Bacterial biosynthetic gene clusters encoding the anti-cancer haterumalide class of molecules: biogenesis of the broad spectrum antifungal and anti-oomycete compound, oocydin A*

Miguel A. Matilla1, Henning Stöckmann2, Finian J. Leeper2, and George P.C. Salmond1*

1From the Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK, CB2 1QW.
2From the Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, UK, CB2 1EW.

*Running title: Identification of oocydin A biosynthetic gene clusters

To whom correspondence should be addressed: George P.C. Salmond, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK, CB2 1QW. Tel: +44 (0)1223 333650; Fax: +44 (0)1223 766108; E-mail: gpcs@mole.bio.cam.ac.uk

Keywords: Oocydin A; haterumalide; Serratia; Dickeya; macrolide; polyketide synthase; halogenation; trans-acyltransferase; anti-cancer, antifungal and anti-oomycete compound.

Background: Oocydin A is an anticancer haterumalide with strong antimicrobial activity against agriculturally important plant pathogenic fungi and oomycetes.

Results: The oocydin A gene cluster has been identified and characterized in four plant-associated enterobacteria.

Conclusion: The ooc gene cluster is organized in three transcriptional units encoding enzymes that belong to a growing class of trans-acyltransferase polyketide synthases.

Significance: Oocydin A has potential agricultural, pharmacological and chemotherapeutic applications.

SUMMARY

Haterumalides are halogenated macrolides with strong antitumor properties, making them attractive targets for chemical synthesis. Unfortunately, current synthetic routes to these molecules are inefficient. The potent haterumalide, oocydin A, was previously identified from two plant-associated bacteria through its high bioactivity against plant-pathogenic fungi and oomycetes. In this study we describe oocydin A (ooc) biosynthetic gene clusters identified by genome sequencing, comparative genomics and chemical analysis in four plant-associated enterobacteria of the Serratia and Dickeya genera. Disruption of the ooc gene cluster ablated oocydin A production and bioactivity against fungi and oomycetes. The ooc gene clusters span between 77- and 80-kb and encode five multimodular polyketide synthase (PKS) proteins, a hydroxy-methylglutaryl-CoA synthase cassette and three flavin-dependent tailoring enzymes. The presence of two free-standing acyltransferase (AT) proteins classifies the oocydin A gene cluster within the growing family of trans-AT PKSs. The amino acid sequences and organization of the PKS domains is consistent with the chemical predictions and functional peculiarities associated with trans-AT PKSs. Based on extensive in silico analysis of the gene cluster, we propose a biosynthetic model for the production of oocydin A, and by extension, for other members of the haterumalide family of halogenated macrolides exhibiting anticancer, antifungal and other interesting biological properties.

INTRODUCTION

Plants have to cope with potentially devastating biotic and abiotic environmental stresses during their life cycle. Among these stresses, fungal and oomycete plant pathogens are responsible for many of the most agriculturally important diseases worldwide (1). Effective chemical control of these infections is extremely difficult and biopesticides (pest management agents based on living microorganisms or natural products) are considered to be one of the most promising methods for rational crop management (2). Many bacteria can synthesise bioactive secondary metabolites, some of which can combat key fungal and oomycete plant-
Identification of oocydin A biosynthetic gene clusters

pathogens and thus these metabolites may have utility in biocatalysis systems (3,4).

Oocydin A is a chlorinated macrolide which was isolated in 1999 from the plant epiphytic bacterial strain Serratia marcescens MSU97 because of its strong bioactivity against plant pathogenic oomycetes (5). The same macrolide, but named haterumalide NA, was also isolated from the sponge Ircinia sp. (6) and subsequently from the rhizosphere bacterium, Serratia plymuthica A153, where it showed ability to inhibit the hyphal growth of the plant pathogenic fungus, Sclerotinia sclerotiorum (7). Furthermore, in 2005, it was also isolated from Serratia liquefaciens No. 1821 on the basis of its anti-hyperlipidemic activity and given the code FR177391 (8). Spectroscopic data for haterumalide NA, oocydin A and FR177391 are identical, although there are differences in their reported optical rotations, so it remains formally possible that the bioactive metabolite from Serratia species is the enantiomer of that derived from the sponge.

Various haterumalides (6,9) and their oxygenated analogues, the biselides (10,11), also show powerfull antitumor activity against human lung, breast, leukemia and colon cancer cells. Interestingly, haterumalide NA/oocydin A exhibited cytotoxicity against P388 leukemia (6) and breast (5) cancer cells. These broad biological impacts made the haterumalides/oocydins very attractive targets for total chemical synthesis, which has been achieved by several research groups (9,12,13). Although there are no reports to date on the biosynthetic route to oocydin A, its structure suggests a polyketide origin. During the biosynthesis of polyketides, the polyketide chain is assembled and elongated while covalently attached to acyl carrier protein (ACP) domains. The elongation is performed by the C-C bond-forming ketosynthase (KS) domains and acyltransferase (AT) domains responsible for introducing (mostly) malonyl or methylmalonyl building units to the ACP during each cycle of elongation. Type I polyketide synthases (PKSs) are usually multidomain proteins in which the domains form an assembly line, consisting of multiple modules, each of which is responsible for one round of chain elongation. The minimal KS-AT-ACP module can be supplemented with one or more of a ketoreductase (KR) domain (converting the β-keto group to a β-hydroxy group) a dehydratase (DH) domain (eliminating water to generate a C=C double bond) and an enoylreductase (ER) domain which reduces double bonds to saturated intermediates (14). According to the textbook model, the number of modules in a polyketide gene cluster is correlated with the number of extension cycles executed by the PKS and therefore with the structure of the eventual secondary metabolite. After the synthesis of the polyketide backbone, the chain is released from the PKS, usually by hydrolysis or cyclisation catalysed by a thioesterase (TE) domain and, in many cases, the polyketide is then modified by a range of tailoring enzymes. These modifications include glycosylations, hydroxylations, acyl transfers, epoxidations and halogenations (14-16). In bacteria, polyketide biosynthesis genes are normally organized in gene clusters with the PKS genes being part of a single operon, which reflects both the coordinate regulation required for the activation of the biosynthetic pathway and the evolutionary origin of the cluster, in most cases, by horizontal gene transfer between microbial genomes (17).

In this study we employed genome sequencing, comparative genomics, mutagenesis and chemical analysis to identify the PKS gene cluster responsible for the biosynthesis of oocydin A. The oocydin A gene cluster is organized in three different transcriptional units and it is present in four different plant-associated enterobacteria, three of them belonging to the genus Serratia. Based on the structure of the gene cluster and analysis of its conserved motifs and catalytic residues, we propose a model for the biosynthesis of oocydin A.

EXPERIMENTAL PROCEDURES

**Bacterial strains, plasmids, phages, culture media and growth conditions**—Bacterial strains and plasmids used in this study are listed in Table 1. Serratia, Dickeya and their derivative strains were routinely grown at 30 °C, unless otherwise indicated, in Luria Broth (LB; 5 g yeast extract 1−1,10 g Bacto tryptone 1−1 and 5 g NaCl 1−1), Potato Dextrose (24 g potato dextrose broth 1−1), enriched Potato Dextrose medium (24 g potato dextrose broth 1−1, 6 g bacto peptone 1−1, 4 g yeast extract 1−1 and 100 mg NaCl 1−1), half-strength tryptic soy medium (15 g tryptic soy broth 1−1), or minimal medium (0.1%, w/v, (NH4)2SO4, 0.41 mM MgSO4, 0.2% (w/v) glucose, 40 mM K2HPO4, 14.7 mM KH2PO4, pH 6.9–7.1). Escherichia coli strains were grown at 37 °C in LB. Escherichia coli DH5α was used for gene cloning. Media for propagation of E.
coli β2163 were supplemented with 300 µM 2,6-diaminopimelic acid (DAPA). When appropriate, antibiotics were used at the following final concentrations (in µg ml⁻¹): ampicillin, 100; kanamycin, 25 (E. coli strains) and 75 (Serratia strains); streptomycin, 50; erythromycin, 200. Sucrose was added to a final concentration of 10% (w/v) when required, to select derivatives that had undergone a second crossover event during marker exchange mutagenesis.

In vitro nucleic acid techniques- A Serratia marcescens MSU97 cosmid library was constructed from high-molecular-weight genomic DNA (35-45 kb) using the pWEB-TNC Cosmid Cloning Kit following the manufacturer’s instructions (Epicentre Biotechnologies). Plasmid DNA was isolated using the Anachem Keyprep plasmid kit. For DNA digestion, the manufacturer’s instructions were followed (New England Biolabs and Fermentas). Separated DNA fragments were recovered from agarose using the Anachem gel recovery kit. Ligation reactions, total DNA extraction and Southern blots were performed by standard protocols (28). DNA digoxigenin-dUTP probes were obtained via PCR following the instructions of the manufacturer (Roche). Competent cells were prepared using calcium chloride and transformations were performed by standard protocols (28). Phusion® high fidelity DNA polymerase (New England Biolabs) was used in the amplification of PCR fragments for cloning. Sequences of these PCR fragments were verified in order to discard amplicons containing mutations. Routine DNA sequencing was carried out at the University of Cambridge DNA Sequencing Facility on an Applied Biosystems 3730xl DNA Analyser.

Genome sequencing and bioinformatics analyses- Genomic DNA sequencing was performed at the DNA Sequencing Facility, Department of Biochemistry (University of Cambridge) using 454 DNA pyrosequencing technology on a Pico Titre Plate for a Roche Genome Sequencer FLX system. The shotgun assemblies were carried out using 454 GS De novo Assembler software (Newbler v2.6). For the Serratia marcescens MSU97 genome sequence, the assembly used 521,156 reads or 204 MB of raw data to give 38X coverage of the genome and resulted in 81 contigs, 68 of which were larger than 500 bp. The average contig size was 77,365 bp and the largest contig was 418,649 bp. The Serratia plymuthica A153 assembly used 308,585 reads or 129 MB of raw data to give a 22X coverage of the estimated genome size and resulted in a total of 36 contigs, 24 of them larger than 500 bp. The average contig size was 230,980 and the largest contig was 1,516,666 bp.

Automated annotation of the bacterial sequences was done using BASys web server (29). antiSMASH was used to analyze the potential secondary metabolite biosynthesis gene clusters present in the genomes (30). Genome comparison analyses were performed employing wgVISTA online tool (31). Open reading frames (ORFs) in the oocydin A gene cluster were automatically predicted using Glimmer 3.0 (32). Blast analyses were made for the functional gene assignment. Protein domain organization was identified using the NCBI conserved domains database (33) and the Pfam database (34). Multiple sequence alignments were carried out with ClustalW2 (European Bioinformatics Institute). Artemis software (Wellcome Trust Sanger Institute) was used to visualize genomic sequences.

Transposon mutagenesis- Random transposon mutagenesis of S. plymuthica A153 using TnKRCPN1 was performed as described below. In a biparental conjugal mating, 500 µL of overnight cultures of E. coli β2163 (pKRCPN1) and S. plymuthica A153 were mixed, collected by centrifugation, resuspended in 30 µL of fresh LB and spotted on an LB agar plate supplemented with 300 µM of DAPA. After overnight incubation at 30 ºC, cells were scraped off the plate and resuspended in 1 ml of LB. Serial dilutions were plated on LB agar medium containing 75 µg ml⁻¹ kanamycin. DAPA was not added to the LB agar medium, to allow counterselection of the E. coli donor. In total, 5000 kanamycin resistant insertion mutants were screened for their inability to inhibit P. ultimum growth. The insertion site of transposon TnKRCPN1 in mutants of interest was determined using random primed PCR following the method described previously (35) and using primers described in Table S5. From the 22 isolated mutants exhibiting no, or reduced, bioactivity, all the insertions were located in the oocydin A gene cluster. Within the isolated mutants, some of the insertions sites were identical, suggesting that the strains were most likely clonal isolates.
Marker exchange mutagenesis- Specific site-directed mutants defective in oocK and oocM were constructed by homologous recombination using derivative plasmids of the suicide vector pKNG101 (20). These plasmids, which are listed in Table 1, were confirmed by DNA sequencing and they carried mutant alleles for the replacement of wild type genes in the chromosome. In all cases, plasmids were transferred to S. plymuthica A153 by triparental conjugation using E. coli CC118lpir and E. coli HH26 (pNJ500) as helper. Mutants defective in oocK and oocM were generated using plasmids pMAMV54 and pMAMV62 respectively (Table 1). All relevant mutations were confirmed by PCR, sequencing and Southern blot analysis. Primers used in this marker exchange mutagenesis are listed in Table S5.

Generalised transduction- The newly isolated generalized transducing phage, ϕMAM1, (Matilla et al., unpublished) was used for transduction of chromosomal mutations using a method similar to that described previously (36). Transductants were selected on plates containing kanamycin and retention of phage sensitivity was confirmed in the transductants.

RNA extraction, cDNA synthesis and reverse transcription PCR (RT-PCR) analyses- RNA was extracted from late exponential (12 h) cultures grown in enriched Potato Dextrose medium using an RNaseasy mini kit (Qiagen) according to the manufacturer’s instructions. RNA concentration was determined spectrophotometrically and RNA integrity was assessed by agarose gel electrophoresis. Genomic DNA contamination was eliminated by treating total RNA with Turbo DNA free (Ambion). The synthesis of cDNA was performed using random hexamers (GE healthcare) and SuperScript II reverse transcriptase (Invitrogen) in a 20 µl reaction with 2.5 µg of total RNA and incubation at 42 ºC for 2 h. A negative-control reaction was also performed, omitting the reverse transcriptase enzyme. Then, the equivalent of 50 ng of total RNA was subjected to PCR amplification using primers to amplify across the junctions (Table S5). Positive and negative control PCRs were performed using genomic DNA and no-RT cDNA samples, respectively, as templates. PCR conditions consisted of 30 cycles of denaturation for 1 min at 94 ºC, annealing for 1 min at 62 ºC, and extension for 40 s at 72 ºC.

Antibacterial, antifungal and antioomycete activity in vitro- Production of antibiotic compounds was tested against E. coli ESS and Bacillus subtilis JH642 using lawn assays as described previously (37). Antagonistic activities of bacterial strains against the fast growing plant-pathogenic oomycete Pythium ultimum were assayed by spotting 5 µL overnight cultures of the selected strains on a potato dextrose agar (PDA) plate. Following incubation for 16 h at 25 ºC the plates were inoculated with 5 mm diameter mycelial plugs taken from a culture of P. ultimum grown on PDA. Plates were incubated at 25 ºC for 3-5 days. For the fungicide assays, indicator top agars of Verticillium dahliae, Thanatephorus cucumeris, Alternaria solani, and Fusarium oxysporum f. sp. lycopersici were prepared by vortexing a 5 mm fungal plug in 10 ml sterile distilled water. Then 15 µl of PDA was added, mixed and 5 µl top lawns were poured into PDA plates. Five microliters of overnight cultures of the selected strains were spotted on the surface of the fungal agar lawn and incubated for 7-10 days at 25 ºC. To determine fungicide/antioomycete levels in bacterial supernatants, culture samples were taken and cells were pelleted by centrifugation (14,000 x g, 10 min) and the supernatant was filtered (0.2 µm). Three hundred microliters of the filter-sterilized supernatant were added to wells cut into the PDA plate and incubated at 25 ºC for 3-5 days. All the experiments were repeated at least five times.

Caenorhabditis elegans virulence assays- The C. elegans virulence assays were adapted from those carried out for Kurz and colleagues (2003; 38). Briefly, NGM plates were inoculated with 50 µL of an overnight culture of the oocynid A-producing wild type strains (S. marcescens MSU97, S. plymuthica A153, S. odorifera 4Rx13 and D. dandatii Ech703) or the oocynid A-deficient S. plymuthica A153 mutants. Plates inoculated with E. coli OP50 were also included as a negative control. The plates were incubated for 16 h at 25 ºC. For each assay, worms previously fed on E. coli OP50 were transferred to the plates inoculated with the strains to test. Fifty L4 stage hermaphrodite DH26 worms (Genotype fer-15(b26)II, sterile at 25 ºC) were used for each strain tested (ten worms per plate). Plates were incubated at 25 ºC and the number of live worms was scored every 24 h. Worms were considered dead when they failed to respond to touch. Survival curves were analyzed using GraphPad Prism software. P values <0.05 were considered statistically significant.
Identification of oocydin A biosynthetic gene clusters

Motility and biofilm formation assays—Motility and biofilm assays were performed at 25 °C in enriched Potato Dextrose medium, conditions where we observed enhanced oocydin A production. Swimming assays were performed on 0.3% agar plates. Biofilm formation assays were carried out in 96-well microtiter plates with shaking at 75 r.p.m. Crystal violet staining and quantification method was performed as previously described (39).

LC-MS studies—For the analysis of the bacterial supernatants, 25 ml of enriched potato dextrose broth was inoculated with the strains to analyze and incubated at 25 °C. After 48 h of incubation cells were pelleted by centrifugation and filtered. Before the extraction, the pH of the culture supernatant was adjusted to 3.8 with citric acid. Six millilitres of the pH-adjusted supernatant were extracted twice with dichloromethane (6 ml). The organic layers were supernatant were extracted twice with citric acid. Six millilitres of the pH-adjusted supernatant were extracted twice with dichloromethane (6 ml). The organic layers were combined, dried over sodium sulphate, and evaporated under reduced pressure. The residue was resuspended in 1 ml of H2O:Acetonitrile (1:1, v/v) and 10 µl was analysed by LC-MS on a Finnigan MAT LCQ instrument using a Phenomenex Kinetex 2.6 u XB-C18 100A column of size 100 x 2.1 mm eluted at 0.3 ml min⁻¹ with a linear gradient over 15 min from 95:5 to 5:95 of water to acetonitrile, each containing 0.1% formic acid.

RESULTS AND DISCUSSION

Identification of the oocydin A gene cluster—In order to identify the oocydin A gene cluster, the genome sequence of Serratia marcescens MSU97 was obtained using 454 pyrosequencing technology, automatically annotated and analyzed. Based on the presence of NRPS and PKS encoding genes, at least five candidate biosynthetic gene clusters were identified in the 5.3 Mb MSU97 draft genome. Among them are gene clusters involved in the biosynthesis secondary metabolites such as siderophores and the red pigmented linear tripyrrole, prodigiosin. A number of studies have reported that genes encoding enzymes responsible for the halogenation of secondary metabolites are closely associated with particular NRPS and PKS gene clusters (40-44). Therefore, we hypothesized that the gene cluster involved in the biosynthesis of oocydin A was likely to encode a halogenase enzyme. During the in silico screening of the MSU97 genome we identified a non-heme chloroperoxidase and a peroxidase which were not associated with any of the candidate biosynthetic gene clusters. However, two flavin-dependent monooxygenases were identified in a 77.5-kb gene cluster involved in the biosynthesis of an unknown polyketide. Flavin-dependent halogenases are the main class of enzyme involved in the halogenation of aliphatic compounds (15,45), so it seemed possible that this was the oocydin A gene cluster, although its complexity made it difficult to predict the structure of the polyketide synthesised.

In a first direct attempt to isolate the oocydin A gene cluster, we constructed a cosmid library of S. marcescens MSU97. Specific primers for five of the MSU97 biosynthetic gene clusters were designed and screening of 2000 clones yielded 90 cosmid clones that hybridized positively to the probes. However, none of these cosmids could induce the production of oocydin A in Escherichia coli, a host generally appropriate for the expression of cloned enterobacterial genes (46-48). Random mutagenesis was then employed as a second strategy in order to isolate S. marcescens MSU97 mutants with reduced antimicrobial activity towards the oomycete Pythium ultimum. However, the MSU97 strain proved to be recalcitrant to various genetic tools that we have used successfully before for the genetic analysis of secondary metabolite production in Serratia and Erwinia (36,47,49). This genetic intractability, coupled with high intrinsic multidrug resistance made a classical mutagenesis approach unfeasible.

Besides S. marcescens MSU97, we had access to another oocydin A producing strain, Serratia plymuthica A153 (7). We postulated that both strains should carry the same gene cluster for oocydin A biosynthesis. Serratia plymuthica A153 is a genetically tractable strain and therefore susceptible to efficient transposon mutagenesis. With the aim of isolating oocydin A defective mutants, a transposon mutant library was screened in dual culture plate bioassays for mutants defective in antimicrobial activity towards the oomycete P. ultimum. We identified six independent transposon insertion mutants showing complete loss of the bioactive properties against the oomycete (Fig. 1E-F and S1A). We also isolated a mutant (MMnO15) that showed reduced anti-Pythium activity (Fig. 1E). All the insertion mutations were transduced back into a S. plymuthica A153 wild type genetic background using a new phage (φMAM1) that we have recently isolated (Matilla et al., unpublished) to ensure that the phenotypes were...
identification of oocydin A biosynthetic gene clusters

Indeed, associated with single insertions and to confirm association between mutation and phenotype. The resulting transductants displayed the same phenotype as the original random mutants confirming genetically that loss of bioactivity was caused by the transposon insertion. Random primed PCR and genome sequence analysis of S. plymuthica A153 confirmed that all the transposon insertions mapped to a gene cluster homologous to the 77.5-kb MSU97 gene cluster (Table S2) that we had proposed as the main candidate for the biosynthesis of oocydin A (Fig. 2A). This corroborated the hypothesis that this particular polyketide gene cluster is also carried by S. plymuthica A153 (Table 2). We then successfully inactivated, in A153, the two candidate halogenase-encoding genes present in the gene cluster (oocK and oocM) using allelic replacement through the insertion of a kanamycin cassette. The resulting mutants were incapable of inhibiting the growth of P. ultimum (Fig. 1F and S1A) as expected.

To confirm the identity of the oocydin A gene cluster and to corroborate its role in the biosynthesis of the macrolide, we analysed by LC-MS the culture supernatants of wild type strains and identified a metabolite with the same molecular weight as oocydin A (Fig. 3A-B). This metabolite showed the characteristic isotope pattern of a monochlorinated compound, with peaks at m/z 493 and 495 [M+Na]+ in a 3:1 ratio as a result of the two isotopes 35Cl and 37Cl (Fig. S2A). The MS/MS fragmentation of [M+Na]+ ion m/z 493 yielded two major ions, at m/z 433 [M+Na-AcOH]+ and 449 [M+Na-CO2]+ (Fig. S2B-C). The same analysis showed that all the non-bioactive mutants displayed a complete loss of oocydin A production (Fig. 3F and S3B-D), confirming the role of the gene cluster in the biosynthesis of the halogenated macrolide. There were no dechloro-analogs of oocydin A detected in the extracts of oocK or oocM mutants which formally could be due to potential polar effects of the kanamycin cassette insertion in the respective genes. The analysis of the MMnO15 extracts showed that oocydin A was produced at significantly low levels (3-5%) as compared to the wild type strain A153 (Fig. 3E). This result suggests that other trans-acting ATs and ERs encoded in the genome of A153 could be catalyzing the required reactions, although less efficiently than OocU, V and W. Taking all these data into account, we have named the genes of this cluster ooc, for oocydin A biosynthetic genes.

The oocydin A gene cluster is present in several plant-associated enterobacteria

In addition to S. marcescens MSU97 and S. plymuthica A153, genome comparison analysis revealed that the ooc gene cluster is also present in the rhizobacterium, Serratia odorifera 4Rx13 (24; GenBank accession no. ADBX01000021; Genomic coordinates 102562-180808) and in the phytopathogenic bacterium Dickeya dadantii Ech703 (GenBank accession NC_012880; Genomic coordinates 1624970-1705054). Chemical analysis of 4Rx13 and Ech703 by LC-MS showed that they also produced oocydin A (Fig. 3C-D and S2D-E) and, as expected, they show very clear anti-Pythium properties (Fig. 1C-D). Interestingly, the genomic context of the ooc gene clusters in S. marcescens MSU97, S. plymuthica A153 and D. dadantii Ech703 is completely different. Therefore, the upstream and downstream ends of the different loci were assigned based on the homologies between the three biosynthetic clusters. The ooc gene clusters of S. plymuthica A153 and S. odorifera 4Rx13 show the highest homology (94.8% at DNA level; Table S1) and they also share the same genomic context. Therefore, these data suggest that the biosynthetic cluster could have been transferred horizontally between different genera and species of bacteria. In fact, although the G+C content of the ooc gene clusters of MSU97 (57.8%), A153 (56.6%), 4Rx13 (56.0%) and Ech703 (52.3%) are similar to the genomic G+C content in the chromosome, sequences reminiscent of mobile genetic elements, such as a bacteriophage P4 integrase, border the ooc gene cluster in D. dadantii Ech703.

The oocydin A gene clusters of A153, 4Rx13 and Ech703 are 78.2-, 78.3- and 80.1-kb respectively, and they are 81.7%, 81.8% and 70.9% identical, respectively, at the DNA level; Table S1) and they also share the same genomic context. Therefore these data suggest that the biosynthetic cluster could have been transferred horizontally between different genera and species of bacteria. In fact, although the G+C content of the ooc gene clusters of MSU97 (57.8%), A153 (56.6%), 4Rx13 (56.0%) and Ech703 (52.3%) are similar to the genomic G+C content in the chromosome, sequences reminiscent of mobile genetic elements, such as a bacteriophage P4 integrase, border the ooc gene cluster in D. dadantii Ech703. The oocydin A gene clusters of A153, 4Rx13 and Ech703 are 78.2-, 78.3- and 80.1-kb respectively, and they are 81.7%, 81.8% and 70.9% identical, respectively, at the DNA level; Table S1) and they also share the same genomic context. Therefore these data suggest that the biosynthetic cluster could have been transferred horizontally between different genera and species of bacteria. In fact, although the G+C content of the ooc gene clusters of MSU97 (57.8%), A153 (56.6%), 4Rx13 (56.0%) and Ech703 (52.3%) are similar to the genomic G+C content in the chromosome, sequences reminiscent of mobile genetic elements, such as a bacteriophage P4 integrase, border the ooc gene cluster in D. dadantii Ech703. The oocydin A gene clusters of A153, 4Rx13 and Ech703 are 78.2-, 78.3- and 80.1-kb respectively, and they are 81.7%, 81.8% and 70.9% identical, respectively, at the DNA level; Table S1) and they also share the same genomic context. Therefore these data suggest that the biosynthetic cluster could have been transferred horizontally between different genera and species of bacteria. In fact, although the G+C content of the ooc gene clusters of MSU97 (57.8%), A153 (56.6%), 4Rx13 (56.0%) and Ech703 (52.3%) are similar to the genomic G+C content in the chromosome, sequences reminiscent of mobile genetic elements, such as a bacteriophage P4 integrase, border the ooc gene cluster in D. dadantii Ech703. The oocydin A gene clusters of A153, 4Rx13 and Ech703 are 78.2-, 78.3- and 80.1-kb respectively, and they are 81.7%, 81.8% and 70.9% identical, respectively, at the DNA level; Table S1) and they also share the same genomic context. Therefore these data suggest that the biosynthetic cluster could have been transferred horizontally between different genera and species of bacteria. In fact, although the G+C content of the ooc gene clusters of MSU97 (57.8%), A153 (56.6%), 4Rx13 (56.0%) and Ech703 (52.3%) are similar to the genomic G+C content in the chromosome, sequences reminiscent of mobile genetic elements, such as a bacteriophage P4 integrase, border the ooc gene cluster in D. dadantii Ech703.
marcescens MSU97 ooc gene cluster are 87.1%, 87.3% and 78.6% identical to those of A153, 4Rx13 and Ech703, respectively. However, there are some interesting differences between them. Firstly, the initial gene of the cluster in MSU97, A153 and 4Rx13 (encoding a putative alpha/beta hydrolase) is not present in Ech703 (Table S4). Secondly, OocN and OocS of the MSU97, A153 and Ech703 ooc gene clusters are divided into 3 and 2 genes, respectively, in S. odorifera 4Rx13 (Table S3), although it is formally possible that the difference might be due to an error in genome sequencing. In accordance with this notion, domains that are encoded by two different proteins in 4Rx13 were found contiguous within the same PKS protein in other bacteria (Table S3). Lastly, the region containing the putative small genes, oocH and oocI, of MSU97 is poorly conserved between the strains showing a 49.4%, 46.5% and 41.7% identity at the DNA level, compared to A153, 4Rx13 and Ech703 respectively. Since this is a region of divergence of two transcriptional units (Fig. 2A) it is reasonable to hypothesise that this region has an important regulatory role(s) in the expression of this complex gene cluster. Interestingly, the G+C content of this intergenic region in MSU97, A153, 4Rx13, and Ech703, which is 42.6%, 41.7%, 42.1% and 37.1% respectively, is significantly lower than the mean G+C content of the gene clusters. OocB, a putative transporter present in the gene cluster of MSU97, A153 and Ech703 (Table S3), was not present in the GenBank draft annotation of the 4Rx13 cluster but the reannotation of the cluster showed that it is, in fact, present.

Bioinformatic analysis of the ooc biosynthetic gene clusters

As described above, the biosynthetic gene cluster of oocydin A spans between 77- and 80-kb and consists of 23 ORFs. The likely function for each gene product was deduced by sequence comparison with proteins of known function and sixteen of the proteins can be assigned possible roles in oocydin A biosynthesis (Table 2 and S2-S4). The oocydin A biosynthetic cluster encodes 5 multifunctional PKS enzymes (OocJ, OocL, OocN, OocR, OocS) which include a total of 16 β-ketoacyl synthase (KS), 7 keto reductase (KR), 20 acyl carrier protein (ACP), 5 dehydratase (DH), 2 enoyl-CoA hydratase (EH), 2 methyltransferase (MT) domains, a thioesterase (TE) and a NRPS condensation (C) domain (Fig. 2B, Table 2 and S2-S4). One of the first observations derived from the analysis was the absence of integrated acyltransferase (AT) domains in all the biosynthetic modules of the five PKS proteins encoded by the ooc gene cluster. This places the oocydin A PKS in the growing class of PKSs that use trans-acting AT domains. In these enzymes each module receives its acyl building block from one or more discrete ATs encoded by separate genes (50). The ooc gene cluster encodes two such proteins, OocV and OocW, containing 2 and 1 AT domains respectively, with all three of the domains containing the catalytic S-H dyad and the highly conserved N-terminal GGGSQ loop (Fig. S4; 51,52). In accordance with the trans-AT nature of the ooc gene cluster, analysis of its interdomain sequences identified regions with homology to truncated AT domains following most of the KS domains. Recently it has been suggested that these regions are the binding sites for the separate AT domains in trans-AT PKSs (53). Two or more free-standing ATs are infrequently found in trans-AT PKSs and, to our knowledge, only nine examples have been described: bacillaene (54), elansolid (55), etnangien (56), kirromycin (57), mupirocin (58), pederin (59), rhizopodin (60), rhizoxin (61) and sorangicin (62). Multiple sequence alignments of OocV and OocW against known trans-ATs revealed that OocV-AT1 and OocW share all the residues characteristic of a malonyl-CoA specific AT whereas OocV-AT2 possesses more active site residue diversity (Fig. S4; 51,52,62). It has been suggested that this residue diversity may be associated with a broader range of substrate specificity or specialized functions (52,62-64). BLAST analyses and multiple sequence alignments also revealed that OocV-AT1 and OocV-AT2 are most similar to BryP-AT1 and BryP-AT2 respectively (Table 2 and S2-S4; Fig. S4), which have been grouped in two different classes of trans-ATs (64). Lopanik et al. (2008) showed that both BryP-AT1 and BryP-AT2 prefer to transfer malonyl-CoA although BryP-AT1 is more active than BryP-AT2. Multiple alignment analyses (Fig. S4) also grouped OocV-AT2 with the stand-alone acyltransferase PedC. A recent study showed that PedC does not exhibit AT activity but possesses acyl hydrolase activity and cleaves acyl groups bound to ACP, suggesting a role in PKS biosynthetic proofreading (65). The number of domains in each of the five Ooc PKS proteins varies between 5 (OocR) and 18
Identification of oocydin A biosynthetic gene clusters

(OocN). Interestingly, the number of KS (16 as domains within the modular proteins and one as a separate ORF) and ACP (22 domains, two of them present in separate ORFs) domains exceeds the number needed for the biosynthesis of oocydin A, which would be 8 and 9 domains respectively. Presumably some of the KS domains are either inactive, and skipped, or are simply used for transferring the acyl intermediate to the next ACP domain without extending it. In accordance with this observation, the catalytic C-H-H triad present in active KS domains, is absent in KS1 (S-H-H), KS5 (C-Q-H), KS10a (C-E-H) and OocF (S-H-H) (Fig. S5-S8, the numbering of domains refers to the module number as shown in Fig. 2B) (66). Furthermore the characteristic HGTGT catalytic motif, which is essential for the decarboxylative condensation, is poorly conserved in KS5 (QGMGL), KS10a (EANGS), KS10b (HGASS), KS12 (HATGG), KS14 (HSTGS) and KS15 (HAA/TGM) which may indicate that these domains are not extending the polyketide chain (Fig. S5-S8; 62). All ACP domains show the highly conserved signature motif GxDS except for ACP7 (GFNS), ACP14 (GLNS), OocG (GANS) and OocO (GFES) (Fig. S5-S8; 67).

All the 7 potential KR domains identified showed the characteristic Rossmann fold for cofactor binding. The S-Y-N conserved triad located in the active site (68) was modified in KR4 (S-Y-A) and KR11 (S-Y/S-A/I) (Fig. S5-S8). Module 3 contains a region of 55 amino acids which shows homology with the Rossmann fold of NAD(P)-binding proteins but it lacks the S-Y-N conserved triad which indicates that it does not function as a ketoreductase. Of the five predicted DH domains found in the PKS, three of them (DH7a, DH8 and DH10) contain the HxxxGxxxxP motif characteristic of DH domains (Fig. S5-S8; 69). Module 1 contains two EH domains showing altered conserved motifs essential for the oxyanion hole that stabilizes the enolate anions (Fig. S9; 70). However, the “oxyanion hole” consensus motifs are present in OocC and OocD (Fig. S9). A TE and a C domain containing the GxSxG (71) and the HHxxxDG (72) conserved motifs, respectively, were identified in the last multimodular PKS of the cluster.

Finally, at least three putative tailoring enzymes are encoded by genes in the cluster and they probably catalyze oxidative/reductive transformations (OocK, OocM and OocU). The cluster also encodes a protein (OocB) 40% identical (60% similar) to the putative efflux protein RRSRL_03865 present in Ralstonia solanacearum UW551 and in other genera such as Enterobacter, Photorhabdus and Burkholderia. This protein might be involved in the efflux of the bioactive molecule to the environment. Perhaps surprisingly, no genes encoding obvious candidate regulatory proteins were found in the ooc gene cluster.

Model for the biosynthesis of oocydin A

It has been observed that trans-AT systems possess numerous functional peculiarities as compared to cis-AT PKSs, in which the AT domains are located within the multidomain PKS. These peculiarities are often translated into unusual chemistry and include the loss of co-linearity found in canonical PKSs, lack of expected domains, unusual domain orders and domain sets and the splitting of modules between two PKS proteins. Additionally, PKS modules can be “skipped” or used more than once during the biosynthetic pathway (16,50). These unusual characteristics make it difficult to establish an obvious predictive correlation between the polyketide structure and the architecture of the trans-AT PKS modules. However, the substrate specificity of KS domains of trans-AT PKSs has been correlated with their sequence and they have been found to be most similar to other KS domains that process similar substrates (50,55). Thus, based on an extensive in silico analysis we propose a mechanism for the biosynthesis of oocydin A as shown in Fig. 2B. In this model functional KS domains have been compared with the nearest characterised homologues in order to determine what acyl groups could be accepted by the respective domains. As has been observed in other trans-AT PKSs, the oocydin A PKS deviates from the textbook PKS model architecture in several features.

The structure of oocydin A shows carboxyl groups at both ends of the linear chain of the polyketide, so its biosynthesis could, in principle, be started at either end. However the organization of the domains makes it much more likely that the starter unit is C-17 to C-19. The predicted DH-MT-FkbH domains in the Load module are uncommon but have some similarity to the DH-KR-FkbH Load module of BryA, involved in the biosynthesis of bryostatin, which functions to load a lactyl starter unit onto an ACP domain (73). However, the structure of oocydin A does not require any such starter unit; the expected starter unit is a malonyl unit instead.
Identification of oocydin A biosynthetic gene clusters

of the glyceric acid unit normally loaded by the FkbH domain (74). Therefore, we propose that the first three domains (DH, MT, FkbH) might be non-functional relics in the process of evolutionary degeneration. In accordance with this, the DH and MT domains present in this unusual DH-MT-FkbH module lack the characteristic HxxxGxxxxP and LExGxGxxG conserved motifs, respectively. Furthermore KS1 has a serine in place of the active site cysteine and so is probably incapable of chain extension. Thus, the proposed biosynthesis would begin with the transfer of an activated malonyl starter unit from malonyl-CoA to the ACP1 domain of OocJ by one of the proposed trans-ATs. It is possible that KS1 catalyses the decarboxylation of this malonyl group to an acetyl group but it is equally likely that this carboxyl is retained and it is the acetic acid unit introduced by the 3-hydroxy-3-methylglutaryl-CoA synthase (HCS) cassette (see below) that is decarboxylated. KS2 would then catalyse the first round of chain extension to give a β-keto-thioester attached to ACP2.

The methyl group at C-17 is attached to a carbon that would initially have been a carbonyl group. Such methyl groups are usually added by HCS cassettes that consist of stand-alone HCS, ketosynthase and ACP proteins (63,73,75). Importantly, these three proteins are contiguous in the cluster (OocE-G). In this mechanism for generating a methyl group, two dehydratase (DH)-like EH domains are also required. Two genes for stand-alone EH proteins (OocC and OocD) that complete the operon encoding the HCS cassette are present in the ooc gene cluster. The similarity between the EH domains of OocD and OocC is only 19.1% (Fig. S9) and it has been hypothesized that such EH domains have different functional roles catalyzing, respectively, the dehydration and decarboxylation required to produce the methyl group (70). Two side-by-side EH domains are also present in module 1 of OocJ. The integration of EHs within a PKS module is unusual, although it has been described in the gene clusters of the related PKSs that make pederin (59) and onnamide A (76) as examples of the biosynthetic versatility of PKSs. The fact that the motifs essential for the oxyanion hole are poorly conserved in EH1a and EH1b and that a mutant defective in the EH OocC (MMnO14) does not produce oocydin A suggest that OocC and OocD are responsible for the required dehydration and decarboxylation. Thus the HCS cassette (or alternatively the DH-like domains in module 1) would function at module 2 by adding a methyl group at C17 (or adding -CH2CO2H if decarboxylation of the starter unit does occur). The domain structures of the other four PKSs (OocL, OocN, OocR and OocS), do not show clearly the order in which they act. However, the structure of oocydin A suggests that they act in the same order as the genetic organisation of their cognate genes, which is normally the case. Thus, the subsequent biosynthetic steps can be largely rationalized by the co-linearity rule with several deviations since some domains have mutant catalytic domains and several modules lack domains that would be predicted from the oocydin A structure. Module 3 appears to contain a fully active KS domain but only has a fragment of a KR domain, which lacks the active site groups and cannot be functional. As oocydin A has no keto groups that have not been reduced, this suggests that this module may be skipped. Module skipping has been described in other trans-AT PKS such as difficidin (54), myxovirescin A (77), leinamycin (78) and chivosazol (79). Following the assembly line, KS5 at the N-terminal end of OocN has a mutated C-H-H motif, so probably does not elongate the chain. However, it may be involved in transferring the chain to the next module without elongating it.

Further along OocN, KS7 is predicted, by homology with other PKS domains, to accept the (S)-β-hydroxy-acyl group that is produced by module 6. Domain 7 has two DH-like domains. In other PKSs, such as the sorangicin one (62), the extra DH domain is proposed to catalyse a cyclisation in which a hydroxyl group attacks the β-position of an α,β-unsaturated thioester in a Michael addition mechanism (interestingly, many domains in the ooc PKS show high similarity to equivalent domains in the sor PKS). In accordance with this model, KS8 has homology to Kss that accept acyl groups where a cyclisation has occurred. In the case of oocydin A, this cyclisation needs to be from a hydroxyl at C-14, as shown in Fig. 2B. It appears therefore that hydroxylation at C-14 occurs while the acyl chain is incomplete and still attached to the PKS (most likely when it was attached to module 4 before the reduction of the keto group at C-15 occurs so that deprotonation of C-14 is facilitated). The enzyme responsible for hydroxylation of C-14 is probably one of the two flavin-dependent monoxygenases OocK and OocM, the genes for which are located between...
PKS-encoding genes, immediately before and after oocL. Since PKS genes are very often contiguous in a cluster, this location of oocK and oocM may suggest that the encoded enzymes act on substrates that are still bound to one of the PKS proteins.

Given that the fragment of a KR domain in module 3 lacks the catalytic triad, the number of active KR domains identified in the ooc biosynthetic cluster is insufficient for the biosynthesis of the macrolide. Therefore, one of the modules is probably used twice, catalysing two rounds of elongation. This unusual feature has mainly been reported in trans-AT systems, such as rhizopodin (60), lankacidin (53) and oxazolomycin (50,80). We propose that module 8 would be used twice, but another possibility is that KS9 is in fact active in chain elongation but then passes the β,keto acyl group back to module 8 for processing by its KR and DH domains (which would be an example of non-linearity often found in trans-AT PKSs).

At this point in assembly an enoyl reductase (ER) domain would be needed but OocN (like all of the other four PKS proteins) lacks this type of domain. However, OocU has high similarity to proteins found in PKS clusters that have been shown to be trans-acting flavin-dependent ERs such as PksE from Bacillus subtilis (81) or BatK from Pseudomonas fluorescens (66). Therefore, we postulate that this ER domain might reduce the α,β,γ,δ-conjugated diene attached to module 8, as shown in Fig. 2B, leaving the remaining double bond in the β,γ-position.

Further along the PKS production line, module 11, which spans OocR and S, contains a MT domain which we suggest adds the methyl group present at C-4. The structure of oocydin A also requires a DH to act at this stage to generate the C-4/C-5 double bond. There is no DH in module 11 but there is one in module 10, so perhaps the β-hydroxy acyl group from module 11 gets passed back for dehydration by that domain. Alternatively a trans-acting DH may effect the dehydration.

The TE and the NRPS C domains of OocS show high similarity to the domains present in the PKS proteins BryX and BryD, respectively, involved in the biosynthesis of bryostatin (73). Interestingly, the last PKS proteins involved in the chain extension of oocydin A and bryostatin both terminate with these C domains. It is not known whether it is the TE or the C domain which catalyses the macroactonisation in bryostatin biosynthesis. By analogy, we propose that either the TE or the NRPS C domain in OocS is the likely candidate for the macroactonisation that releases the acyl group from the PKS. After the release of the polyketide from the PKS, the final acetylation (shown in Fig. 2B) can be catalysed by a separate acyltransferase. However, there is a possibility that one of the KS14, KS15, TE or C domains of OocS catalyse this acetylation. KS's, TE's and C domains all catalyse acyl transfers of some complexion, so in principle, O-acylation would be only a small change of function.

Finally, the chlorine atom at C-8 should be introduced by a chlorinase. Chlorinating enzymes have been classified into highly specific halogenases requiring dioxygen and either a reduced flavin or α-ketoglutarate as co-substrates, and less specific haloperoxidases that use hydrogen peroxide, often P450 enzymes (15,45). Since there are no predicted P450 enzymes in the ooc gene cluster, the main candidates for the halogenase are the flavin-dependent monoxygenases OocK and M. In agreement with this, the mechanism of action of flavin-containing halogenases is just a slight modification of that of monoxygenases (82) and crystalization of some of these halogenases has revealed a flavin monoxygenase domain in most of them (83). The chlorination could be a post-PKS biosynthesis step (Fig. 2B) or could happen on a PKS-bound intermediate, perhaps in module 7 at the β-keto-thioester stage, when C-8 is easy to deprotonate.

The oocydin A cluster is formed by three transcriptional units.

The genetic organization of the oocydin A cluster suggests the presence of at least three transcriptional units. To further investigate this hypothesis, transcript analysis by RT-PCR was performed on cultures of S. marcescens MSU97. To select the conditions where the oocydin A genes were being expressed, MSU97 culture samples were taken at different points along the growth curve. Filter-sterilized supernatants from these samples were added to holes punched in Pythium ultimum bioassay plates. Oomycete growth inhibition was observed in the supernatants after 12 h of growth in enriched Potato Dextrose medium (5) and these growth conditions were therefore used for the RT-PCR studies. For this analysis, primers were designed to cover the region between the 3’ end of the upstream gene and the 5’ end of the contiguous downstream gene (Fig. 5A). Two smaller
transcriptional units, one consisting of oocA and oocB and a second consisting of oocG, oocF, oocE, oocD and oocC were identified. In addition, RT-PCR products were detected across all the intergenic regions containing oocJ-W, indicating the existence of a large polycistronic transcript (Fig. 5B). Unexpectedly, prominent PCR products were also obtained for the region covering the genes oocG-I (Fig. 5B). This might indicate the presence of divergent overlapping messenger RNAs in this region, consistent with the proposed regulatory role of this locus in the ooc cluster. Our analysis confirmed the presence of the three predicted transcriptional units in the oocynin A cluster, although the absence of additional internal promoters can not be assumed.

Biological properties of oocynin A

Oocynin A has been shown to be very active against plant pathogenic oomycetes belonging to Phytophthora and Pythium genera (5). However, the activity of oocynin A against plant pathogenic fungi such as Sclerotinia sclerotiorum was controversial since two different studies observed contradictory results (5,7). These apparently contradictory findings perhaps could be explained by the concentration of bioactive molecule present in each bioassay since oocynin A had fungistatic or fungicidal effects, depending on the concentration (7). With the aim of determining the biological roles of oocynin A, we characterized phenotypically the strain S. plymuthica A153 and its non-oocynin A producing mutants. The other oocynin A producers, S. marcescens MSU97, S. odorifera 4Rx13 and D. dadantii Ech703 were also analyzed. Perhaps surprisingly, given the number of enzymes that it encodes, inactivation of the large ooc gene cluster had no detectable effect on the growth rate. Neither was there any detectable impact on motility or biofilm formation of S. plymuthica A153 (data not shown). Besides the very high bioactivity that oocynin A shows against the oomycete, Pythium ultimum, we also tested its bioactive properties against fungal plant pathogens. In contrast to the report by Strobel et al. (1999) we observed that the halogenated molecule was also active against the fungus, Verticillium dahliae, since all the producing strains show bioactivity against the fungus whereas the A153 mutants defective in the ooc gene cluster had lost the anti-Verticillium activity (Fig. 1H and S1B). Interestingly, as we had observed previously in the P. ultimum inhibition assays, MMnO15 also showed reduced bioactivity against V. dahliae (Fig. 1H). In our study we have also verified that oocynin A inhibits the growth of other fungal plant pathogens including Alternaria solani, Fusarium oxysporum and, to a lesser degree, Thanatephorus cucumeris (Fig. S1C-E). However, oocynin A was not active against the ascomycetes Saccharomyces cerevisiae and Schizosaccharomyces pombe (data not shown). Serratia plymuthica A153 and S. marcescens MSU97 also show antibacterial activity against Gram-positive and Gram-negative bacteria (Fig. S10). However, the non-oocynin A producing mutants of A153 showed the same antibacterial properties (Fig. S11), whereas these activities were not detected in the other oocynin A producers Ech703 and 4Rx13 (Fig. S10), showing that, in accordance with previous observations (8), oocynin A is not responsible for the antibacterial properties.

A similar result was observed with Caenorhabditis elegans. The nematode C. elegans has been used previously as a model system for the in vivo identification of bacterial virulence factors (84). A153 and MSU97 showed a dramatic virulence in C. elegans assays (Fig. S12). However, the non-oocynin A producing mutants killed C. elegans at the same rate as the wild type strain (Fig. S13), whereas 4Rx13, Ech703 show reduced, or no, virulence against the nematode (Figs. S12), implying that oocynin A is unlikely to be responsible for the observed virulence seen in MSU97 and A153 (Fig. S12).

Concluding remarks

Oocynin A is an antifungal and anti-oomycete halogenated macrolide that has been shown also to possess antitumor properties. Therefore this molecule, or analogues, may have considerable utility and promise for crop disease biocontrol, and/or in medical chemotherapy (5,6,9,11). In this work we describe for the first time the identification of the ooc biosynthetic gene cluster in four different strains of plant-associated enterobacteria, including bacteria from Serratia and Dickeya genera. The ooc biosynthetic gene cluster consists of 23 genes (or 22 genes in D. dadantii Ech703) organized in three transcriptional units. The analysis of the domains and module organization in the PKS proteins has revealed that these oocynin A biosynthetic enzymes belong to the subclass of trans-AT PKSs. In accordance with this, uncommon domain orders and splits of modules between
PKS proteins were found in the ooc cluster products, similar to other trans-AT PKSs involved in biosynthesis of bryostatin (73), rhizopodin (60), rhizoxin (61) and sorangicin (62). A HCS cassette, required to modify the oocydin A PKS-bound intermediate, was also identified in the ooc gene cluster.

The broad spectrum of biological activities of oocydin A has made it an attractive compound for chemical studies. Although its chemical synthesis has been accomplished, the overall yield of only 1.3% for the best current synthesis remains inefficient (9). Thus, our definition of the genes responsible for the biosynthesis of oocydin A and our proposed biosynthetic model now provide an excellent opportunity to investigate the enzymatic mechanisms involved and the nature of the biosynthetic machinery. By a combination of molecular genetics and (bio)chemistry we can now test the model for biosynthesis, and confirm the assembly pathway. Furthermore, given the distribution of the ooc cluster in different enterobacterial genera and in non-conserved genetic contexts, we are now in a position to investigate how the production of this curious bioactive molecule is regulated. Future research will provide information that enables knowledge-based strategies for enhancing productivity and for generating novel analogue haterumalides by synthetic biology methods. Such novel analogues might have important agricultural, pharmacological and chemotherapeutic applications.

REFERENCES

1. Fisher, M.C., Henk, D.A., Briggs, C.J., Brownstein, J.S., Madoff, L.C., McCraw, S.L., and Gurr, S.J. (2012) Emerging fungal threats to animal, plant and ecosystem health. Nature 484, 186-194.
2. Chandler, D., Bailey, A.S., Tatchell, G.M., Davidson, G., Greaves, J., and Grant, W.P. (2011) The development, regulation and use of biopesticides for integrated pest management. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 366, 1987-1998.
3. Lugtenberg, B., and Kamilova, F. (2009) Plant-growth-promoting rhizobacteria. Annu. Rev. Microbiol. 63, 541-556
4. Raaijmakers, J.M., Paulitz, T.C., Steinberg, C., Alabouvette, C., and Moënne-Loccoz, Y. (2009) The rhizosphere: a playground for soilborne pathogen and beneficial microorganisms. Plant Soil 321, 341-61.
5. Strobel, G., Li, J.Y., Sugawara, F., Koshino, H., Harper, J., and Hess, W.M. (1999) Oocycin A, a chlorinated macrocyclic lactone with potent anti-oomycete activity from Serratia marcescens. Microbiology 145, 3557-3364.
6. Takada, N., Sato, H., Suenaga, K., Arimoto, H., Yamada, K., Ueda, K., and Uemura, D. (1999) Isolation and Structures of Haterumalides NA, NB, NC, ND, and NE, Novel Macrolides from an Okinawan Sponge Ircinia sp. Tetrahedron Lett. 40, 6309-6312.
7. Thaning, C., Welch, C.J., Borowicz, J.J., Hedman, R., and Gerhardson, B. (2001) Suppression of Sclerotinia sclerotiorum apothecial formation by the soil bacterium Serratia plymuthica: identification of a chlorinated macrolide as one of the causal agents. Soil Biol. Biochem. 33, 1817-1826.
8. Sato, B., Nakajima, H., Fujita, T., Takase, S., Yoshimura, S., Kinoshita, T., and Terano, H. (2005) FR177391, a new anti-hyperlipidemic agent from Serratia. I. Taxonomy, fermentation, isolation, physico-chemical properties, structure elucidation and biological activities. J. Antibiot. 58, 634-639.
9. Ueda, M., Yamaura, M., Ikeda, Y., Suzuki, Y., Yoshizato, K., Hayakawa, I., and Kigoshi, H. (2009) Total synthesis and cytotoxicity of haterumalides NA and B and their artificial analogues. J. Org. Chem. 74, 3370-7.
10. Teruya, T., Shimogawa, H., Suenaga, K., and Kigoshi, H. (2004) Biselides A and B, novel macrolides from the Okinawan ascidian Didemnidae sp. Chem. Lett. 33, 1184-1185.
11. Teruya, T., Suenaga, K., Maruyama, S., Kurotaki, M., and Kigoshi, H. (2005) Biselides A–E: novel polyketides from the Okinawan ascidian Didemnidae sp. Tetrahedron 61, 6561–6567.
Identification of oocydin A biosynthetic gene clusters

12. Kigoshi, H., Kita, M., Ogawa, S., Itoh, M., and Uemura, D. (2003) Enantioselective synthesis of 15-epi-haterumalide NA methyl ester and revised structure of haterumalide NA. Org. Lett. 5, 957-960.
13. Hoye, T.R., and Wang, J. (2005) Alkyne haloallylation [with Pd(II)] as a core strategy for macrocycle synthesis: a total synthesis of (-)-haterumalide NA/(-)-oocydin A. J Am. Chem. Soc. 127, 6950-6951.
14. Sattely, E.S., Fischbach, M.A., and Walsh, C.T. (2008) Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. Nat. Prod. Rep. 25, 757-793.
15. Neumann, C.S., Fujimori, D.G., and Walsh, C.T. (2008) Halogenation strategies in natural product biosynthesis. Chem. Biol. 15, 99-109.
16. Hertweck, C. (2009) The biosynthetic logic of polyketide diversity. Angew. Chem. Int. Ed. Engl. 48, 4688-4716.
17. Fischbach, M.A., and Walsh, C.T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem. Rev. 106, 3468-3496.
18. Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z., and Graham, M.W. (1989) Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucleic Acids Res. 17, 3469-3478.
19. Herrero, M., de Lorenzo, V., and Timmis, K.N. (1990) Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. J. Bacteriol. 172, 6557-6567.
20. Kaniga, K., Delor, I., and Cornielis, G.R. (1991) A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the blaA gene of Yersinia enterocolitica. Gene 109, 137-141.
21. Demarre, G., Guéroult, A.M., Matsumoto-Mashimo, C., Rowe-Magnus, D.A., Marliere, P., and Mazel, D. (2005) A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and their cognate Escherichia coli host strains. Res. Microbiol. 156, 245-255.
22. Bainton, N. J., Stead, P., Chhabra, S. R., Bycroft, B. W., Salmond, G. P., Stewart, G. S., and Williams, P. (1992). N-(3-Oxohexanoyl)-L-Homoserine lactone regulates carbapenem antibiotic production in Erwinia carotovora. Biochem. J. 288, 997-1004.
23. Hökeberg, M., Gerhardson, B., and Johnsson, L. (1997) Biological control of cereal seed-borne diseases by seed bacterization with greenhouse-selected bacteria. Eur. J. Plant Pathol. 103, 25-33.
24. Berg, G., Roskot, N., Steidle, A., Eberl, L., Zock, A., and Smalla, K. (2002) Plant-dependent genotypic and phenotypic diversity of antagonistic rhizobacteria isolated from different Verticillium host plants. Appl. Environ. Microbiol. 68, 3328-3338.
25. Dennis, J.J., and Zylstra, G.J. (1998) Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of Gram-negative bacterial genomes. Appl. Environ. Microbiol. 64, 2710-2715.
26. Roberts, K.J. (2010) Quorum sensing in the mouse intestinal pathogen Citrobacter rodentium. Ph.D. thesis. University of Cambridge, Cambridge, United Kingdom.
27. Gränter, N. J. (1983). A broad-host-range cloning vector transposable to various replicons. Gene 21, 133-143.
28. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
29. Van Domselaar, G.H., Stothard, P., Shrivastava, S., Cruz, J.A., Guo, A., Dong, X., Lu, P., Szafron, D., Greiner, R., and Wishart, D.S. (2005) BASys: a web server for automated bacterial genome annotation. Nucleic Acids Res. 33, W455-459.
30. Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A., Weber, T., Takano, E., and Breitling, R. (2011) antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. 39, W339-346.
Identification of oocydin A biosynthetic gene clusters

31. Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* **32**, W273-279.

32. Delcher, A.L., Harmon, D., Kasif, S., White, O., and Salzberg, S.L. (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**, 4636-4641.

33. Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Feng, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.J., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mulkohandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., and Zhang, D. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**, 225-9.

34. Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A., and Finn, R.D. (2012) The Pfam protein families database. *Nucleic Acids Res.* **40**, D290-301.

35. Fineran, P.C., Everson, L., Slater, H., and Salmond, G.P. (2005) A GntR family transcriptional regulator (PigT) controls gluconate-mediated repression and defines a new, independent pathway for regulation of the tripyrrole antibiotic, prodigiosin, in *Serratia Microbiology* **151**, 3833-3845.

36. Coulthurst, S.J., Williamson, N.R., Harris, A.K., Spring, D.R., and Salmond, G.P. (2006) Metabolic and regulatory engineering of *Serratia marcescens*: mimicking phage-mediated horizontal acquisition of antibiotic biosynthesis and quorum-sensing capacities. *Microbiology* **152**, 1899-1911.

37. Slater, H., Crow, M., Everson, L., and Salmond, G.P. (2003) Phosphate availability regulates biosynthesis of two antibiotics, prodigiosin and carbapenem, in *Serratia* via both quorum-sensing-dependent and -independent pathways. *Mol. Microbiol.* **47**, 303-320.

38. Kurz, C.L., Chauvet, S., André, E., Aurouze, M., Vallet, I., Michel, G.P., Uh, M., Celli, J., Filloux, A., De Bentzmann, S., Steinmetz, I., Hoffmann, J.A., Finlay, B.B., Gorvel, J.P., Ferrandon, D., and Ewbank, J.J. (2003) Virulence factors of the human opportunistic pathogen *Serratia marcescens* identified by *in vivo* screening. *EMBO J.* **22**, 1451-60.

39. O’Toole, G.A., and Kolter, R. (1998) Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Mol. Microbiol.* **28**, 449-461.

40. Cadel-Six, S., Dauga, C., Castets, A.M., Rippka, R., Bouchier, C., Tandeau, de Marsac, N., and Welker, M. (2008) Halogenase genes in nonribosomal peptide synthetase gene clusters of *Microcystis* (cyanobacteria): sporadic distribution and evolution. *Mol. Biol. Evol.* **25**, 2031-2041.

41. Foulston, L.C., and Bibb, M.J. (2010) Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13461-13466.

42. Fujimori, D.G., Hrvatin, S., Neumann, C.S., Strieker, M., Marahiel, M.A., and Walsh, C.T. (2007) Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16498-16503.

43. Reeves, C.D., Hu, Z., Reid, R., and Kealey, J.T. (2008) Genes for the biosynthesis of the fungal polyketides hypothemycin from *Hypomyces subiculosus* and radicicol from *Pochonia chlamydosporia*. *Appl. Environ. Microbiol.* **74**, 5121-5129.

44. Xiao, Y., Li, S., Niu, S., Ma, L., Zhang, G., Zhang, H., Zhang, G., Ju, J., and Zhang, C. (2011) Characterization of tiacumicin B biosynthetic gene cluster affording diversified tiacumicin analogues and revealing a tailoring dihalogenase. *J. Am. Chem. Soc.* **133**, 1092-1105.

45. Wagner, C., El Omari, M., and König, G.M. (2009) Biohalogenation: nature's way to synthesize halogenated metabolites. *J. Nat. Prod.* **72**, 540-553.

46. McGowan, S.J., Sebaihia, M, Porter, L.E., Stewart, G.S., Williams, P., Bycroft, B.W., and Salmond, G.P. (1996) Analysis of bacterial carbapenem antibiotic production genes reveals a novel beta-lactam biosynthesis pathway. *Mol. Microbiol.* **22**, 415-426.

47. Harris, A.K., Williamson, N.R., Slater, H., Cox, A., Abbasi, S., Foulds, I., Simonsen, H.T., Leeper, F.J., and Salmond, G.P. (2004) The *Serratia* gene cluster encoding biosynthesis of the red antibiotic, prodigiosin, shows species- and strain-dependent genome context variation. *Microbiology* **150**, 3547-3560.

48. Ramsay, J.P., Williamson, N.R., Spring, D.R., and Salmond, G.P. (2011) A quorum-sensing molecule acts as a morphogen controlling gas vesicle organelle biogenesis and adaptive flotation in an enterobacterium. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14932-14937.
Identification of oocydin A biosynthetic gene clusters

49. Williamson, N.R., Simonsen, H.T., Ahmed, R.A., Goldet, G., Slater, H., Woodley, L., Leeper, F.J., and Salmond, G.P. (2005) Biosynthesis of the red antibiotic, prodigiosin, in *Serratia*: identification of a novel 2-methyl-3-n-amyl-pyrole (MAP) assembly pathway, definition of the terminal condensing enzyme, and implications for undecylprodigiosin biosynthesis in *Streptomyces*. *Mol. Microbiol.* **56**, 971-89.

50. Piel, J. (2010) Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* **27**, 996-1047.

51. Keatinge-Clay, A.T., Shelat, A.A., Savage, D.F., Tsai, S.C., Miercke, L.J., O’Connell, J.D. 3rd, Khosla, C., and Stroud, R.M. (2003) Catalysis, specificity, and ACP docking site of *Streptomyces coelicolor* malonyl-CoA:ACP transacylase. *Structure* **11**, 147-54.

52. Yadav, G., Gokhale, R.S., and Mohanty, D. (2003) Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* **328**, 335-363.

53. Dickschat, J.S., Vergnolle, O., Hong, H., Garner, S., Bidgood, S.R., Dooley, H.C., Deng, Z., Leadlay, P.F., and Sun, Y. (2011) An additional dehydratase-like activity is required for lankacidin antibiotic biosynthesis. *Chembiochem* **12**, 2408-2412.

54. Irschik, H., Kopp, M., Weissman, K.J., Buntin, K., Piel, J., and Müller, R. (2010) Analysis of the sorangicin gene cluster reinforces the utility of a combined phylogenetic/retrobiosynthetic analysis for deciphering natural product assembly by *trans*-AT PKS. *ChemBioChem* **11**, 1840-1849.
65. Jensen, K., Niederkrüger, H., Zimmermann, K., Vagstad, A.L., Moldenhauer, J., Brendel, N., Frank, S., Pöplau, P., Kohlhaas, C., Townsend, C.A., Oldiges, M., Hertweck, C., and Piel, J. (2012) Polyketide proofreading by an acyltransferase-like enzyme. Chem. Biol. 19, 329-339.
66. Mattheus, W., Gao, L.J., Herdevijn, P., Landuyt, B., Verhaegen, J., Masschelein, J., Volkaert, G., and Lavigne, R. (2010) Isolation and purification of a new kalimantacin/batumin-related polyketide antibiotic and elucidation of its biosynthesis gene cluster. Chem. Biol. 17, 149-159.
67. Aparicio, J.F., Molanar, I., Schwecke, T., Koning, A., Haydock, S.F., Khaw, L.E., Staunton, J., and Leadlay, P.F. (1996). Organization of the biosynthetic gene cluster of rapamycin in Streptomyces hygroscopicus: analysis of the enzymatic domains in the modular polyketide synthase. Gene 169, 9-16.
68. Reid, R., Piagentini, M., Rodrigue, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D.V., Hutchinson, C.R., and McDaniel, R. (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide syntheses. Biochemistry 42, 72-79.
69. Keatinge-Clay, A. (2008) Crystal structure of the erythromycin polyketide synthase dehydratase. J. Mol. Biol. 384, 941-53.
70. Gu, L., Jia, J., Liu, H., Håkansson, K., Gerwrick, W.H., and Sherman, D.H. (2006) Metabolic coupling of dehydration and decarboxylation in the curacin A pathway: functional identification of a mechanistically diverse enzyme pair. J. Am. Chem. Soc. 128, 9014-9015.
71. Konz, D., and Marahiel, M.A. (1999). How do peptide synthetases generate structural diversity? Chem. Biol. 6, R39-R48.
72. Marahiel, M.A. (1997). Protein templates for the biosynthesis of peptide antibiotics. Chem. Biol. 4, 561-567.
73. Sudek, S., Lopanik, N.B., Waggoner, L.E., Hildebrand, M., Anderson, C., Liu, H., Patel, A., Sherman, D.H., and Haygood, M.G. (2007) Identification of the putative bryostatin polyketide synthase gene cluster from "Candidatus Endobugula sertula", the uncultivated microbial symbiont of the marine bryozoan Bugula neritina. J. Nat. Prod. 70, 67-74.
74. Sun, Y., Hong, H., Gillies, F., Spencer, J.B., and Leadlay, P.F. (2008) Glyceryl-S-acyl carrier protein as an intermediate in the biosynthesis of tetronate antibiotics. Chembiochem. 9, 150-156.
75. Geders, T.W., Gu, L., Mowers, J.C., Liu, H., Gerwick, W.H., Håkansson, K., Sherman, D.H., and Smith, J.L. (2007) Crystal structure of the ECH2 catalytic domain of CurF from Lyngbya majuscula. Insights into a decarboxylase involved in polyketide chain beta-branching. J. Biol. Chem. 282, 35954-35963.
76. Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N., and Matsunaga, S. (2004) Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge Theonella swinnoei. Proc. Natl. Acad. Sci. U.S.A. 101, 16222-16227.
77. Simunovic, V., Zapp, J., Rachid, S., Krug, D., Meiser, P., and Müller, R. (2006) Myxovirescin A biosynthesis is directed by hybrid polyketide synthases/nonribosomal peptide synthetase, 3-hydroxy-3-methylglutaryl-CoA synthases, and trans-acting acyltransferases. Chembiochem. 7, 1206-1220.
78. Tang, G.L., Cheng, Y.Q., and Shen, B. (2006) Polyketide chain skipping mechanism in the biosynthesis of the hybrid nonribosomal peptide-polyketide antitumor antibiotic leinamycin in Streptomyces atroolivaceus S-140. J. Nat. Prod. 69, 387-93.
79. Perlova, O., Gerth, K., Kaiser, O., Hans, A., and Müller, R. (2006) Identification and analysis of the chivosazol biosynthetic gene cluster from the myxobacterial model strain Sorangium cellulosum So ce56. J. Biotechnol. 121, 174-191.
80. Zhao, C., Coughlin, J.M., Ju, J., Zhu, D., Wendt-Pienkowski, E., Zhou, X., Wang, Z., Shen, B., and Deng, Z. (2010) Oxazolomycin biosynthesis in Streptomyces albus JA3453 featuring an
Identification of oocydin A biosynthetic gene clusters

"acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. J. Biol. Chem. 285, 20097-20108.

81. Bumpus, S.B., Magarvey, N.A., Kelleher, N.L., Walsh, C.T., and Calderone, C.T. (2008) Polyunsaturated fatty-acid-like trans-enoyl reductases utilized in polyketide biosynthesis. J. Am. Chem. Soc. 130, 11614-11616.

82. Ross-Anderson, J.L., and Chapman, S.K. (2006) Molecular mechanisms of enzyme-catalysed halogenation. Mol. BioSyst. 2, 350-357

83. Podzelnitska, K., Latimer, R., Bhattacharya, A., Vining, L.C., Zechel, D.L., and Jia, Z. (2010) Chloramphenicol biosynthesis: the structure of CmlS, a flavin-dependent halogenase showing a covalent flavin-aspartate bond. J. Mol. Biol. 397, 316-331.

84. Coulthurst, S.J., Kurz, C.L., and Salmond, G.P.C. (2004) luxS mutants of Serratia defective in autoinducer-2-dependent 'quorum sensing' show strain-dependent impacts on virulence and production of carbapenem and prodigiosin. Microbiology 150, 1901-1910.

Acknowledgements-We thank members of the Salmond and Leeper groups for helpful discussions. We thank G. Strobel, C. Thaning, R. Cooper and Bio-Iliberis Research and Development for the generous gift of microbial strains.

FOOTNOTES

*M.A.M. was supported by a Spanish Government postdoctoral research contract BVA- 2009-0200.
†The nucleotide sequences of the Serratia marcescens MSU97 and Serratia plymuthica A153 oocydin A gene clusters have been submitted to the GenBank Data Bank with accession numbers JX315603 and JX315604 respectively.

1To whom correspondence should be addressed: Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK, CB2 1QW. Tel: +44 (0)1223 333650; Fax: +44 (0)1223 766108; E-mail: gpcs2@cam.ac.uk.

2The abbreviations used are: ACP, acyl carrier protein; AT, acyltransferase; DH, dehydratase; EH, enoyl-CoA hydratase; ER, enoylreductase; HCS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMG, 3-hydroxy-3-methylglutaryl; KR, ketoreductase; KS, ketosynthase; MT, methyltransferase; NRPS, nonribosomal peptide synthetase; Ooc, oocydin A; PKS, polyketide synthase; TE, thioesterase.

FIGURES AND LEGENDS

Figure 1. Antifungal and anti-oomycete activities of the oocydin A producing strains. Bioactivities against Pythium ultimum (A-F) and Verticillium dahliae (G-H) are shown. Pythium ultimum bioassays were performed by inoculating 5 mm diameter mycelial plugs in the centre (E-F) or at four different positions at the edge (A-D) of the bacterial-inoculated plates, as described in experimental procedures. Pythium ultimum inoculation points are indicated with a black dot. For the fungicide assays, an indicator Verticillium dahliae top agar was prepared as described in experimental procedures and five microlitres of overnight cultures of the selected strains were spotted on the surface of the fungal indicator agar lawn. The bioassays were repeated at least five times and representative results are shown. Pythium ultimum and Verticillium dahliae pictures were taken after 72 h and 96 h of incubation at 25 ºC, respectively.

Figure 2. Proposed biosynthetic pathway for oocydin A. A. Genetic organization of the ooc biosynthetic gene cluster in Serratia plymuthica A153. The same genetic organization was found in S. marcescens MSU97, S. odorifera 4Rx13 and D. dadantii Ech703. Location of the Tn-KRCPN1 transposon insertions and marker exchange mutants are indicated by black and red arrows respectively. Modular polyketide synthase genes are shown in red. B, Model for the biosynthesis of oocydin A showing the structures of some of the proposed intermediates. T, acyl carrier protein; AT, acyltransferase; C, NRPS condensation domain; Cl, chlorinase; DH, dehydratase; EH, enoyl-CoA
Identification of oocydin A biosynthetic gene clusters

hydratase; FH, FkbH-like domain; HCS, 3-hydroxy-3-methylglutaryl-CoA synthase cassette; KR, ketoreductase; KS, ketosynthetase; MT, methyltransferase; TE, thioesterase. Domains in grey are proposed to be inactive or not involved in chain-extension. Domains in greyish colors may be active.

Figure 3. LC-MS analysis of culture supernatant extracts of oocydin A producing strains and mutants in the ooc gene cluster. A, Serratia marcescens MSU97 wild type strain; B, Serratia plymuthica A153 wild type strain; C, Serratia odorifera 4Rx13 wild type strain; D, Dickeya dadantii Ech703 wild type strain; E, Serratia plymuthica A153 MmnO15 mutant strain; F, Serratia plymuthica A153 MmnO13 mutant strain. The chromatograms show Selective Ion Monitoring at m/z 493 ± 0.5. The peaks at 11.4 min correspond to oocydin A (m/z 493.1 [M+Na]+). The chromatograms of the MSU97 (A) wild type A153 (B) and Ech703 (D) strains were scaled down 4-, 20- and 2-times, respectively, as compared with the ooc mutants MmnO15 (E) and MmnO13 (F). Fragmentation patterns of oocydin A are shown in Fig. S2. Chemical analyses of other mutants in the ooc gene cluster are shown in Fig. S3.

Figure 4: DNA homology between the oocydin A gene cluster of Serratia marcescens MSU97 and the oocydin A gene cluster of the other producing strains. The alignments represent the percentage of homology between the ooc gene cluster of S. marcescens MSU97 and the ooc gene cluster of S. plymuthica A153 (A), S. odorifera 4Rx13 (B) and D. dadantii Ech703 (C). Alignments were performed using wgVISTA (31).

Figure 5: Analysis of the transcriptional units in the oocydin A gene cluster by RT-PCR. A, Schematic representation of the ooc gene cluster in Serratia marcescens MSU97. Lines labelled 1-22 above the gene cluster represent the regions amplified in the RT-PCR shown in (B). Numbers below the arrows represent the intergenic distance between contiguous genes, while negative numbers indicate overlapping genes. Black arrows represent the three transcriptional units found in the ooc gene cluster. B, Transcript analysis by RT-PCR using primers designed to span the intergenic region between two adjacent genes. For each region, three PCR analyses were carried out: Plus symbol (+), RT-PCR on cDNA; Minus symbol (-), negative control with no reverse transcriptase; gDNA, positive control with genomic DNA as template.
Identification of oocydin A biosynthetic gene clusters

Table 1. Bacteria, oomycete, fungi, plasmids and phages used in this study

| Bacteria/fungi/oomycete/ phage/plasmid(s) | Genotype or relevant characteristics | Reference or source |
|------------------------------------------|-------------------------------------|---------------------|
| Escherichia coli DH5α                    | supE4lacU169lacZΔM15 hsdR17 (rK mK - recA1 endA1 gyrA96 thi-1 relA1) | 18                  |
| E. coli CC118pir                         | araD, Δara, leu), ΔlacZ74, pHeA20, gapK, thi-1, rpsE, rpoB, argE, recA1, λpir | 19                  |
| E. coli HH26                             | Mobilizing strain for conjugal transfer | 20                  |
| E. coli EP1100-1TR                       | F’ merA Δmrr-hsdRMS-mcrBC (Ø80lacZ ΔM15) ΔlacX74 recA1 | Epicentre          |
| E. coli β1263                           | F’ RP4-2-Tc::Mu DapA4::erm-pir, Km’Em’ | 21                  |
| E. coli ESS                             | β-Lactam super sensitive indicator strain | 22                  |
| Bacillus subtilis JH642                  | pHeA1 trpC2 | J.A. Hoch          |
| Serratia plymuthica A153                | Wild type, rhizosphere isolate | 23                  |
| MMnO2                                    | A153 transposon mutant oocS::Tn-KRCPN1; Km’ | This study         |
| MMnO4                                    | A153 transposon mutant oocQ::Tn-KRCPN1; Km’ | This study         |
| MMnO9                                    | A153 transposon mutant oocV::Tn-KRCPN1; Km’ | This study         |
| MMnO13                                   | A153 transposon mutant oocJ::Tn-KRCPN1; Km’ | This study         |
| MMnO14                                   | A153 transposon mutant oocC::Tn-KRCPN1; Km’ | This study         |
| MMnO15                                   | A153 transposon mutant oocU::Tn-KRCPN1; Km’ | This study         |
| MMnO18                                   | A153 transposon mutant oocQ::Tn-KRCPN1; Km’ | This study         |
| OocK                                     | A153 oocK::Km, constructed by marker exchange, Km’ | This study         |
| OocM                                     | A153 oocM::Km, constructed by marker exchange, Km’ | This study         |
| S. marcescens MSU97                      | Wild type, plant epiphyte, pigmented | 5                   |
| S. odorifera 4Rx13                       | Wild type, rhizosphere isolate | 24                  |
| Dickeya dadianii Ech703                  | Wild type, plant pathogen | CP001654          |
| Fungi/oomycete strains                   |                                      |                     |
| Pythium ultimum                          | Wild type, plant pathogen | R. Cooper          |
| Verticillium dahliae 5368                | Wild type, plant pathogen | R. Cooper          |
| Thanatephorus cucumeris                  | Wild type, plant pathogen | STCC 2813         |
| Alternaria solani                        | Wild type, plant pathogen | STCC 2997        |
| Fusarium oxysporum                      | Wild type, plant pathogen | STCC 2715        |
| Saccharomyces cerevisiae                 | Wild type | S.G. Oliver       |
| Schizosaccharomyces pombe               | Wild type | J. Mata          |
| Phages/Plasmids                          |                                      |                     |
| ϕMAM1                                   | Generalized transducing phage for S. plymuthica A153 | Matilla et al. (unpublished) |
| pWEB-TNC                                 | Ap’, Cm’; Cosmid cloning vector | Epicentre           |
| pKN101                                   | Sm’, oriR6K mob sacBR,             | 20                  |
| pUC18Not                                 | Ap’, identical to pUC18 but with two NotI sites flanking pUC18 polyninker | 19                  |
| p34S-Km3                                 | Km’, Ap’, Km3 antibiotic cassette | 25                  |
| pKCPRN1                                  | Km’, Tc’; Derivative of pDS1028uidA with the uidA and cat genes replaced with lacZ and km genes. | 26                  |
| pMAMV52                                  | Ap’, 2.2 kb EcoRI/PstI PCR product containing oocK was inserted into the EcoRI/PstI sites of pUC18Not | This study         |
| pMAMV53                                  | Ap’, Km’, 0.95 kb Smal fragment containing km3 cassette of p34S- Km3 was inserted into HindIII site of oocK in pMAMV52 | This study         |
| pMAMV54                                  | Sm’, Km’, 3.2 kb NotI fragment of pMAMV53 was cloned at the same site in pKNG101 | This study         |
| pMAMV60                                  | Ap’, 2.0 kb EcoRI/PstI PCR product containing oocM was inserted into the EcoRI/PstI sites of pUC18Not | This study         |
| pMAMV61                                  | Ap’, Km’, Replacement of 0.26 kb BamHI fragment internal to oocM of pMAMV60 for a 0.96 kb BamHI km3 cassette of p34S- Km3 | This study         |
| pMAMV62                                  | Sm’, Km’, 2.7 kb NotI fragment of pMAMV61 was cloned at the same site in pKNG101 | This study         |
| pNJ5000                                  | Tc’; Mobilizing plasmid used in marker exchange | 27                  |

*Ap, ampicillin; Km, kanamycin; Cm, chloramphenicol; Rif, rifampin; Sm, streptomycin; Tc, tetracycline; Em, erythromycin; STCC, Spanish Type Culture Collection.*
Table 2. Deduced functions of ORFs in *Serratia plymuthica* A153 ooc biosynthetic gene cluster

| Protein | Size (aa) | Proposed function* | Sequence similarity (Protein, Origin) | Identity/ Similarity (%) | GenBank accession no. |
|---------|-----------|--------------------|---------------------------------------|--------------------------|-----------------------|
| OocA    | 295       | Alpha/beta hydrolase | BY123_C008640, Burkholderia sp. Y123 | 28/48                    | AE93010               |
| OocB    | 201       | Efflux protein      | RSMK1349, Ralstonia solanacearum MoI2K | 40/60                    | YP_002254246          |
| OocC    | 249       | Enoyl-CoA hydratase | BatE, Pseudomonas fluorescens         | 62/75                    | ADD82946              |
| OocD    | 258       | Enoyl-CoA hydratase | Pks3, Burkholderia pseudomallei 406e  | 50/69                    | ZP_04967830           |
| OocE    | 420       | HMG-CoA synthase    | PksG, Paenibacillus polymyxa E81     | 66/82                    | YP_003871376          |
| OocF    | 417       | Ketosynthase        | KAOT1_04360, Kordia algicida OT-1     | 43/61                    | ZP_02163873           |
| OocG    | 83        | Acyl carrier protein| AcpK, Streptomyces griseoflavus Tu7000 | 43/69                    | ZP_07315130           |
| OocH    | 44        | Hypothetical protein| -                                      | -                       | -                     |
| OocI    | 53        | Hypothetical protein| -                                      | -                       | -                     |
| OocJ    | 3657      | -                   |                                        |                          |                       |
| OocK    | 434       | Flavin-dependent monooxygenase | PedG, symbiont bacterium of *Paederus fuscipes* | 61/78                    | AAS47561              |
| OocL    | 2403      | PKS (T-KS-T-KS-KR-T) | BaeL, *Bacillus amyloliquefaciens* FZB42 | 46/66                    | YP_001421293          |
| OocM    | 377       | Flavin-dependent monooxygenase | ORF2, Nostoc sp. "*Peltigera membranacea" | 54/69                    | ADA9238               |
| OocN    | 6614      | PKS (T-KS-T-KS-DH-KR-T-KS-DH-KR-T-KS-KS-T-KS) | -                                      |                          |                       |
| OocO    | 89        | Acyl carrier protein | BSn5_20710, *Bacillus subtilis* BSn5  | 32/60                    | ZP_07900867           |
| OocP    | 389       | Hypothetical protein | BuboB_26000, Burkholderia ubonensis Bu | 25/40                    | YP_02381200           |
| OocQ    | 119       | Hypothetical protein | SpBuddy_0991, *Sphaerochaeta globus* st. Buddy | 25/42                    | YP_00247002           |
| OocR    | 2264      | PKS (T-KS-T-KS-KR)  | BBR47_39880, *Brevibacillus brevis* 100599 | 33/51                    | YP_002773469          |
| OocS    | 5341      | PKS (T-KS-T-KS-T-KS-DH-KR-T-KS-DH-KR-T-KS-KS-KS-TE-KS-C) | -                                      |                          |                       |
| OocT    | 276       | Hypothetical protein | KrcA_1408, *Kiedonobacter racemifer* DSM 44963 | 26/46                    | ZP_06972179           |
| OocU    | 482       | Flavin-dependent nitroreductase | BatK, *Pseudomonas fluorescens*    | 53/73                    | ADD82952              |
| OocV    | 624       | Acyl transferase (AT1-AT2) | -                                      |                          |                       |
| OocW    | 372       | Acyl transferase    | PedD, *Burkholderia pseudomallei* 1710b | 49/65                    | YP_337765             |

*AT, acyl transferase; C, NRPS condensation domain; HMG, hydroxymethylglutaryl; DH, dehydratase; EH, enoyl-CoA hydratase; KS, keto synthase; KR, keto reductase; MT, methyl transferase; T, acyl carrier protein; TE, thioesterase.*
Identification of oocydin A biosynthetic gene clusters

Figure 1
Identification of oocydin A biosynthetic gene clusters

Figure 2

A

B

Proteins

Modules

OocJ

OocL

OocN

OocR

OocS

Load

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

R = CO₂ or H

OocC-G

hydroxylation

OocK or M

reduced by ER OocU

Cl, O₂

AT

AspCoA
Figure 3

Identification of oocydin A biosynthetic gene clusters

(A) Relative Intensity (%)
(B) Relative Intensity (%)
(C) Relative Intensity (%)
(D) Relative Intensity (%)
(E) Relative Intensity (%)
(F) Relative Intensity (%)

0 2 4 6 8 10 12 14 16 18

Relative Intensity (%) vs. t (min)
Identification of oocynin A biosynthetic gene clusters

Figure 4

A

B

C

Downloaded from http://www.jbc.org by guest on March 23, 2020
Identification of oocydin A biosynthetic gene clusters

Figure 5

A

B

+, RT
-: No RT
c: gDNA

Downloaded from http://www.jbc.org/ by guest on March 23, 2020
Bacterial biosynthetic gene clusters encoding the anti-cancer haterumalide class of molecules: biogenesis of the broad spectrum antifungal and anti-oomycete compound, oocydin A
Miguel A. Matilla, Henning Stockmann, Finian J. Leeper and George P.C. Salmond

J. Biol. Chem. published online September 24, 2012

Access the most updated version of this article at doi: 10.1074/jbc.M112.401026

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

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

Supplemental material:
http://www.jbc.org/content/suppl/2012/09/24/M112.401026.DC1