Characterization of the Mupirocin Biosynthesis Gene Cluster from *Pseudomonas fluorescens* NCIMB 10586

A. Kassem El-Sayed,[1,3] Joanne Hothersall,[1]
Sian M. Cooper,[1] Elton Stephens,[1]
Thomas J. Simpson,[2] and Christopher M. Thomas[1,*]
[1]School of Biosciences
University of Birmingham
Edgbaston
Birmingham B15 2TT
[2]School of Chemistry
University of Bristol
Cantock's Close
Bristol BS8 1TS
United Kingdom

Summary

The polyketide antibiotic mupirocin (pseudomonic acid) produced by *Pseudomonas fluorescens* NCIMB 10586 competitively inhibits bacterial isoleucyl-tRNA synthase and is useful in controlling *Staphylococcus aureus*, particularly methicillin-resistant *Staphylococcus aureus*. The 74 kb mupirocin biosynthesis cluster has been sequenced, and putative enzymatic functions of many of the open reading frames (ORFs) have been identified. The mupirocin cluster is a combination of six larger ORFs (*mmpA-F*), containing several domains resembling the multifunctional proteins of polyketide synthase and fatty acid synthase type I systems, and individual genes (*mupA-X* and *macpA-E*), some of which show similarity to type II systems (*mupB*, *mupD*, *mupG*, and *mupS*). Gene knockout experiments demonstrated the importance of regions in mupirocin production, and complementation of the disrupted gene confirmed that the phenotypes were not due to polar effects. A model for mupirocin biosynthesis is presented based on the sequence and biochemical evidence.

Introduction

Analysis of genes for polyketide biosynthesis has revealed rules that add to our ability to predict how certain compounds are synthesized and to manipulate the biosynthetic genes to generate novel variants with potentially new and useful properties [1]. By analyzing additional biosynthetic pathways in diverse species, the rules can be made more general and the repertoire of tools for combinatorial biosynthesis expanded.

Mupirocin is a mixture of polyketides (pseudomonic acids) produced by *P. fluorescens* NCIMB 10586. Pseudomonic acid A, which accounts for 90% of the mixture, consists of a C_{17} unit (monic acid [MA]) thought to be derived from an unsaturated polyketide containing a pyran ring, and a C_9 saturated fatty acid (9-hydroxynona-

*Correspondence: c.m.thomas@bham.ac.uk
[3]Present address: Botany Department, Faculty of Science (Damietta), University of Mansoura, Damietta El-Gededa, PO Box 34517, Egypt.
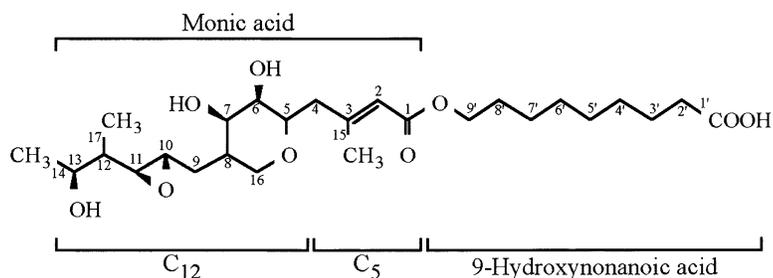
noic acid [9-HN]) [2–4] (Figure 1). Pseudomonic acid B (8%) has an additional hydroxyl group at C8 [4], pseudomonic acid C (<2%) has a double bond in place of the epoxide group at C10-C11 [5], and pseudomonic acid D (<2%) has an unsaturated fatty acid side chain with an alkene group at C4′-C5′ [6]. Novel pseudomonic acid analogs have also been isolated from marine organisms, e.g., *Altermonas* sp. SANK 73390 [7].

Mupirocin is used to control *S. aureus*, particularly methicillin-resistant *S. aureus* (MRSA), when other antibiotics are ineffective. It competitively inhibits bacterial isoleucyl-tRNA synthase (IleRS) [8] by interaction of the methyl end of MA, which has the same carbon skeleton as isoleucine, with the amino acid binding site of IleRS, while the rest of the MA part and/or C_9 side chain interacts with the ATP binding site [9]. Formation of Ile.tRNA is blocked, thus inhibiting protein synthesis. Despite low affinity for mammalian IleRS [10], mupirocin can only be used topically because it is quickly inactivated in the serum by hydrolysis of the ester linkage between MA and 9-HA. Ninety-five percent is also serum bound, resulting in poor bioavailability [11]. Resistance to mupirocin is also spreading due to a gene for a mupirocin-insensitive IleRS [for review, see 12]. Manipulation of mupirocin biosynthesis may make more versatile variants that are active against resistant strains and may produce novel structures with completely different activities.

Polyketide biosynthesis occurs by a process similar to fatty acid biosynthesis. A simple carboxylic acid starter unit, typically acetate or propionate activated as a CoA thioester, is transferred to the cysteine group of a β-ketoacyl synthase (KS), and a dicarboxylic acid extender unit, usually a malonate or methyl malonate CoA thioester, is transferred to the thiol group of the phosphopantetheine arm of an acyl carrier protein (ACP). These are joined by a decarboxylative condensation catalyzed by KS and remain covalently attached to ACP as a thioester. The β-carbonyl group of the resulting intermediate may be partially or fully reduced by the activities of β-ketoacyl reductases (KR), dehydratases (DH), and enoyl reductases (ER) acting sequentially or left unreduced. The growing chain is transferred from the ACP back to a KS, and further rounds of elongation and reduction occur to achieve the full-length polyketide chain, which may be further modified by the activity of tailoring enzymes. The completed chain is typically released from the polyketide synthase (PKS) by hydrolysis catalyzed by a terminal thioesterase (TE). The catalytic sites of PKSs may be encoded either by distinct proteins (type II) or by domains within large multifunctional proteins (type I). These may be used iteratively, as for aromatic polyketides, or, in the case of complex polyketides, may be modular, with each module encoding a specific round of condensation and reduction. In some systems, a methyl group may be introduced at the β-ketothioester stage by the action of a *C*-methyl transferase (MeT), which transfers the methyl group from *S*-adenosyl-methionine (SAM) onto the activated meth-

Figure 1. Structure of Pseudomonic Acid A

Pseudomonic acid consists of a C_{17} unit (monic acid) derived from an unsaturated polyketide containing a pyran ring, and a C_9 saturated fatty acid (9-hydroxynonanoic acid).

ylene of the β-ketothioester. In fungal type I systems, the MeT is encoded as part of the PKS gene.

Mupirocin biosynthesis has previously been studied by classical methods such as feeding of isotopically labeled substrates. This showed that pseudomonic acid A is derived from a linear combination of acetate-derived units, except for C16 and C17, which are derived from methyl of SAM, and C15, which has been proposed to be derived from acetate via 3-hydroxy-3-methylglutarate (HMG) [13, 14]. The oxygen atoms attached to C1, C5, C7, C13, and C9 are acetate-derived, and the C1 ester linkage is proposed to be formed from condensation of MA and 9-HA subunits [15].

Transposon mutagenesis, gene cloning, and reverse genetics identified a >65 kb region involved in mupirocin biosynthesis [16]. In this paper we describe the sequence of the cluster, its predicted enzyme activities, and a model for the mupirocin biosynthetic pathway that is largely supported by the sequence.

Results and Discussion

Cloning and Sequencing of a Mupirocin Biosynthesis Cluster

Deletion derivatives of two previously described plasmids, pBROC130/131 [16], were constructed and used for DNA sequencing. A strategy involving suicide vector pAKE100 or pCAW5.3 was then used to clone chromosomal fragments extending upstream and downstream from this region. Fragments from already cloned chromosomal DNA were inserted into a suicide vector, and the derivative was transferred into *P. fluorescens* by biparental mating. Cointegrants with the plasmid inserted by homologous recombination into the chromosome were selected. Chromosomal DNA was isolated and fragmented by digestion with a restriction enzyme that does not cut in the vector or the DNA already cloned. The resulting DNA fragments were circularized by ligation at dilute DNA concentrations and transformed into *Escherichia coli* HB101. Clones containing the suicide vector and the largest upstream or downstream chromosomal fragments were selected and either subcloned or used directly for sequencing. This procedure was used repeatedly in a step-wise fashion to perform a chromosome walk. A cointegrant (*P. fluorescens* 33.2:5.8) constructed previously [16] was used to clone the chromosomal DNA at the other end of the mupirocin cluster. Finally, the gap between the two blocks was obtained by PCR. Details of the clones encompassing the mupirocin biosynthesis cluster are available as supplemental data

(see http://www.chembiol.com/cgi/content/full/10/5/419/DC1 or write to chembiol@cell.com for a PDF).

The whole chromosomal segment was sequenced in both directions, and the sequence of the 74 kb mupirocin biosynthesis cluster has been deposited with GenBank (AF318063). The GC content ranges from 50% to 68% and is on average 61%, typical of other Pseudomonads (60%) [17]. Open reading frames (ORFs) were identified using a genomic codon preference table of *P. fluorescens* (www.kazusa.or.jp/codon) and GCG program [18]. All the ORFs identified use ATG (86%), GTG, (11%), or TTG (3%) as the start codon preceded by good ribosome binding sites. All ORFs showed ~60% or higher GC content except for *mupI* (51%) and *mupR* (50%), encoding quorum regulators of *mup* gene expression [19]. Thus, these genes may have been relatively recently recruited. Alignments of the DNA sequence and amino acid sequence of all six frames were performed using FASTA and BLAST to identify putative functions of the ORFs, with the predictions shown in Table 1 and Figure 2. The first 40.5 kb contains two individual genes (*mupA* and *mupB*) and four large ORFs (*mmpA, B, C, and D*) whose predicted products are referred to as mupirocin multifunctional proteins (Mmp).

We previously identified a putative MupR box upstream of *mupA* that may activate transcription from the *mupA* promoter [19]. Transcription from this promoter may proceed across the whole *mup* gene cluster since all except three of the ORFs identified (*mupF*, *mupP*, and *mupI*) run in the same direction as *mupA*. No potential internal promoters were identified by similarity either to consensus sequences for any of the known sigma factors or to the putative MupR binding site.

Upstream of the putative *mupA* promoter region, we found genes for tRNA$_{Val}$ and tRNA$_{Asp}$. Southern blotting revealed that *mup* DNA was not present in *P. fluorescens* SBW25, which does not produce mupirocin, but DNA on the other side of the tRNA cluster was present in SBW25. This interpretation has recently been confirmed by alignments with a contiguous sequence from the ongoing *P. fluorescens* SBW25 genome sequencing project (www.sanger.ac.uk/Projects/P_fluorescens/). This suggests that the *mup* cluster may be an insertion integrated by recombination at tRNA genes as observed for a number of phage and mobile elements [20].

Domains Typical of Polyketide Production Encoded in the *mup* Cluster

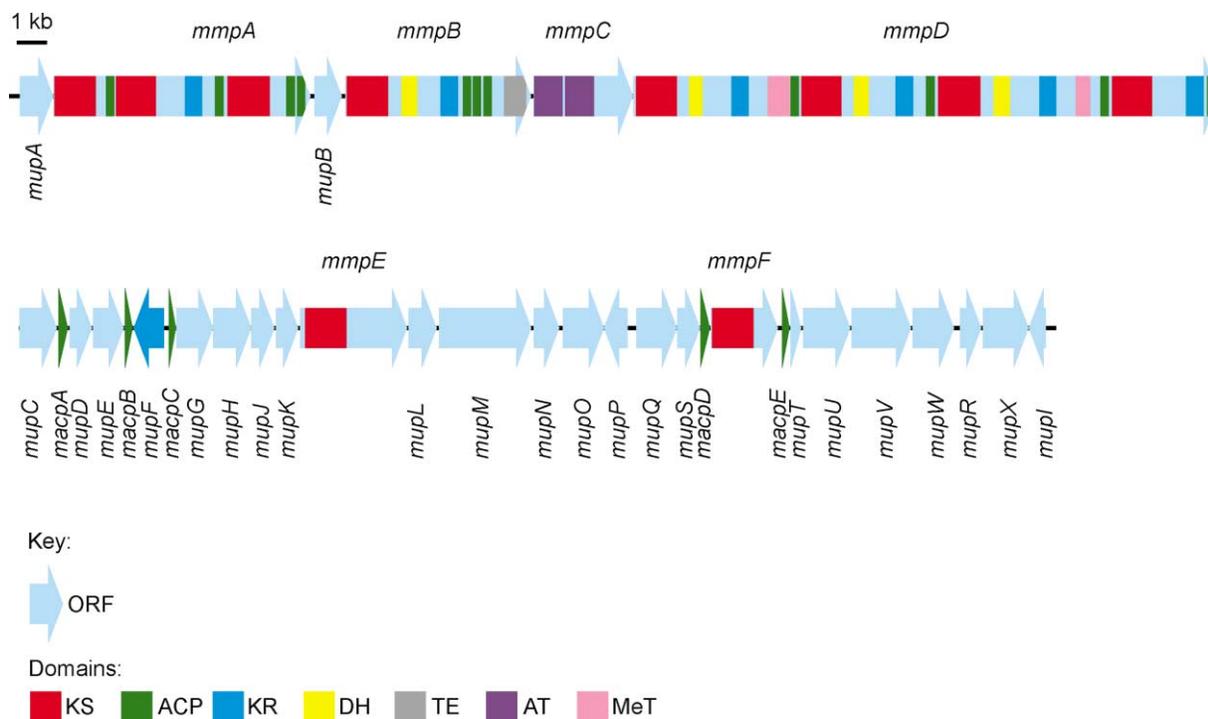The cluster contains various individual genes (*mupE, F, J, K, Q*, and *U* and *macpA, B, C, D*, and *E*) and six

Figure 2. Organization of Mupirocin Biosynthesis Cluster

multifunctional genes (*mmpA, B, C, D, E*, and *F*) with predicted type I PKS and fatty acid synthase (FAS) functions. In general, the closest relatives of the *mup* domains are found in genes for polyketide biosynthesis in *Bacillus subtilis* [21] or erythromycin biosynthesis in *Saccharopolyspora erythraea*, [22] as listed in Table 1. Some of the individual genes (*mupB, mupD, mupG*, and *mupS*) show similarity to PKS genes of type II systems. It therefore seems that the mupirocin PKS is a combination of both type I and type II PKS systems. Those genes and domains typical of polyketide production identified within the *mup* cluster are summarized below.

Acyltransferases
Only two potential acyltransferase (AT) domains were identified, both located in MmpC. Both AT1 and AT2 show similarity to AT from type I systems such as DEBS, responsible for production of the 6-deoxyerythronolide aglycone of erythromycin in *S. erythraea* (27%–29% identity at the amino acid level). AT2 shows higher identity (40%) with FabD, which possesses malonyl-CoA-ACP transacylase activity [23]. AT domains of PKS and FAS systems can be divided into two main groups according to six conserved motifs. The first includes malonyl-CoA- and acetyl-CoA-dependent AT, and the second includes methylmalonyl-CoA- and propionyl-CoA-dependent AT [24–26]. AT1 and AT2 show similarity to both types but are more similar to each other than to other members of either of the subgroups.

β-Ketoacyl Synthases
Ten potential β-ketoacyl synthase (KS) domains of the PKS type I system were identified. KS domains contain an active site Cys within a taCsssl motif and two His residues at ~135 and 175 aa downstream of the Cys active site [24]. All *mup* KS domains contain a Cys resi-

due within a (t/a/s)(a/m)Csssl motif, except KS1 (amCagg). The two His residues are also present, except for the first His residue of KS1 and KS9, which are Gln and Ser, respectively.

β-Ketoacyl Reductases
Six potential β-ketoacyl reductase (KR) domains were identified. KR domains possess a potential NADP(H) binding motif. KR1 matches a motif (GxGxxAxxxA) present in the rapamycin and erythromycin KRs [22, 24], while Ala replaces the first Gly of KR2. The KR3-KR6 motifs (GxGxxGxxxT/A/C) are similar to the niddamycin KR motif [27].

Dehydratases
Four potential dehydratase (DH) domains were identified. DH1 and DH3 possess a motif (HxxxGxxxxP) found in other DH domains of erythromycin, rapamycin, and niddamycin [22, 28]. The DH4 motif contains Ala instead of Pro, while DH2 contains Asp and Ala instead of Gly and Pro, respectively. We predict that DH2 is unlikely to be active (see Figure 3A).

Methyltransferases
Two potential methyltransferase (MeT) domains were identified within MmpD. Both MeT1 and MeT2 possess the conserved motif (LexGxGxG) found in MeT domains involved in *C*-methylation of lovastatin by *Aspergillus terreus* [29] and yersiniabactin by *Yersinia pestis* [30].

Acyl Carrier Proteins
Sixteen potential acyl carrier proteins (ACPs) were identified: 11 within Mmps and 5 present as individual genes. The characteristic ACP motif GxDS [24] with the essential serine was present in Mmp ACP1-11 and MacpA and B, but MacpC, D, and E have the serine residue within GANS, GLSS, and KINS, respectively. An unusual but not unique feature of the *mup* pathway is the exis-

Table 1. Organization of the Mupirocin Biosynthesis Gene Cluster

Gene/mmp	Function[a]/Domain	Location (bp)[b]	Size (aa)	GC (%)	Closest Relative (% identity; protein; organism)	Accession No.
mupA	FMNH$_2$-dependent oxygenase	600–1709	369	59	27, Y4VJ; *X. luminescens*	P07740
mmpA		1706–10246	2846	63		Q05470
	KS1	1787–3172		61	37; KS2 Pks1; *B. subtilis*	Q05470
	ACP1	3533–3799		65	38; ACP PksM; *B. subtilis*	P40872
	KS2	3842–5185		62	47; KS2 Pks1; *B. subtilis*	Q05470
	KR1	6161–6739		64	45; KR1 Pks1; *B. subtilis*	Q05470
	ACP2	7172–7438		64	43; ACP2 Pks1; *B. subtilis*	Q05470
	KS3	7589–9001		63	62; KS2 Pks1; *B. subtilis*	Q05470
	ACP3	9551–9820		61	31; ACP3 Pks1; *B. subtilis*	Q05470
	ACP4	9857–10132		62	27; ACP4 Pks1; *B. subtilis*	Q05470
mupB	3-oxoacyl-ACP synthase	10281–11186	301	61	25; FabH; *P. aeruginosa*	P20582
mmpB		11328–17558	2076	64		Q05470
	KS4	11397–12770		62	61; KS2 Pks1; *B. subtilis*	Q05470
	DH1	13206–13745		65	33; PksK; *B. subtilis*	P40803
	KR2	14610–15188		63	47; KR PksK; *B.subtilis*	P40803
	ACP5	15351–15623		64	34; ACP3 Pks1; *B.subtilis*	Q05470
	ACP6	15633–15905		61	31; ACP; *R. sphaeroides*	P12784
	ACP7	15915–16187		60	31; ACP, *S. rimosus*	P43677
	TE	16668–17426		65	26; TE Ery3; *S. erythraea*	Q03133
mmpC		17632–20964	1110	62		P03132
	AT1	17632–18591		62	28; AT1 Ery2; *S. erythraea*	P73242
	AT2	18604–19569		61	40; MCT FabD; *Synechocystis* sp.	
	Function unknown	20128–20964		62		
mmpD		20961–40526	6521	65		
	KS5	21063–22430		64	43; KS2 Pks1; *B. subtilis*	Q05470
	DH2	22833–23303		68	25; DH1 Pks1; *B. subtilis*	Q05470
	KR3	24258–24821		63	40; KR PksM; *B. subtilis*	P40872
	MeT1	25476–26210		64	52; PksM; *B. subtilis*	C69679
	ACP8	26229–26498		62	35; ACP2 Ery3; *S. erythraea*	Q03133
	KS6	26583–27917		63	49; KS3 Pks1; *B. subtilis*	Q05470
	DH3	28302–28844		64	29; PksK; *B. subtilis*	P40803
	KR4	29703–30293		65	42; KR PksK; *B. subtilis*	P40803
	ACP9	30720–30998		59	36; ACP2 Pks1; *B. subtilis*	Q05470
	KS7	31098–32504		64	59; KS2 Pks1; *B. subtilis*	Q05470
	DH4	32955–33515		65	32; PksM; *B. subtilis*	P40872
	KR5	34467–35030		62	49; KR PksM; *B. subtilis*	P40872
	MeT2	35694–36422		64	52; PksM; *B. subtilis*	C69679
	ACP10	36477–36749		60	32; ACP1 Ery3; *S. erythraea*	Q03133
	KS8	36870–38216		65	48; KS3 Pks1; *B. subtilis*	Q05470
	KR6	39339–39914		64	35; KR1 Pks1; *B. subtilis*	Q05470
	ACP11	40257–40523		66	36; ACP4 Pks1; *B. subtilis*	Q05470

(continued)

Table 1. Continued

Gene/mmp	Function[a]/Domain	Location (bp)[b]	Size (aa)	GC (%)	Closest Relative (% identity; protein; organism)	Accession No.
mupC	NADH/NADPH oxidoreductase	40523–41818	431	62	27; YqjM; *B. subtilis*	P54550
macpA	ACP	41870–42175	101	54	32; ACP; *Synechocystis* sp.	P20804
mupD	3-oxoacyl-ACP reductase	42216–42959	247	63	37; NodG; *Rhizobium* sp.	P72332
mupE	Enoyl reductase	42983–44005	340	63	29; Adh3; *B. stearothermophilus*	P42328
macpB	ACP	44062–44313	83	54	32; ACP; *A. aeolicus*	O67611
mupF	KR	45344–44334	336	63	28; DfrA; *C. chinesis*	P51103
macpC	ACP	45506–45742	78	60	30; ACP Acp5; *B. napus*	P08971
mupG	3-oxoacyl-ACP synthase I	45729–46967	412	62	53; PksF; *B. subtilis*	P40804
mupH	HMG-CoA-synthase	46964–48229	421	59	62; PksG; *B. subtilis*	P40830
mupJ	Enoyl-CoA-hydratase	48226–48993	255	60	51; PksH; *B. subtilis*	P40805
mupK	Enoyl-CoA-hydratase	49047–49790	247	60	62; PksI; *B. subtilis*	P40802
mmpE		49802–53377	1191	63		
	KS9	49934–51331		62	45; KS2 Pks1; *B. subtilis*	Q05470
	Hydroxylase	52166–53374		64	27; TcmG; *S. glaucescens*	P39888
mupL	Putative hydrolase	53423–54361	312	60	26; XylF; *P. putida*	P23106
mupM	Isoleucyl tRNA synthase	54444–57536	1030	59	41; Syi; *M. thermoautotrophica*	P56690
mupN	Putative regulatory DNA binding protein	57597–58448	233	62	34; HetI; *Anabaena* sp.	P37695
mupO	Cytochrome P450	58574–59938	454	59	25; Cp51; *M. tuberculosis*	P77901
mupP	Function unknown	60615–59821	264	64		
mupQ	Acyl-CoA-synthase	60912–62267	451	59	30; LucI; *L. cruciata*	P13129
mupS	3-oxoacyl-ACP reductase	62264–63037	257	60	35; NodG; *Rhizobium* sp.	P72332
macpD	ACP	63031–63351	106	60	30; ACP; *P. purpurea*	P51280
mmpF		63348–65585	745	62		
	KS10	63375–64769		62	39; KS2 PksK; *B. subtilis*	P40803
macpE	ACP	65730–65972	80	59	21; ACP; *C. paradoxa*	P48078
mupT	Ferredoxin dioxygenase	65995–66345	116	64	39; NdoA; *P. putida*	P23082
mupU	Acyl-CoA-synthase	66415–67992	525	59	27; LcfA; *H. influenzae*	P46450
mupV	Oxidoreductase	68027–70015	662	57	37; F4re; *M. organophilum*	P80951
mupW	Dioxygenase	70055–71446	463	56	31; BnzA; *P. putida*	P08084
mupR	N-AHL-responsive transcriptional activator	71638–72342	234	50	41; LasR; *P. aeruginosa*	P25084
mupX	Amidase/hydrolase	72380–73918	512	59	37; Ami2; *M. tuberculosis*	Q11056
mupI	N-AHL synthase	74486–73911	191	51	54; LasI; *P. aeruginosa*	P33883

[a]Putative function.
[b]Nucleotide position in the DNA cluster, accession number AF318063.

tence of a tandem ACP doublet (ACP3 and 4) and a tandem ACP triplet (ACP5, 6, and 7). They may hold multiple substrates prior to condensation or provide increased throughput at rate-limiting steps in the pathway. Tandem ACPs have also been observed in various type I iterative fungal PKS, e.g., WA, a naphthopyrone synthase of *Aspergillus nidulans* in which both ACP function but only one is required [31].

Thioesterase
The C-terminal end of MmpB showed significant similarity with thioesterase (TE) domains and contains the motif GxSxG found in other TE domains but does not contain a second conserved motif GxH [22, 32].

Potential Auxiliary and Tailoring Functions Specific to Mupirocin Biosynthesis
Other putative gene functions were identified within the *mup* cluster (Table 1). MupQ and MupU show similarity to proteins that activate their substrates by covalently binding AMP to a carboxylic side group, e.g., long chain fatty acid-CoA synthases. Mup A shows similarity to a subunit of LuxA of *Vibrio harveyi*, and this could represent part of an $FMNH_2$-dependent oxygenase. MupB and MupG show similarity to FabH (KASIII) and to FabB (KASI), respectively, which are 3-oxoacyl-ACP synthases. Two 3-oxoacyl-ACP reductases, MupD and MupS, were also identified. MupH shows similarity to many proteins possessing HMG-CoA synthase activity and is suggested to play a role in introducing the methyl group at C15 (see below). MupJ and MupK both show good similarity to enoyl-CoA hydratases, and MupE to shows good similarity enoyl reductases. Several other potential oxygenases and reductases were also identified within the second half of the cluster: MupO is proposed to be a cytochrome P450, MupW a dioxygenase, MupT a ferredoxin dioxygenase, MupC an NADH/NADPH oxidoreductase, and MupF a ketoreductase. MupL is a putative hydrolase, and MupV a putative oxidoreductase. However, their possible functions are unclear. The C-terminal of MmpE shows similarity with hydroxylases, and MupX with amidases and hydrolases. MupN shows similarity with regulatory binding proteins such as Lpa14 from *B. subtilis*, which is required for the expression and regulation of genes for synthesis of cyclic lipopeptides surfactin and iturin A [33]. MupR and MupI have been described before [19] and are referred to above.

Resistance to Mupirocin Is Encoded within the Cluster
The target for mupirocin is isoleucyl t-RNA synthetase (IleRS). MupM showed significant similarity with several mupirocin-resistant IleRS (see Table 1). Another IleRS (IleS; P18330) has previously been cloned from *P. fluorescens* NCIMB 10586 and shown to confer mupirocin resistance [9]. MupM and IleS both contain the consensus motifs for type I aminoacyl-tRNA synthetases [34] (HIGH and KMSKS; HYGH and KMSKR, respectively), but only show 29% identity to each other. Thus there may be two distinct mupirocin-resistant IleRS genes within *P. fluorescens* NCIMB 10586, one within the *mup*

biosynthesis cluster and one elsewhere in the chromosome.

The ability of MupM to confer mupirocin resistance was demonstrated by pAKE900, which contains *mupM* under the control of the *lacZ* operon, raising the MIC of *E. coli* DH5α to mupirocin from 30 to 250 $\mu g \cdot ml^{-1}$.

Gene Knockouts across the *mup* Region Generate the Mup⁻ Phenotype
Gene knockouts were constructed to confirm regions needed for mupirocin production (Table 2). Derivatives of suicide vectors pAKE603/604 with pairs of arms at least 500 bp in size flanking in-frame deletions within selected genes were introduced into *P. fluorescens* by biparental mating and homologous recombination. Deletions were constructed either by restriction digestion of a chromosomal fragment or ligation of two PCR fragments flanking the region to be deleted. Cointegrants with an inserted plasmid were passaged to medium containing 5% sucrose to select for the products of plasmid excision by homologous recombination. These colonies were screened by PCR to identify those with a mutation. Knockout phenotype was determined by bioassay for mupirocin production (Table 2). Selected knockouts were tested for complementation by the wild-type gene expressed in *trans* to confirm that the phenotypes observed were due to the specific deletion in that gene and not polar effects.

KRI was deleted from MmpA (10586ΔKR1); TE was deleted from MmpB (10586ΔTE); and AT2 was deleted from MmpC (10586ΔAT2). MupT was also deleted (10586ΔMupT). Each of these deletions resulted in loss of detectable antibiotic activity by comparison with the negative control SBW25 by bioassay and HPLC analysis. Wild-type *P. fluorescens* NCIMB 10586 has a peak with a retention time of 20.2 min (Figure 4A), which was purified and shown to have antibiotic activity. This is comparable to the retention time of mupirocin standard (20.1 min) (Figure 4C). 10586ΔKR1 did not have a peak in this region (Figure 4B) and is typical of all the mutants reported herein (data not shown). None of the mutants appeared to excrete detectable biosynthetic intermediates. The most obvious possibility was that MA might accumulate. The MA standard runs at 9.7 min, and no peak appeared here in any of the mutants. However, it should be noted that the profiles initially analyzed were of culture supernatant rather than cell lysates, so if MA accumulated in the cytoplasm, we would not have detected it. The main purpose of the HPLC analysis was to confirm the disappearance of the mupirocin peak. Knockouts of MupR and MupI were shown to be essential for mupirocin biosynthesis [19].

KS1 may be inactive due to differences from other KS in the consensus motifs (amCagg cf taCsssl; and Gln cf His at 135 aa). However, both a deletion of this region (10856ΔKS1) and a point mutation of the proposed KS1 active site (10586KS1,C196A) abolished detectable mupirocin production.

Complementation of ΔAT2 was carried out by placing a cloned fragment encoding MmpC in *trans* in *P. fluorescens* NCIMB10586ΔAT2. To do this, a 3.7 kb EcoRI-ClaI fragment containing *mmpC* was cloned into an IncQ
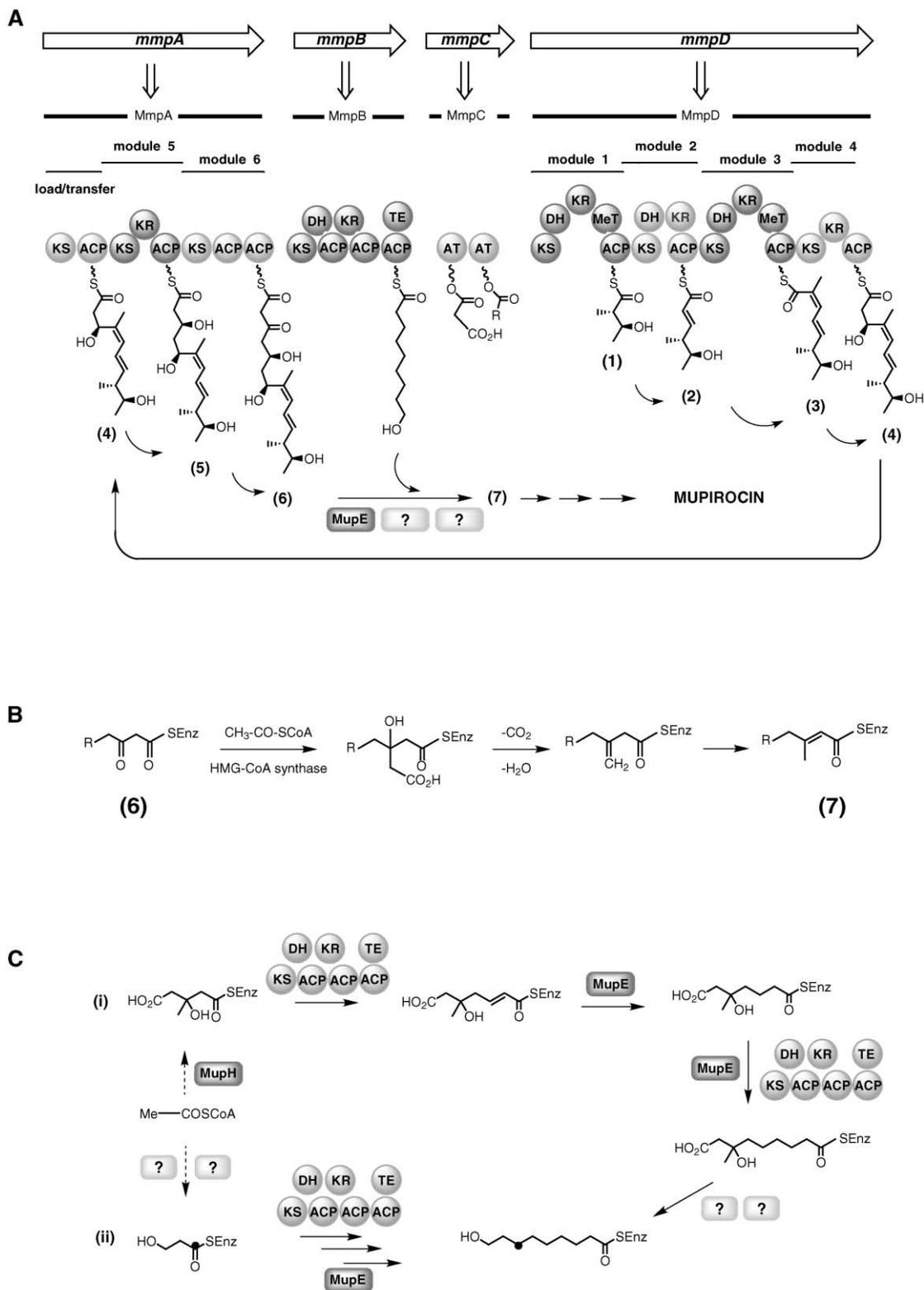
A

B

C

Figure 3. Models for Biosynthesis of Mupirocin Components

(A) Mupirocin biosynthesis model.

(B) HMG-based model for incorporation of the C15 methyl group in monic acid.

(C) 9-hydroxynonanoic biosynthesis model. (Ci) HMG starter unit formed from condensation of acetate and acetoacetate catalyzed by MupH and extended by two malonate condensations followed by removal of 3-methyl and hydroxyl groups. (Cii) Alternatively, a 3-hydroxyproprionate starter unit may be extended by three malonate condensations. ACP = acyl carrier protein; AT = acyl transferase; DH = dehydratase; HMG = 3-hydroxy-3-methylglutaric acid; KS = ketosynthase; MupE = enoyl reductase; MupH = HMG-CoA synthase; TE = thioesterase.

Table 2. Summary of Strains/Plasmids for Gene Knockout Experiments

| Strain | Mutation | | | | | Complementation | | | |
	Genotype[a]	Plasmid	Vector	Description[b]	Phenotype	Plasmid	Vector	Description	Phenotype
10586ΔKS1	MmpAΔ28-488	pJHΔKS1	pAKE604	1 kb PCR fragment, Δ1383 bp KS1	Mup⁻	-	-	-	-
10586KS1,C196A	MmpAC196A	pJHKS1C196A	pAKE604	1007 bp PCR fragment, KS1 C196A	Mup⁻	-	-	-	-
10586ΔKR1	MmpAΔ1495-1662	pJHΔKR1	pAKE604	1058 bp PCR fragment, Δ504 bp within KR1	Mup⁻	-	-	-	-
10586ΔTE	MmpBΔ1781-2033	pJHΔTE	pAKE604	1 kb PCR fragment, Δ759 bp TE	Mup⁻	-	-	-	-
10586ΔAT2	MmpCΔ337-548	pAKEΔAT2	pAKE604	1.4 kb BamHI-HindIII, Δ641 bp SmaI-SalI, 14 bp SmaI-SalI linker maintains frame	Mup⁻	pJH3	pJH10	3.7 kb EcoRI-ClaI containing mmpC	Mup⁺
10586ΔMupT	MupTΔ32-88	pSCΔT	pAKE604	1165bp PCR fragment, Δ171 bp within mupT	Mup⁻	pSCT	pJH10	351 bp PCR fragment of mupT	Mup⁺

[a] Amino acid coordinates relative to ORF.
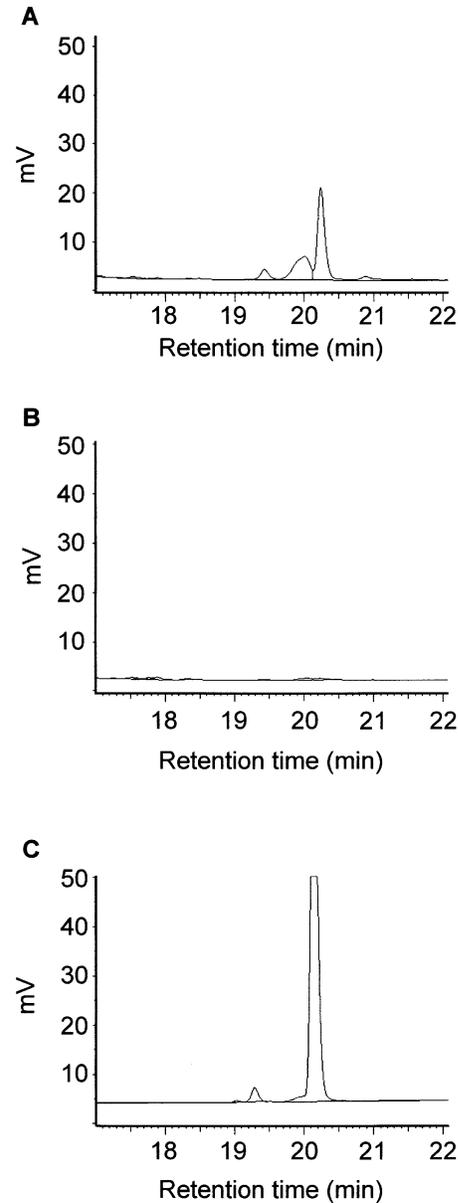[b] Clone fragment, deletion, linker size.

Figure 4. HPLC Analysis of Mupirocin Production
(A) *P. fluorescens* (NCIMB 10586).
(B) NCIMBΔKR1.
(C) Mupirocin standard (100 ng).

plasmid under the control of the *tac* promoter (pJH3). The phenotype of ΔAT2 was restored to wild-type by pJH3 without induction by IPTG but not after induction with ≥0.1 mM IPTG, so excess expression of this gene may block biosynthesis. Similarly, 10586ΔMupT was complemented by a 351 bp PCR fragment of *mupT* (pSCT) after induction by IPTG (0.5 mM). These results indicate that the phenotypes of the knockout result from inactivation of the specific genes targeted.

A Possible Pathway for Biosynthesis of Mupirocin
Monic Acid

It has been demonstrated that mupirocin is derived mainly from acetate and proposed that separate PKS

and FASs are involved in its assembly from MA and 9-HA via an esterification [4, 15]. Based on other PKS/FAS systems, a plausible biosynthetic pathway for MA is proposed in Figures 3A and 3B.

Unlike the erythromycin biosynthesis pathway, which is colinear with the gene order in the cluster [35], the *mup* genes do not appear to follow a simple progression and lack an obvious loading domain, which should catalyze the covalent attachment of an activated starter unit to the first KS domain of a PKS. No AT domains were identified within the condensation modules themselves, so the AT domains of *mmpC* may perform the task of loading and transferring the growing polyketide chain to each KS, or other unidentified transferase activities may be involved.

We predict that biosynthesis starts with production of the C_{12} pentaketide moiety (**4**, from Figure 3A) by MmpD, which then acts as a primer for biosynthesis of the remainder of the MA polyketide precursor (**6**) directed by MmpA, and possibly MupH (see below). The location of MmpC (AT1 and AT2) immediately before MmpD may be significant in this regard. Indeed, if the segment from *mmpA* to *mmpD* were looped out as a circle, the proposed pathway would be colinear. MmpB, with additional enzyme activities provided from other parts of the *mup* gene cluster, may synthesize 9-HN.

MmpD should encode sequential modules that catalyze the first four condensations. The following description refers to carbons by number in the conventional diagrams of mupirocin structure (Figure 1). Starter and extender units could be activated by MupU and MupQ, which show similarity to acyl-CoA synthases that activate substrates by linking their carboxyl group to the phosphoryl group of AMP. The activated substrate is then transferred to Coenzyme A (CoA). Mupirocin biosynthesis could start with one of the ATs transferring an activated acetyl starter unit from CoA to the 4′-phosphopantetheine arm of ACP8 and then to the thiol group of the active cysteine of KS5 (both in the proposed module 1 of MmpD) in a decarboxylative step. An activated malonyl extender could then be transferred to the vacant ACP8. Alternatively, an activated acetyl unit may be transferred directly to KS5. KS5 is proposed to catalyze the first decarboxylative condensation step of acetate and malonate, followed by addition of a methyl group from SAM at C12 by MeT1 (in the proposed module 1 of MmpD) and ketoreduction catalyzed by KR3, in the same module, producing a C_5 structure (**1**). DH2, in the same module, is probably nonfunctional due to its putative consensus active site (HxxxDxxxxA) differing from other DHs (HxxxGxxxxP). Such nonfunctional domains have been observed in several modular PKSs, e.g., RAPS 1 and 2 in rapamycin biosynthesis in *Streptomyces hygroscopicus* [24] and in *pltD* from the pyoluterin biosynthesis cluster in *P. fluorescens* Pf-5 [36]. The second cycle of condensation is predicted to be catalyzed by MmpD KS6 (module 2) and followed by ketoreduction and then dehydration catalyzed by KR4 and DH3, respectively (of the same module), producing a C_7 structure (**2**). The third condensation reaction corresponds to MmpD KS7 (module 3), followed by addition of a methyl group from SAM at C8 by MeT2, reduction, and dehydration of the keto group by KR5 and DH4, respectively (again within the same module), producing

a C_{10} structure (**3**). The fourth condensation step is catalyzed by MmpD KS8 (module 4), in the same module, followed by reduction by KR6, producing a C_{12} structure (**4**).

Further assembly of MA may be continued by MmpA. A fifth condensation of malonate catalyzed by MmpA KS2 (module 5) followed by ketoreduction by KR1, in the same module, should produce a C_{14} structure (**5**). The sixth condensation catalyzed by MmpA KS3 (module 6) and leaving the keto unreduced should produce a C_{16} β-ketothioester (**6**). However, this does not predict a function for KS1. Although KS1 may not have a completely normal function since it has Gln instead of His in its proposed active site, it is necessary for biosynthesis rather than being redundant, as indicated by the fact that mutation of KS1 gives a Mup⁻ phenotype. Thus, KS1 may be acting as a pseudoloading module either to facilitate the transfer of the pentaketide intermediate between MmpD and MmpA or even to help load the starter of the process at MmpD KS5.

The incorporation of a methyl group at C15 could be via condensation of an acetate unit at the β-keto group of structure (**6**) by the putative HMG-CoA synthase (MupH) (Figure 3B). Similar reactions have been observed during the biosynthesis of, for example, virginiamycin [37], myxopyronins [38], and aurantins [39]. The resulting carboxylic acid group would undergo decarboxylation followed by dehydration, possibly involving activation of the hydroxyl group, for example, by phosphorylation, producing a terminal alkene CH_2 group. The double bond would rearrange in conjugation with the thioester carbonyl, presumably by an isomerase-dependent reaction producing the C_{17} αβ-unsaturated thioester (**7**) in which the 15-methyl group is derived from the methyl carbon of the cleaved acetate unit.

Further modifications are required to generate MA from structure (**7**), which would include oxidation and reduction to create the epoxide group and tetrahydropyran ring. These reactions could be catalyzed by MupO (P450), MupC (NADH oxidoreductase), MupT (ferredoxin dioxygenase), and/or MupW (dioxygenase).

9-Hydroxynonanoic Acid

It is much more difficult to predict the source of 9-HN, which is a C_9 fully saturated fatty acid chain. This may be synthesized by a single PKS/FAS with full reduction of the keto groups by ketoreduction, dehydration, and enoyl reduction. One possibility is that MupH (HMG-CoA synthase) may catalyze the condensation of acetate and acetoacetate to generate HMG, which would be chain-extended by two condensations of malonate moieties followed by removal of the 3-methyl and hydroxyl groups and reduction of the distal carboxyl group (Figure 3Ci). However, studies with doubly ^{13}C-labeled HMG showed no intact incorporation [15]. Formally, 9-HN could be formed from a 3-hydroxypropionate starter unit by three malonate condensations catalyzed by a dedicated FAS (Figure 3Cii). Preliminary evidence has been obtained for this by synthesis and feeding of [1-^{13}C]-3-hydroxypropionate, which resulted in specific enrichment (2.5-fold) of C7′ [40].

If 3-hydroxypropionate is involved, there are several ORFs of unknown function that could be involved. One candidate PKS/FAS for iterative incorporation of the

Table 3. Plasmids Used or Constructed During This Study

Plasmid	Size (kb)	Replicon	Genotype/Phenotype	Reference/ Source
pAKE100	4.9	pMB1	Ap^r, Km^r, oriT	[19]
pAKE603	5.9	pMB1	Ap^r, oriT, lacZα, sacB	This study
pAKE604	7.2	pMB1	Ap^r, Km^r, oriT, lacZα, sacB	[19]
pBROC130	18.7	ColE1	Ap^r, Km^r, Tn5	[16]
pBROC131	15.6	ColE1	Ap^r	[16]
pCAW5.3	4.9	pMB1	Ap^r, Km^r, oriT	[16]
pDM1.2	14.5	IncQ	Sm^r, Tc^r	[53]
pGBT400	11.6	IncQ	Sm^r, oriT, tacp	[54]
pJH10	14.5	IncQ	pOLE1 IncC1 deleted, with EcoRI-SacI polycloning site, Tc^r from pDM1.2	This study
pOLE1	12.7	IncQ	pGBT400 with IncC1, Sm^r, oriT, tacp	[54]
pUC18	2.7	pMB1	Ap^r, lacZα, polylinker	[55]

three malonates could be MmpB, which contains KS4, KR2, DH1, ACP5-7, and TE. However, MmpB does not contain an enoyl reductase (ER) domain. MupE shows sequence similarity to ER and therefore could catalyze the final reduction step of 9-HN. Precedence for this possibility is found in lovastatin biosynthesis, where the proper functioning of the lovD nonaketide synthase (LNKS) requires the accessory lovC protein [41]. The unusual ACP triplet within MmpB may be involved in receiving 3-hydroxypropionate and three malonate molecules or other intermediates. It is also possible that the organism makes use of the ER used in endogenous fatty acid metabolism, as has been suggested for prodigiosin biosynthesis in *Streptomyces coelicolor* [42].

The TE domain at the end of MmpB would fit with this model for 9-HN biosynthesis being involved with release of either 9-HN or mupirocin or with esterification of 9-HN and MA. Alternatively, the C$_9$ moiety may be derived from a pathway outside of the cluster. It could also involve other enzymes from within the cluster, e.g., *mupB* and *mupG* (3-oxoacyl-ACP synthases), and *mupD* and *mupS* (3-oxoacyl-ACP reductases) for which roles have not been assigned.

Biosynthesis of Other Pseudomonic Acids

Pseudomonic acid B contains a hydroxyl group at C8, which may be a result of further oxidative modification. Pseudomonic acid C has an alkene at C10-C11 instead of an epoxide, which may be due to lack of processing by MupO (putative P450 oxygenase). Pseudomonic acid D has an unsaturated C$_9$ chain (alkene at C4'-C5'), which may arise from the second malonate condensed with 3-hydroxy-propionate not being fully reduced by ER.

Significance

The gene cluster we have sequenced may be a phenotypic island that could carry mupirocin biosynthesis and resistance from one strain to another. There seems to be an excess of coding capacity in the *mup* cluster for synthesis of the whole of mupirocin and its relatives. There also seem to be some notable features: its initial lack of an obvious linear relationship between gene order and biosynthetic pathway; the provision of only two AT domains, which are separate

from the type I modules; the multiple tandem ACPs; and the single TE despite two predicted products. Our gene knockouts so far have shown all predicted genes to be necessary for mupirocin production, and this is not due to polar effects, since complementation by the intact ORF has been possible where attempted. Current characterization of products of mutant strains or subcloned segments of the pathway should allow a tighter link between the biosynthetic pathway and gene sequence to be established. Nevertheless, a plausible order of biosynthetic steps and genes can be constructed at least for MA and provides a working model to underpin future experimental design. The apparent wealth of tailoring enzymes suggests great potential for *mup* genes as part of the repertoire currently available for combinatorial genetics in polyketide drug discovery.

Experimental Procedures

Bacterial Strains and Plasmids
P. fluorescens NCIMB 10586 [16] was used as the wild-type mupirocin producer, and SBW25 [17] was used as a nonproducer of mupirocin. *E. coli* strains DH5α [43] and NEM259 (SupE SupF *hsdR met trpR*, N.E. Murray, University of Edinburgh) were used for plasmid transformation and propagation; S17-1 [44] was used to mobilize suicide and expression plasmids into *P. fluorescens*; and HB101 [45] was used for transformation to recapture suicide plasmids containing large chromosomal fragments. *B. subtilis* 1064 [46] was used in bioassays to monitor mupirocin production. Other bacterial strains constructed in this study are shown in Table 2, and plasmids used or constructed during this work are listed in Tables 2 and 3.

Growth and Culture Conditions
E. coli strains were grown at 37°C in L broth [47] and L agar (L broth supplemented with 1.5%, w/v agar) supplemented with appropriate antibiotics. *P. fluorescens* strains were grown at 30°C in L broth and L agar. For plasmid maintenance and selection of antibiotic-resistant transformants, media were supplemented with appropriate antibiotic concentrations as follows: 100 μg ampicillin ml^{-1}, 50 μg kanamycin ml^{-1}, 150 and 300 μg penicillin G ml^{-1} in liquid and solid medium, respectively, 600 μg mupirocin ml^{-1}, and 15–25 μg tetracycline.HCL ml^{-1}.

DNA Isolation and Manipulation
Plasmid DNA extraction was performed by the alkaline SDS method of Birnboim and Doly [48] with slight modifications [49], or with Wizard Plus SV Mini Preps DNA Purification Systems (Promega). Genomic DNA was isolated from *P. fluorescens* using a yeast DNA

mini-prep kit (Nucleon, Tepnel Life Sciences). DNA was digested with appropriate restriction enzymes (MBI Fermentas). DNA was analyzed by standard agarose gel electrophoresis [50]. Extraction of DNA from agarose was performed using GeneClean kit (Bio101). Sticky ended DNA was blunt-ended using DNA polymerase I, large (Klenow) fragment (New England Biolabs). Ligations were performed using T4 DNA ligase [50]. PCR fragments were cloned into a T-tailed vector pGEM-TEasy (Promega). Competent *E. coli* cells were transformed with plasmid DNA using the method of Cohen et al. [51].

Suicide Mutagenesis of *P. fluorescens*

Biparental mating was carried out to mobilize the suicide plasmid derivatives and expression vectors from *E. coli* S17-1 to *P. fluorescens* as follows. A mixture of late exponential phase cultures of *E. coli* S17-1, containing the relevant plasmid (0.5 ml), and *P. fluorescens* (0.5 ml) were filtered on to a 0.45 μm sterile Millipore filter. The filter was placed on L agar overnight at room temperature. The mating mixture was resuspended in sterile saline solution (1 ml), and aliquots (100 μl) were spread on either L agar, supplemented with the appropriate antibiotic to select for the presence of the plasmid and mupirocin to kill the *E. coli* donor strain, or on M9 minimal media, which will not support *E. coli* S17-1 growth, and supplemented with the appropriate antibiotic to select for the presence of the plasmid. To isolate strains in which the suicide plasmid had excised from the chromosomal DNA, the cointegrant clone was incubated in L broth at 30°C without antibiotic selection. Serial dilutions were plated on L agar supplemented with sucrose (5%), and colonies were replica plated onto L agar Km. Colonies that were Sucr and Kms were selected.

Sequence Analysis

DNA sequencing was carried out using the Big Dye Terminator kit (PE-ABI), which is based on the chain termination method [52]. The sequencing reactions were separated on either an ABI 377 (Alta Bioscience, Birmingham University) or an ABI 3700 (Genomics Facility, Birmingham University) DNA Analyser. The DNA sequences were analyzed using GCG programs of the Wisconsin package [18]. The sequence of the mupirocin biosynthesis cluster was deposited with GenBank and has been assigned the accession number AF318063.

Bioassay for Mupirocin Production

10 μl samples of 16 hr *P. fluorescens* cultures, grown to the same OD$_{600}$, were spotted onto L agar and incubated at 30°C for 16 hr. The bioassay plates were then overlaid with molten L agar seeded with a 16 hr *B. subtilis* culture (150 μl·ml^{-1}) and triphenyl tetrazolium chloride (0.025%), incubated at 37°C for 16 hr, and scored for the presence of clear zones around the *P. fluorescens* patch.

HPLC Analysis

Seed cultures of *P. fluorescens* were grown in 25 ml L broth, at 25°C, 200 rpm for 24 hr. 25 ml MPM (2.3 g·liter^{-1} yeast extract, 1.1 g·liter^{-1} glucose, 2.6 g·liter^{-1}Na$_2$HPO$_4$, 2.4 g·liter^{-1} KH$_2$PO$_4$, 5.0 g·liter^{-1} (NH$_4$)$_2$SO$_4$) was inoculated with 5% of the seed culture and grown at 22°C, 200 rpm for 40 hr. Cells were removed by centrifugation at 5000 × g for 10 min and the supernatant stored at −20°C. 200 μl samples were filtered (0.2 μm) prior to injection.

Presence of mupirocin in the supernatant was determined by HPLC with a reverse-phased C$_{18}$ Supelco Discovery column (15 cm × 4.6 mm, 5 μM), UV detection at 233 nm, and a mobile phase gradient (5%–70%) of acetonitrile and TFA (0.01%), with H$_2$O and TFA (0.01%), at a flow rate of 1 ml·min^{-1}.

Acknowledgments

References

1. Staunton, J., and Wilkinson, B. (2001). Combinatorial biosynthesis of polyketides and nonribosomal peptides. Curr. Opin. Chem. Biol. *5*, 159–164.
2. Fuller, A.T., Mellows, G., Woodford, M., Banks, G.T., Barrow, K.D., and Chain, E.B. (1971). Pseudomonic acid: an antibiotic produced by *Pseudomonas fluorescens*. Nature *234*, 416–417.
3. Chain, E.B., and Mellows, G. (1974). Structure of pseudomonic acid, an antibiotic from *Pseudomonas fluorescens*. J. Chem. Soc. Chem. Commun., 847–848.
4. Chain, E.B., and Mellows, G. (1977). Pseudomonic acid. Part 1. The structure of pseudomonic acid A, a novel antibiotic produced by *Pseudomonas fluorescens*. J. Chem. Soc. Perkin Trans. 1, 294–309.
5. Clayton, J.P., O'Hanlon, P.J., and Rogers, N.H. (1980). The structure and configuration of pseudomonic acid C. Tetrahedron Lett. *21*, 881–884.
6. O'Hanlon, P.J., Rogers, N.H., and Tyler, J.W. (1983). The chemistry of pseudomonic acid. Part 6. Structure and preparation of pseudomonic acid D. J. Chem. Soc. Perkin Trans. 1, 2655–2665.
7. Shiozawa, H., Shimada, A., and Takahashi, S. (1997). Thiomarinols D, E, F, and G, new hybrid antimicrobial antibiotics produced by a marine bacterium. Isolation, structure, and antimicrobial activity. J. Antibiot. (Tokyo) *50*, 449–452.
8. Hughes, J., and Mellows, G. (1978). Inhibition of isoleucyl-transfer ribonucleic acid synthase in *Escherichia coli* by pseudomonic acid. Biochem. J. *176*, 305–318.
9. Yanagisawa, T., Lee, J.T., Wu, H.C., and Kawakami, M. (1994). Relationship of protein-structure of isoleucyl-transfer-RNA synthetase with pseudomonic acid resistance of *Escherichia coli*—proposed mode of action of pseudomonic acid as an inhibitor of isoleucyl-transfer-RNA synthetase. J. Biol. Chem. *269*, 24304–24309.
10. Hughes, J., and Mellows, G. (1980). Interaction of pseudomonic acid A with *Escherichia coli* B isoleucyl-tRNA synthase. Biochem. J. *191*, 209–219.
11. Clayton, J.P., Luk, K., and Rogers, N.H. (1979). The chemistry of pseudomonic acid. Part 2. The conversion of pseudomonic acid A into monic acid A and its esters. J. Chem. Soc. Perkin Trans. 1, 308–313.
12. Cookson, B.D. (1998). The emergence of mupirocin resistance: a challenge to infection control and antibiotic prescribing practice. J. Antimicrob. Chemother. *41*, 11–18.
13. Feline, T.C., Jones, R.B., Mellows, G., and Phillips, L. (1977). Pseudomonic acid. Part 2. Biosynthesis of Pseudomonic acid A. J. Chem. Soc. Perkin Trans. 1, 309–318.
14. Mantle, P.G., and MacGeorge, K.M. (1991). Pseudomonic acid biosynthesis. The putative role of 3-hydroxy-3-methylglutarate. J. Chem. Soc. Perkin Trans. 1, 255–258.
15. Martin, F.M., and Simpson, T.J. (1989). Biosynthetic studies on pseudomonic acid (mupirocin), a novel antibiotic metabolite of *Pseudomonas fluorescens*. J. Chem. Soc. Perkin Trans. 1, 207–209.
16. Whatling, C.A., Hodgson, J.E., Burnham, M.K.R., Clarke, N.J., Franklin, F.C.H., and Thomas, C.M. (1995). Identification of a 60 kb region of the chromosome of *Pseudomonas fluorescens* NCIB 10586 required for the biosynthesis of pseudomonic acid (mupirocin). Microbiology *141*, 973–982.
17. Rainey, P.B., and Bailey, M.J. (1996). Physical and genetic map of the *Pseudomonas fluorescens* SBW25 chromosome. Mol. Microbiol. *19*, 521–533.
18. Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. *12*, 387–395.
19. El-Sayed, A.K., Hothersall, J., and Thomas, C.M. (2001). Quorum-sensing-dependent regulation of biosynthesis of the polyketide antibiotic mupirocin in *Pseudomonas fluorescens* NCIMB 10586. Microbiology *147*, 2127–2139.

20. Hayashi, T., Matsumoto, H., Ohnishi, M., and Terawaki, Y. (1993). Molecular analysis of a cytotoxin-converting phage φCTX, of *Pseudomonas aeruginosa*: structure of the *att-cos-ctx* region and integration into serine tRNA gene. Mol. Microbiol. *7*, 657–667.

21. Albertini, A.M., Caramori, T., Scoffone, F., Scotti, C., and Galizzi, A. (1995). Sequence around the 159 region of the *Bacillus subtilis* genome: the *pksX* locus spans 33.6kb. Microbiology *141*, 299–309.

22. Donadio, S., and Katz, L. (1992). Organisation of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccaropolyspora erythraea*. Gene *111*, 51–60.

23. Bouquin, N., Tempete, M., Holland, I.B., and Seror, S.J. (1995). Resistance to trifluoroperazine, a calmodulin inhibitor, maps to the *fabD* locus in *Escherichia coli*. Mol. Gen. Genet. *246*, 628–637.

24. Aparicio, J.F., Molanar, I., Schwecke, T., Koning, A., Haydock, S.F., Khaw, L.E., Staunton, J., and Leadlay, P.F. (1996). Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. Gene *169*, 9–16.

25. Marsden, A.F., Wilkinson, B., Cortes, S., Dunster, N.J., Staunton, J., and Leadlay, P.F. (1994). Stereospecific acyl transfers on the erythromycin producing polyketide synthase. Science *263*, 378–380.

26. Haydock, S.F., Aparicio, J.F., Molanar, I., Schwecke, T., Koning, A., Marsden, A.F.A., Galloway, I.S., Staunton, J., and Leadlay, P.F. (1995). Divergent structure motifs correlated with substrate specificity of (methyl)malonyl-CoA: acyl carrier protein transacylase domains in modular polyketide synthase. FEBS Lett. *374*, 246–248.

27. Kakavas, S.J., Katz, L., and Stassi, D. (1997). Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*. J. Bacteriol. *179*, 7515–7522.

28. Bevitt, D.J., Cortes, J., Haydock, S.F., and Leadlay, P.F. (1992). 6-Deoxyerythronolide B synthase 2 from *Saccaropolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. Eur. J. Biochem. *204*, 39–49.

29. Kennedy, J., Auclair, K., Kendrew, S.G., Park, C., Vederas, J.C., and Hutchinson, C.R. (1999). Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science *284*, 1368–1372.

30. Gehring, A., Mori, I., Perry, R.D., and Walsh, C.T. (1998). The nonribosomal peptide synthase HMWPZ forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of *Yersinia pestis*. Biochemistry *37*, 11637–11650.

31. Fujii, I., Watanabe, A., Sankawa, U., and Ebizuka, Y. (2001). Identification of Clasein cyclase domain in fungal polyketide synthase WA, a naphthopyroine synthase of *Aspergillus nidulans*. Chem. Biol. *8*, 189–197.

32. Witkowski, A., Naggert, J., Wessa, B., and Smith, S. (1991). A catalytic role for histidine-237 in rat mammary gland thioesterase II. J. Biol. Chem. *266*, 18514–18519.

33. Huang, C.C., Ano, T., and Shoda, M. (1993). Nucleotide sequence and characteristics of the gene, *lpa-14*, responsible for biosynthesis of the lipopeptide antibiotics iturin A and surfactin from *Bacillus subtilis* RB14. J. Ferment. Bioeng. *76*, 445–450.

34. Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990). Partition of transfer-RNA synthases into 2 classes based on mutually exclusive sets of sequence motifs. Nature *347*, 203–206.

35. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J., and Katz, L. (1991). Modular organisation of genes required for complex polyketide biosynthesis. Science *252*, 675–679.

36. Nowak-Thompson, B., Chaney, N., Wing, J.S., Gould, S.J., and Loper, J.E. (1999). Characterization of the Pyoluteorin biosynthetic gene cluster of *Pseudomonas fluorescens* Pf-5. J. Bacteriol. *181*, 2166–2174.

37. Kingston, D.G.I., Kolpak, M.X., LeFevere, J.W., and Borup-Grochtmann, I. (1983). Biosynthesis of antibiotics of the virginiamycin family. 3. Biosynthesis of virginiamycin M1. J. Am. Chem. Soc. *105*, 5106–5110.

38. Kohl, W., Irschik, H., Reichenbach, H., and Hofle, G. (1984). Antibiotics from gliding bacteria. 22. The biosynthesis of myxopyronin-A, an antibiotic from *Myxococcus fulvus*, strain MX-F50. Liebigs Ann. Chem. *6*, 1088–1093.

39. Nakagawa, A., Konda, Y., Hatano, A., Harigaya, Y., Onda, M., and Omura, S. (1988). Structure and biosynthesis of novel antibiotics, aurantinin-A and aurantinin-B produced by *Bacillus aurantinus*. J. Org. Chem. *53*, 2660–2661.

40. Sugden, M.J. (1992). The biosynthesis of pseudomonic acid, Ph.D. Thesis, University of Bristol, Bristol, United Kingdom.

41. Auclair, K., Kennedy, J., Hutchinson, C.R., and Vederas, J.C. (2001). Conversion of cyclic nonaketides to lovastatin and compactin by a *lovC* deficient mutant of *Aspergillus terreus*. Bioorg. Med. Chem. Lett. *11*, 1527–1531.

42. Cerdeno, A.M., Bibb, M.J., and Challis, G.C. (2001). Analysis of the prodiginin biosynthesis gene cluster of *Streptomyces coelicolor* A3(2): new mechanism for chain initiation and termination in modular multienzymes. Chem. Biol. *8*, 817–829.

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

44. Simon, R., Priefer, U., and Puhler, A. (1983). A broad host range mobilization system for *in vivo* genetic-engineering-transposon mutagenesis in gram-negative bacteria. Biotechnology *1*, 784–791.

45. Boyer, H.W., and Roulland-Dussoix, D. (1969). A complementation analysis of restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. *41*, 459–472.

46. Moir, A., Lafferty, E., and Smith, D.A. (1979). Genetic analysis of spore germination mutants of *Bacillus subtilis 168*, the correlation of phenotype with map location. J. Gen. Microbiol. *111*, 165–168.

47. Kahn, M., Kolter, R., Thomas, C.M., Figurski, D.H., Meyer, R., Remaut, E., and Helinski, D.R. (1979). Plasmid cloning vehicles derived from plasmids ColE1 and RK2. Methods Enzymol. *68*, 268–280.

48. Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. *7*, 1513–1523.

49. Smith, C.A., and Thomas, C.M. (1983). Deletion mapping of *kil* and *kor* functions in the *trfA* and *trbB* regions of broad-host-range plasmid RK2. Mol. Gen. Genet. *190*, 245–254.

50. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: a Laboratory Manual, 2nd edition, (Cold Spring Harbour, NY: Cold Spring Harbour Laboratory).

51. Cohen, S.N., Chang, A.C.Y., and Hsu, H. (1972). Non-chromosomal antibiotic resistance in bacteria; genetic transformation of *Escherichia coli* by R factor DNA. Proc. Natl. Acad. Sci. USA *69*, 2110–2114.

52. Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA-sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 4563–4567.

53. Macartney, D.P., Williams, D.R., Stafford, T., and Thomas, C.M. (1997). Divergence and conservation of the partitioning and global regulation functions in the central control region of the IncP plasmids RK2 and R751. Microbiology *143*, 2167–2177.

54. Jagura-Burdzy, G., Kostelidou, K., Pole, J., Khare, D., Jones, A., Williams, D.R., and Thomas, C.M. (1999). IncC of broad-host-range plasmid RK2 modulates KorB transcriptional repressor activity *in vivo* and operator binding *in vitro*. J. Bacteriol. *181*, 2807–2815.

55. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of M13mp18 and pUC19 vectors. Gene *33*, 103–119.

Accession Numbers

The sequence of the mupirocin biosynthesis cluster was deposited with GenBank and has been assigned the accession number AF318063.