Microbiology (2008), **154**, 2070–2083 DOI 10.1099/mic.0.2008/016444-0

Genetic and functional characterization of the gene cluster directing the biosynthesis of putisolvin I and II in *Pseudomonas putida* strain PCL1445

Jean-Frédéric Dubern,[1,2] Eric R. Coppoolse,[3] Willem J. Stiekema[3] and Guido V. Bloemberg[1,4]

Correspondence
Guido V. Bloemberg
bloemberg@immv.uzh.ch

[1]Institute of Biology, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands

[2]Institute of Infection, Immunity & Inflammation, Centre for Biomolecular Sciences, University of Nottingham, Nottingham NG27 2RD, UK

[3]Wageningen University, Laboratory of Bioinformatics, Genome Informatics Group, PO Box 8128, 6700 ET Wageningen, The Netherlands

[4]Institute of Medical Microbiology, University of Zürich, Gloriastr. 32, CH-8006 Zürich, Switzerland

Received 4 January 2008
Revised 10 March 2008
Accepted 19 March 2008

Pseudomonas putida PCL1445 secretes two cyclic lipopeptides, putisolvin I and putisolvin II, which possess a surface-tension-reducing ability, and are able to inhibit biofilm formation and to break down biofilms of *Pseudomonas* species including *Pseudomonas aeruginosa*. The putisolvin synthetase gene cluster (*pso*) and its surrounding region were isolated, sequenced and characterized. Three genes, termed *psoA*, *psoB* and *psoC*, were identified and shown to be involved in putisolvin biosynthesis. The gene products encode the 12 modules responsible for the binding of the 12 amino acids of the putisolvin peptide moiety. Sequence data indicate that the adenylation domain of the 11th module prioritizes the recognition of Val instead of Leu or Ile and consequently favours putisolvin I production over putisolvin II. Detailed analysis of the thiolation domains suggests that the first nine modules recognize the D form of the amino acid residues while the two following modules recognize the L form and the last module the L or D form, indifferently. The *psoR* gene, which is located upstream of *psoA*, shows high similarity to *luxR*-type regulatory genes and is required for the expression of the *pso* cluster. In addition, two genes, *macA* and *macB*, located downstream of *psoC* were identified and shown to be involved in putisolvin production or export.

INTRODUCTION

Pseudomonas putida PCL1445 was isolated from soil heavily polluted with polyaromatic hydrocarbons (Kuiper *et al.*, 2001) and produces two surface-active compounds, putisolvin I and putisolvin II, which have been identified as cyclic lipopeptides (CLPs) with a novel structure consisting of a 12 aa peptide linked to a hexanoic lipid chain and containing an ester linkage between the ninth serine residue and the C-terminal carboxyl group (Kuiper *et al.*, 2004). Putisolvins I and II were shown (i) to reduce the surface tension of the medium, (ii) to increase the formation of an emulsion with toluene, (iii) to stimulate swarming motility and (iv) to inhibit biofilm formation, and to degrade existing biofilms (Kuiper *et al.*, 2004). It was hypothesized that biosurfactants produced by PCL1445 alter the cell-surface hydrophobicity and may influence the interaction between the individual cells in biofilms (Kuiper *et al.*, 2004).

The diverse roles and functions of CLPs seem to be determined by their structure. However, this structure–function relationship is still not well understood.

CLPs are versatile molecules with antimicrobial, cytotoxic and surfactant properties. CLPs are produced by members of the genera *Bacillus*, *Serratia*, *Burkholderia* and *Pseudomonas* (Neu, 1996; Lindum *et al.*, 1998; Huber *et al.*, 2002; Raaijmakers *et al.*, 2006). The CLP viscosinamide produced by a *Pseudomonas* sp. isolated from the sugar beet rhizosphere has antibiotic properties towards root-pathogenic fungi (Nielsen *et al.*, 1999). CLPs of *Bacillus* sp.

Abbreviations: BAC, bacterial artificial chromosome; CLP, cyclic lipopeptide; NRPS, nonribosomal peptide synthetase.

The GenBank/EMBL/DDBJ accession number for the nucleotide sequence of the *P. putida* PCL1445 *oprM-psoR-psoA-psoB-psoC-macA-macB* DNA region is DQ151887.

were reported to permeabilize membranes, resulting in pore formation and haemolysis (Grangemard *et al.*, 2001) and to play a role in bacterial attachment to surfaces (Neu, 1996). Lipopeptides produced by *Serratia* (Lindum *et al.*, 1998) and *Burkholderia* (Huber *et al.*, 2002) were shown to be essential for the stimulation of swarming motility and thus could contribute to the regulation of biofilm formation (Huber *et al.*, 2002).

Structural identification showed one difference between putisolvin I and II. This consists of a Val-11 in putisolvin I and an Ile-11 or Leu-11 in putisolvin II, which accounts for a mass difference of 14 Da (Kuiper *et al.*, 2004). Other known *Pseudomonas* lipopeptides, such as viscosinamide (Nielsen *et al.*, 1999), syringomycin (Hutchison *et al.*, 1995), amphisin (Sørensen *et al.*, 2001), tensin (Henriksen *et al.*, 2000) and arthrofactin (Roongsawang *et al.*, 2003), have a shorter amino acid moiety and the fatty acid chain is longer. Cyclization also appears to be different for the previously described lipopeptides, in which an ester linkage is formed between the C-terminal carboxyl group and the side-chain of the first or third amino acid. In contrast, this linkage involves the ninth amino acid residue in the case of putisolvin I and II (Kuiper *et al.*, 2004). The lipopeptides produced by *Pseudomonas syringae* and *Pseudomonas tolaasii* (Nutkins *et al.*, 1991) have a longer fatty acid chain than putisolvins I and II.

CLPs are produced nonribosomally by large, multifunctional peptide synthetases via a thiotemplate mechanism (Marahiel *et al.*, 1997; Stachelhaus *et al.*, 1999). The genes encoding multimodular nonribosomal peptide synthetases (NRPSs) for syringomycin and syringopeptin (Kleinkauf & von Döhren, 1996; Guenzi *et al.*, 1998; Scholz-Schroeder *et al.*, 2003) of *P. syringae* pv. *syringae* B301D and arthrofactin of *Pseudomonas* sp. MIS38 (Roongsawang *et al.*, 2003) have been cloned and characterized. NRPSs recognize, activate, modify and link the amino acid intermediates to the product (Marahiel *et al.*, 1997). The order and number of the modules of an NRPS protein are, in many cases, collinear to the amino acid sequence of the corresponding peptide moiety of the final CLP molecule ('collinearity rule'). Each module of the NRPSs can be further subdivided into domains, each of which exhibits a single enzymic activity. The adenylation (A) domain is responsible for amino acid recognition and adenylation at the expense of ATP. The thiolation (T) or peptidyl carrier protein domain is the attachment site of the 4′-phosphopantetheine cofactor and serves as a carrier of thioesterified amino acid intermediates. The condensation (C) domain catalyses peptide bond formation between two consecutive amino acids. Modifying domains such as the epimerization (E) domain catalyse the conversion of L-amino acids to their D-isomers. Finally, cyclization and release of the peptide product are catalysed by the C-terminal thioesterase (Te) domain.

P. putida strain PCL1445 was indicated to produce putisolvin I and II via a NRPS (Kuiper *et al.*, 2004). Relatively little is known about the genetic network involving the perception of external factors and the signal-transduction pathways that drive transcription of the CLP biosynthetic genes. The GacS/GacA two-component system appears to serve as a master switch for the production of CLPs in several *Pseudomonas* genera (Bender *et al.*, 1999; Koch *et al.*, 2002; Dubern *et al.*, 2005). Our previous results showed that the DnaK heat-shock protein plays an important role in putisolvin biosynthesis at the transcriptional level and that it is positively regulated by GacA/GacS (Dubern *et al.*, 2005). Putisolvins I and II are strongly upregulated at low temperatures and require DnaK (Dubern *et al.*, 2005). In addition, we showed that a quorum-sensing system consisting of *ppuI*, *rsaL* and *ppuR* regulates the expression of putisolvin production and biofilm formation (Dubern *et al.*, 2006).

This study aimed at the isolation, sequencing and characterization of the putisolvin biosynthetic gene(s) and characterization of the surrounding genes for their possible regulatory role in putisolvin production.

METHODS

Bacterial strains and growth conditions. Bacterial strains used in this study are listed in Table 1. *P. putida* strains were grown in King's medium B (King *et al.*, 1954) or in a defined BM medium (Lugtenberg *et al.*, 1999) supplemented with 2 % (v/v) glycerol (BDH) at 28 °C. *Escherichia coli* strains were grown in Luria–Bertani medium (Sambrook & Russell, 2001) at 37 °C. Media were solidified with 1.8 % agar (Invitrogen). The antibiotics kanamycin, tetracycline or carbenicillin were added, when necessary, to final concentrations of 50, 40, and 100 μg ml^{-1}, respectively.

Construction and screening of Tn*5* library. Transposon mutants were generated by tri-parental mating using pRL1063a, which harbours a Tn*5* transposon carrying the promoterless *luxAB* reporter genes (Wolk *et al.*, 1991), and the helper plasmid pRK2013 (Schnider *et al.*, 1995). Transposants were initially screened for the decreased ability to flatten a droplet of water on Parafilm using cells of a single colony as described previously (Dubern *et al.*, 2005). Culture supernatants of the selected mutants, obtained after growth overnight in KB medium, were analysed for the presence of surfactant production using the drop-collapsing assay (Jain *et al.*, 1991).

To isolate the chromosomal DNA region flanking the Tn*5*, total genomic DNA was isolated and digested with *Eco*RI, the restriction site of which is not present in pRL1063a (Wolk *et al.*, 1991). DNA fragments were circularized by self-ligation and selected for kanamycin resistance, resulting in plasmids maintained via OriV (p15A) and containing genomic DNA regions flanking the Tn*5luxAB*. All DNA techniques were performed as described by Sambrook & Russell (2001). Sequencing of the plasmids was performed by BaseClear. DNA sequences were analysed with the software packages provided by the NCBI (National Center for Biotechnology Information) BLAST network server.

Bacterial artificial chromosome (BAC) DNA analysis. A BAC library of *P. putida* strain PCL1445 was constructed by Bio S&T, using pIndigoBAC-5 (EPICENTRE) as a cloning vector for high-molecular-mass inserts (Table 1).

To isolate a genomic fragment containing the putisolvin biosynthetic gene(s) the resulting fragment library was screened for the presence of

Table 1. Strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Reference or source
Strains		
P. putida		
PCL1445	Wild-type; colonizes grass roots and produces the biosurfactants putisolvin I and II	Kuiper *et al.* (2004)
PCL1436	Tn*5luxAB* derivative of PCL1445; mutated in *psoA*, a lipopeptide synthetase homologue	Kuiper *et al.* (2004)
PCL1626	PCL1445 derivative mutated in *psoR*, a *luxR* transcriptional regulator homologue	This study
PCL1630	Tn*5luxAB* derivative of PCL1445; mutated in *psoA*	This study
PCL1633	Tn*5luxAB* derivative of PCL1445; mutated in *psoB*	Dubern *et al.* (2005)
PCL1634	Tn*5luxAB* derivative of PCL1445; mutated in *psoA*	This study
PCL1644	PCL1445 derivative mutated in the *macA* homologue; constructed by single homologous recombination using pMP7595	This study
PCL1645	PCL1445 derivative mutated in the *macB* homologue; constructed by single homologous recombination using pMP7597	This study
E. coli		
DH5α	*endA1 gyrSA96 hrdR17* ($r_K^- m_K^-$) *supE44 recA1*; general-purpose host strain used for transformation and propagation of plasmids	Hanahan (1983)
Plasmids		
pIndigoBAC-5	BAC cloning vector (*Hin*dIII-Cloning Ready) for high-molecular-mass DNA insert library derived from pBeloBAC11 *lacZ*, Cm^r	EPICENTRE
pBluescript	General-purpose cloning vector, Cb^r	Stratagene
pRL1063a	Plasmid harbouring a promoterless Tn*5luxAB* transposon, Km^r	Wolk *et al.* (1991)
pRK2013	Helper plasmid for tri-parental mating, Km^r	Schnider *et al.* (1995)
pME6010	Cloning vector which is maintained in *Pseudomonas* strains without selection pressure, Tc^r	Heeb *et al.* (2000)
pME3049	Cloning vector, used for homologous recombination, Tc^r Hg^r	Ditta *et al.* (1980)
pMP5285	pME3049 derivative, lacking the Hg^r gene, used for single homologous recombination, Km^r	Kuiper *et al.* (2004)
pMP5539	pMP5537 derivative harbouring *psoA*::*gfp* transcriptionally active harbouring P_{tac} DsRed, Gm^r Tc^r	Dubern *et al.* (2005)
pMP5540	pMP5538 derivative harbouring *psoA*::*gfp* transcriptionally inactive fused to pMP4669 harbouring P_{tac} DsRed, Gm^r Tc^r	Dubern *et al.* (2005)
pMP7570	pMP5285 containing a 0.58 kb *Eco*RI–*Eco*RI PCR fragment of the central part of *psoR* gene from PCL1445, Km^r	This study
pMP7595	pMP5285 containing a 0.61 kb *Eco*RI–*Eco*RI PCR fragment of the central part of *macA* gene from PCL1445, Km^r	This study
pMP7596	pME6010 containing a 3.6 kb *Eco*RI–*Eco*RI PCR fragment with P_{tac} *macA-macB* genes from PCL1445, used for complementation, Tc^r	This study
pMP7597	pMP5285 containing a 0.5 kb *Eco*RI–*Eco*RI PCR fragment of the central part of *macB* gene from PCL1445, Km^r	This study
pMP7598	pIndigoBAC-5 (*Hin*dIII-Cloning Ready) containing a 152 860 bp chromosomal fragment of PCL1445 with *psoA-psoB-psoC* gene cluster, Cm^r	This study

the entire lipopeptide biosynthetic gene cluster by colony blotting (Sambrook & Russell, 2001) using four different probes. The probes were obtained by PCR using PCL1445 chromosomal DNA as a template and primers derived from sequences of regions flanking Tn*5* mutants oMP907 (5′-GCATGCAAGCGATGAAAGCAGATGACCC-AG-3′) and oMP908 (5′-GCATGCGTCGGCAGGTCCTTCTGATT-GATC-3′), resulting in a 0.75 kb product PCR1, oMP872 (5′-ACCTCAGTGAATGGACCCTTG-3′), and oMP873 (5′-GAGCTG-TTTTTCACGTTCAGC-3′), resulting in a 0.58 kb product PCR2, oMP893 (5′-ATCACATCGCCATGGACCACA-3′), and oMP894 (5′-ACAAGGCCTTCAAGGACCTGCT-3′), resulting in a 0.65 kb product PCR3, oMP891 (5′-ATCGATGAATTCACCGGCCTGG-3′), and oMP892 (5′-CCCAGCAACTGGTTGGCGATCT-3′), resulting in a 0.35 kb product, PCR4.

The labelling and detection of the DNA probes were performed using the AlkPhos Direct Labelling and Detection system (GE). Hybridization of DNA was detected by the emission of light after applying a Fuji medical X-ray film.

Sequencing of the BAC DNA of pMP7598 was performed by Macrogen.

For insert size determination BAC DNA was isolated with the Qiagen large construct kit and digested to completion with *Spe*I (New England Biolabs). This digest was separated by pulsed-field gel electrophoresis using the Bio-Rad CHEF system on a 1 % SeaKEM LE agarose gel in 0.5 × Tris/borate/EDTA buffer, with switch-times of 10–40 s for 44 h with a set angle of 120° in a field of 5 V cm^{-1} at 10 °C, along with a Lambda concatemer PFGE marker (New England

Biolabs) and the 1 kb plus DNA ladder (Invitrogen) supplemented with a complete *Xho*I digest of Lambda (Fermentas).

BAC DNA that was isolated with the Qiagen large construct kit was sheared by nebulization, fractionated by gel electrophoresis, and fragments 2.5–3.0 kb in size were cloned into the dephosphorylated *Eco*RV site of pBlueScript I SK$^+$ (Stratagene) using standard methods (Sambrook & Russell, 2001). Shotgun templates were prepared from XL2 transformants (Stratagene) and sequenced using the DYEnamic ET Terminator Cycle Sequencing kit (GE Healthcare Life Sciences). For gap closure, PCR products were amplified with custom-made primers using a regular PCR protocol with 4 % (v/v) DMSO. Typically a 20 µl PCR contained 0.5 µl 5 µM forward and 0.5 µl 5 µM reverse custom primer, 1 µl 2.5 mM dNTPs, 2 µl 25 mM $MgCl_2$, 2 µl 10× GoldStar buffer (200 mM Tris/HCl pH 9.0, 5 mM $MgCl_2$), 0.8 µl DMSO, 0.2 µl 5 units µl^{-1} Goldstar (Eurogentec) polymerase, and 1 µl 10 µg µl^{-1} template DNA from a gap-spanning clone or the BAC. Sequencing reactions were carried out in a 10 µl reaction mixture with 0.5 µl Amerdye (GE), 1 µl sequence primer, 3.5 µl sequence buffer (200 mM Tris/HCl pH 9.0, 10 mM $MgCl_2$), and 5 µl template DNA. Sequence PCRs were analysed on a 3730 XL DNA analyser (Applied Biosystems).

Using the PREGAP4 module of the Staden package 2004, raw trace data were processed into assembly-ready sequences. Sequences were base called by the PHRED base caller (Ewing & Green, 1998; Ewing *et al.*, 1998). Clipping was performed to remove sequencing vector, cloning vector and poor-quality sequences. Processed sequences were subsequently assembled with GAP4, with a sequence percentage mismatch threshold of 8 %, and parsed into the GAP4 assembly database. The GAP4 contig editor interface was used for editing and finishing. Consensus calculations with a quality cutoff score of 40 were performed from within GAP4 using a probabilistic consensus algorithm based on the expected error rates output by PHRED.

To manage the sequence, assembly and scaffolding of the data, TOPAAS (Peters *et al.*, 2006) was used. The software is implemented on a SUN V440 server running Solaris 2.9. Primers for gap closing are automatically designed on contig end sequences, using Primer3 as a core primer design program. Maximum distance of primer positions to contig ends is set to 400 bp.

Protein-encoding ORFs were predicted using Glimmer 2.0 (Delcher *et al.*, 1999) and GeneMark.hmm for Prokaryotes version 2.4 (Lukashin & Borodovsky, 1998) using the *Pseudomonas fluorescens* Pf01 genome as a reference. The predicted genes were used in a BLAST search (Altschul *et al.*, 1990) to assign putative gene-names and function by homology transfer.

The DNA sequence was analysed with VectorNTI Advanced 10.1.1 software (Invitrogen) and BLAST programs (www.ncbi.nlm.nih.gov/BLAST/) (Altschul *et al.*, 1990). Amino acid sequences of the NRPS involved in putisolvin biosynthesis were compared and analysed with NRPS-PKS web-based software (www.nil.res.in/nrps-pks.html) (Ansari *et al.*, 2004).

Construction of *psoR*, *macA* and *macB* mutant strains. The *P. putida* PCL1445 *psoR* mutant, PCL1626, was constructed by single homologous recombination. A 0.58 kb internal fragment of the *psoR*-homologous gene of strain PCL1445 was obtained by PCR using primers oMP872 (5′-ACCTCAGTGAATGGACCCTTG-3′) and oMP873 (5′-GAGCTGTTTTTCACGTTCAGC-3′), cloned into the pGEM-T Easy Vector System I (Promega) and transferred as an *Eco*RI–*Eco*RI insert to pMP5285 (Kuiper *et al.*, 2001), resulting in pMP7570. Plasmid pMP7570 was conjugated to *P. putida* PCL1445 by tri-parental mating using *E. coli* containing pRK2013 as a helper strain (Schnider *et al.*, 1995). Strain PCL1626 was obtained as a kanamycin-resistant colony. The insertion of the suicide construct in *psoR* was confirmed by sequence analysis of the suicide plasmid that was recovered from the genomic DNA of PCL1626 using *Cla*I.

P. putida PCL1445 *macA* and *macB* mutants were constructed using a similar mutagenesis strategy. The *macA* fragment for the construction of the suicide plasmid pMP7595 resulted from a PCR using primers oMP1152 (5′-ACGTTGGAGGGCATCAAG-3′) and oMP1153 (5′-GTTTGTCTTCGCCCAGGATA-3′), with chromosomal DNA of strain PCL1445 as a template. Single homologous recombination in *macA* resulted in strain PCL1644. The *macB* fragment of PCL1445 for the construction of the suicide plasmid pMP7597 resulted from a PCR using primers oMP1154 (5′-AAGGGCTATGTGCTCAAGGA-3′) and oMP1155 (5′-CGTTGCAGCATCAGTTTTGT-3′), with chromosomal DNA of strain PCL1445 as a template. Single homologous recombination in *macB* resulted in strain PCL1645.

Complementation of *macA* and *macB* mutants of PCL1445. Complementation of mutants PCL1644 (*macA*) and PCL1645 (*macB*) was carried out using pMP7596, a shuttle vector derived from pME6010 (Heeb *et al.*, 2000) in which a 3.6 kb fragment containing *macA* and *macB* of strain PCL1445 was inserted. This insert was obtained by PCR using primers oMP1156 (5′-GGGGAATTCTT-GACAATTAATCATCGGCTCGTATAATGTGTGGAATTGTGAGCG-GATAACAATTTTCACACAGGAAACAGCTAAATGGAAAAGTCCA-GATTCCG-3′) containing the sequence of a P_{tac} promoter and oMP1158 (5′-CGATGCTGGAGAAACACATC-3′). pMP5518 was transferred to strains PCL1644 and PCL1645 by tri-parental mating as described above and transformants were selected on KB agar medium supplemented with tetracycline (40 µg ml^{-1}).

Quantification of fluorescence in *gfp* reporter strains. Expression of *psoA*::*gfp* (Dubern *et al.*, 2005) was quantified using an HTS7000 Bio Assay Reader (Perkin Elmer). Bacterial strains were grown to an OD_{620} of 2.0 and diluted to OD_{620} 0.6. Fluorescence of the diluted cultures was quantified using a white 96-well microtitre plate containing 200 µl culture aliquots. Fluorescence of the cultures was determined at an excitation wavelength of 485 nm and an emission wavelength of 520 nm.

Quantification of putisolvin production. Putisolvin (biosurfactant) production in KB or BM culture medium was quantified by two different methods. First, the decrease of surface tension between culture medium and air was measured using a Du Nouy ring (K6 Krüss) as previously described (Kuiper *et al.*, 2004).

Secondly, putisolvin production was analysed by HPLC. Briefly, 7 ml of a KB culture supernatant was extracted with 1 vol. ethyl acetate (Fluka) as described previously (Kuiper *et al.*, 2004). Ethyl acetate extracts obtained from 7 ml culture supernatant were prepared as described previously (Dubern *et al.*, 2005) and separated by HPLC (Jasco), using a reverse-phase C8 5 µm Econosphere column (Alltech), a PU-980 pump system (Jasco), an LG-980-02 gradient unit (Jasco) and an MD 910 detector (Jasco) (Dubern *et al.*, 2005). Chromatograms were analysed in the wavelength range between 195 nm and 420 nm. The amount of putisolvins produced was determined as the area of the peak detected in micro-absorbance units (µAU) at 206 nm.

RESULTS

Structural organization of *pso* and its regulatory region

To isolate a chromosomal fragment of strain PCL1445 containing the entire gene cluster directing putisolvin biosynthesis, a *Hin*dIII BAC library of PCL1445 was

constructed and screened for the presence of the entire lipopeptide biosynthetic cluster by colony blotting using four different probes, as shown in Fig. 1: PCR1 of 0.75 kb, PCR2 of 0.58 kb, PCR3 of 0.65 kb, and PCR4 of 0.35 kb. This resulted in one positive clone, pMP7598, which was used for sequencing. Analysis of its 152 860 bp nucleotide sequence allowed identification of a 41 131 bp region containing three large ORFs, which we designated *psoA* (6358 bp), *psoB*, (22 422 bp) and *psoC* (11 321 bp) (Fig. 1). The putative promoter regions −35 (5′-TCATAT-3′) and −10 (5′-TTCGAG-3′) are found in the nucleotide sequences from 411 to 419 and from 434 to 439, respectively. The translational start of *psoA* is located 12 bp downstream of a putative ribosome-binding site (RBS), AGAAGG. The *psoA* gene encodes a protein with a predicted molecular mass of 233 kDa, which consists of 2199 aa. The translational start of the *psoB* gene starts from the GTG located 21 bp upstream of the TGA stop codon of the *psoA* gene. The *psoB* gene encodes a protein with a predicted molecular mass of 808 kDa (7474 aa). The translation of *psoC* starts from the GTG codon located 9 bp upstream of *psoB* and ends at the TAA stop codon at position 40 130. The *psoC* gene encodes a protein with a predicted molecular mass of 409 kDa (3773 aa).

Sequence analysis of the region upstream of *psoA* identified the presence of an ORF (*orf1*) transcribed in the opposite direction to *psoA* (Fig. 1). The predicted protein encoded by *orf1* showed 61 % homology at the amino acid level with an as yet uncharacterized transcriptional regulator from the LuxR family encoded by the *psyr2575* gene in *P. syringae* pv. *syringae* B728a (Feil *et al.*, 2005) and 30 % homology with the putative DNA-binding protein SalA characterized in *P. syringae* pv. *syringae* B301D (Kitten *et al.*, 1998). Thus *orf1* appears to code for a novel protein possibly involved in the transcriptional regulation of CLPs of *P. putida* and was tentatively named *psoR*. The predicted promoter region, composed by a −35 box (5′-TAGAAG-3′) and a −10 box (5′-ATGTAG-3′), separated by 13 nucleotides and located 321 bp upstream from the *psoR* translational start, shows similarity with a TyrR-regulating promoter (Yang *et al.*, 2004).

An additional ORF was identified downstream of *orf1* (Fig. 1). The predicted ORF2 protein was most similar to outer-membrane proteins associated with secretion systems in Gram-negative bacteria. The highest degree of similarity to ORF2 was shown by OprM of *P. aeruginosa* (Nakajima *et al.*, 2000), which shares 63 % identity at the amino acid level.

Analysis of the sequence downstream of *psoC* revealed the presence of two ORFs that are predicted to encode homologues of MacA and MacB of *E. coli* (Fig. 1). The *macA* gene of *P. putida* is 1.15 kb in length and has a similarity of 85 % at amino acid level with *macA* of *Pseudomonas entomophila* L48, 76 % similarity with *Pseudomonas* sp. MIS38 and *P. syringae* pv. *tomato* DC3000, and 75 % similarity with *P. fluorescens* Pf-5. The *macB* gene of *P. putida* is 2.01 kb in length and has a similarity of 88 % at amino acid level with *macA* of *P. entomophila* L48, 82 % similarity with *P. syringae* pv. *tomato* DC3000, and 80 % similarity with *P. fluorescens* Pf-5 and *Pseudomonas* sp. MIS38.

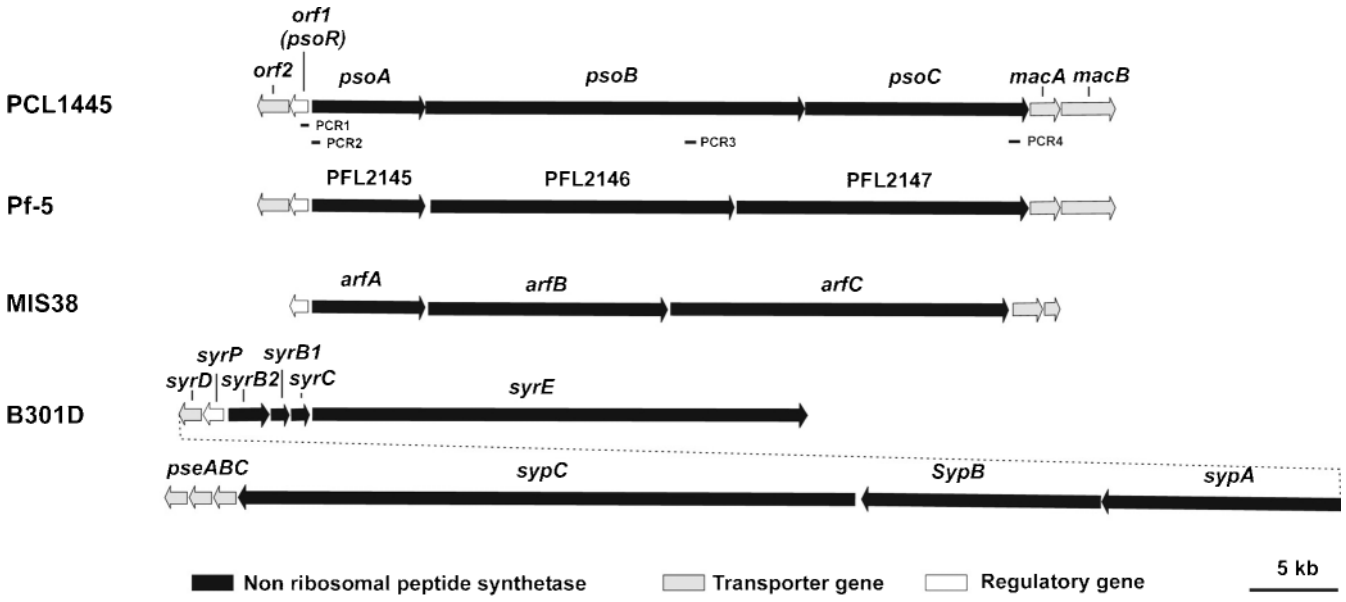

Fig. 1. Schematic representation of the putisolvin synthetic cluster and surrounding ORFs in *P. putida* PCL1445, and comparison with other previously characterized CLP synthetic clusters of *P. fluorescens* Pf-5, *Pseudomonas* sp. MIS38 and *P. syringae* B301D. The oligonucleotides PCR1–PCR4 used as probes to isolate a genomic fragment of PCL1445 harbouring the putisolvin synthetic cluster by colony blotting are shown.

No typical *E. coli* σ^{70} promoter was found upstream of *psoB* or *psoC* and no ρ-independent terminator-like structure was found in the region downstream of *psoC*. The *psoC* and *macA* genes are separated by only 66 bp and no putative promoter sequence could be predicted in the region upstream of *macA*. Moreover, the intergenic region between *macA* and *macB* spans only 5 bp. Consequently, we hypothesize that *psoA*, *psoB*, *psoC*, *macA* and *macB* could form a single operon with a promoter element upstream of *psoA*.

Isolation and characterization of putisolvin biosynthetic mutants

In order to isolate mutants in the biosynthetic *pso* cluster of PCL1445, 2400 Tn*5luxAB* transposants were screened for loss of surfactant activity as judged by the drop-collapsing assay, using cells derived from a single colony. This resulted in strains PCL1436 (Kuiper *et al.*, 2004), PCL1630, PCL1633 and PCL1634 (Dubern *et al.*, 2005).

Sequence analysis of the chromosomal regions flanking the Tn*5luxAB* showed the sites of insertion. The Tn*5* in strain PCL1634 was inserted in *psoA*, 3534 bp downstream from its transcriptional start. The insertion of the Tn*5* in PCL1633 was found to be positioned in *psoB*, 18 618 bp downstream of the predicted transcriptional start of the *pso* gene cluster. The Tn*5* in strain PCL1436 was inserted in *psoC*, 29 301 bp downstream of the start site. The Tn*5* in PCL1630 was located in *psoC*, 39 707 bp downstream from the predicted transcriptional start of *psoA* (Fig. 2).

To assess the role of *psoA*, *psoB* and *psoC* in putisolvin synthesis, biosurfactant production of PCL1436 (*psoC*), PCL1630 (*psoC*), PCL1633 (*psoB*) and PCL1634 (*psoA*) grown in liquid KB under standard conditions to stationary phase (28 °C with vigorous agitation) was quantified by the Du Nouy ring method. Culture supernatant of all four mutants was not able to decrease the surface tension between culture medium and air (54 mN m^{-1}) when compared to PCL1445 (31 mM m^{-1}). HPLC analysis confirmed that the production of putisolvins I and II was lacking in PCL1633 (*psoB*) (Fig. 4b), PCL1436 (*psoC*) (Kuiper *et al.*, 2004), PCL1630 (*psoC*) and PCL1634 (*psoA*) (data not shown).

Functional analysis of the putisolvin synthetases

Based on the analogy to other NRPSs, PsoA, PsoB and PsoC are composed of two (1 and 2), seven (3 to 9), and three (10 to 12) modules, respectively (Fig. 2b). A detailed analysis of PsoA, PsoB and PsoC revealed that the 12 modules have in common the three major domains known as the condensation (C), adenylation (A), and thiolation (T) domains (Marahiel *et al.*, 1997) (Fig. 2b).

Most of the consensus amino acid residues identified within other NRPS modules were highly conserved in the NRPS modules of PsoA, B and C (Table 2). The condensation domains within PsoA, PsoB and PsoC are thought to catalyse the formation of the amide bond between consecutively bound thioesterified intermediates of putisolvins. The conserved motif HHxxD(H/G) was identified in each of the 12 condensation domains (Table 2).

The condensation domains are followed by amino acid adenylation domains containing 12 conserved sequences previously described by Marahiel *et al.* (1997) (Fig. 2b, Table 3). The putative substrate-binding pocket (Stachelhaus *et al.*, 1999) was identified for each of the 12 amino acid activating modules of PsoA, PsoB and PsoC (Fig. 2b; Table 3). Using NRPS-PKS web-based software (see Methods), the primary structure of the 12 A-domains was compared to that of the phenylalanine A-domain (PheA) of gramicidin synthetase, the crystal structure of which has been determined (Stachelhaus *et al.*, 1999). The signature sequences consists of 10 selectivity-conferring amino acid residues embedded in the substrate-binding pocket at positions 235, 236, 239, 278, 299, 301, 322, 330, 331 and 527 bp of each of the 12 A-domains (Table 3). The specificity-conferring codes of PsoA-M1, PsoB-M3 and M4 were similar to the Leu-conferring code of bacitracin (47 %), lichensin (47 %) and surfactin (43 %) (Table 3). The signature sequence for PsoB-M8 conferred selectivity preferably for Leu, while structural analysis of putisolvin showed that the eighth amino acid residue to be incorporated was most likely an Ile (Table 3) (Kuiper *et al.*, 2004). The signature of PsoB-M7 had the highest similarity with Val of gramicidin (58 %) and surfactin (55 %) (Table 3). Interestingly, PsoC-M11, which differentiates putisolvin I and II, had a higher selectivity for Val than for Leu or Ile and had a weaker selectivity for Val than did PsoB-M7 (Table 3). Structural analysis of putisolvin I and II revealed that PsoA-M2 and PsoB-M5 have highest selectivity for Glu and Gln, respectively (Kuiper *et al.*, 2004). Analysis of the signature sequence of the two domains shows highest identity with Gln of tyrocidin (59 %). The conferring code of PsoB-M6, PsoB-M9 and PsoC-M12 matches 70 % with Ser of syringomycin synthetase. This correlates with structural analysis of putisolvin I and II showing highest selectivity for Ser PsoB-M6, PsoB-M9 and PsoC-M12 domains (Kuiper *et al.*, 2004).

Additional domains in NRPSs may include an epimerization (E) domain, responsible for the L or D configuration of an amino acid, as was found, for example, for the surfactin synthetic cluster in *Bacillus* spp. (Peypoux *et al.*, 1999; Sieber & Marahiel, 2005). None of the 12 modules in the *pso* synthetic cluster harbours an internal E-domain for the conversion of the L to the D form, suggesting that an external racemase is responsible for the L to D conversion. This suggestion was further supported by specific sequence motifs identified in the T-domains of the *pso* biosynthetic genes (Fig. 3). The T-domains that are associated with an E-domain have a highly conserved (F[F/Y]XXGGDSIKA[I/L]Q) motif, in which the aspartate residue in front of serine

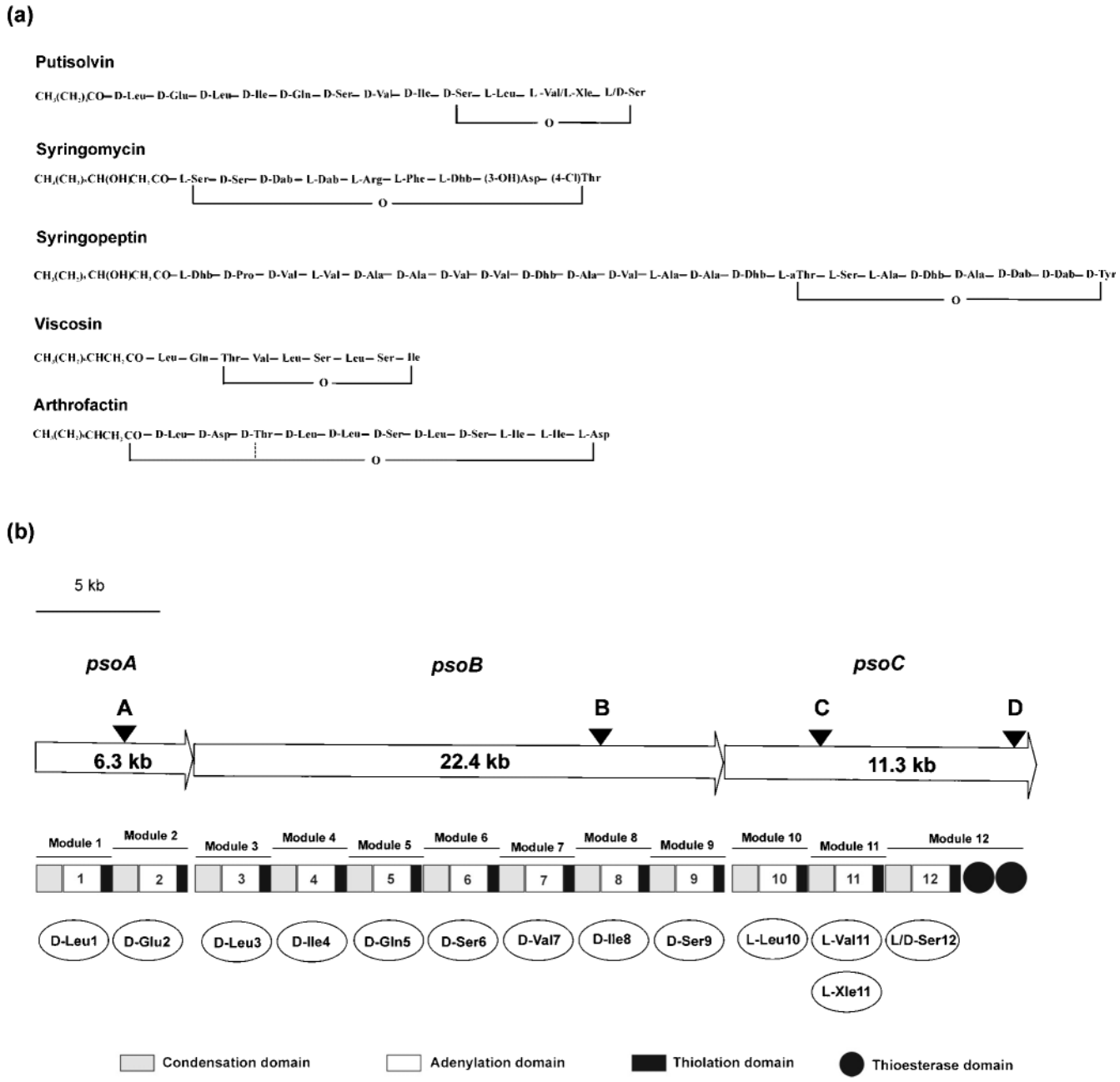

Fig. 2. Schematic representation of the amino acid modules present in PsoA, PsoB and PsoC and the predicted D and L forms of the corresponding amino acids. (a) Primary structure of putisolvin and the other classes of CLPs. Putisolvin harbours a lactone structure between the carboxyl group of the C-terminal L/D-Ser and the hydroxyl group of the ninth residue D-Ser. (b) Modular organization of the *pso* biosynthetic genes. The domain organizations of the NRPs are illustrated within the individual *pso* genes. The individual amino acid residues are incorporated by module 1 to module 12 in a collinear manner with the structure of the putisolvin molecule. The locations of the transposon insertion in *psoA* of mutant strain PCL1634 (A), in *psoB* of mutant strain PCL1633 (B), and in *psoC* of mutant strains PCL1436 (C) and PCL1630 (D) are indicated above the gene organization.

(the 4′-phosphopantethenine-binding site) especially seems to play a important role in the proper interaction between T- and E-domains (Linne *et al.*, 2001). In the putisolvin biosynthetic genes, the T-domains contain a different conserved sequence (FFELGGHSLLA[V/M]), which was further classified by comparison with the arthrofactin synthetic cluster into T(L) and T(D) domains, responsible for transferring L- or D-amino acids (Fig. 3).

Table 2. Highly conserved motifs of the putisolvin synthetase modules

Domain*	Motif	Conserved residues	
		Peptide synthetase†	Putisolvin synthetase
Condensation	C1	SxAQxR(L/M)(W/Y)xL	LAPLQ(Q/E)G
	C2	LHExLRTxF	-
	C3	MHHxISDG(W/V)S	HHxSAx(D/Q)
	C4	YxD(F/Y)AVW	Y
	C5	(I/V)GxFVNT(Q/L)(C/A)xR	GxFxNxL(P/R)
	C6	(H/N)QD(Y/V)PFE	HExx(P/S)
	C7	RDxSRNPL	(P/A)LFxx(L/M)L
Adenylation	A1	L(T/S)YxEL	L(S/D)YxxL
	A2	LKAGxAYL(V/L)P(L/I)D	LKAG(G/A)AYVPxD
	A3	LAYxxYTSG(S/T)TGxPKG	LAY(V/L)xYTSGSTGxPKG
	A4	FDxS	FDx(S/M)
	A5	NxYGPTE	N(N/V)YGPTE
	A6	GELxIxGxG(V/L)ARGYL	G(E/Q)LxxGGxxVARGY(L/F)
	A7	Y(R/K)TGDL	YR(T/S)GDL
	A8	GRxDxQVKIRGxRIELGEIE	GRNDxGVK (I/L)RGxRVELxEIE
	A9	LPxYM(I/V)P	PxYMVP
	A10	NGK(V/L)DR	NGK(I/L)DR
Thiolation‡	T	DxFFxxLGG(H/D)S(L/I)	D(H/N)FFELGGHSLLA(M/V)
Thioesterase	Te1	GxSxG	GWSFG
	Te2	GxSxG	GHSFG

*See Fig. 2(b).
†Marahiel *et al.* (1997).
‡See Fig. 3.

Analysis of the sequence downstream of the T-domain conserved core motif region revealed that highly conserved amino acid sequences present in the T(D) domains of arthrofactin synthetase were present in PsoA-T1, PsoA-T2, PsoB-T3, PsoB-T4, PsoB-T5, PsoB-T6, PsoB-T7, PsoB-T8 and PsoB-T9 (Fig. 3). Similar analysis of the T(L) domains showed that a conserved amino acid sequence present in the T(L) domains of the Arf synthetic cluster matches those of PsoC-T10 and PsoC-T11. Most interestingly, the amino acid sequence downstream of the conserved core motif of PsoC-T12 lacks five of the nine highly conserved amino acids present in the nine other T(D) domains of the Pso synthetic cluster and did not show any similarity with any T(L) domain (Fig. 3).

The C-terminal region of PsoC consists of 553 aa and contains two putative thioesterase (TE) domains designated PsoTe1 and PsoTe2. The two TE-domains are separated by a short sequence of 25 aa and contain 260 and 182 aa residues, respectively. Both the PsoTe1 and the PsoTe2 domain harbour a highly conserved signature sequence (GWSFG) (amino acids 12 883 to 12 887) and (GHSFG) (amino acids 13 169 to 13 173) identical to the conserved motifs previously described in ArfTe1 and ArfTe2, respectively (Roongsawang *et al.*, 2003). However, PsoTe2 has only one (GXSXG), although two were identified in ArfTe2 (Roongsawang *et al.*, 2003). Both PsoTe1 and PsoTe2 showed highest similarity with SyrE (35 %) from *P. syringae* B301D and contain additional conserved amino acid residues Ser (residues 12 885 in Te1 and 13 171 in Te2), Asp (residues 12 912 in Te1 and 13 198 in Te2) and His (residues 13 054 in Te1 and 13 253 in Te2), which form a catalytic triad of TE-domains in SrfA-C, ArfCTe1 and ArfTe2 (Roongsawang *et al.*, 2003).

The *psoR*, *macA* and *macB genes* are involved in putisolvin biosynthesis

To investigate whether *psoR*, *macA* and *macB* are involved in putisolvin production, insertion mutants of PCL1445 were constructed by single homologous recombination using suicide plasmids pMP7570, pMP7595 and pMP7597 (see Methods), resulting in strains PCL1626 (*psoR*), PCL1644 (*macA*) and PCL1645 (*macB*), respectively. The proper integration of pMP7570, pMP7595 and pMP7597 was confirmed by sequence analysis (data not shown).

Biosurfactant production by PCL1626 (*psoR*) grown in KB medium under standard conditions to stationary phase was quantified by the Du Nouy ring assay. In contrast to culture supernatant of the wild-type, which decreases the surface tension between culture medium and air, culture supernatant of PCL1626 (*psoR*) was not able to decrease the surface tension (53 mN m^{-1}), indicating that putisol-

Table 3. Homology and selectivity-conferring code of the amino-acid-specific adenylation domains (A-domains) of the putisolvin biosynthetic gene cluster compared to A-domains extracted from other NRPSs

A-domain*	Selectivity-conferring amino acid at position†:										Amino acid	Identity (%)‡
	235	236	239	278	299	301	322	330	331	527		
PsoA-M1	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	
PsoB-M3	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	
PsoB-M4	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	
PsoB-M8	**D**	**A**	**W**	L	**L**	**G**	**N**	**V**	**N**	**K**	Ile	
PsoC-M10	**D**	**A**	**W**	**S**	**R**	**L**	**N**	**V**	**V**	**K**	Leu	
bacit001_A_003	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	51
micro002_A_001	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	48
liche002_A_003	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	47
surfa004_A_003	**D**	**A**	**W**	**F**	**L**	**G**	**N**	**V**	**N**	**K**	Leu	43
bacit001_A_005	**D**	G	F	**F**	F	**G**	V	**V**	Y	**K**	Ile	38
liche003_A_001	**D**	G	F	**F**	F	**G**	V	**V**	Y	**K**	Ile	38
PsoB-M7	**D**	**A**	**L**	W	**I**	**G**	**G**	**T**	**F**	**K**	Val	
PsoC-M11	**D**	**A**	**L**	F	**I**	**G**	**G**	**T**	**F**	**K**	Val/Leu/Ile	
gram002_A_002	**D**	**A**	F	W	**I**	**G**	**G**	**T**	**F**	**K**	Val	58
surfa005_A_001	**D**	**A**	F	W	**I**	**G**	**G**	**T**	**F**	**K**	Val	55
bacit001_A_005	**D**	G	F	**F**	F	**G**	V	V	Y	**K**	Ile	42
liche003_A_001	**D**	G	F	**F**	F	**G**	V	V	**F**	**K**	Ile	42
surfa004_A_003	**D**	**A**	F	**F**	L	**G**	C	V	**F**	**K**	Leu	41
liche001_A_002	**D**	**A**	F	**M**	L	**G**	M	V	**F**	**K**	Leu	41
PsoA-M2	**D**	**A**	**W**	K	**F**	**G**	**V**	**V**	**D**	**K**	Glu	
PsoB-M5	**D**	**A**	**L**	Q	V	**G**	**V**	**V**	**D**	**K**	Gln	
tyroc003_A_002	**D**	**A**	**W**	Q	**F**	**G**	**L**	**I**	**D**	**K**	Gln	59
liche001_A_002	**D**	**A**	**W**	G	**F**	**G**	**L**	**I**	**D**	**Q**	Gln	55
ituri001_A_002	**D**	**A**	**Q**	D	L	**G**	**V**	**V**	**D**	**K**	Gln	53
myc002_A_003	**D**	**A**	**Q**	D	L	**G**	**V**	**V**	S	**K**	Gln	53
PsoB-M6	**D**	**V**	**W**	**H**	**L**	**S**	**L**	**V**	**D**	**K**	Ser	
PsoB-M9	**D**	**V**	**W**	**H**	**L**	**S**	**L**	**V**	**D**	**K**	Ser	
PsoC-M12	**D**	**V**	**W**	**H**	**L**	**S**	**L**	**V**	**D**	**K**	Ser	
syr001_A_002	**D**	**V**	**W**	**H**	**L**	**S**	**L**	**V**	**D**	**K**	Ser	70
ituri003_A_002	**D**	**V**	**W**	**H**	F	**S**	**L**	I	**D**	**K**	Ser	66
myc003_A_001	**D**	**V**	**W**	**H**	F	**S**	**L**	**I**	**D**	**K**	Ser	64
cdap003_A_001	**D**	**V**	**W**	**H**	F	**S**	H	A	A	**K**	Ser	63

*bacit, bacitracin; cdap, daptomycin; gram, gramicidin; ituri, iturin A; liche, lichensin; micro, microcystin; myc, mycosubtilin; surfa, surfactin; syr, syringomycin; tyroc, tyrocidin.

†Conserved residues of the substrate-binding pockets of A-domains with assigned functions are in bold.

‡Homology of the whole A-domain, 507 aa, compared to the *pso* biosynthetic genes.

vin production was abolished (Fig. 4a). Quantification of putisolvins in culture supernatant using HPLC showed a 70 % decrease for PCL1644 (*macA*) and PCL1645 (*macB*) when compared to the wild-type (Fig. 4b). Introduction of pMP7596 harbouring *macA* and *macB* genes restored putisolvin production to the wild-type level in culture supernatant of strains PCL1644 (*macA*) and PCL1645 (*macB*) (Fig. 4b).

Influence of *psoR* on the expression of the *pso* biosynthetic genes

PsoR of strain PCL1445, which is predicted to be 260 aa in length, revealed homology to prokaryotic regulatory proteins (Fig. 5a). Database searches with BLASTP demonstrated that approximately 60 aa in the C-terminal region of PsoR shared significant similarity with similar regions of known regulatory proteins. The C terminus of PsoR exhibited 35 % identity to GerE (Cutting & Mandelstam, 1986), 34 % identity to DctR (Hamblin *et al.*, 1993) and 32 % identity to FixJ (Anthamatten & Hennecke, 1991) (Fig. 5a).

Further analysis of the C-terminal regions of PsoR identified a three-element fingerprint that provides a signature for the HTH DNA-binding motif of LuxR bacterial regulatory proteins (Bairoch, 1993). Moreover, three highly conserved residues in the N-terminal regions

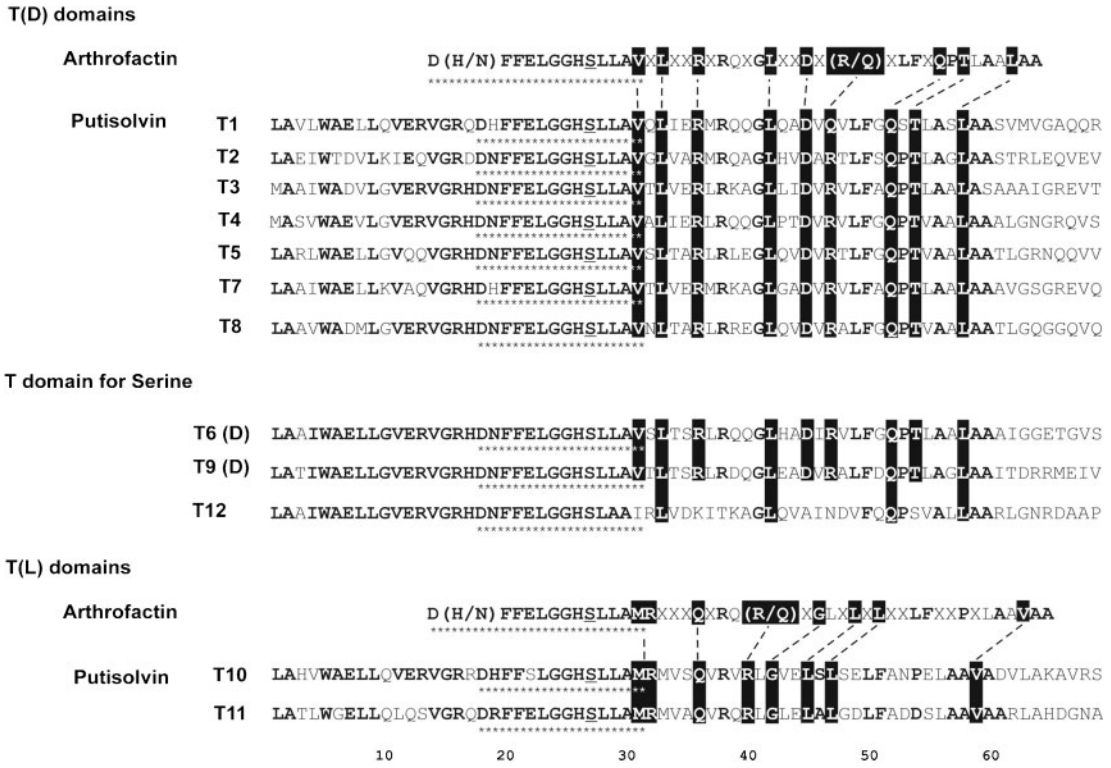

Fig. 3. Alignment of the predicted amino acid sequence of the 12 T-domains in the putisolvin synthetic cluster with the T(D) and T(L) consensus of arthrofactin synthetase. The asterisks show the conserved core motif. The conserved amino acid residues found downstream of the core motif are boxed.

of members of the response regulator subfamily corresponding to Asp206, Glu207 and Lys244 in FixJ (Parkinson & Kofoid, 1992) were detected in PsoR (Fig. 5a). However, five highly conserved amino acids in the N-terminal regions of members of an autoinducer-binding subfamily (Fuqua *et al.*, 1996) corresponding to Trp59, Tyr69, Asp79, Pro80 and Gly121 of LuxR were not detected in PsoR.

To determine whether PsoR regulates putisolvin expression in PCL1445, a *psoA*::*gfp* transcriptional fusion (Dubern *et al.*, 2005) was introduced into PCL1626 (*psoR*). The expression of the *gfp* reporter strongly decreased in the *psoR* mutant when compared to the wild-type strain (Fig. 5b).

DISCUSSION

The genomic arrangement of the *psoA*, *psoB* and *psoC* genes and their surrounding genes is similar to that of at least three other CLP synthetases of *Pseudomonas*, including syringopeptin of *P. syringae* pv. *syringae* B301D (Scholz-Schroeder *et al.*, 2003), arthrofactin of *Pseudomonas* MIS38 (Roongsawang *et al.*, 2003), and a predicted lipodecapeptide produced by *P. fluorescens* Pf-5 (Paulsen *et al.*, 2005), suggesting that the genomic organization of CLP synthetic gene clusters is well conserved among *Pseudomonas* species. As for already identified CLP gene clusters in other *Pseudomonas* strains (Raaijmakers *et al.*, 2006), the NRPS genes of *P. putida* PCL1445 are physically linked. Putisolvin synthetases have a modular structure and each module contains the typical feature of a NRPS, consisting of an A-, C- and T-domain (Fig. 2). The order of the 12 modules present in the genetic structure (Fig. 2) indicates that putisolvin synthesis by the *psoA*, *psoB* and *psoC* genes obeys the collinearity rule.

Analysis of the homologies and selectivity-conferring code of the amino-acid-specific A-domains showed a higher affinity of the PsoA-M11 A-domain for Val than for Leu or Ile (Table 2). This finding at the genetic level supports our previous work, in which it was shown first that putisolvin I is produced at a higher level (1.6-fold) than putisolvin II (Kuiper *et al.*, 2004; Dubern *et al.*, 2005, 2006), and secondly that the ratio between putisolvin I and putisolvin II was affected by the availability of L-Val, L-Leu or L-Ile in the culture medium (Dubern & Bloemberg, 2006). Most interestingly, studies on the environmental regulation of putisolvin biosynthesis showed that environmental conditions that do not favour bacterial growth, including 'low temperature', and which PCL1445 is most likely to encounter in its natural environment, have a strong positive effect on the level of putisolvin produced (Dubern *et al.*, 2005). In addition, Stachelhaus *et al.*

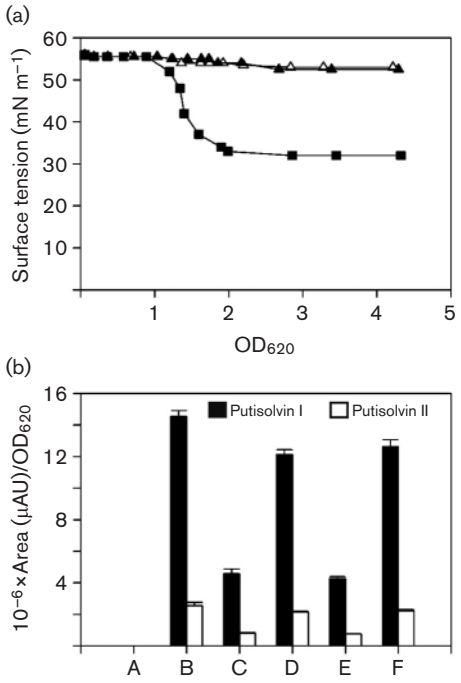

Fig. 4. Effect of mutations in *psoB*, *psoR*, *macA* and *macB* on the production of putisolvins of *P. putida* PCL1445. (a) Quantification of surface tension decrease by culture supernatants of *P. putida* strains PCL1445 (■), PCL1633 (*psoB*) (△) and PCL1626 (*psoR*) (▲) grown to stationary phase in KB medium. (b) C8-reverse-phase HPLC analysis of putisolvin production by *P. putida* strain PCL1445 and its mutants PCL1633 (*psoB*), PCL1644 (*macA*) and PCL1645 (*macB*) grown to stationary phase in KB medium at 11 °C; A, mutant strain PCL1633 (*psoB*); B, PCL1445; C, PCL1644 (*macA*); D, PCL1644 (*macA*) harbouring pMP7596 (P_{tac} *macA-macB*); E, PCL1645 (*macB*); F, PCL1645 (*macB*) harbouring pMP7596 (P_{tac} *macA-macB*). Mean values of triplicate cultures are given; the error bars indicate SEM.

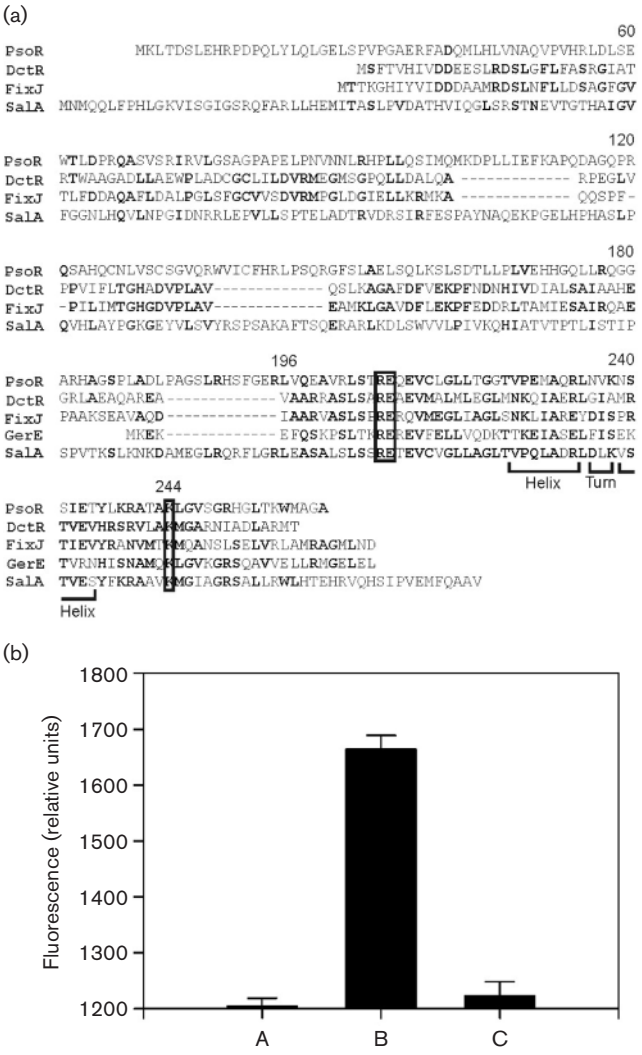

Fig. 5. Role of PsoR in the regulation of putisolvin biosynthesis. (a) Alignment of the predicted amino acid sequence of PsoR with homologous proteins. Amino acids that are shared among two of more proteins are indicated in bold. Dots indicate gaps introduced to optimize alignments. The three conserved amino acid residues among regulatory proteins of the FixJ subfamily are boxed. When amino acids 204–265 of PsoR were used for BLASTP analysis, *P*-values for the alignments were 1.5×10^{-6} for DctR, 2.9×10^{-5} for FixJ and 4.0×10^{-5} for GerE. The percentage of identity for amino acids 196–244 of PsoR to the corresponding regions of the other three proteins was 34 % (DctR), 32 % (FixJ), 35 % (GerE), 36 % (SalA). (b) Expression of *psoA* of *P. putida* PCL1445. Expression was determined using the *psoA*::*gfp* reporter in PCL1445 and PCL1626 (*psoR*) by measuring fluorescence from cells containing the *psoA* promoter fused to *egfp* (pMP5539). pMP5540 containing the *psoA* promoter in the transcriptionally inactive orientation was used as a control vector; A, PCL1445 harbouring the control vector pMP5540; B, PCL1445 harbouring pMP5539; C, PCL1626 (*psoR*) harbouring pMP5539. Mean values of duplicate cultures are given; the error bars indicate the range of results.

(1999) provided *in silico* evidence that certain signature sequences in the A-domains contain adaptable or so-called 'wobble-like' positions which can result in a high variability of amino acid usage. Our experimental data raise the question of a possible connection between flexibility with respect to amino acid usage and environmental conditions. In particular, it will be interesting to investigate the flexibility of module 11 for Val, Leu or Ile usage at different temperatures through the analysis of the ratio between putisolvin I and II at different temperatures.

No internal E-domain for the conversion from L to D form was detected in the *pso* synthetic cluster. This is similar to previous studies on syringopeptin (Scholz-Schroeder *et al.*, 2003), arthrofactin (Roongsawang *et al.*, 2003), and the predicted lipodecapeptide of strain Pf-5 (Paulsen *et al.*, 2005) in which no internal E-domains were found. This observation suggests the presence of an external racemase, which recently was further supported by the discovery of specific sequence motifs in the T-domains of the *arf*

synthetic template (Roongsawang *et al.*, 2003). Based on comparative analysis of the different conserved sequence motifs present within the T-domains of putisolvin synthetase and the arthrofactin synthetic template, the present study shows that the first nine T-domains in *pso* are responsible for transferring D-amino acids and T10 and T11 for transferring L amino acids in the putisolvin structure. Interestingly, analysis of the last T-domain, T12, did not allow determination of whether L- or D-Ser was incorporated (Fig. 3). This finding suggests the existence of a flexibility of module 12 for the transfer of a D- or L-serine at that particular position.

An ORF located upstream of *psoA* was identified as a member of the LuxR family of regulatory proteins based on homology analysis, and referred to as *psoR* (Fig. 1). Sequence analysis revealed the presence of HTH DNA-binding motifs at the C terminus of PsoR (Fig. 5a). The HTH motif has been observed in many regulatory proteins (Pabo & Sauer, 1992), which are divided into more than 10 groups, including the LuxR, AraC and MarR families. PsoR appears to be most closely related to members of the LuxR regulatory family, such as DctR (Hamblin *et al.*, 1993) and FixJ (Anthamatten & Hennecke, 1991). An approximately 60 aa region of the C terminus containing the four helices and their turns, which is called a three-element fingerprint, provides the signature for the HTH motif of the LuxR family of bacterial regulatory proteins. The observation that the PsoR protein exhibits the highest similarity to DctR and FixJ and contains the three-element fingerprint suggests that it is a member of the LuxR family (Fig. 5a). Despite this homology, PsoR lacks five highly conserved residues at the N terminus characteristic of the LuxR subfamily, which is composed of autoinducer-binding regulators activated by homoserine lactones (Fuqua *et al.*, 1996). The second major subfamily of transcriptional regulators is composed of the response regulators of two-component signal transduction systems, such as FixJ (Anthamatten & Hennecke, 1991) and DctR (Hamblin *et al.*, 1993). Three highly conserved residues (Arg, Glu, Lys) characteristic of the response regulators were found in the PsoR sequence, suggesting that PsoR may be closely related to this subfamily of regulators (Fig. 5a).

A mutation in PsoR abolished putisolvin production in the wild-type strain (Fig. 4a), giving the first evidence of its regulatory role in putisolvin biosynthesis. The decrease of expression of the *psoA*::*gfp* fusion in *psoR* mutant when compared to the wild-type (Fig. 5b) suggests that the effect of *psoR* on putisolvin production can be accounted for by its effect on *psoA* transcriptional activity, although PsoR may regulate other genes involved in putisolvin production as well. This evidence was supported by the identification of a nucleotide consensus sequence similar to those involved in the regulation of response regulators such as TyrR (Yang *et al.*, 2004). In *E. coli*, the TyrR protein can act both as a repressor and as an activator of transcription. The regulation of transcription is determined by the position and nature of the recognition site (TyrR boxes) associated with each of the promoters (Pittard *et al.*, 2005).

Sequence analysis revealed almost immediately downstream (66 bp) of *psoC* the presence of two ORFs, which encode homologues of MacA and MacB of *E. coli* (Kobayashi *et al.*, 2001). This genetic organization is similar to that of other lipopeptide synthetic clusters including those of *P. fluorescens* Pf-5 (Paulsen *et al.*, 2005) and *Pseudomonas* sp. MIS38 (Roongsawang *et al.*, 2003), in which *macA* and *macB* were not characterized. The MacA and MacB proteins, along with a TolC outer-membrane protein, form a macrolide-specific ABC-type efflux carrier in which MacB is an integral membrane protein with four transmembrane segments and one nucleotide-binding domain, while MacA belongs to a membrane fusion protein (MFP) family with a signal-like sequence at its N terminus (Kobayashi *et al.*, 2001; Rouquette-Loughlin *et al.*, 2005). Interestingly, Kobayashi *et al.* (2001) showed that *macA* and *macB* genes conferred resistance against macrolides composed of 14- and 15-membered lactones but no or weak resistance against 16-membered ones. Mutations of *macA* or *macB* of strain PCL1445 resulted in reduced putisolvin production (Fig. 4b). This preliminary characterization of the mutants in *macA* and *macB* suggests that these two ORFs are (directly or indirectly) involved in putisolvin production and are possibly coregulated with *psoA–C*. In a recent report (Dubern *et al.*, 2006), we showed that the *psoA* promoter region includes a putative *lux* box that could be involved in the regulation of putisolvin biosynthesis by a quorum-sensing mechanism. It will therefore be interesting to know whether the *psoA-psoB-psoC-macA-macB* genes are organized in an operon and to further investigate the relationship between *psoR* and the previously described *ppu* quorum-sensing system (Dubern *et al.*, 2006).

ACKNOWLEDGEMENTS

We thank T. Hesselink (Greenomics, The Netherlands) for help with the random shotgun BAC sequencing and assembly. We thank E. Lagendijk for technical assistance. This research was supported by grant number 700.50.015. from the Council for Chemical Sciences of the Netherlands Foundation for Scientific Research (NWO/CW).

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Ansari, M. Z., Yadav, G., Gokhale, R. S. & Mohanty, D. (2004). NRPS-PKS: a knowledge-based resource for analysis of NRPS/PKS megasynthases. *Nucleic Acids Res* **32**, 405–423.

Anthamatten, D. & Hennecke, H. (1991). The regulatory status of the FixL-like and Fix-J-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*. *Mol Gen Genet* **225**, 38–48.

Bairoch, A. (1993). The PROSITE dictionary of site and patterns in proteins, its current status. *Nucleic Acids Res* **21**, 3097–3103.

Bender, C. L., Alarcon-Chalaidez, F. & Gross, D. C. (1999). *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol Rev* **63**, 266–292.

Cutting, S. & Mandelstam, J. (1986). The nucleotide sequence and the transcription during sporulation of *gerE* gene of *Bacillus subtilis*. *J Gen Microbiol* **132**, 3013–3024.

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

Ditta, G., Stanfield, S., Corbin, D. & Helinski, D. R. (1980). Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc Natl Acad Sci U S A* **77**, 7347–7351.

Dubern, J.-F. & Bloemberg, G. V. (2006). Influence of environmental conditions on putisolvins I and II production in *Pseudomonas putida* strain PCL1445. *FEMS Microbiol Lett* **263**, 169–175.

Dubern, J.-F., Lagendijk, E. L., Lugtenberg, B. J. J. & Bloemberg, G. V. (2005). The heat shock genes *dnaK, dnaJ*, and *grpE* are involved in regulation of putisolvin biosynthesis in *Pseudomonas putida* PCL1445. *J Bacteriol* **187**, 5967–5976.

Dubern, J.-F., Lugtenberg, B. J. J. & Bloemberg, G. V. (2006). The *ppuI-rsaL-ppuR* quorum sensing system regulates biofilm formation of *Pseudomonas putida* PCL1445 by controlling biosynthesis of the cyclic lipopeptides putisolvin I and II. *J Bacteriol* **188**, 2898–2906.

Ewing, B. & Green, P. (1998). Basecalling of automated sequencer traces using PHRED. II. Error probabilities. *Genome Res* **8**, 186–194.

Ewing, B., Hillier, L., Wendle, M. C. & Green, P. (1998). Base-calling of automated sequencer traces using PHRED. I. Accuracy assessment. *Genome Res* **8**, 175–185.

Feil, H., Feil, W. S., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M. & other authors (2005). Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. *Proc Natl Acad Sci U S A* **102**, 11064–11069.

Fuqua, C., Winans, S. C. & Greenberg, E. P. (1996). Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum sensing transcriptional regulators. *Annu Rev Microbiol* **50**, 727–751.

Grangemard, I., Wallach, J., Maget-Dana, R. & Peypoux, F. (2001). Lichenysin: a more efficient cation chelator than surfactin. *Appl Biochem Biotechnol* **90**, 199–210.

Guenzi, E., Galli, G., Grgurina, I., Gross, D. C. & Grandi, G. (1998). Characterization of the syringomycin synthetase gene cluster: a link between prokaryotic and eukaryotic peptide synthetases. *J Biol Chem* **273**, 32857–32863.

Hamblin, M. J., Shaw, J. G. & Kelly, D. J. (1993). Sequence analysis and interposon mutagenesis of a sensor-kinase (DctS) and response-regulator (DctR) controlling synthesis of the high-affinity C4-dicarboxylate transport system in *Rhodobacter capsulatus*. *Mol Gen Genet* **237**, 215–224.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580.

Heeb, S., Itoh, Y., Nishijyo, T., Schnider, U., Keel, C., Wade, J., Walsh, U., O'Gara, F. & Haas, D. (2000). Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. *Mol Plant Microbe Interact* **13**, 232–237.

Henriksen, A., Anthoni, U., Nielsen, T. H., Sørensen, J., Christophersen, C. & Gajhede, M. (2000). Cyclic lipoundecapeptide tensin from *Pseudomonas fluorescens* strain 96.578. *Acta Crystallogr C* **56**, 113–115.

Huber, B., Riedel, K., Kothe, M., Givskov, M., Molin, S. & Eberl, L. (2002). Genetic analysis of function involved in the late stages of biofilm development in *Burkolderia cepacia* HIII. *Mol Microbiol* **46**, 411–426.

Hutchison, M. L., Tester, M. A. & Gross, D. C. (1995). Role of biosurfactants and ion-channel-forming activities of syringomycin in transmembrane ion flux – a model for the mechanism of action in the plant-pathogen interaction. *Mol Plant Microbe Interact* **8**, 610–620.

Jain, D. K., Thompson, D. L., Lee, H. & Trevors, J. T. (1991). A drop-collapsing test for screening surfactant-producing microorganisms. *J Microbiol Methods* **13**, 271–279.

King, E. O., Ward, M. K. & Raney, D. E. (1954). Two simple media for the demonstration of pyocyanin and fluorescin. *J Lab Clin Med* **44**, 301–307.

Kitten, T., Kinscherf, T. G., McEvoy, J. L. & Willis, D. K. (1998). A newly identified regulator is required for virulence and toxin production in *Pseudomonas syringae*. *Mol Microbiol* **28**, 917–929.

Kleinkauf, H. & von Döhren, H. (1996). A nonribosomal system of peptide biosynthesis. *Eur J Biochem* **236**, 335–351.

Kobayashi, N., Nishino, K. & Yamagushi, A. (2001). Novel macrolide-specific ABC-type efflux transporter in *Escherichia coli*. *J Bacteriol* **183**, 5639–5644.

Koch, B., Nielsen, T. H., Sørensen, D., Andersen, J. B., Christophersen, C., Molin, S., Givskov, M., Sørensen, J. & Nybroe, O. (2002). Lipopeptide production in *Pseudomonas* sp. strain DSS73 is regulated by components of sugar beet seed exudates via Gac two-component regulatory system. *Appl Environ Microbiol* **68**, 4509–4516.

Kuiper, I., Bloemberg, G. V., Noreen, S., Thomas-Oates, J. E. & Lugtenberg, B. J. (2001). Increased uptake of putrescine in the rhizosphere inhibits competitive root colonization by *Pseudomonas fluorescens* strain WCS365. *Mol Plant Microbe Interact* **14**, 1096–1104.

Kuiper, I., Lagendijk, E. L., Pickford, R., Derrick, J. P., Lamers, G. E. M., Thomas-Oates, J. E., Lugtenberg, B. J. J. & Bloemberg, G. V. (2004). Characterization of two *Pseudomonas putida* lipopeptide biosurfactants, putisolvin I and II, which inhibit biofilm formation and break down existing biofilms. *Mol Microbiol* **51**, 97–113.

Lindum, P. W., Anthoni, U., Christophersen, C., Eberl, L., Molin, S. & Givskov, M. (1998). *N*-Acyl-L-homoserine lactone autoinducer control production of an extracellular lipopeptide biosurfactant required for swarming motility of *Serratia liquefaciens* MG1. *J Bacteriol* **180**, 6384–6388.

Linne, U., Doekel, S. & Marahiel, M. A. (2001). Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases. *Biochemistry* **40**, 15824–15834.

Lugtenberg, B. J. J., Kravchenko, L. V. & Simons, M. (1999). Tomato seed and root exudates sugars: composition, utilization by *Pseudomonas* biocontrol strain and role in rhizosphere colonization. *Environ Microbiol* **1**, 439–446.

Lukashin, A. V. & Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* **26**, 1107–1115.

Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem Rev* **97**, 2651–2673.

Nakajima, A., Sugimoto, Y., Yoneyama, H. & Nakae, T. (2000). Localization of the outer membrane subunit OprM of resistance-nodulation-cell division family multi-component efflux pump in *Pseudomonas aeruginosa*. *J Biol Chem* **275**, 30064–30068.

Neu, T. R. (1996). Significance of bacterial surface-active compounds in interaction of bacteria with interfaces. *Microbiol Rev* **60**, 151–166.

Nielsen, T. H., Christophersen, C., Anthoni, V. & Sørensen, J. (1999). Viscosinamide, a new cyclic depsipeptide with surfactant and antifungal properties produced by *Pseudomonas fluorescens* DR54. *J Appl Microbiol* **87**, 80–90.

Nutkins, J. C., Mortishire-Smith, R. J., Packman, L. C., Brodey, C. L., Rainey, P. B., Johnstone, K. & Williams, D. H. (1991). Structure

determination of tolaasin, an extra-cellular lipodepsipeptide produced by the mushroom pathogen *Pseudomonas tolaasii* Paine. *J Am Chem Soc* **113**, 2621–2627.

Pabo, C. O. & Sauer, R. T. (1992). Transcriptional factors: structural families and principles of DNA recognition. *Annu Rev Biochem* **61**, 1053–1095.

Parkinson, J. S. & Kofoid, E. C. (1992). Communication modules in bacterial signaling proteins. *Annu Rev Genet* **26**, 71–112.

Paulsen, I. T., Press, C. M., Ravel, J., Kobayashi, D. Y., Myers, G. S. A., Mavrodi, D. V., DeBoy, R. T., Seshadri, R., Ren, Q. & other authors (2005). Complete genome sequence of the plant commensal *Pseudomonas fluorescens* Pf-5. *Nat Biotechnol* **23**, 873–878.

Peters, S. A., van Haarst, J. C., Jesse, T. P., Woltinge, D., Jansen, K., Hesselink, T., van Staveren, M. J., Abma-Henkens, M. H. C. & Klein-Lankhorst, R. M. (2006). TOPAAS, a tomato and potato assembly assistance system for selection and finishing of bacterial artificial chromosomes. *Plant Physiol* **140**, 805–817.

Peypoux, F., Bonmatin, J. M. & Wallach, J. (1999). Recent trends in the biochemistry of surfactin. *Appl Microbiol Biotechnol* **51**, 553–563.

Pittard, J., Camakaris, H. & Yang, J. (2005). The TyrR regulon. *Mol Microbiol* **55**, 16–26.

Raaijmakers, J. M., de Bruijn, I. & de Kock, J. D. (2006). Cyclic lipopeptide production by plant-associated *Pseudomonas* spp.: diversity, activity, biosynthesis, and regulation. *Mol Plant Microbe Interact* **19**, 699–710.

Roongsawang, N., Hase, K., Haruki, M., Imanaka, T., Morikawa, M. & Kanaya, S. (2003). Cloning and characterization of the gene cluster encoding arthrofactin synthetase from *Pseudomonas* sp. MIS38. *Chem Biol* **10**, 869–880.

Rouquette-Loughlin, C. E., Balthazar, J. T. & Shafer, W. M. (2005). Characterization of the MacA-MacB efflux system in *Neisseria gonorrhoeae*. *J Antimicrob Chemother* **56**, 856–860.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Schnider, U., Keel, C., Voisard, C., Defago, G. & Haas, D. (1995). Tn5-directed cloning of *pqq* genes from *Pseudomonas fluorescens* CHAO: mutational inactivation of the genes results in overproduction of the antibiotic pyoluteorin. *Appl Environ Microbiol* **61**, 3856–3864.

Scholz-Schroeder, B. K., Soule, J. D. & Gross, D. C. (2003). The *sypA, sypB* and *sypC* synthetase genes encode twenty-two modules involved in the nonribosomal peptide synthesis of syringopeptin by *Pseudomonas syringae* pv. *syringae* B301D. *Mol Plant Microbe Interact* **16**, 271–280.

Sieber, S. A. & Marahiel, M. A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem Rev* **105**, 715–738.

Sørensen, D., Nielsen, T. H., Christophersen, C., Sørensen, J. & Gajhede, M. (2001). Cyclic lipoundecapeptide amphisin from *Pseudomonas* sp. strain DSS73. *Acta Crystallogr C* **57**, 1123–1124.

Stachelhaus, T., Mootz, H. D. & Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **6**, 493–505.

Wolk, C. P., Cai, Y. & Panoff, J. M. (1991). Use of a transposon with luciferase as a reporter to identify environmentally responsive genes in a cyanobacterium. *Proc Natl Acad Sci U S A* **88**, 5355–5359.

Yang, J., Hwang, J. S., Camakaris, H., Irawaty, W., Ishihama, A. & Pittard, J. (2004). Mode of action of the TyrR protein: repression and activation of the *tyrP* promoter of *Escherichia coli*. *Mol Microbiol* **52**, 243–256.

Edited by: M. A. Kertesz