The *Serratia* gene cluster encoding biosynthesis of the red antibiotic, prodigiosin, shows species- and strain-dependent genome context variation


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

The prodigiosin biosynthesis gene cluster (*pig* cluster) from two strains of *Serratia* (*S. marcescens* ATCC 274 and *Serratia* sp. ATCC 39006) has been cloned, sequenced and expressed in heterologous hosts. Sequence analysis of the respective *pig* clusters revealed 14 ORFs in *S. marcescens* ATCC 274 and 15 ORFs in *Serratia* sp. ATCC 39006. In each *Serratia* species, predicted gene products showed similarity to polyketide synthases (PKSs), non-ribosomal peptide synthases (NRPSs) and the Red proteins of *Streptomyces coelicolor* A3(2). Comparisons between the two *Serratia pig* clusters and the *red* cluster from *Str. coelicolor* A3(2) revealed some important differences. A modified scheme for the biosynthesis of prodigiosin, based on the pathway recently suggested for the synthesis of undecylprodigiosin, is proposed. The distribution of the *pig* cluster within several *Serratia* sp. isolates is demonstrated and the presence of cryptic clusters in some strains shown. The *pig* cluster of *Serratia marcescens* ATCC 274 is flanked by *cueR* and *copA* homologues and this configuration is demonstrated in several *S. marcescens* strains, whilst these genes are contiguous in strains lacking the *pig* cluster.

The red pigment, prodigiosin (Pig), is a member of the prodiginines produced by some *Serratia* species, actinomycetes and a few other bacteria (Grimont & Grimont, 1978). Prodigiosin is a linear tripyrrole and a typical secondary metabolite, appearing only in the later stages of bacterial growth (Williams et al., 1971). The actinomycete *Streptomyces coelicolor* A3(2) produces a closely related linear tripyrrole, undecylprodigiosin, and a cyclic derivative, butyl-\(\beta\)-cycloheptylprodiginine in a 2:1 ratio (Tsao et al., 1985).

A bifurcated pathway has been proposed for the synthesis of prodigiosin culminating in the enzymic condensation of the terminal products of the two pathways, 4-methoxy-2,2-bipyrrrole-5-carboxaldehyde (MBC) and the monopyrrole, 2-methyl-3-n-amylypyrrole (MAP). The precursors for prodigiosin were shown to be acetate, serine, alanine, methionine and proline (Williams, 1973). Recently, the mechanism of proline incorporation into a pyrrole moiety has been shown biochemically and a pathway for synthesis of undecylprodigiosin proposed (Cerdeño et al., 2001; Thomas et al., 2002).

Prodiginines have no defined role in the physiology of producing strains, but have been reported to have antifungal, antibacterial and antiprotozoal/antimalarial activities
(Domains, 1995; Williams & Quadri, 1980). Consequently, prodigiosins are of interest because they might have potential clinical utility. Prodigiosins and synthetic derivatives have been shown to have potent and specific immunosuppressive activity, with novel targets clearly distinct from other drugs (D'Alessio et al., 2000; Mortellaro et al., 1999; Tsuji et al., 1990, 1992). Montaner et al. (2000) showed that prodigiosin extracted from S. marcescens could induce apoptosis in haematopoietic cancer cell lines including acute T-cell leukaemia, myeloma and Burkitt’s lymphoma, with little effect on non-malignant cell lines. Prodigiosin has also been shown to induce apoptosis in human primary cancer cells in the case of B and T cells in B-cell chronic lymphocytic leukaemia samples (Cámpas et al., 2003).

Previous attempts to reconstitute prodigiosin production in Escherichia coli have been unsuccessful. Dauenhauer et al. (1984) isolated a S. marcescens genomic clone capable of condensing the two prodigiosin precursors, MAP and MBC, to form prodigiosin. However, no sequence data were reported for this clone. We were able to demonstrate expression of prodigiosin in an heterologous host (Thomson et al., 2000). The pig gene cluster from Serratia sp. ATCC 39006 (Serratia 39006) was expressed in Erwinia carotovora subsp. carotovora (Ecc; 25 out of 36 strains tested), though it was not expressed in several other members of the Enterobacteriaceae, including E. coli (Thomson et al., 2000). In Serratia 39006 the production of prodigiosin is regulated by multiple factors, including a quorum-sensing system, via the LuxIR homologues, Smul and SmuR (Slater et al., 2003; Thomson et al., 2000). Interestingly, prodigiosin production became N-(3-oxohexanoyl)-L-homoserine lactone (OHL) dependent when expressed in Ecc, despite the latter producing a different signalling molecule from that made by Serratia 39006 (Thomson et al., 2000).

Initially it was thought that Serratia 39006 was an S. marcescens strain (Thomson et al., 1997). It is now clear that this Serratia strain is atypical and does not fall into the marcescens classification. We now refer to this strain by the ATCC nomenclature, strain 39006. In this paper we report the organization of the prodigiosin biosynthetic gene (pig) clusters in Serratia 39006 and in an S. marcescens strain, ATCC 274 (Sma 274) and we define the genomic context of the pig cluster in several S. marcescens strains. We discuss a pathway for prodigiosin biosynthesis based on predicted protein functions of the genes of the pig clusters and comparison with the recently proposed pathway for undecylprodigiosin (Cerdeño et al., 2001).

**METHODS**

**Bacterial strains, plasmids and culture conditions.** The bacterial strains and plasmids used in this study are shown in Table 1. E. coli strains were cultured at 37 °C and Erwinia and Serratia strains were cultured at 30 °C in Luria–Bertani medium. Serratia strains were also cultured on peptone glycerol media (PGM, 1% glycerol, 0.5% peptone, 1-6% agar) to enhance pigmentation. Antibiotic final concentrations were as follows, unless otherwise stated: 100 µg ml⁻¹ ampicillin (Ap), 50 µg ml⁻¹ kanamycin (Km), 50 µg ml⁻¹ spectinomycin (Sp), 100 µg ml⁻¹ streptomycin (Sm) and 10 µg ml⁻¹ tetracycline (Tc).

**DNA manipulations.** All DNA manipulations were performed as described in Sambrook et al. (1989). Genomic DNA from Serratia strains was extracted with the DNeasy Tissue kit (Qiagen). DNA fragments were isolated from agarose gels using a gel extraction kit (Qiagen). Isolation of cosmids and plasmids was performed with the Qiagen Miniprep kit. Oligonucleotide primers were purchased from Sigma-Genosys. Double-stranded sequencing was performed at the DNA Sequencing facility, Department of Biochemistry, University of Cambridge, using an ABI 3700 DNA analyser.

**Isolation and sequencing of the biosynthetic cluster of Serratia sp. ATCC 39006 and S. marcescens ATCC 274.** The cosmid pNRT104 carrying the pig genes of Serratia 39006 was identified previously (Thomson et al., 2000), and this was then subcloned to the smallest insert size (pTON245) that would still produce pigment in an Ecc host. The insert was further subcloned into pDAH330 (McGowan et al., 1995) and primer walking was used to sequence the resultant plasmids.

Chromosomal DNA of Sma 274 was partially digested with Sau3AI and phosphorylated and then ligated into the BamHI sites of SuperCos-1. The ligation products were packaged with a Gigapack III XL λ packaging kit (Stratagene) and used to transduce E. coli XL-1 Blue as per the manufacturer’s instructions. A prodigiosin-producing clone (pPIG4) was identified.

A transposon-based approach using the EZ::TN <TET-1> system (Epicentre) was employed to sequence pPIG4. Transposition in vitro was performed on pPIG4 and the transposition mix was used to transform E. coli XL-1 Blue by electroporation. Transformants were selected on LBA Tc plates and non-pigmented colonies selected for further analysis. Transposon-specific primers were used to obtain the sequence of the flanking DNA (forward primer 5'-CTCTAGA-GTCGACCTGAGGCCATGAAAGC-3' and reverse primer 5'-TAA-ATTGACTGAAAATCTAGAAA-3'). The sequences were assembled using the Staden package (Staden, 1996). Primer walking was used to join contigs and cover problematic areas. In total, 27,794 bp covering the pig cluster were determined with double-stranded coverage of each base in the consensus. ORFs were determined using GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/whmm_info.html) and Glimmer 2.0 (Delcher et al., 1999) and annotation was done using Artemis (Rutherford et al., 2000).

**Southern blotting using S. marcescens ATCC 274 and Serratia sp. ATCC 39006 pig clusters as probes.** Genomic DNA from Serratia strains was digested overnight with Aell (New England Biolabs). The digested DNA was transferred to a nylon membrane (Hybond-N, Amersham). The DNA probe was generated through a Bfod–Ascl digest of pPIG4 from which three fragments (9.5, 5.6, 5.3 kb) containing the pig genes were gel extracted (Qiagen) and DIG labelled by random prime (Roche). The DNA probe for the Serratia 39006 cluster was generated by DIG labelling pTON245 using random prime and this was used to probe MIlal-digested genomic DNA of S. marcescens strains.

**PCR amplification of cueR–pigA and pigN–copA.** The regions between cueR and pigA and pigN and copA were amplified by PCR using the primers cueR (5'-TGGTAAAACGGAAATGC-3'), PE1 (5'-GCAAACCTCGAGGGAATTGCC-3'), ab77 (5'-GAAACACTTAAACCTGACG-3') and PE2 (5'-GCGATTTCATGCGACGACG-3').
Table 1. Bacterial strains and plasmids used in this study

The strains shown in bold are pigmented and the strains in normal font are non-pigmented. The *Serratia marcescens* strains 1–19 were kindly donated by Hazel Aucken of the Public Health Laboratory Service (PHLS), UK.

<table>
<thead>
<tr>
<th>Bacterial strains</th>
<th>Strain</th>
<th>Note</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Serratia</strong></td>
<td>Sma 1695</td>
<td>Clinical</td>
<td>Blood</td>
</tr>
<tr>
<td>2</td>
<td>Sma 0006</td>
<td>Environmental</td>
<td>Food</td>
</tr>
<tr>
<td>3</td>
<td>Sma 1047</td>
<td>Clinical</td>
<td>Blood</td>
</tr>
<tr>
<td>4</td>
<td>Sma 0026</td>
<td>Environmental</td>
<td>Insect</td>
</tr>
<tr>
<td>5</td>
<td>Sma 4028</td>
<td>Clinical</td>
<td>Urine</td>
</tr>
<tr>
<td>6</td>
<td>Sma 3078</td>
<td>Clinical</td>
<td>Stool</td>
</tr>
<tr>
<td>7</td>
<td>Sma 0038</td>
<td>Environmental</td>
<td>Reservoir</td>
</tr>
<tr>
<td>8</td>
<td>Sma 1024</td>
<td>Clinical</td>
<td>Urine</td>
</tr>
<tr>
<td>9</td>
<td>Sma 0022</td>
<td>Environmental</td>
<td>Insect</td>
</tr>
<tr>
<td>10</td>
<td>Sma 3127</td>
<td>Clinical</td>
<td>Wound drain</td>
</tr>
<tr>
<td>11</td>
<td>Sma 0035</td>
<td>Environmental</td>
<td>Fig</td>
</tr>
<tr>
<td>12</td>
<td>Sma 3888</td>
<td>Clinical</td>
<td>Blood</td>
</tr>
<tr>
<td>13</td>
<td>Sma 0028</td>
<td>Environmental</td>
<td>Insect</td>
</tr>
<tr>
<td>14</td>
<td>Sma 3255</td>
<td>Clinical</td>
<td>Urine</td>
</tr>
<tr>
<td>15</td>
<td>Sma 0040</td>
<td>Environmental</td>
<td>Reservoir</td>
</tr>
<tr>
<td>16</td>
<td>Sma 2595</td>
<td>Clinical</td>
<td>Blood</td>
</tr>
<tr>
<td>17</td>
<td>Sma 0030</td>
<td>Environmental</td>
<td>Plant</td>
</tr>
<tr>
<td>18</td>
<td>Sma 3127</td>
<td>Clinical</td>
<td>Wound drain</td>
</tr>
<tr>
<td>19</td>
<td>Sma 3078*</td>
<td>Clinical</td>
<td>Stool</td>
</tr>
<tr>
<td>20</td>
<td>Sma ATCC 274</td>
<td>Environmental</td>
<td>Lab stock</td>
</tr>
<tr>
<td>21</td>
<td>Sma S6</td>
<td>Clinical</td>
<td>Yang <em>et al.</em> (1990)</td>
</tr>
<tr>
<td>22</td>
<td>Sma Db11</td>
<td>Environmental</td>
<td>Insect (Flyg <em>et al.</em>, 1980)</td>
</tr>
<tr>
<td><strong>Serratia sp. ATCC 39006</strong></td>
<td>Sma ATCC 274 MUT12</td>
<td>Biosynthetic mutant</td>
<td>Harris (2003)</td>
</tr>
<tr>
<td><strong>Erwinia</strong></td>
<td><em>Erwinia carotovora</em> subsp. <em>atroseptica</em>, strain SCRI 1043</td>
<td>Wild-type, restrictionless, carbapenem-producing strain</td>
<td>Lab stock</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td>SuperCos</td>
<td>ColE1 replicon, Ap', Km'</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pPIG4</td>
<td>SuperCos vector with ~41 kb insert containing the <em>Sma</em> 274 <em>pig</em> genes</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>pNRT104</td>
<td>psf6 vector with ~35 kb insert containing the <em>Serratia</em> 39006 <em>pig</em> genes</td>
<td>Thomson <em>et al.</em> (2000)</td>
<td></td>
</tr>
<tr>
<td>pTGN245</td>
<td>pHCP19 vector, containing the <em>pig</em> cluster (24 kb) from <em>Serratia</em> 39006</td>
<td>This study</td>
<td></td>
</tr>
</tbody>
</table>

*Strain Sma 19 is a spontaneous non-pigmented variant of Sma 6.*
RESULTS AND DISCUSSION

Organization of the *pig* clusters from *Serratia* sp. ATCC 39006, *S. marcescens* ATCC 274 and comparison with the undecylprodigiosin biosynthetic genes of *Str. coelicolor* A3(2)

We previously identified a cosmid (pNRT104) containing *Serratia* 39006 genomic DNA which encoded synthesis of prodigiosin in the heterologous host Ecc (Thomson *et al*., 2000). This cosmid was further subcloned and sequenced by primer walking to identify a cluster of 21 634 bp containing 15 candidate prodigiosin biosynthetic genes. To identify the *pig* genes in Sma 274, a cosmid library was constructed and used to transform *E. coli* XL-1 Blue. Approximately 0·1 % of the Ap+ colonies recovered after transduction produced a red pigment, presumed to be prodigiosin. A prodigiosin-producing clone (pPIG4) was selected. This represents the first successful reconstitution of expression of prodigiosin in *E. coli*. A transposon-based approach was used to generate pPIG4 *E. coli* XL-1 Blue non-pigmented transformants, and sequencing revealed a cluster of 20 960 bp which consisted of 14 putative prodigiosin biosynthetic genes.

The *Serratia* pig clusters contain 14 candidate genes common to both strains and are arranged *pigA* through to *pigN* (Fig. 1a). As described below, we have putatively assigned the products of five genes to the biosynthesis of MBC (red) and four to the MAP (blue) pathway. Four are unassigned, with the remaining protein (PigL) proposed to be involved in the post-translational modification of some of the proteins in the cluster (discussed later). The order of the genes is conserved between the two different *Serratia* species and the corresponding 14 predicted proteins are similar in size and share significant amino acid

---

**Fig. 1.** (a) Comparison of the prodigiosin biosynthetic clusters (*pig* cluster) from *Serratia* 39006, Sma 274, and the undecylprodigiosin biosynthetic cluster (red cluster) from *Str. coelicolor* A3(2). Lines between the *Str. coelicolor* red and *Serratia* pig clusters trace the homologue genes relative positions in the two clusters. Regulatory genes of the red cluster are shown in black, those thought to direct the synthesis of the bipyrrole moiety are coloured red, and those responsible for the monopyrrole moiety are shown in blue. Genes of no assigned function are shown in white and those considered not to be part of the cluster are coloured grey. The black arrows indicate transcriptional units of the red cluster (Cerdeño *et al*., 2001). (b) Proposed pathway for the biosynthesis of prodigiosin. This scheme is similar to that proposed for undecylprodigiosin (Cerdeño *et al*., 2001) and we have, as far as possible, assigned putative *pig* gene products to each of the steps. Proteins not assigned a role in this pathway are PigL, putative phosphopantetheinyltransferase and PigB, PigD, PigK and PigM. Each colour represents a different substrate whose incorporation into prodigiosin and its precursors is shown by the same colour. ? indicates no particular enzyme could be assigned to these roles.
identities between the species. However, the relative genomic contexts of the clusters and their regulation are very different.

The *Serratia* 39006 pig gene cluster contains an additional gene, tentatively designated pigO. RT-PCR and primer extension has confirmed that there is transcriptional read-through between pigN and pigO (Slater et al., 2003), consistent with pigO being part of the pig operon in that strain. However, the pigO gene has been inactivated and the pigment produced by the pigO mutant has been shown to be prodigiosin, the same as wild-type *Serratia* 39006 and Sma 274, by LC-MS analysis (unpublished). These data suggest that pigO is not involved in the biosynthesis of prodigiosin, but could have a regulatory role. The *Serratia* 39006 pig cluster (pigA–O) is transcribed as a polycistronic message from a promoter upstream of pigA (Slater et al., 2003), and so we presume that the pig cluster (pigA–pigN) of Sma 274 is also transcribed as a polycistronic message.

The only substantial gap in both pig clusters is between pigC and pigD. There is a 184 bp gap between pigC and pigD in *Serratia* 39006, suggesting a possibility of two transcriptional units. However, transcription levels of the genes either side of this gap were shown to be similar in studies using lacZ reporter fusions to pigA and pigH (Crow, 2001). The Sma 274 pig cluster has a 64 bp gap between pigC and pigD, smaller than that in *Serratia* 39006, with no other significant intergenic gaps.

Table 2 shows the relatedness of each *Serratia* 39006 Pig protein to its homologue in Sma 274. Although the 39006 and Sma 274 clusters contain the same genes and are orientated identically, they are only 68 % identical, on average, at the DNA level. There is also wide variation in the level of identity between proteins encoded by genes in the pig clusters. For example, the *Serratia* PigE homologues are 85-2 % similar (90-4 % similar) whilst the PigL homologues are only 52-7 % identical (56-1 % similar). This implies different functional constraints altering the rates of molecular evolution in different genes within the same cluster. Another striking difference between the two pig clusters is the G+C content of the genes. Each *Serratia* 39006 pig gene has a lower G+C content than its Sma 274 counterpart. The *Serratia* 39006 pig genes have a mean G+C content of 51-4 % whilst the mean G+C content of the Sma 274 pig genes is 59-8 % per gene. This compares to a mean of 58-4 % across the whole genus (Grimont &

Table 2. Deduced functions and homologues of the gene products from the pig clusters

Putative functions were assigned based on the results of BLASTP searches and comparison with the red cluster. The relatedness of each Sma 274 Pig protein to its homologue in *Serratia* 39006 is shown in the comparison of *Serratia* Pig proteins percentage sequence identity column. Sma 274 has no homologue of PigO. The number of amino acid residues (aa) in each of the Pig proteins is listed. In the Red homologue percentage sequence identity column, values on the left refer to *Serratia* 39006 and values on the right refer to the percentage sequence identity for the corresponding Sma 274 protein. PigL contains motifs which are conserved within the phosphopantetheinyl transferases (Lambalot et al., 1996). PigD and pigE are the only genes from the *Serratia* pig clusters not to contain a homologue in the red cluster.
Grimont, 1984) and 59% for the whole genome of S. marcescens Db11 (http://www.sanger.ac.uk/Projects/S_marcescens/). The significance of this difference in G+C content is unclear, although it may reflect the fact that Serratia 39006 is an atypical strain.

Biochemical, syntrophic and genetic analyses have also been done using the undecylprodigiosin-producing strain, Str. coelicolor A3(2) (Cerdeño et al., 2001; Coco et al., 1991; Feitelson et al., 1985; Rudd & Hopwood, 1980). Undecylprodigiosin synthesis is encoded by the red gene cluster, as shown in Fig. 1(a) (White & Bibb, 1997). A biosynthetic pathway for undecylprodigiosin synthesis has been proposed recently, based on the predicted functions of the proteins in the red cluster (Cerdeño et al., 2001).

Table 2 shows the predicted functions of pig gene products based on BLAST searches (http://www.ncbi.nlm.nih.gov/blast/) and the Red homologues of each of the Pig proteins. Comparison of the three prodigiosin gene clusters proved extremely interesting. The red cluster is larger (23 genes) than the pig clusters (14–15 genes), possibly reflecting the greater complexity of the undecylprodigiosin products made by the former (Tsao et al., 1985). Twelve genes are conserved between all three clusters, as might be expected of pathways making such similar products. However, their orientation and regulation seem to have undergone pronounced variation. The Serratia pig clusters are orientated and transcribed unidirectionally (Slater et al., 2003). In contrast, the Red cluster is arranged in four transcriptional units and is regulated by RedZ and RedD (Fig. 1a). The Red-dependent promoter region lies between RedQ and RedP and contains two divergent promoters directing transcription in opposite directions. RedZ, a response regulator, is thought to bind upstream of the RedD gene (White & Bibb, 1997). The dissimilar arrangement of these homologous genes, in clusters that produce chemically similar products, raises questions about the origins and evolution of prodigiosin biosynthetic gene clusters in both Gram-positive and Gram-negative bacteria. It is unclear whether there is a species-related functional significance to the dissimilar arrangement of the genes or whether the differences have arisen by random events.

Differences between the other genes (see Fig. 1a) present in these operons might explain differences in the final molecules produced, or the biochemical routes to their production. pigD and pigE are the only pig genes that do not have homologues in the red cluster. Their respective homologues, pdorfL and pdorfM, are part of a gene cluster encoding the synthesis of the carbon tetrachloride dechlorination agent pyridine-2,6-bis(thiocarboxylic acid) by Pseudomonas stutzeri (Lewis et al., 2000). pdorfL and pdorfM are contiguous, as are pigD and pigE in the pig cluster. PigE is also similar to a putative aminotransferase (SC6A5.18, 461 aa) from Str. coelicolor A3(2), encoded by a gene unlinked to the red gene cluster (38% identical to PigE from Serratia 39006 and 39% identical to PigE from Sma 274; both are 50% similar over 460 aa). Indeed, it is possible that genes encoding proteins that take part in the synthesis of prodigiosin and undecylprodigiosin may not be associated with the gene clusters. RedR, RedQ and RedP do not have homologues encoded by the pig cluster, yet are thought to direct the condensation of acetate to form three of the carbon atoms of the pyrrole ring and the undecyl side chain of undecylprodigiosin (Cerdeño et al., 2001). The genes encoding the corresponding homologues could be present at another location on the Serratia chromosome, or other enzymes may perform the necessary catalytic roles in prodigiosin biosynthesis in Serratia. However, as we have shown that prodigiosin can be made in recombinant E. coli and Erwinia, this would necessitate functional activity of the corresponding E. coli and Erwinia enzymes to supplement the activity of the enzymes encoded by the pig cluster.

The twelve Red proteins with homologues encoded in the pig cluster (the ‘core’ prodigiosin enzymes) are similar in size to their Pig counterparts. An exception to this is PigB (670 aa), whose homologue RedS (Sc3f7.05c) is a much smaller protein (146 aa) and is similar to only the N-terminal region of PigB. The rest of PigB shows sequence identity with amine oxidases. The PigI and PigJ proteins are also significantly smaller than their respective RedM and RedX homologues.

**Biosynthesis of prodigiosin**

Cerdeño et al. (2001) have suggested a likely sequence of steps for the biosynthesis of the closely related undecylprodigiosin in Str. coelicolor A3(2) and have assigned putative Red proteins to most of the steps. Undecylprodigiosin is thought to be formed from combination of the identical bipyrrole unit, MBC, with a slightly different monopyrrole (2-undecylpyrrole). We now suggest a scheme for the biosynthesis of prodigiosin, shown in Fig. 1(b), which is similar to that proposed for undecylprodigiosin and we have, as far as possible, assigned putative pig gene products to each of the steps.

**Biosynthesis of MBC**

We propose that the biosynthesis of MBC follows the same general chemical pathway in Serratia as in Str. coelicolor A3(2). Recent studies have shown that RedM, RedO and RedW from the red cluster, which have homologues in both pig clusters, are sufficient for the in vitro synthesis of a pyrrole group from L-proline (Thomas et al., 2002). RedM from Str. coelicolor, which is homologous to adenylation domains in non-ribosomal peptide synthetases (NRPSs), has been shown to activate the carboxy group of proline in vitro and transfer the prolyl group to RedO, a homologue of peptidyl carrier protein (PCP) domains of NRPSs (Cerdeño et al., 2001). The pyrrole group is then oxidized by RedW, which shares sequence identity with flavin-dependent acyl-CoA dehydrogenases. Homologues of these three proteins have been shown to be required for the incorporation of L-proline into the pyrrole moieties of other pyrrole-containing molecules.
compounds (Thomas et al., 2002; Xu et al., 2002). Thus PigI, which is homologous to RedM, is similarly presumed to adenylate proline and transfer the prolyl group to PigG, a homologue of RedO (Table 2). The prolyl-PigG is then presumably oxidized by PigA, which is a homologue of RedW. It is not known whether PigA performs both oxidations required to convert the prolyl group to a pyrrole, whether a second enzyme is also involved, or whether the second oxidation is non-enzymic (Thomas et al., 2002). PigA has good sequence identity with a wide range of acyl-CoA dehydrogenases. Sequence alignment with human isovaleryl-CoA dehydrogenase, the crystal structure of which has been determined, showed that the majority of amino acids around the flavin and the phosphopantetheinyl moiety of CoA are conserved but the amino acids involved in binding the adenosyl group of CoA are not conserved (Tiffany et al., 1997). This supports the suggestion that the substrate for PigA is a prolyl PCP rather than prolyl-CoA.

In the case of Str. coelicolor, it was proposed that the next enzyme in the biosynthesis of MBC is RedL, which appears to be a hybrid NRPS/polyketide synthase (PKS), consisting of the following six domains: an adenylation domain, an acyl carrier protein (ACP), a keto-synthase, an acyltransferase, another ACP, and a pyridoxal phosphate-dependent domain thought to be a serine transferase (Cerdeño et al., 2001). There is no similar multifunctional enzyme among the Pig gene products and so we assume that the various reactions are all catalysed by smaller monofunctional enzymes or by distinct mechanisms.

The next step in MBC biosynthesis in Serratia should, by analogy with that proposed for Str. coelicolor, be condensation of prolyl-PigG with malonyl-CoA or a malonyl ACP. As there is no apparent ACP in the pig cluster, we have shown malonyl-CoA as the substrate in Fig. 1(b), but it is quite possible that an ACP from outside the pig cluster is used, such as the fatty acid synthase ACP. The keto-synthase which catalyses this condensation is also unknown. There is one putative keto-synthase in the cluster (PigJ), which could catalyse this reaction, but PigJ has highest homology with RedX, which has been assigned to the biosynthesis of the monopyrrole moiety of undecylprodigiosin. Accordingly, we have assigned PigJ to the biosynthesis of MAP in Fig. 1(b) and not assigned any of the Pig proteins to the keto-synthase reaction in the biosynthesis of MBC.

In Str. coelicolor, it was proposed that formation of pyrrole ring B of undecylprodigiosin is catalysed by the terminal pyridoxal phosphate (PLP)-dependent domain of RedL. In the pig cluster there are two gene products with homology to aminotransferases, PigE and PigH. PigH is homologous to RedN, which is proposed to be involved in the biosynthesis of the monopyrrole moiety, so here we suggest that PigE is the one involved in MBC biosynthesis in Serratia, although it only has low homology (35% identity) with the putative PLP-dependent domain of RedL (Cerdeño et al., 2001).

The remaining two steps required for the formation of MBC are (i) oxidation of the hydroxymethyl group to an aldehyde and (ii) methylation of the hydroxyl group on ring B. PigF shows sequence identity to S-adenosylmethionine-dependent methyltransferases, including RedL, and is assumed to be responsible for the methylation (Cerdeño et al., 2001; Summers et al., 1992). In Str. coelicolor the gene product proposed for the oxidation step was RedK, which shows sequence identity to NAD(P)-dependent dehydrogenases. No homologue of RedK was found in the pig cluster and so the enzyme that catalyses this reaction is uncertain. One possibility is PigB, the majority of which shows sequence identity to FAD-dependent amine oxidas. Another possibility is that PigB catalyses one of the two oxidations required for the conversion of the prolyl-PigG to the corresponding pyrrole.

**Biosynthesis of MAP**

Four of the Pig proteins share homology with Red proteins proposed to be involved in the biosynthesis of 2-undecylpyrrole. The first of these, PigI, is homologous to RedX. Both of these proteins appear to have two keto-synthase-like domains, but in each case only the C-terminal domain can be active as the N-terminal domain has the active-site cysteine replaced by an aspartate residue. In PigI, unlike RedX, the amino acids flanking this aspartate are poorly conserved when compared with the active site of a true keto-synthase. It is proposed that the apparently active C-terminal domain of PigI catalyses the condensation of a malonyl thioester with a hexanoyl thioester to give a 3-keto-octanoyl thioester. The hexanoyl group is probably derived from fatty acid biosynthesis or degradation and may be introduced as either hexanoyl-CoA or a hexanoyl ACP. PigH, thought to be the next enzyme in the pathway, appears to have two ACP-like domains at its N-terminus, as does its homologue RedN. One possibility is that one or other of these ACP domains is malonylated (by an unknown acyltransferase) and this malonyl group then condenses with the hexanoyl thioester in the active site of PigI to give a 3-keto-octanoyl thioester of PigH. In this pathway we assign PigE and PigH to their respective reactions in MBC and MAP biosynthesis; however, both proteins are putative aminotransferases and therefore their roles could be exchanged.

The remainder of PigH (other than the two N-terminal ACP domains) shows sequence similarity with PLP-dependent enzymes, in particular 8-aminoo-7-oxononanoate synthase (AONS), one of the enzymes of biotin biosynthesis. In the AONS-catalysed reaction, glycine is decarboxylated and the remaining aminomethyl group attacks a CoA thioester, releasing free CoA (Alexeev et al., 1998; Webster et al., 2000). The reaction proposed for PigH is similar, except that the amino acid decarboxylated is alanine and the thioester has a β-keto group. It is this ketone group which is attacked by the α-carbon of the decarboxylated alanine while the thioester is attacked by the amino group, giving the five-membered pyrrolinone ring. Red N, the
homologue of PigH in Str. coelicolor, was proposed to catalyse a similar reaction, except that the amino acid involved is glycine and the \( \alpha \)-carbon attacks the thioester while the amine attacks the ketone. Sequence alignment of PigH with AONS from E. coli, for which a crystal structure is known, shows complete conservation of the amino acids surrounding the PLP in the active site (Webster et al., 2000). The lysine which forms the Schiff’s base with the PLP is predicted to be Lys-485 in PigH, and the aspartate which is hydrogen bonded to N-1 of PLP is predicted to be Asp-451.

The final step required for the formation of MAP is reduction of the amide group. The equivalent step in Str. coelicolor was proposed to be catalysed by RedH and RedF, and the homologous proteins in Serratia are PigC and PigN. PigC (like RedH) appears to have three domains. The N-terminal and C-terminal domains share sequence identity with the first two domains of pyruvate phosphate dikinase, which are the ATP-binding domain and a swivelling phosphoryl transfer domain, respectively. By homology, the phosphorylation site on the phosphoryl transfer domain of PigC is His-840 (Cerdeño et al., 2001; Herzberg et al., 1996). In PigC these two domains are separated by a domain that has no significant sequence identity to any protein of known function. This is presumably the binding domain for the substrate that is to be phosphorylated. From a chemical point of view, it makes sense that phosphorylation of the amide occurs, because without some form of activation, amides are generally too stable to be reduced by a NAD(P)H-dependent reductase. PigN is homologous to RedF and has low homology to known NAD(P)-dependent oxidoreductases (Prosite entry PS0299). It is therefore proposed to be the reductase that generates MAP.

### Biosynthesis of prodigiosin and other Pig proteins

The remaining pig gene products that have not been included in Fig. 1(b) are PigB, PigD, PigK, PigL and PigM. Of these, PigL has sequence identity (and contains conserved motifs) with phosphopantetheinyltransferases (PPTases) (Lambalot et al., 1996; Reuter et al., 1999). PigL is therefore presumed to transfer the 4'-phosphopantetheinyl moiety from CoA to the hydroxyl of the conserved serine residues in PigG and/or the ACP domains of PigH, converting these enzymes from the inactive apo- to the active holo- form. As mentioned above, PigB is a putative amine oxidase and is possibly involved in one or other of the proposed oxidation steps. Thus the remaining candidates for the enzyme which couples MBC and MAP to give prodigiosin are PigD, PigK and PigM. PigD is not likely to be the enzyme for this step as it has no homologue in the Red cluster. There are currently no significant sequence identities of PigK or PigM with enzymes of known function, so it is not possible to predict which is more likely to catalyse the final coupling reaction.

<table>
<thead>
<tr>
<th>Host</th>
<th>pPIG4</th>
<th>pTON245</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli XL-1 Blue</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>E. coli DH5x</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Serratia 39006 HSPIG28</td>
<td>+*</td>
<td>+</td>
</tr>
<tr>
<td>Sma 274 MUT 12</td>
<td>+</td>
<td>ND</td>
</tr>
<tr>
<td>Ecc</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

*Expression of pigment in Serratia 39006 HSPIG28 biosynthetic mutant complemented with pPIG4 is not to wild-type levels.

### Heterologous expression of prodigiosin genes in Serratia, Erwinia and E. coli strains and mutants

The two Pig\(^+\) cosmids (pPIG4 and pTON245) were used to transform electrocompetent cells of various Erwinia, E. coli and Serratia strains and the results are shown in Table 3. Surprisingly, pPIG4 from Sma 274 was not capable of heterologous expression in Ecc, despite encoding pigment production in E. coli XL-1 Blue and E. coli DH5x. The reverse was true for the Serratia 39006-derived cosmid, pTON245: it encoded prodigiosin synthesis in Ecc, but not in E. coli. The Sma 274 pPIG4 cosmid could complement biosynthetic mutants in both clusters, although complementation was not to wild-type levels in the Serratia 39006 biosynthetic mutant. The reasons for the genus/species-dependent expression of the two recombinant cosmids are not known.

### Differences in pig cluster regulation

Although the global arrangement of the pig clusters is very similar in Serratia 39006 and Sma 274, the regulation of the two Serratia pig clusters is markedly different. In Serratia 39006, transcription of the pig cluster is under quorum-sensing control via the small/R genes and the N-acyl homoserine lactones (AHLs) N-butanoyl-L-homoserine lactone (BHL) and N-(hexanoyl)-L-homoserine lactone (HHL) (Thomson et al., 2000). Sma 274 has been tested for production of AHLs and the presence of small/R homologues (data not shown), and no AHL-dependent quorum-sensing system has been detected (data not shown). Recently, luxS, which is responsible for the production of Autoinducer-2 (AI-2), has been shown to regulate prodigiosin production via extracellular signalling in Sma 274 (Coulthurst et al., 2004). No evidence of an AHL-dependent quorum-sensing system controlling pigment biosynthesis has been found in any of the other 22 S. marcescens strains tested, of which seven are pigmented (data not shown), implying that quorum-sensing control of prodigiosin production is both highly strain...
dependent and comparatively rare. Interestingly, a recent report has described a strain of *S. marcescens* in which pigment is affected by production of OHHL, another AHL (Horng *et al.*, 2002). Where AHLs are made by *S. marcescens* strains, BHL is usually the molecule produced. Therefore, there is considerable strain-dependent variation in both the ability to make AHLs and in the chemical nature of the molecule made.

Another conspicuous difference between the *pig* clusters is the intergenic region upstream of *pigA*. In Sma 274 this region is 488 bp compared to 922 bp in *Serratia* 39006. In *Serratia* 39006, transposon insertions in the region 200-320 bp upstream of the *pigA* transcriptional start site cause hyperpigmentation. This has been termed the ‘Rop1’ class of mutation and has been hypothesized to be due to disruption of a binding site for a repressor(s) (Slater *et al.*, 2003). The *Serratia* 39006 promoter has been mapped (Slater *et al.*, 2003). Based on this, corresponding −35 and −10 regions are proposed for Sma 274 in the region upstream of *pigA* (Fig. 2).

A long-documented phenomenon is hyperpigmentation of *S. marcescens* strains when grown on PGM (Morrison, 1966). Recent work in *Serratia* 39006 has confirmed that phosphate (P₁) limitation is a cause of prodigiosin operon upregulation (Slater *et al.*, 2003). In *Serratia* 39006, a mutation in the Pst transport system, the major route for inorganic phosphate uptake under phosphate-limiting conditions, physiologically mimics the phenotype of P₁ limitation. This notion is consistent with the identification of the Pho box covering the –35 and −10 regions of the *pig* cluster promoter of *Serratia* 39006 (Fig. 2) (Slater *et al.*, 2003). Hyperpigmentation due to growth on PGM is widely reported in various *S. marcescens* strains (Grimont & Grimont, 1978) and so the Pho box may be present in the 5′ regulatory region of all *S. marcescens* *pig* clusters. Indeed, a predicted Pho box has been identified 5′ of *pigA* in Sma 274 (Fig. 2). The Pho box consensus CTGTCAT A(AT)A(TA) CTGT(CA)A(CT) is composed of two 7 bp direct repeats separated by a 4 bp segment that is part of the −35 region (VanBogelen *et al.*, 1996). The two predicted Pho boxes in *Serratia* 39006 and Sma 274 are 4/7 identical to each other in the 7 bp repeated region and 3/4 in the middle region, but the 4 bp segment is not part of the −35 region in either.

---

**Fig. 2.** Promoter regions of the prodigiosin biosynthetic clusters (*pig* clusters) from *Serratia* 39006 (top) and *S. marcescens* ATCC 274 (bottom). (a) Promoter region of the 39006 *pig* cluster with the −10 and −35 regions underlined. The arrows show the predicted Pho box. (b) Predicted promoter region of the Sma 274 *pig* cluster with the −10 and −35 regions underlined and a predicted Pho box italicized and denoted with arrows. (c) Predicted *copA* promoter. (d) Putative transcriptional terminator sequence. The ‘Rop’ region is 200 to 320 bp upstream of the translational start of *pigA* in *Serratia* 39006.

http://mic.sgmjournals.org

---

The prodigiosin biosynthetic cluster of *Serratia*
Southern blotting using *S. marcescens* ATCC 274 and *Serratia* sp. ATCC 39006 *pig* clusters as probes

The Sma 274 *pig* cluster cosmid (pPIG4) was digested, DIG labelled and used to probe genomic DNA of several pigmented and non-pigmented *Serratia*. DNA from 12 of the 23 *S. marcescens* strains tested hybridized to the 274 *pig* cluster probe (Fig. 3). All of the pigmented strains have bands, many corresponding to the restriction pattern of Sma 274 from which the probe is generated, suggesting that restriction sites are conserved between many of the *pig* clusters. Interestingly, the prodigiosin biosynthetic operon is also present in some non-pigmented *S. marcescens* strains. These ‘cryptic’ *pig* clusters are found in six non-pigmented *S. marcescens* isolates (Sma 8, 9, 13, 10, 16 and 21). These strains were tested on LBA and PGM at 25, 30 and 37 °C to see if pigment production could be induced. They were also tested to see if cross-feeding from other pigmented strains could provide any intermediates or signal molecules that may be required for pigment synthesis, but with the exception of Sma 10, which appeared slightly pink on PGM, no pigment was induced in any of the experiments (data not shown).

The *Serratia* 39006 *pig* genes were DIG labelled using random prime (Roche) and used to probe genomic DNA from the same *S. marcescens* strains. The restriction pattern of all of the strains that hybridized was more similar to the Sma 274 restriction pattern than the *Serratia* 39006 restriction pattern. This is consistent with the observation that *Serratia* 39006 is not in the species *marcescens* and is quite different from the other *S. marcescens* strains. Five strains (8, 9, 10, 13 and 16) were shown to harbour cryptic *pig* clusters when probed with both cluster-specific probes. Of these strains, three are clinical isolates and two are environmental. It has long been observed that non-pigmented antibiotic-resistant strains of *S. marcescens* are much more common in the clinical setting than in the natural environment (Aucken & Pitt, 1998; Carbonell et al., 2000; Grimont & Grimont, 1978). There is no obvious reason for the prevalence of non-pigmented clinical strains,

---

**Fig. 3.** Distribution of the *pig* cluster within pigmented and non-pigmented *S. marcescens* strains. Southern hybridization of restriction enzyme digested chromosomal DNAs of *S. marcescens* strains isolated from different sources using (a) *Aci*I-digested genomic DNA and probed with a Sma 274 *pig* cluster-specific probe and (b) *Mlu*I-digested genomic DNA and probed with a *Serratia* 39006 *pig* cluster-specific probe. Numbers are as in Table 1, with the pigmented strains underlined.
though multiple theories have been presented (Blazevic, 1970; Cooksey et al., 1976; Ding & Williams, 1983; Grimont & Grimont, 1978; Holland & Dale, 1978). Presumably, there is some pressure counter-selecting expression of prodigiosin in a clinical environment where fitness of Serratia strains may be improved by loss of prodigiosin production due to biosynthetic or regulatory mutations.

**Relative genomic contexts of the pig clusters**

Flanking the Serratia 39006 pig cluster are genes transcribed in the same direction as the pig genes. The product of the 5’ gene (orfY) shows sequence similarity to NADH oxidases and contains the pyridine nucleotide-disulphide oxidoreductase dimerization domain. The 3’ gene encodes a hypothetical protein with no known significant homologues. In contrast to Serratia 39006, the genes flanking the pig gene cluster in Sma 274 (Fig. 1a) encode proteins similar to CueR (68% identity and 76% similarity to YbbI, accession number AAK06655) and CopA (69% identity and 80% similarity to CopA, accession number Q59385) of E. coli (Rensing et al., 2000; Stoyanov et al., 2001). CueR is the transcriptional regulator of CopA. CopA is a member of the MerR family of transcriptional regulators, which occur in a range of bacterial genera and respond to a wide variety of external stimuli (Amable-Cuevas & Demple, 1991; Brocklehurst et al., 1999; Chiu et al., 1999; Stoyanov et al., 2001; Summers, 1992). CopA is a copper-transporting P-type ATPase efflux pump and a central component of copper homeostasis in E. coli (Rensing et al., 2000). As shown in Fig. 4, the copA and cueR genes and their homologues are typically contiguous and divergently transcribed. It is therefore interesting that the complete pig gene cluster interrupts these genes in strain Sma 274.

We examined 22 S. marcescens isolates to see whether the cueR/copA genes always flank the pig cluster, or are adjacent – as more typically found in other bacteria. The regions between cueR and pigA and pigN and copA were amplified by PCR using the primers cueR, PE1, ab77 and PE2. As shown in Fig. 4, this PCR can generate a cueR–copA product if these homologues are adjacent to each other, or can generate two products (cueR–pigA and pigN–copA) if cueR and copA flank the pig cluster. The pigmented strains had the cueR–copA homologues arranged at either end of the pig cluster, as in Sma 274, with the exception of strain Sma 11 which generated no PCR products (Fig. 4). The cueR and copA homologues also flanked the pig cluster in four S. marcescens strains containing a cryptic pig locus. In Sma 17, 12 and Db11 the 173 bp PCR product was sequenced and found to be similar to the intergenic region between cueR and copA, showing that the cueR–copA genes are adjacent and divergent in these S. marcescens strains, as more commonly found in other bacteria. This may suggest an ancestral state of adjacent cueR–copA genes, with the pig cluster having recombined into this intergenic locus comparatively recently in some strain lineages. These observations could suggest that the pig cluster may be under CueR-mediated control and/or that one or more of the Pig enzymes requires copper, or a related metal, as a cofactor. The CueR/CopA system is involved in copper efflux (Changela et al., 2003; Petersen & Møller, 2000; Stoyanov et al., 2001). It is therefore intriguing that prodigiosin has been shown to effect Cu-dependent DNA cleavage (Melvin et al., 2001; Park et al., 2003) and so it is conceivable that intracellular copper availability must be regulated in concert with prodigiosin biosynthesis to avoid DNA damage.

In E. coli K-12 the cueR–copA intergenic region contains the copA promoter where CueR binds and promotes transcription of copA (Stoyanov et al., 2001). This entire intergenic region between cueR and copA is very similar to the sequence found in the intergenic region between pigN and copA in Sma 274 (Fig. 2). The promoter is highly conserved between E. coli K-12 and Sma 274, with 32/40 bases identical over the whole region. Sequence analysis of the PCR products from strains Db11, Sma 12 and 17, in which the cueR–copA genes are adjacent to each other rather than either side of the pig cluster, had an identical predicted copA promoter region to Sma 274 (data not shown). Database searches have shown this promoter to be highly conserved in many bacteria that contain cueR and copA homologues, such as E. coli O157:H7, Yersinia pestis, Salmonella typhi and Sal. typhimurium (data not shown). Despite the location of the pig cluster, the presence of the copA promoter, the putative transcriptional terminator (Fig. 2) and the 183 bp gap between pigN and copA might suggest that the copA gene is regulated independently of the pig genes in Sma 274.

**Summary**

Our results show that the genes for prodigiosin biosynthesis in Serratia lie in a large operon, at least for strains from two different Serratia species. Although some of the encoded ‘core’ enzymes are similar to those found in the red cluster of Str. coelicolor, there are some clear differences between the biosynthetic components from the Gram-negative and Gram-positive clusters. We have assigned a role for the protein of each gene where a putative function can be predicted based on sequence identity with proteins of known function and comparison with the proposed pathway for undecylprodigiosin biosynthesis. We are currently investigating the biosynthesis of prodigiosin further to confirm the function of each protein in the biosynthetic pathway. The transcriptional organization is retained in the two Serratia clusters, but is completely different in Streptomyces. Furthermore, the genomic context is quite different in the two species of Serratia investigated in this study. Among S. marcescens strains there is considerable heterogeneity. Some strains have no pig genes or have cryptic pig genes. Where a pig cluster was present in the S. marcescens strains, the pig genes were flanked by cueR and copA, genes involved in metal (copper)-dependent regulation and efflux. This observation begs the question of whether or not there is a physiological connection between copper (or other metal) homeostasis and prodigiosin biosynthesis.
Fig. 4. PCR amplification of the cueR to pigA intergenic region and the pigN to copA intergenic region from S. marcescens strains 1–21. Two PCR products are amplified (B, cueR–pigA and C, pigN–copA) if cueR and copA flank the pig cluster as is the case for Sma 274 (lane 20). PCR product C is amplified if cueR–copA are contiguous. The cosmid PPG4 contains the pig genes and flanking DNA (50 kb total) from Sma 274. Numbers are as in Table 1 with pigmented strains underlined, except number 23 which is Erwinia carotovora subsp. atroseptica strain SCRI 1043 (http://www.sanger.ac.uk/Projects/E_carotovora/). Lane 24 is a negative control.
ACKNOWLEDGEMENTS

We thank Martin Welch and members of the Salmond group for helpful discussions. We would also like to thank Hazel Aucken for kindly supplying the *Serratia marcescens* environmental and clinical isolates. This study was supported generously by the BBSRC, UK. A. K. P. H., A. C. and S. A. were supported by BBSRC studentships.

REFERENCES


Regulated by CarR, a homolog of the Lux transcriptional activator.


