacteriostatic effect was measured by CFU counting following a 15-h incubation at 30 °C. A bactericidal kinetic was assayed with *M. lysodeikticus*. Ten microliters of purified peptide, at a concentration 10 times over the MIC value, was mixed with 90 μl of an exponential phase culture of *M. lysodeikticus* at a starting A_{600} of 0.01 in PB nutrient medium and incubated at 30 °C. Aliquots of 10 μl were plated after 0 min, 2 min, 10 min, 90 min, and 15 h of incubation on nutrient agar plates. The number of CFUs was established after a 15-h incubation at 30 °C. As a control, the bacterial culture was incubated with 10 μl of sterile water.

NMR Spectroscopy—NMR samples of *Cg*-Def were prepared either in a 95:5 H_2O:D_2O mixture (v/v) or in 99.98% D_2O to yield a 0.6–0.8 mM solutions. The pH was adjusted to the desired values by addition of deuterated HCl or NaOD and measured at room temperature with a 3-mm electrode. They are given uncorrected for the deuterium isotopic effect. Proton chemical shifts were referenced with respect to sodium 4,4-dimethyl-4-silapentane-1-sulfonate according to the IUPAC recommendations. ^{1}H NMR experiments were performed on a Bruker Avance 600 spectrometer equipped with a triple resonance cryoprobe, and spectra were recorded in the temperature range of 10–30 °C. In all experiments, the carrier frequency was set at the center of the spectrum at the water frequency. Double-quantum filtered-correlated spectroscopy (DQF-COSY) (17, 18), z-filtered total-correlated spectroscopy (z-TOCSY) (19, 20), and nuclear Overhauser effect spectroscopy (NOESY) (21) spectra were acquired in the phase-sensitive mode using the States-TPPI (time proportional phase incrementation) method (22). For spectra recorded in H_2O, the water resonance was suppressed by the WATERGATE method (23), except for the DQF-COSY spectra where a low power irradiation was used. The z-TOCSY spectra were obtained with a mixing time of 90 ms and NOESY spectra with mixing times of 100, 150, and 300 ms. Data were processed by using the XWINNMR software. The full sequential assignment was achieved using the general strategy described by Wüthrich (24). To identify the slow exchanging amide protons, the sample was dissolved in D_2O at 25 °C, and first a TOCSY (80 mn) and then an NOESY (180 mn) experiment was recorded for their identification.

Structure Calculation and Analysis—The NOESY cross-peaks were measured from two NOESY spectra (pH 3.25, at 20 and 25 °C) with a mixing time of 150 ms, and subsequently divided into five classes, according to their intensities. Very strong, strong, medium, weak, and very weak NOEs were then converted into 2.5, 3.0, 3.5, 4.0, and 5.0 Å distance constraints, respectively. The ϕ angle restraints were derived from the 3JNH-CαH coupling constants, and the χ1 angle restraints were derived from the combined analysis of the 3JHα-Hβ,β' coupling constants and intra-residues NOEs, respectively. To calculate the three-dimensional structures, distance and dihedral angle restraints were used as input in the DYANA program that uses simulated annealing combined with molecular dynamics in torsion angle space (25). In the first stage of the calculation, an initial ensemble of 20 structures 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. 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°. Finally, to calculate the *Cg*-Def three-dimensional structure, 456 NOE-derived distances and 13 dihedral constraints, including the cis conformation of the Cys^4–Pro^5 amide bond ($\omega = 0°$) and the disulfide bonds were used as input. A calculation of 60 conformers was carried out, and the resulting 10 structures with a minimum of restrained violations (no violation of >0.3 Å) were analyzed with INSIGHT 97 (Molecular Sim-

ulation Inc., San Diego, CA). The Ramachandran analysis was performed with PROCHECK (26), and the limits of the secondary structure elements and the van der Waals surfaces were determined with STRIDE (27). The chemical shifts and coordinates of *Cg*-Def are deposited in the BioMagResBank (BMRB entry: 6849) and the Protein Data Bank (PDB entry: 2B68), respectively.

Analyses of C. gigas Mantle by RP-HPLC and MALDI-TOF MS—Mantle tissue from 200 bacteria-challenged *C. gigas* oysters was collected, rapidly frozen with liquid nitrogen, and ground to fine powder. Then, the mantle sample was diluted in 10% acetic acid, homogenized using an ultra-turax, and left at 4 °C under gentle stirring for 15 h. Extracts were centrifuged at 8000 × *g* for 20 min at 4 °C, and the supernatant was prepurified by solid-phase extraction on Sep-Pak C_{18} cartridges (Waters Associates) equilibrated with acidified water (0.05% trifluoroacetic acid). After washing with acidified water, peptides were eluted with 100% acetonitrile containing 0.05% trifluoroacetic acid. The Sep-Pak fraction was then lyophilized, reconstituted in ultrapure water, and loaded onto a C_8-reversed-phase high performance liquid chromatography (RP-HPLC) on an UP10C_8 column (Interchrom modulocart uptisphere 10 C_8, 250 × 10 mm), and elution was performed using a 0–55% acetonitrile gradient developed over 30 min at a flow rate of 5 ml/min. Fractions were hand collected, lyophilized, resuspended in 100 μl of ultrapure water, and then assayed for antimicrobial activity. Finally, the active fractions were purified in a third step by C_{18} RP-HPLC, on a Symmetry Shield RP_{18} column (Waters, 5 μm, 4.6 × 250) using a 0–45% acetonitrile gradient developed over 45 min at a flow rate of 1 ml/min. The corresponding fractions were then analyzed by MALDI-TOF MS.

Bacterial Challenge and Quantification of Cg-def Gene Expression—Adult *C. gigas* oysters were stimulated by bath with heat-killed bacteria (5×10^8 bacteria/liter). The stimulation was performed with three bacterial species, *M. lysodeikticus*, *Vibrio splendidus*, and *V. anguillarum*. The bacterial strains were grown separately overnight, in saline peptone water at 25 °C for *Vibrio* strains and in Luria-Bertani (LB) medium for *M. lysodeikticus* (30 °C). Bacteria were collected by centrifugation (10,000 × *g*, 5 min) and suspended in fresh growth medium. Bacteria concentration was calculated from the optical density at 550 nm (1 unit of A_{550} corresponds to 5×10^8 bacteria/ml). Mantle tissue samples (100–300 mg) were then collected at two times post-stimulation (24 and 48 h) and washed in sterile seawater, cut into small pieces, and incubated overnight at 4 °C in TRIzol reagent (1 ml/100 mg of tissue). Total RNAs were extracted following manufacturer instructions (Invitrogen) and treated with DNase Turbo (Ambion). The experiments were done in triplicate and, to minimize individual variability, at least ten oysters were used in each experimental condition.

Following heat denaturation (70 °C for 5 min), reverse transcription was performed using 1 μg of total RNA prepared with 50 ng/μl oligo$(dT)_{12–18}$ in a 50-μl reaction volume containing 1 mM dNTPs, 1unit/μl RnaseOUT™ (Invitrogen), and 200 units/μl murine mammary tumor virus reverse transcriptase in reverse transcriptase buffer. PCR amplifications were performed with the LightCycler™ (Roche Applied Science) in the presence of SYBR-Green™ (Master SYBR Green™) with the following primers: Defm1, 5′-GATGGTTTCTGCAGACATGG-3′ and Defm2, 5′-CACAGTAGCCCGCTCTACAAC-3′ as sense and antisense primers, respectively. The gene encoding the elongation factor (EF, GenBank™ AB122066) was used as internal control. For EF the forward and reverse primers were EF (5′-ATGCACCAAGGCTGCACAGAAAG-3′) and EFR (5′-TCCGACGTATTTCTTTGCGATGT-3′), respectively. Samples were run under the following conditions: 95 °C (10 min); then 39 cycles of 95 °C (10 s), 62 °C (5 s), and 72 °C (15 s). For further expression level analysis, the crossing points were determined for each transcript using the LightCycler software. Specificity of RT-PCR product was analyzed on agarose gel and melting curve analysis. The copy ratio of each analyzed cDNA was determined as the mean of three replicates. The relative expression ratio of *Cg-def* was calculated based on the crossing points deviation of each RT-PCR product of RNA extracted from stimulated oyster *versus* the appropriate control sample and expressed in comparison to the reference gene *EF*. The relative expression ratio of *Cg*-Def was calculated based on the delta-delta method for comparing relative expression results (28).

RESULTS

Cg-def cDNA Cloning and Primary Structure Analyses—The complete cDNA of *C. gigas* defensin (*Cg*-Def) was obtained by 5′ amplification of cDNA ends-PCR on a *C. gigas* mantle edge cDNA library, as described under "Experimental Procedures" (GenBank™ AJ565499). The oyster *Cg-def* cDNA contained 323 bp, comprising a coding region of 195-bp, a 110-bp 3′ untranslated region containing a polyadenylation consensus sequence (AATAAA) at position 66–72 of the 3′UTR, and a poly(A) tail. The 195-bp coding region encoded a 65-amino acid propeptide (GenBank™ CAJ19280). The deduced amino acid sequence starts with a 22-residue signal peptide, and the cleavage site for signal peptidase is most likely located after the alanine residue preceding the glycine in position 23 as predicted by SignalP 3.0 software (data not shown). *Cg*-Def is not synthesized as a precursor with a C-terminal extension nor an N-terminal pro-region as observed with the Mediterranean mussel or the *Drosophila* defensins, respectively. The amino acid sequence of the mature peptide was aligned with other defensins from the "arthropod defensin family" available in GenBank™ and ExPASy, that contains defensins from arthropods, mollusks, and fungi (Fig. 1, *A* and *B*). As already observed for the mussel defensins MGD-1 and MGD-2, *Cg*-Def is an original member of this family due to the presence of two extra cysteine residues. *Cg*-Def shares 50% of identity with MGD-1 from *M. galloprovincialis*, but has a unique doublet of lysine residue at the C-terminal part of the molecule (Fig. 1*A*). Less identity was shared with defensin from the scorpion *Androctonus australis* (46%) and the tick *Ixodes scapularis* (44%). The selected members of the "arthropod defensin family" belonging to the so-called primitive defensin family show a high degree of conservation with both *Cg*-Def and the different mussel defensins (Fig. 1*A*) (29). Interestingly, very recently, a 40-amino acid AMP named plectasin (GenBank™ number CAI83768) with marked homologies to the arthropod and mollusk defensins has been isolated from the saprophytic fungus *Pseudoplectania nigrella*.

Genomic Organization and Analysis of the Promoter Region—*Cg-def* gene (EMBL AM050547) harbors a unique intron whose splice junction follows the AG-GT rule (Fig. 1*C*). The genomic organization of *Cg-def* is similar to that of the mussel and scorpion defensin genes (29). It also displays the features of the other arthropod and mollusk defensin genes studied so far that demonstrate a genetic relatedness: (i) the exon encoding the mature defensin is never split by any intron, (ii) the intron flanking the exon encoding the mature defensin shows for all genes a strict level of phase conservation (phase 1) (29). Analysis of the 3.6-kb genomic DNA sequence upstream from the putative ATG start codon identified several consensus sequences for transcription factors commonly observed in promoters of cellular housekeeping genes (*Sp1*), genes expressed during early embryogenesis (*AP2*, *Gata1*, and *Gata2*), development (*Zeste*), cell cycle regulation (*H1TF2*, *H4TF1*, and *H4TF2*), and cell differentiation (*Gata2*, *Ik1*, *Ik2*, *Ik3*, *IRF2*, and *Oct1*, *-2*, *-4*, *-6*, and *-9*) and genes expressed in a tissue-specific manner (*Pit-1* and *SRF*).

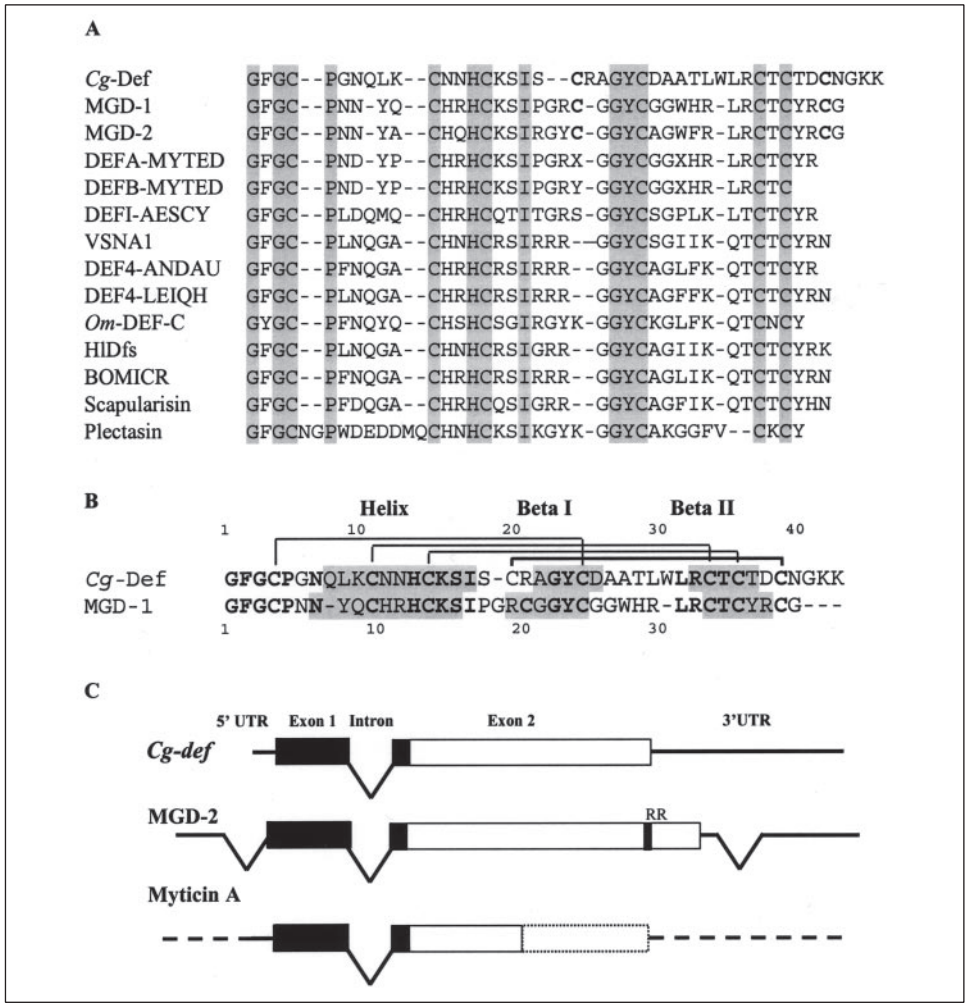

FIGURE 1. **Sequence alignment and genomic organization of representative of the defensin family AMP.** *A*, amino acid sequence alignment of defensins from arthropods, mollusks, and a fungus. Oyster *C. gigas* (*Cg*-Def); mussels *M. edulis* (DEFA-MYTED (P81610) and DEFB-MYTED (P8161)) and *M. galloprovincialis* (MGD-1 (P80571) and MGD-2 (AAD52660)); tick *Dermacentor variabilis* (varisin, VSNA1 (AAO24323)), *Ornithodoros moubata* (*Om*-DEF-C (AB041814)), *Hemaphysalis longicornis* (HlDfs (BAD93183)), *Boophilus microplus* (BOMICR (AAO48943)), and *Ixodes scapularis* (IXSCA, scapularisin (AAV74387)); dragonfly *Aeschna cyanea* (DEFI-AESCY (P80154)); scorpion *Androctonus australis* (DEF4-ANDAU (P56686)) and *Leiurus quinquestriatus* (DEF4-LEIQH (P41965)); and the saprophytic fungus *Pseudoplectania nigrella* (plectasin (CAI83768)). The nomenclature derived from the SwissProt data base and the accession numbers are in parentheses. Identical residues are *shaded*. The two additional cysteine residues found in the *Cg*-Def, MGD-1 and MGD-2 are in *bold*. *B*, *Cg*-Def and MGD-1 sequence alignment manually modified according to the structure comparison (this work). The similar elements of structure are *shaded gray* and reported above. Conserved residues are in *bold*. Notice the shift for one cysteine (Cys^{20}/Cys^{21}) out of the eight. *C*, organization of *Cg-def* gene and the related genes coding MGD-2 and myticin A from the mussel *Mytilus galloprovincialis*. Boxes represent the coding region; the signal sequence appears in *black*, and the mature peptide is in *white*. *Gray lines* represent the untranslated regions of the exons. Potential sites of cleavage (mono or dibasic (RR) and (KK)) involved in the splitting of the mature peptide (*white*) from the C-terminal peptide (*spotted*) are indicated. Accession numbers: *Cg*-Def, CAJ19280; *Cg*-Def genomic sequence, AJ582630. Gene organization of MGD-2b is from Mitta *et al.* (39), and incomplete structure of Myticin A (*dotted box*) is deduced from Mitta *et al.* (51).

Production, Folding, and Purification of Recombinant Cg-Def—To produce sufficient amounts of *Cg*-Def for detailed studies (5-mg scale) of its biological activity and its three-dimensional structure, *Cg*-Def was expressed in *E. coli* Rosetta (DE3) pLysS cells transformed with the pET-28a/*Cg*-Def construct. Recombinant His_6-tagged fusion protein was purified by affinity chromatography from the bacteria cell lysates. After chemical cleavage with CNBr and refolding, *Cg*-Def was purified by RP-HPLC, and its purity was judged by analytical RP-HPLC, MALDI-TOF, and ESI-MS mass spectrometry analyses (Fig. 2, *A* and *B*). The oxidative folding of *Cg*-Def was assessed under two different conditions. First, after complete reduction with dithiothreitol, purified denatured *Cg*-Def was refolded during 48 h at room temperature in the presence of a Tris-HCl, 100 mM, pH 8, buffer. The refolding process was followed by RP-HPLC revealing at 48 h an additional more hydrophobic peak eluting ~35 min (Fig. 2*A*, *asterisk*). By ESI-MS analyses, this hydrophobic fraction was found to contain a molecule with a molecular mass of 4,634.10 Da (see Fig. 2*B1*). The value measured by ESI-MS is in perfect agreement with the calculated average molecular mass of 4,634.34 Da. This suggests that the recombinant *Cg*-Def has its eight cysteine residues paired. Alternatively, as described by Wu *et al.* (15), the use of *N*,*N*-dimethylformamide and urea was shown to enhance 5-times the folding efficiency of *Cg*-Def (data not shown). Using this protocol, the purified recombinant *Cg*-Def was also found to have a molecular mass measured by ESI-MS that is in perfect agreement to the calculated average molecular mass (see Fig. 2*B1*). The purified refolded peptide was then submitted to antimicrobial assays and NMR spectroscopy studies.

Antimicrobial Activity of Cg-Def—The antimicrobial activity of the recombinant *Cg*-Def was determined against a panel of microorganisms, including Gram-positive and Gram-negative bacteria and filamentous fungi. The MIC values obtained are reported in Table 1. The peptide was active at very low concentration against most of the Gram-positive bacteria tested, including *M. lysodeikticus*, *S. aureus*, and the marine bacteria *Brevibacterium stationis* and *Microbacterium maritypicum*. However, *Cg*-Def was not active at 20 μM against *B. megaterium*. *Cg*-Def showed no activity at 20 μM against the Gram-negative bacteria *V. alginolyticus* and *S. typhimurium*. At higher concentration (35 μM), the peptide was active against *E. coli*. *Cg*-Def displayed antifungal activity against *F. oxysporum* at relatively high concentrations (9 μM; Table 1). Experiments were conducted to examine the bactericidal effects of *Cg*-Def. The bactericidal activity of the peptide was assessed by plating cultures, and the number of CFUs was counted after overnight incubation at 30 °C. *Cg*-Def exerted bactericidal effects against all the Gram-positive bacteria tested, excepted *M. maritypicum* (Table 1). When *M. lysodeikticus* was incubated with *Cg*-Def at concentration 10 times higher than the MIC value, all bacteria were killed in few minutes (Table 2).

To determine if the high salinity of the seawater might modify the efficacy of *Cg*-Def *in vivo*, the effect of the peptide on bacterial growth was tested *in vitro* at NaCl concentrations ranging from 85 mM to 1 M. *Cg*-Def is highly active against *M. lysodeikticus* and *B. stationis* even at 600 mM NaCl, a concentration closed to the value measured in seawater. Interestingly, the MIC value is unchanged throughout the NaCl con-

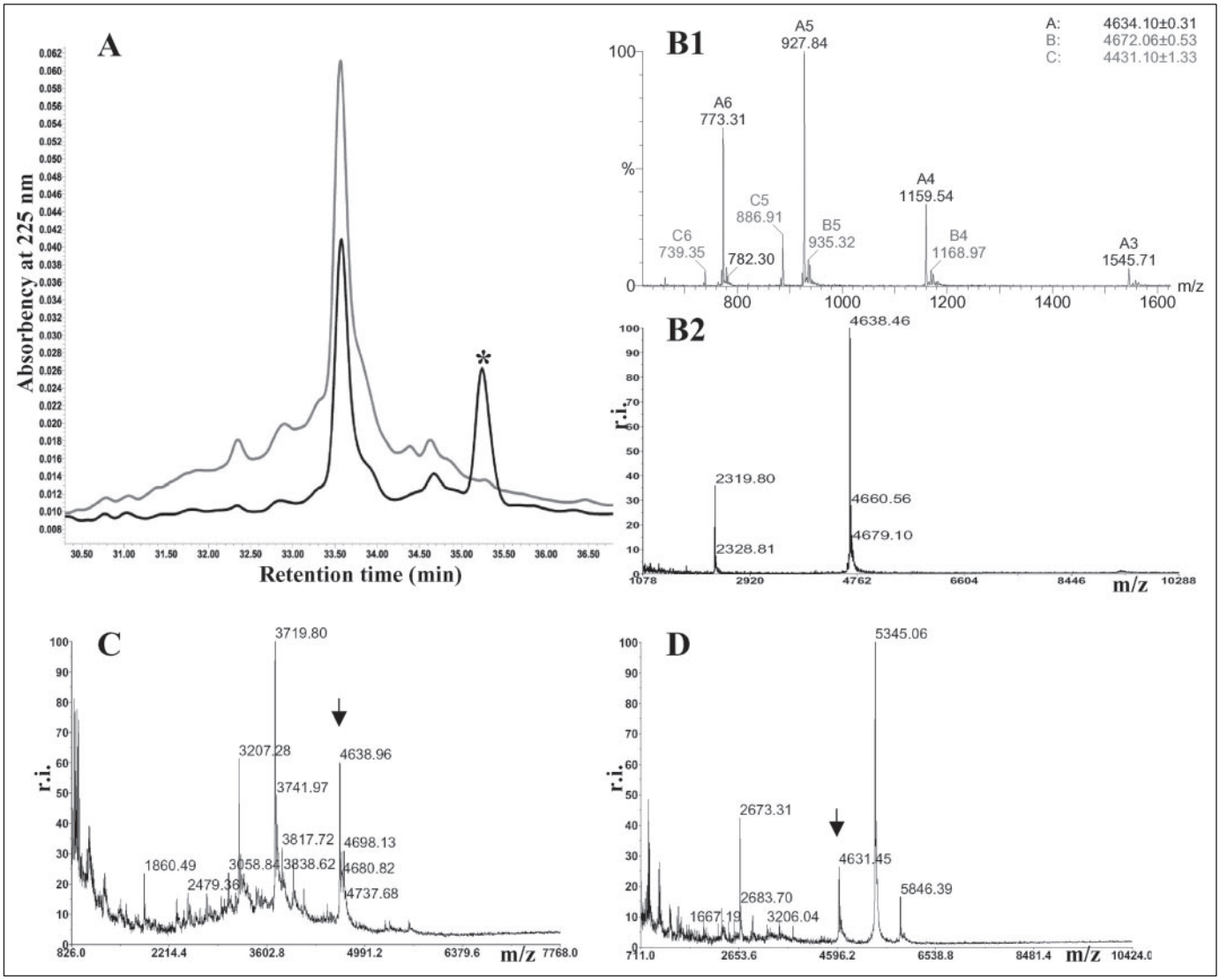

FIGURE 2. **Quality control of recombinant *C. gigas* defensin (*Cg*-Def) and detection of the natural form in acidic extracts of mantle tissue.** *A*, following reduction, refolding of the recombinant peptide was performed at room temperature in the presence of Tris-HCl buffer. The refolding was monitored by RP-HPLC using an acetonitrile gradient in acidified water (for details see "Experimental Procedures"). The *gray* and *black lines* correspond to the linear and to the refolded (*asterisk*) recombinant peptides, respectively. The molecular mass of the refolded peptide was measured by ESI-MS (*B1*) and MALDI-TOF MS (*B2*). The molecular mass (average value) measured at 4,634.10 Da is in perfect agreement with the calculated average molecular mass at 4,634.34 Da. A difference of 3 Da was observed between the molecular masses measured using MALDI-TOF MS (*m/z* 4638.46, see *B2*) *versus* the one detected in ESI-MS at 4,634.10 Da (*m/z* 4,635.10). *C* and *D*, following purification by RP-HPLC of an acidic extract of mantle tissue, the two fractions active against the tested microorganisms were analyzed by MALDI-TOF MS. The first fraction (see the *arrowhead* in *C*) eluting at the same retention time as the recombinant-refolded *Cg*-Def contains ions of *m/z* 4,638.96 corresponding to the one measured by MALDI-MS for the recombinant peptide. In the second bioactive fraction (see the *arrowhead* in *panel D*), ions of *m/z* 4,631.45 were detected.

centration range tested (85–600 mM) (0.01 μM for *M. lysodeikticus* and 0.2 μM for *B. stationis*; Table 3). At 1 M NaCl, *M. lysodeikticus* did not grow in the control, and a low bacterial growth can be observed for the marine bacteria *B. stationis*. In contrast to most AMPs, *Cg*-Def seems to conserve its antibacterial activity at such a high salt concentration (30, 31).

NMR Structural Study—The two-dimensional NMR spectra (TOCSY, DQF-COSY, and NOESY) of *Cg*-Def were recorded at different temperatures, ranging from 17 to 32 °C. The identification of all the spin systems of *Cg*-Def was obtained by analysis and comparison of DQF-COSY, TOCSY, and NOESY spectra according to the strategy described by Wüthrich (24). Nevertheless, we noticed that the two-dimensional spectra showed, for each residue of the D38CNGK42 sequence, two spin systems approximately equally populated. This suggests a heterogeneity for this C-terminal part inferred to result from the partial deamidation to yield the Asp^{40} and $isoAsp^{40}$ mixture. The Asn-Gly sequence (Asn^{40}-Gly^{41}) has been shown to experience the highest propensity for the deamidation process through the succinimide intermediate (32). Indeed, such a deamination process is known to occur and depend upon primary sequence, three-dimensional structure, and solution properties such as pH, temperature, ionic strength, and buffer ions (33). The two spin systems were assigned to the initial molecule (Asn^{40}) and to the Asp^{40} analogue. These two similar spin systems were identified by recording TOCSY spectra at several pH values in the 2–5 pH range to monitor the ionization state change of the carboxyl group. The spin system sensitive to the ionization state change was assigned to Asp^{40}. The other one, unaffected in this pH range, was assigned to Asn^{40}. Only a small amount of the third expected compound ($isoAsp^{40}$) was formed. It was detected in the TOCSY experiment but not fully characterized in the NOESY.

Chemical shifts of the recombinant *Cg*-Def are reported in the Table S1 provided as supplementary materials. Both amide and α protons

TABLE 1

Antimicrobial activity of recombinant *Cg*-Def

MIC values are expressed as the interval of concentration $[a] - [b]$, where $[a]$ is the highest concentration tested at which microbial growth can be observed, and $[b]$ is the lowest concentration that causes 100% growth inhibition (50).

Microorganisms	MIC	Bactericidal effect of *Cg*-Def
	μM	
Bacteria		
Gram-positive		
M. lysodeikticus	0.005–0.01	Yes
S. aureus	1.25–2.5	Yes
B. stationis	0.1–0.2	Yes
M. maritypicum	0.5–1	No
B. megaterium	>20	
Gram-negative		
E. coli 363	35–20	ND[a]
V. alginolyticus	>20	
S. typhimurium	>20	
Filamentous fungi		
F. oxysporum	9–4.5	
B. cinerea	>20	
P. crustosum	>20	

[a] ND, not determined.

TABLE 2

Kinetic of killing of *Cg*-Def on *M. lysodeikticus*

Cg-Def at a final concentration (0.1 μM) 10 times higher than the MIC value or water (control) was added to an exponential growth phase culture of *M. lysodeikticus* in phosphate-buffered saline medium. Aliquots were removed at various times, and the number of CFUs/ml (10^{-4} CFU.ml^{-1}) was determined after incubation of the LB agar plates during overnight at 37 ° C.

Time of incubation	Control	*Cg*-Def
	CFU	
0 min	40.7	40.7
2 min	41.6	0
10 min	49.2	0
90 min	66.4	0
15 h	ND[a]	0

[a] ND, not determined.

resonances were spread in a large range of chemical shifts, and some atypical values were measured. When compared with statistical values, among the α proton resonances, those of Ala[27] (1.87 ppm) and of Trp[31] (3.17 ppm) were upfield shifted by ~2.5 and 1.5 ppm, respectively. The large spreading of amide signals is illustrated by the two extreme values measured for Ala[27] (9.84 ppm) and for Ile[18] (6.91 ppm). Altogether, the large spreadings of chemical shifts suggest a highly constrained structure, in agreement with the presence of the four disulfide bonds. This is also supported by several large non-equivalencies such as for the Gly[3] and Gly[23] α protons resonances for which non-equivalencies of 1.24 and 1.29 ppm were measured, respectively. Similarly, non-equivalencies of 1.21 and 1.02 ppm were also observed for ββ′ signals of Cys[4] and Cys[11], respectively. The large difference of chemical shifts measured between the α (at 4.14 ppm) and the β (at 3.08 ppm) protons of Thr[35] residue and in a lesser extent for Thr[37] (at 4.61 and 4.05 ppm) also suggests that they belong to highly constrained parts of the molecule. By comparison, chemical shifts of Thr[29] (at 4.45 and 4.52 ppm) were very close to that usually observed for a less constrained sequence. The NOESY data showed several d_{NN} and $d_{\alpha\alpha}$ NOEs suggesting that the molecule both contains a helical and a β-stranded parts. Moreover, the strong d_{α}Cys[4]-$_{\alpha}$Pro[5] NOE and the absence of the d_{α}Cys[4]-$_{\delta\delta'}$Pro[5] NOE unambiguously indicated a cis conformation for the Cys[4]–Pro[5] amide bond. Among the amide protons found in slow exchange, eight of them were successive and span the 13–21 sequence. The amide protons of Ala[22], Thr[35], Cys[36], and Thr[37] residues were also found in slow exchange.

Solution Structure of Cg-Def—A stereoview of a family of 10 *Cg*-Def conformers is reported in Fig. 3. The global structure of *Cg*-Def was well

TABLE 3

Effect of sodium chloride on *Cg*-Def antimicrobial activity

MIC values were determined on the bacterial strains *M. lysodeikticus* and *B. stationis* using liquid growth assay. In this table, only the lowest concentration of *Cg*-Def that causes 100% growth inhibition is indicated. Final A_{600} of control *M. lysodeikticus* and *B. stationis* cultures without *Cg*-Def in the same corresponding medium is indicated in parentheses.

Final NaCl concentration in poor broth	MIC in (A_{600} after 24 h in the control experiment without peptide)	
	M. lysodeikticus	*B. stationis*
	μM	
85 mM[a]	0.01 (0.6)	0.2 (0.55)
300 mM	0.01 (0.4)	0.2 (0.5)
600 mM[b]	0.01 (0.15)	0.2 (0.35)
1000 mM	-[c]	0.06 (0.1)

[a] Control poor broth.
[b] NaCl concentration close to the value measured in seawater.
[c] No bacterial growth in the control experiment

defined and displays the cystine stabilized-α-β motif that consists of an helical structure and two β strands cross-linked by three disulfide bonds (34) and sometimes by a fourth disulfide bond (35–37). The backbone atoms of the ten conformers were overlaid for the well defined regions spanning residues 4–39 resulting in a pairwise average r.m.s.d. of 0.43 ± 0.09 and 0.98 ± 0.16 Å for the backbone and heavy atoms, respectively. The Ramachandran plot of all residues (except for the glycine and proline residues) of the ten best conformers indicated that 89.7% were located in the most favored and the additional allowed regions, and 5.8% in the generously allowed regions. A total of 4.4% of the residues were found in the disallowed regions. Limits of secondary structure elements indicated that the *Cg*-Def structure mainly consisted of a helical part (8–18 residues), two β-strands (residues 22–26 and 33–37), and three loops (residues 1–7, 19–21, and 27–32 for loops L1, L2, and L3, respectively) (Figs. 1*B* and 3). Three successive type IV turns were identified in L1 (1–4, 3–6, and 4–7) and L3 (27–30, 28–31, and 29–32) loops, and a type I turn was found for the 26–29 sequence. It is interesting to note that the Cys[4]–Pro[5] amide bond adopted the unusual cis conformation. Finally, the global fold is highly constrained by the 4–25, 11–34, 15–36, and 20–39 disulfide bonds. The *Cg*-Def helix exhibits an amphiphilic character with a hydrophobic side contributing to the hydrophobic core of the molecule and a hydrophilic side accessible to the solvent, including mainly, Lys[10], Asn[13], His[14], Lys[16], and Ser[17] residues. Nevertheless, the Leu[9] and Ile[18] hydrophobic residues are solvent-exposed. The hydrophobic core mainly includes the four disulfide bonds, the Ala[22] side chain and is extended to the L1 (Pro[5]) and L3 hydrophobic loops (Ala[27], Ala[28], Leu[30], Trp[31], and Leu[32]). As expected, the positively (Lys[10], Lys[16], Arg[21], Arg[33], Lys[42], and Lys[43]) and negatively charged (Asp[26] and Asp[38]) residues are solvent-exposed.

The *Cg*-Def structure was compared with that of MGD-1 (PDB entry: 1FJN) whose three-dimensional structure was previously determined (37). Their superimposition showed that the helical and β-strands elements were very conserved (Fig. 4).

Cg-def mRNA and Peptide Tissue Localization—To investigate the tissue distribution of *Cg-def* expression, we analyzed by real-time RT-PCR mRNA content in various oyster tissues. Low or no *Cg-def* mRNA amount was measured in most of the organs analyzed, including hemocytes, heart, digestive gland, and gills, whereas a high mRNA level, ~20- to 60-fold the level measured in the remaining animal tissues, was detected in mantle (Fig. 5). In non-stimulated *C. gigas* oysters, the mantle appears as the main tissue expressing *Cg-def* transcripts.

To demonstrate the presence of native *Cg*-Def in mantle, an acidic extract of this tissue was prepared, subjected first to RP-HPLC and then to antibacterial activity screening. Following solid-phase extraction of the acidified extracts from oyster mantle, RP-HPLC

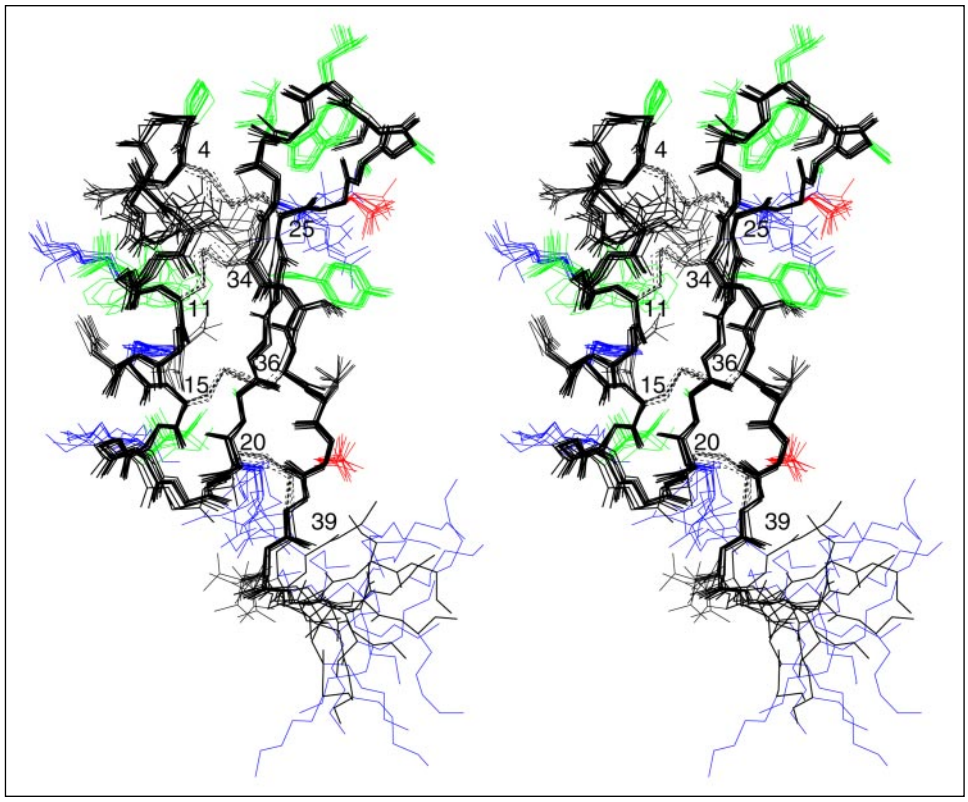

FIGURE 3. **Structure of *Cg*-Def.** Stereoview of the ten best conformers of *Cg*-Def. The 4–39 heavy atoms of the backbone were used for the superimposition. The mean pairwise r.m.s.d. is of 0.43 ± 0.09 Å. The four disulfide bridges are labeled and displayed as *dashed lines*. Hydrophobic, positively, and negatively charged side chains are colored in *green*, *blue*, and *red*, respectively.

revealed the presence of two fractions with antimicrobial activity. These two fractions were subjected to molecular mass fingerprint analysis by MALDI-TOF MS. To ascertain the molecular mass measured, the well folded recombinant *Cg*-Def was used as external calibrant (see Fig. 2*B2*). A difference of 3 Da was observed between the molecular mass measured by ESI-MS (4,634.10 Da, *panel B1*) and the one measured by MALDI-TOF MS (4,637.46 Da, *arrow* in *panel B2*). The less hydrophobic fraction (see Fig. 2*C*) contains a component with a molecular mass at 4637.96 corresponding to that measured by MALDI-TOF MS for the recombinant folded *Cg*-Def. This peptide with antimicrobial activity against *M. lysodeyticus* might likely correspond to the native form of *Cg*-Def. Interestingly, the second active fraction (more hydrophobic), that has not yet been identified, showed a close but different molecular mass of 4631.45 Da (see Fig. 2, *arrow* in *panel D*).

Real-Time PCR Analyses of Cg-def Transcript Levels after Oyster Bacterial Challenges—To determine the expression pattern of *Cg-def* during bacterial challenge, two batches of oysters were selected. In the first one, oysters were stimulated by bath with killed bacteria (see "Experimental Procedures") and in the second one, non-stimulated oysters were used as control. Then, quantitative RT-PCR analyses were performed with total RNA extract from mantle collected at two times post challenge (24 and 48 h). No striking discrepancies were observed for *Cg-def* expression between naive and bacteria challenged oysters at the different times analyzed (Fig. 6). These preliminary analyses revealed that *Cg-def* mRNA are present in naive *C. gigas* oyster and that the level of transcript was unaffected by the bacterial-challenge performed.

DISCUSSION

Efficient host defense mechanisms are needed to neutralize microbial invasions to which living organisms are exposed. In higher vertebrates, innate and adaptive immunity are present. In contrast, invertebrate and plant defense against pathogens takes place exclusively through mechanisms that are part of the innate immunity. The involvement of AMPs in natural resistance to infection is sustained by their strategic location in phagocytes, in body fluids, and at the epithelial level, *i.e.* at interfaces between the organisms and its environment. In oyster, even if antimicrobial activities have been detected in the hemolymph of some species (38), no AMP has been fully characterized despite many attempts to purify such molecules by biochemical approaches from hemolymph and other tissues (13). In this report, and for the first time in an oyster, we describe the characterization of an AMP isolated from mantle tissue. Using molecular approaches, a new member of the widespread defensin family, which is present in animal and plant kingdoms, was identified. Usually, three disulfide bonds characterize this defensin family (Fig. 1*A*). The amino acid sequence of the mature *C. gigas* defensin displays interesting homologies with defensins from the "arthropod family," including the mussel's defensins. Moreover, as observed with the defensins MGD-1, MGD-2, and MGD-2b from *M. galloprovincialis*, *Cg*-Def has four disulfide bonds. This additional bridge was proposed to render the peptide more stable in high osmolarity media such as in seawater (37). However, this fact is not common to all marine mollusk defensins, because the mussel *M. edulis* defensins do not bear this additional disulfide bond (Fig. 1*A*) (11). Interestingly, *Cg*-Def possesses an N-terminal pre-segment presumed to be a signal sequence for translocation to the lumen of the rough endoplasmic reticulum but has neither a C-terminal extension nor an N-terminal pro-region as observed in the mussel or the *Drosophila* defensins, respectively (39). This property would suggest a different way of processing for the oyster defensin.

To investigate its biological properties, *Cg*-Def was produced in *E. coli* and refolded *in vitro*. This recombinant approach was efficient, yielding enough material (~1 mg of well folded pure peptide per liter of culture) for functional assays and structural characterization. A challenging aspect for the production of *Cg*-Def in *E. coli* was the presence of

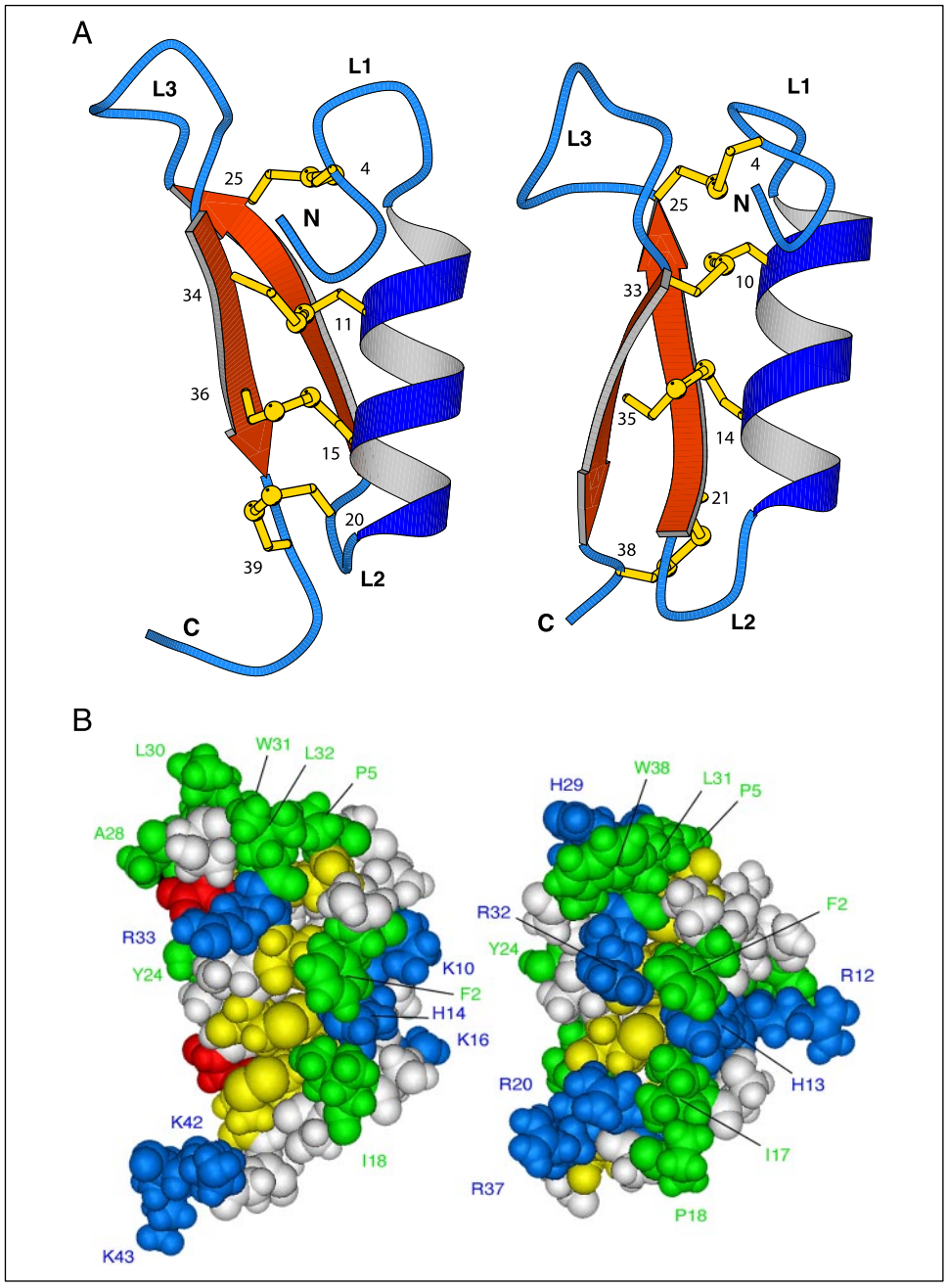

FIGURE 4. **Structure comparison of *Cg*-Def and MGD-1.** *A*, global -fold comparison of the *Cg*-Def (*left*) and MGD-1 (*right*) structures. The structures prepared using MOLSCRIPT show the arrangement of the four disulfide bonds (in *yellow*). *B*, with a similar orientation, comparison of their Van der Waals surface showing the location of the hydrophobic (*green*) and positively charged (*blue*) residues. For this orientation, the Trp^{31} side chain of *Cg*-Def is hidden. Most of the labeled residues are conserved in the two peptides. Aspartic and cysteine residues are in *red* and *yellow*, respectively.

eight cysteine residues in the mature peptide. Indeed, the correct refolding of proteins with several disulfide bonds is a general problem. As discussed by Harder *et al.*, only few reports describe the production of AMP in bacteria, a fact that reflects the difficulty of expressing and folding such cysteine rich molecules in a bacterial host system (6). Previous reports described the production of human α- and β-defensins in *E. coli* (6, 40). We report here for the first time the efficient production in a bacterial host of an AMP with four cysteine bridges that can be properly refolded *in vitro*. The activity spectrum of the recombinant *Cg*-Def was evaluated against a panel of bacteria and fungi. Consistent with studies on invertebrate defensins (2), *Cg*-Def was active *in vitro* against Gram-positive bacteria but showed no or limited activities against Gram-negative bacteria and fungi. Most interestingly, *Cg*-Def kept its antibacterial properties in the presence of high NaCl concentrations (up to 1 M). This favors the hypothesis that *Cg*-Def retains its activity in *in vivo* conditions (seawater) and would play a role in the antimicrobial defense of the oyster *C. gigas*. For these reasons, *Cg*-Def represents a model of choice for structure-activity relationship studies to design a lead antibacterial drug to treat infections of bacteria in a salt-rich environment (41). This is particularly interesting for cystic fibrosis, for which it has been suggested that the primary genetic defect increases the salt content of fluid lining the airway, which reversibly inactivates antimicrobial molecules (42, 43).

As expected from the disulfide bridges array, the global fold of *Cg*-Def includes the CS-$\alpha\beta$ motif observed in the defensins isolated from mussels (37), insects (34), and plants (44). The *Cg*-Def and MGD-1 sequences were aligned, and their structures were compared (Figs. 1*B* and 4). This alignment revealed 50% of identity (23 residues with three gaps), including eight cysteines (45). *Cg*-Def and MGD-1 share the fourth disulfide bond that contributes to the stability and to the rigidity of their three-dimensional structure. The identical residues were mainly gathered in four segments (residues 1–5, 14–18, 23–25, and 32–36).

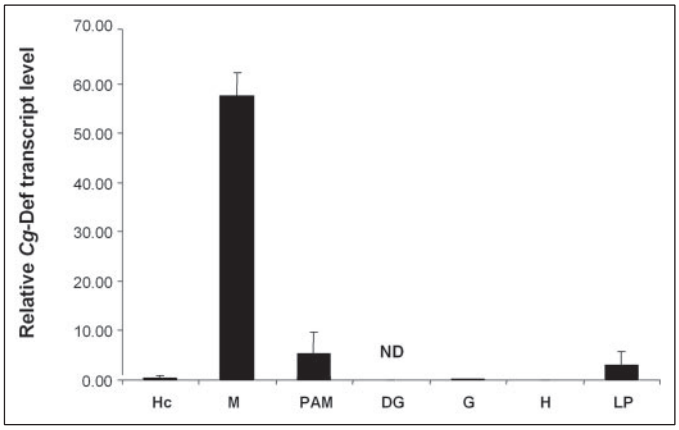

FIGURE 5. **Tissue expression of *Cg-def* mRNA analyzed by real-time quantitative RT-PCR.** Each value is the mean of four pools of four animals. *Hc*, hemocytes; *M*, mantle; *PAM*, posterior adductor muscle; *DG*, digestive gland; *G*, gills; *H*, heart; *LP*, labial palps. *ND*, not detected. *Bars* represent the relative *Cg-def* transcript levels normalized to GAPDH transcript levels.

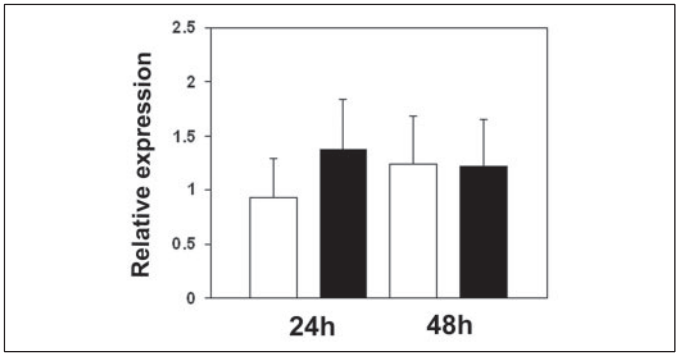

FIGURE 6. ***Cg-def* mRNA expression in oyster mantle following a bacterial challenge.** Results are means ± S.E. of three independent experiments realized on a pool of ten oysters from non-stimulated (*white*) and stimulated (*black*) oysters, at 24 and 48 h post-infection. *Bars* represent the relative *Cg-def* transcript levels normalized to elongation factor transcript levels.

Beyond their sequence comparison the *Cg*-Def structure was compared with that of MGD-1 (PDB entry: 1FJN) (37). Their superimposition showed that the helical and β-strand elements were rather conserved and the best fit was obtained for the backbone superimposition of the residues 1–4, 7–17, 21–25, and 34–36 of *Cg*-Def with residues 1–4, 6–16, 21–25, and 33–35 of MGD-1 (23 residues). An r.m.s.d. of 1.04 Å was measured for the backbone. Moreover, the two structures share the Cys^4–Pro^5 *cis*-amide bond and a similar disulfide bond pattern. Nevertheless, some differences were observed for the Cys^{20}–Cys^{39} disulfide bond and for the length of the three loops. Although the three equivalent disulfide bonds (4–25, 11–34, and 15–36 for *Cg*-Def and 4–25, 10–33, and 14–35 for MGD-1) well overlaid, the fourth one (Cys^{20}–Cys^{39} for *Cg*-Def and Cys^{21}–Cys^{38} for MGD-1) is significantly different mainly due to the translation of Cys^{20} by one residue toward the center of the molecule. This is certainly due to the shorter L2 loop (one residue less) joining the helix and the first β-strand. The consequence is a shift of this disulfide bridge that locates on the opposite side of the β-sheet ("*above*" in *Cg*-Def and "*below*" in MGD-1 as displayed in Fig. 4). Conversely, the L1 and L3 loops of *Cg*-Def display one additional residue. Whereas the sequences of the L1 loops, including the Cys^4–Pro^5 *cis*-amide bond are well conserved (Fig. 4), the sequences of the L3 loops are totally different. This latter is much more hydrophobic (^{27}AATLWL^{32} for *Cg*-Def) than that of MGD-1 (^{27}GWHRL^{31}) that contains two positively charged residues. Nevertheless, due to the Trp positions in the sequence (at the end for *Cg*-Def and at the beginning for MGD-1), it appeared that they have opposite location with regard to the loop. In contrast, the Leu^{32}(*Cg*-Def)/Leu^{31}(MGD-1) side chains were similarly located (Fig. 4*B*). Because it has been shown that the L3 loop of MGD-1 is responsible for a large part of the antibacterial activity (46), this significant structural difference is worth being noted.

It is commonly admitted that the surface distribution of hydrophobic and hydrophilic side chains in AMPs is essential for their antimicrobial activity. The alignment showed that several of them (Phe^2, Pro^5, Ile^{18}, Tyr^{24}, Leu^{32}, and His^{14}, Lys^{16}, and Arg^{33}) were conserved (Fig. 1 and see below). However, the two antimicrobial sequences mainly differ by the presence of two aspartic residues (Asp^{26} and Asp^{38}) and by the addition of a dibasic peptide (Lys^{42} and Lys^{43}) at the C terminus. Although, the *Cg*-Def and MGD-1 surfaces display significant differences mainly due to the two aspartic acids, the hydrophobic L3 loop and, by the "above"/"below" locations of the Cys^{20}–Cys^{39}/Cys^{21}–Cys^{38} disulfide bonds, several hydrophobic ($Phe^{2/2}$, $Pro^{5/5}$, Leu^9/Tyr^8, $Ile^{18/17}$, $Tyr^{24/24}$, $Leu^{32/31}$, and charged residues (Lys^{10}/Arg^{12}, $His^{14/13}$, Arg^{21}/Lys^{15}, $Arg^{33/32}$, Lys^{42}/Arg^{20}, and Lys^{43}/Arg^{37}) are similarly located at the surface giving rise to a comparable location of hydrophobic and hydrophilic clusters (Fig. 4*B*).

Our results showed that in non-stimulated animals, the mantle appears as the main tissue expressing *Cg*-Def. Although we were not able to isolate by HPLC significant amount of pure defensin from the oyster, we clearly detected its presence by mass spectrometry analyses of mantle tissue (Fig. 2*C*). The results presented here are in contrast with those reported for the mollusk, *M. galloprovencialis*. In the mussel, AMPs are only produced in hemocytes where they are stored and released following bacterial challenges (45). In the well studied insect model, *Drosophila melanogaster*, AMPs are predominantly produced in the fat body (the functional equivalent of mammalian liver) and secreted into the blood in response to a microbial challenge, which characterizes the systemic response in insects. Otherwise, the expression of these peptides is also locally regulated in the surface epithelia of several *Drosophila* tissues (47), as also observed for the expression of peptides in mammalian epithelia. For example, *Drosophila* defensin is expressed in the labelar glands and in the seminal receptacle and spermatheca (for a review see Imler and Bulet (48)). Additionally, in mammals, AMPs are constitutively produced by blood cells (49). Apart from insects, most of the AMPs reported in the different groups of invertebrates have been isolated from hemocytes (13). Our preliminary results would suggest that *Cg*-Def is continuously expressed in the oyster mantle, *i.e.* at epithelial surface, indicating that this AMP has an important function in host protection against environmental microbes. Indeed, in oyster, the mantle is a site of intense exposure to external environment that represents a dynamic ecosystem for numerous bacterial species, including commensals and potential pathogens. A more detailed study of *Cg-def* gene expression in response to various microbial challenges will be performed to gain insight into the role of *Cg*-Def in oyster immunity. In addition, AMPs isolation and characterization must be investigated from other oyster tissues. These data on AMPs are of great interest to understand how the oyster immune system interacts with the commensal flora and how it responds to environmental pathogens.

Acknowledgments—We are very grateful to Dr. A. J. Ouellette for excellent advice during defensin recombinant expression. We also thank B. Romestand, J. de Lorgeril, G. Desserre, and M. Leroy for helpful discussion and assistance.

REFERENCES

1. Yang, D., Biragyn, A., Kwak, L. W., and Oppenheim, J. J. (2002) *Trends Immunol.* **23,** 291–296

2. Bulet, P., Stocklin, R., and Menin, L. (2004) *Immunol. Rev.* **198,** 169–184
3. Boman, H. G. (2003) *J. Intern. Med.* **254,** 197–215
4. Brogden, K. A. (2005) *Nat. Rev. Microbiol.* **3,** 238–250
5. Boulanger, N., Lowenberger, C., Volf, P., Ursic, R., Sigutova, L., Sabatier, L., Svobodova, M., Beverley, S. M., Spath, G., Brun, R., Pesson, B., and Bulet, P. (2004) *Infect. Immun.* **72,** 7140–7146
6. Harder, J., Bartels, J., Christophers, E., and Schroder, J. M. (2001) *J. Biol. Chem.* **276,** 5707–5713
7. Izadpanah, A., and Gallo, R. L. (2005) *J. Am. Acad. Dermatol.* **52,** 381–390; quiz 391–382
8. Raj, P. A., and Dentino, A. R. (2002) *FEMS Microbiol. Lett.* **206,** 9–18
9. Sima, P., Trebichavsky, I., and Sigler, K. (2003) *Folia. Microbiol.* (*Praha*) **48,** 123–137
10. Dimarcq, J.-L., Bulet, P., Hétru, C., and Fofmann, J. A. (1998) *Biopolymers* **47,** 465–477
11. Charlet, M., Chernysh, S., Philippe, H., Hétru, C., Hoffmann, J. A., and Bulet, P. (1996) *J. Biol. Chem.* **271,** 21808–21813
12. Mitta, G., Vandenbulcke, F., and Roch, P. (2000) *FEBS Lett.* **486,** 185–190
13. Bachère, E., Gueguen, Y., Gonzalez, M., de Lorgeril, J., Garnier, J., and Romestand, B. (2004) *Immunol. Rev.* **198,** 149–168
14. Satchell, D. P., Sheynis, T., Kolusheva, S., Cummings, J., Vanderlick, T. K., Jelinek, R., Selsted, M. E., and Ouellette, A. J. (2003) *Peptides* **24,** 1795–1805
15. Wu, Z., Powell, R., and Lu, W. (2003) *J. Am. Chem. Soc.* **125,** 2402–2403
16. Hetru, C., and Bulet, P. (1997) in *Methods in Molecular Biology* (Shafer, W. M., ed) Vol. 78, pp. 35–49, Humana Press Inc., Totowa, NJ
17. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117,** 479–485
18. Derome, A. E., and Williamson, M. P. (1990) *J. Magn. Reson.* **88,** 177–185
19. Bax, A., and Davis, G. D. (1985) *J. Magn. Reson.* **65,** 355–360
20. Rance, M. (1987) *J. Magn. Reson.* **74,** 557–564
21. Macura, S., Huang, Y., Sutter, D., and Ernst, R. R. (1981) *J. Magn. Reson.* **43,** 259–281
22. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson.* **85,** 393–399
23. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR* **2,** 661–665
24. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & sons, New York
25. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298
26. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
27. Frishman, D., and Argos, P. (1995) *Proteins* **23,** 566–579
28. Livak, K. J., and Schmittgen, T. D. (2001) *Methods* **25,** 402–408
29. Froy, O., and Gurevitz, M. (2003) *Trends Genet.* **19,** 684–687
30. Landon, C., Thouzeau, C., Labbe, H., Bulet, P., and Vovelle, F. (2004) *J. Biol. Chem.* **279,** 30433–30439
31. Zaballos, A., Villares, R., Albar, J. P., Martinez, A. C., and Marquez, G. (2004) *J. Biol. Chem.* **279,** 12421–12426
32. Li, B., Gorman, E. M., Moore, K. D., Williams, T., Schowen, R. L., Topp, E. M., and Borchardt, R. T. (2005) *J. Pharm. Sci.* **94,** 666–675
33. Robinson, N. E. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 5283–5288
34. Cornet, B., Bonmatin, J. M., Hetru, C., Hoffmann, J. A., Ptak, M., and Vovelle, F. (1995) *Structure* **3,** 435–448
35. Savarin, P., Romi-Lebrun, R., Zinn-Justin, S., Lebrun, B., Nakajima, T., Gilquin, B., and Menez, A. (1999) *Protein Sci.* **8,** 2672–2685
36. Guijarro, J. I., M'Barek, S., Gomez-Lagunas, F., Garnier, D., Rochat, H., Sabatier, J. M., Possani, L., and Delepierre, M. (2003) *Protein Sci.* **12,** 1844–1854
37. Yang, Y. S., Mitta, G., Chavanieu, A., Calas, B., Sanchez, J. F., Roch, P., and Aumelas, A. (2000) *Biochemistry* **39,** 14436–14447
38. Anderson, R. S., and Beaven, A. E. (2001) *Aquat. Living Resour.* **14,** 343–349
39. Mitta, G., Hubert, F., Dyrynda, E. A., Boudry, P., and Roch, P. (2000) *Develop. Comp. Immunol.* **24,** 381–393
40. Ouellette, A. J., Satchell, D. P., Hsieh, M. M., Hagen, S. J., and Selsted, M. E. (2000) *J. Biol. Chem.* **275,** 33969–33973
41. Friedrich, C., Scott, M. G., Karunaratne, N., Yan, H., and Hancock, R. E. (1999) *Antimicrob. Agents Chemother.* **43,** 1542–1548
42. Goldman, M. J., Anderson, G. M., Stolzenberg, E. D., Kari, U. P., Zasloff, M., and Wilson, J. M. (1997) *Cell* **88,** 553–560
43. Aarbiou, J., Rabe, K. F., and Hiemstra, P. S. (2002) *Ann. Med.* **34,** 96–101
44. Da Silva, P., Jouvensal, L., Lamberty, M., Bulet, P., Caille, A., and Vovelle, F. (2003) *Protein Sci.* **12,** 438–446
45. Mitta, G., Vandenbulcke, F., Hubert, F., and Roch, P. (1999) *J. Cell. Sc.* **112,** 4233–4242
46. Romestand, B., Molina, F., Richard, V., Roch, P., and Granier, C. (2003) *Eur. J. Biochem.* **270,** 2805–2813
47. Tzou, P., Ohresser, S., Ferrandon, D., Capovilla, M., Reichhart, J. M., Lemaitre, B., Hoffmann, J. A., and Imler, J. L. (2000) *Immunity* **13,** 737–748
48. Imler, J. L., and Bulet, P. (2005) *Chem. Immunol. Allergy* **86,** 1–21
49. Lehrer, R. I., and Ganz, T. (1999) *Curr. Opin. Immunol.* **11,** 23–27
50. Casteels, P., Ampe, C., Jacobs, F., and Tempst, P. (1993) *J. Biol. Chem.* **268,** 7044–7054
51. Mitta, G., Hubert, F., Nöel, T., and Roch, P. (1999) *Eur. J. Biochem.* **264,** 1–9

Characterization of a Defensin from the Oyster *Crassostrea gigas*: RECOMBINANT PRODUCTION, FOLDING, SOLUTION STRUCTURE, ANTIMICROBIAL ACTIVITIES, AND GENE EXPRESSION

Yannick Gueguen, Amaury Herpin, André Aumelas, Julien Garnier, Julie Fievet, Jean-Michel Escoubas, Philippe Bulet, Marcelo Gonzalez, Christophe Lelong, Pascal Favrel and Evelyne Bachère

J. Biol. Chem. 2006, 281:313-323.
doi: 10.1074/jbc.M510850200 originally published online October 24, 2005
