CHEMBIOCHEM

DOI: 10.1002/cbic.200900565

Phosphopantetheinylation and Specificity of Acyl Carrier Proteins in the Mupirocin Biosynthetic Cluster

Jennifer A. Shields,[a, c] Ayesha S. Rahman,[a] Christopher J. Arthur,[b] John Crosby,[b] Joanne Hothersall,[a] Thomas J. Simpson,[b] and Christopher M. Thomas*[a]

Acyl carrier proteins are vital for the biosynthesis of fatty acids and polyketides. The mupirocin biosynthetic cluster of *Pseudomonas fluorescens* encodes eleven type I ACPs embedded in its multifunctional polyketide synthase (PKS) proteins plus five predicted type II ACPs (mAcpA-E) that are known to be essential for mupirocin biosynthesis by deletion and complementation analysis. MupN is a putative Sfp-type phosphopantetheinyl transferase. Overexpression of three type I and three type II mupirocin ACPs in *Escherichia coli*, with or without *mupN*, followed by mass spectroscopy revealed that MupN can modify both mupirocin type I and type II ACPs to their holo-form. The endogenous phosphopantetheinyl transferase of *E. coli* modified mAcpA but not mAcpC or D. Overexpression of the type II ACPs in *macp* deletion mutants of the mupirocin producer *P. fluorescens* 10586 showed that they cannot substitute for each other while hybrids between mAcpA and mAcpB indicated that, at least for mAcpB, the C-terminal domain determines functional specificity. Amino acid alignments identified mACPs A and D as having C-terminal extensions. Mutation of these regions generated defective ACPs, the activity of which could be restored by overexpression of the *macp* genes on separate plasmids.

Introduction

Acyl carrier proteins (ACPs) are generally small (70–100 amino acids), acidic proteins that tether acyl chains during metabolic processes including the biosynthesis of polyketides, fatty acids and N-acyl homoserine lactones.[1] By definition, type I ACPs are found as part of multifunctional megaproteins whereas type II ACPs exist as discrete proteins. Structural studies indicate that all ACPs adopt a common tertiary structure consisting of four α helices enclosing a hydrophobic core.[1–6] In type II ACPs this core has been proposed to accommodate the acyl chain during biosynthesis.[7] By contrast recent studies of the rat type I fatty acid synthase (FAS) ACP[8] and the actinorhodin ACP from *Streptomyces coelicolor*[9] provided evidence that the acyl/polyketide chain does not interact with the ACP.

While ACPs are not generally interchangeable between polyketide and fatty acid synthases[10,11] they can be exchanged between different type II polyketide synthases (PKSs).[12] Surprisingly, the overexpressed ACP domain from the type I fungal norsolorinic acid synthase can function, albeit with low efficiency, in the type II actinorhodin minimal PKS.[13] Within a type I PKS, however, ACPs are not necessarily interchangeable due to the specificity of a particular KS domain for its cognate ACP[14] though our understanding of the rules governing specificity with respect to protein partners and substrate are still rudimentary.[15]

ACPs are dependent on post-translational modification to the active holo-form, in which a 20 Å phosphopantetheine prosthetic group is attached to a conserved serine residue in a reaction catalysed by a phosphopantetheinyl transferase (PPTase).[1,16] Acyl intermediates are tethered to the terminal thiol of this phosphopantetheine moiety, and it is this chain that is used to deliver substrates into the active sites of the KS and other catalytic domains. PPTases are therefore essential enzymes for both primary and secondary metabolism.[1]

PPTases are divided into three types. The first is the AcpS-type, of which the 125-residue AcpS of *Escherichia coli* is the archetype. These PPTases are normally responsible for the modification of ACPs of primary metabolism,[16,17] although some AcpSs are capable of modifying heterologous type II PKS and FAS ACPs.[18–21] The *Bacillus subtilis* AcpS has been crystallised with and without its substrates and shown to be a trimer with three active sites formed at the interface between pairs of monomers.[22] The second group of PPTases is the Sfp-type, based on the surfactin synthase PPTase of *Bacillus subtilis*. These are typically twice as large as their AcpS-like relatives, functioning as a pseudo-dimer, and generally modify carrier proteins (CPs) involved in secondary metabolic pathways.[17,23] Sfp is the only enzyme available for modification of the 43 carrier proteins that belong to the PKS and non-ribosomal peptide synthase (NRPS) pathways in *B. subtilis*, and was also shown to be capable of complementing a deletion in the

[a] *J. A. Shields, A. S. Rahman, J. Hothersall, Prof. C. M. Thomas*
School of Biosciences, University of Birmingham
Edgbaston, Birmingham B15 2TT (UK)
Fax: (+44) 121-414-5925
E-mail: c.m.thomas@bham.ac.uk

[b] *C. J. Arthur, J. Crosby, T. J. Simpson*
School of Chemistry, University of Bristol
Cantock's Close, Bristol, BS8 1TS (UK)

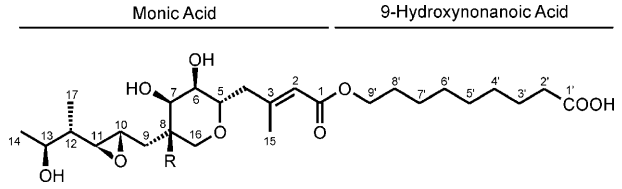
[c] *J. A. Shields*
Present address: Marks & Clerk LLP, Patent and Trade Mark Attorneys
Alpha Tower, Suffolk Street Queensway, Birmingham, B1 1TT (UK)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900565.

E. coli AcpS.[17] The third type of PPTases are integrated as domains in multifunctional type I proteins. These are usually found in fungal type I FASs, although analysis of enediyne PKS sequences, which have a type I domain structure, revealed the presence of a C-terminal domain similar to a phosphopantetheinyl transferase, proposed to activate the ACP of the PksEs.[24] A further example of an integrated PPTase was also identified in an unknown PKS in *Saccharopolyspora erythraea*.[25]

The *Pseudomonas aeruginosa* genome was found to contain a single Sfp-like PPTase, PcpS, which can phosphopantetheinylate CPs from both primary and secondary metabolism[26] and is essential for viability.[27] Analyses of genome sequences suggest that the orthologue of this protein is also the only PPTase found in other *Pseudomonas* species, including *P. fluorescens*,[28] although comparison of biochemical properties of the enzymes from *P. syringae* and *P. aeruginosa* and others suggest that the substrate specificity of the enzymes might vary significantly. Hence, some Sfp-type PPTases have broad substrate specificity and accept a variety of CPs, whilst others, such as EntD of *E. coli*, are dedicated to a single biosynthetic cluster.[16]

Mupirocin (also known as pseudomonic acid, PA) is a polyketide antibiotic produced by *Pseudomonas fluorescens* NCIMB 10586 (Scheme 1). It is active against Gram-positive bacteria,



Scheme 1. Structure of mupirocin, which is comprised of four pseudomonic acids (PA), PA-A accounts for 90% of the mixture. PA-A: R = H; PA-B: R = OH; PA-C: R = H, C10/C11 *E*-alkene; PA-D: R = H, C-4′/5′ *E*-alkene.

and is used topically to treat bacterial skin infections and intranasally to remove methicillin-resistant *Staphylococcus aureus* (MRSA). The mupirocin biosynthetic cluster encodes four large type I PKSs, thought to catalyse formation of the polyketide/fatty acid backbone, plus an unusually large number of modifying enzymes (Figure 1),[29] which are all essential for normal pseudomonic acid-A (PA-A) biosynthesis.[30] The *mup* genes include eleven type I ACPs integrated into the PKS and putative FAS modules, plus a further five putative type II ACPs (mAcpA to mAcpE) in the tailoring region as well as a putative Sfp-type PPTase, MupN. We recently reported that deletion of *mupN* abolished PA-A production and decreased antibiotic activity to 11% of wild-type level.[30] This paper demonstrates that MupN is capable of converting representatives of both the type I and type II mAcps to the holo-form in vivo and describes studies on the function and specificity of the tailoring mACPs.

Results and Discussion

Putative type II acyl carrier proteins in the mup cluster

Annotation of the *mup* cluster identified five of the tailoring ORFs as putative type II ACPs, mAcpA-E.[29] However, they generally share less than 30% identity with typical type II PKSs/FASs ACPs and with each other (Figure 2). mAcps A (101 aa) and D (106 aa) are significantly longer than the others which are of more "normal" length consisting of 83 (mAcpB), 78 (mAcpC) and 80 aa (mAcpE). mAcps C and E are less acidic, with *pI*s of 5.69 and 7.30, respectively, compared to the *pI*s of the other mAcps which range from 4.01–4.66.

The closest relatives of the mACPs identified by Blast searches were scattered across diverse bacteria; this suggests acquisition from multiple genetic sources, possibly as part of functional sets of genes, and that each is likely to have a different role in mupirocin biosynthesis. For mAcpA, the closest relatives (35–38% identity mainly in *Rhizobium* and related species) were ACPs linked to putative fatty acid biosynthesis gene clusters, while relatives of mAcpB were found with similar levels of identity (35–38%) in both bacteria and eukaryotes, most commonly as part of large multifunctional proteins involved in polyunsaturated fatty acid biosynthesis. For mAcpC there are many relatives (the closest being from *Sorangium cellulosum* with 59% identity) that often occur in PKS clusters, generally as part of cassettes with hydroxy-methyl-glutarate (HMG) CoA-synthase (HCS) genes that are involved in creation of species as diverse as methyl and vinyl chloride side chains as well as cyclopropane rings.[31–34] The closest homologue of mAcpD, DifC of *Bacillus amyloliquefaciens* FZB 42 (32% identity), is thought to be involved in synthesis of the polyketide difficidin,[35] and this cluster also contains homologues of MupS (a putative 3-oxo-acyl-ACP reductase) and MupQ (a putative acyl-CoA synthase). A blast of the mAcpE sequence produced no hits.

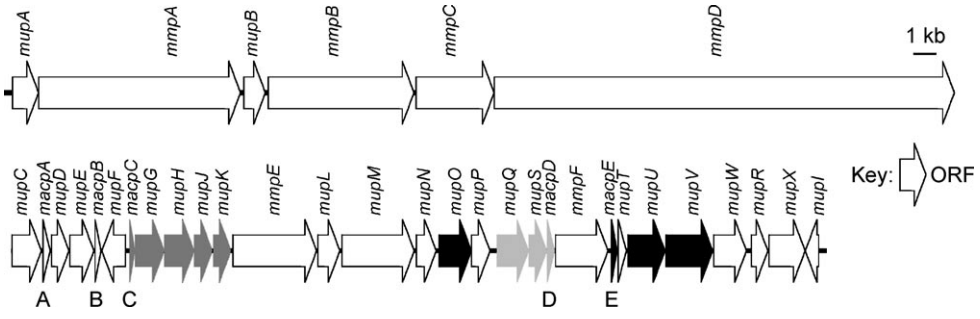


Figure 1. The mupirocin biosynthetic cluster. ORFs encoding proteins that are predicted to work together are indicated by shading in grey/black. The positions of the tailoring mAcps are shown by upper case letters A–E.

An Sfp-type PPTase in the mupirocin tailoring region

The MupN protein has strong similarity to Sfp-like PPTases although it is longer (283 aa) than typical members of this family which are usually about 240 aa. Its closest relative is Mta of *Stigmatella aurantiaca* (39% identity) which is also longer than average, with 277 aa, while

Figure 2. Multiple alignment of tailoring mAcps with other type II PKS and FAS ACPs. Identical residues are highlighted in black whilst similar residues are highlighted in grey. * indicates the position of the conserved active site serine residue. Ec : *E. coli* FAS ACP; Bs: *B. subtilis* FAS ACP; Act: ACP of actinorhodin PKS (*S. coelicolor*); Fren: ACP of frenolicin PKS (*S. roseofulvus*); Oxytet: ACP of oxytetracycline PKS (*S. rimosus*).

MupN is a more distant relative of Sfp itself (31% identity). It lacks most of the P1 motif found in Sfp-type but not AcpS-type PPTases—the highest conservation is in the region from aa 100 to 190; this region also has strong similarity to AcpS-like PPTases and contains the consensus motifs P2 ($X^{V}/_{I}G^{V}/_{I}D$) and P3 ($^{P}/_{W}$ $^{S}/_{C}/_{T}XKE^{S}/_{A}hhK$).[16,36] On this basis MupN can be classified as belonging to the W/KEA subfamily of Sfp-like PPTases.[37] As *mupN* is located within the mupirocin tailoring region, and is the only PPTase identified in the cluster, it is possible that it modifies the type I ACPs of the PKS modules as well as the type II tailoring ACPs. Its *pI* (7.28) lies almost exactly between the *pI*s of Sfp (5.6) and AcpS (9.6).

Because *mupN* is essential for mupirocin production,[30] it was of interest that the primary putative PPTases encoded by *Pseudomonas fluorescens* SBW25 (PFLU 4722), Pf0-1 (PfL01 4243) and Pf-5 (PFL 4473),[28] each of which show approximately 75 to 76% sequence identify to each other by pairwise comparisons, belong to the F/KES sub-family of Sfp-like PPTases.[37] Thus MupN shows only 30% sequence identity to the PPTase of SBW25 (which we expected to be the closest relative based on the similarity of DNA sequences that flank the *mup* cluster in *P. fluorescens* NCIMB10586) and no significant similarity to the other two PPTases.

In vivo conversion of ACPs to Holo-form by MupN

To determine whether MupN activates representatives of both type I and type II mupirocin ACPs, the *mup* ACPs were produced in *E. coli* in the absence or presence of MupN. The ACPs were purified as N-terminal His_6 fusion proteins and their masses determined by electrospray mass spectrometry (ES-MS) to check for post-translational addition of the phosphopantetheine arm (340 Da). Type I ACPs 3, 5 and 6 and type II mAcps A, C and D were successfully purified as soluble proteins whereas the remaining ACPs were either poorly expressed, gave a complex product profile or were insoluble when overproduced in *E. coli* and therefore could not be easily characterized. The data obtained for the soluble ACPs is presented in Figures S1 and S2 in the Supporting Information and summarised in Table 1.

In addition to the expected species type II mAcps A and D and type I ACPs 3, 5 and 6 showed masses apparently due to addition to the histidine tag (after loss of the N-terminal methionine) of a previously reported 6-phosphogluconyl group (258 Da) followed by dephosphorylation to leave an extra 178 Da.[38] The spectrum of mAcpA showed peaks corresponding to apo- and holo-mAcpA lacking the N-terminal methionine; this indicates that the endogenous *E. coli* AcpS is capable of post-translationally phosphopantetheinylating this ACP (approximately 50% conversion was observed). When MupN was coexpressed with mAcpA, greater than 80% was present as the holo-ACP. In contrast, mAcps C and D showed predominantly a single mass for each of the proteins correlating well with the expected masses for the apo-forms of these ACPs, again lacking the N-terminal methionine (Table 1 and Figure S1). Coproduction of MupN with these ACPs resulted in about 50% conversion to holo-form. Therefore, MupN clearly does activate representatives of the type II ACPs.

For the type I ACPs, the spectrum of ACP3 (from MmpA) in the absence of MupN contained peaks corresponding to the expected mass of the apo-protein with loss of methionine. In the presence of MupN the corresponding phosphopantetheinylated peaks are observed as well as about 30% remaining apo-ACP3. For both ACP5 and ACP6 (both from MmpB) a peak corresponding to the apo-form minus methionine was observed and this was completely converted to the holo-form in the presence of MupN. MupN appears, therefore, to be very efficient in activating representatives of the type I ACPs.

Tailoring mAcps to have specific roles

No other polyketide biosynthetic cluster has yet been reported to have as many tailoring ACPs as the *mup* cluster. To determine whether any of the tailoring mAcps can substitute for

Table 1. Summary of expected and observed masses in mass spectroscopic analysis of purified ACPs in the absence and presence of MupN.

	Type I mAcps								Type II mAcps							
	ACP3			ACP5			ACP6		mAcpA			mAcpC			mAcpD	
	Expected	Observed		Expected	Observed		Expected	Observed	Expected	Observed		Expected	Observed		Expected	Observed	
		−mupN	+mupN		−mupN	+mupN		−mupN	+mupN		−mupN	+mupN		−mupN	+mupN		−mupN	+mupN
Apo	14004			13172			13413			13803			11153			13847		
Apo-fM[a]	13873	13870	13870	13041	13037		13282	13279		13654	13655	13656	11004	11005	11004	13698	13699	13700
Holo	14344			13512			13753			14143			11493			14187		
Holo-fM[a]	14213		14210	13381		13377	13622		13619	13994	13995	13995	11344		11345	14038		14040

[a] Specific removal of fMet from the N terminus of polypeptides is common in *E. coli*, resulting in a loss of 149 Daltons.

each other we overexpressed each mAcp *in trans* in each *macp* deletion strain and tested for restoration of antibiotic production. We previously showed that *in trans* expression of the cognate *macp* in each deletion strain restores mupirocin biosynthesis to wild-type levels.[30,39] The *macp*s were cloned into the IncQ vector pJH10 under the control of the *tac* promoter, and mobilized into *P. fluorescens* by biparental mating as described in Experimental Section. The ability of noncognate *macp*s to overcome in-frame deletions and restore PA production was assessed by bioassay and HPLC.

For the strains Δ*macpA-D*, overexpression of any noncognate *macp in trans* failed to restore antibiotic biosynthesis, and the zones of inhibition produced were no larger than that of the negative control carrying pJH10. In the case of the Δ*macpE* strain, which generates a significant but diffuse clearing zone due to the switch of production from PA-A to PA-B,[39] HPLC confirmed that none of the noncognate mAcps resulted in restoration of PA-A production from PA-B. Thus the mAcps are not interchangeable.

mAcpE is predicted to interact with MupO, U and V (putative cytochrome P450, acyl-CoA synthase and oxidoreductase respectively) based on the observation that deletion of each gene individually results in the same phenotype.[39] mAcpC is predicted to interact with MupG and MupH as part of the HMG-synthase cassette which introduces the C-15 methyl group.[30,40] In vitro evidence for this has been provided from studies of the overexpressed HCS cassette proteins in both the bacillaene and myxovirescin systems.[32,41] mAcpD is predicted to work with MupQ and MupSΔ because a similar cluster of genes, *difC, difD* and *difE* is encoded by *B. amyloliquefaciens* FZB 42.[35] Although the functions and interacting partners for mAcps A and B are still unknown it thus seems likely that each type II mAcp is dedicated to a specific step in biosynthesis.

A hybrid mAcp retains the function and specificity of its C-terminal domain

To investigate the determinants of mAcp specificity, two hybrid proteins were constructed, the first consisting of the N terminus of mAcpA and the C terminus of mAcpB, while the second was the converse of this. The two halves of the hybrids were amplified and joined at a unique ClaI site close to the active-site serine residue, within the GIDS motif that is conserved between the two proteins. According to previously determined ACP structures, the conserved serine is located at the N terminus of helix 2,[2,42–46] while the three previous residues are predicted to lie in the flexible $\alpha1\alpha2$ loop region. By choosing this conserved region as the junction of the hybrid, the effect of structural alterations of the two termini should be minimised. The *macpA--macpB* and *macpB--macpA* hybrids expressed from pASRAB11 and pJS221, respectively, were tested in strains 10586Δ*macpA* and 10586Δ*macpB* to determine whether they could complement the deletions. Bioassay revealed that the *macpB-macpA* hybrid did not restore antibiotic production in either deletion strain even in the presence of IPTG. By contrast the *macpA-macpB* hybrid also failed to restore biosynthesis in Δ*macpA*, but did complement the Δ*macpB* deletion, restoring mupirocin production to WT levels in the presence of 0.1 mM IPTG (Figure 3).

This result suggests that the region determining mAcp specificity lies in the C-terminal half. The reduced activity of the hybrid compared to the normal mAcp, as indicated by the greater induction required to observe complementation, could be due to reduced expression (transcription and translation) of the hybrid gene, more rapid degradation of a less compact protein due to folding problems, reduced binding strength with normal partner proteins or reduced enzymatic activity. Complete loss of function in the mAcpB-A hybrid could result from extreme versions of any one, or a combination, of these. Both functional and nonfunctional hybrids have previously been constructed. For example, a fusion protein created from the N terminus of the tetracenomycin ACP from *Streptomyces glaucescens* and the C terminus of the actinorhodin synthase ACP from *Streptomyces coelicolor* gave rise to antibiotic production when expressed in *S. coelicolor*.[12] These proteins have higher sequence identity (40%) than between mAcpA and mAcpB studied here (26.6%) but because the reciprocal hybrid was not constructed the specificity in that system is unclear. Unfavourable protein–protein interactions within a hybrid ACP

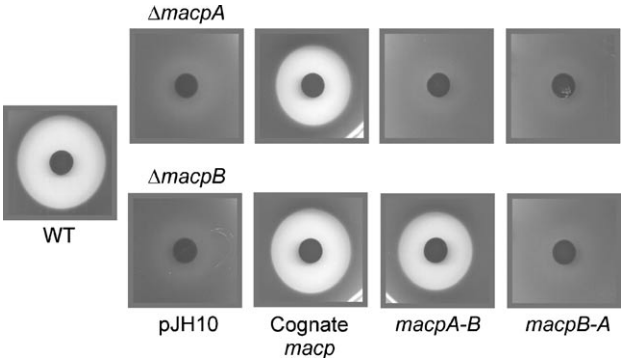


Figure 3. Biological activity of hybrid mAcps. Ability of pJH10 derivatives encoding mAcpAB or mAcpBA hybrids to complement Δ*macpA* or Δ*macpB* was assessed by bioassay as described in Experimental Procedures. WT or in-frame deletion mutants with empty pJH10 vector were used as positive and negative controls respectively.

Table 2. Effect of C-terminal mutations in mAcpA assessed by their ability to complement *in-frame* deletions in our standard bioassay.

	Area of Zone of Inhibition (% of WT)[a] IPTG [mM]		
	0	0.1	0.5
WT	100	100	100
pJH10	0.6 ± 1.0	1.3 ± 2.0	2.6 ± 3.3
QIAARAAASA	79.8 ± 12.3	70.5 ± 10.1	56.6 ± 18.4
QIAARAA	39.9 ± 2.5	17.8 ± 0.7	33.1 ± 13.8
QIAA	41.0 ± 1.0	43.6 ± 0.8	49.0 ± 7.4
Q	1.8 ± 1.0	2.9 ± 2.1	5.3 ± 4.6
QIAARAAA**A**A	45.3 ± 0.3	49.3 ± 8.9	42.7 ± 7.7
Q**A**AARAAASA	10.9 ± 3.2	91.0[b] ± 7.7	84.2[b] ± 7.7

[a] Ability of the mutants was compared to that of the empty expression vector pJH10 (negative control) and the WT *macp* (positive control). [b] denotes a significant change in the inhibition zone upon addition of IPTG.

have also been suggested when an engineered synthase comprised of erythromycin and rapamycin modules failed to produce any triketide lactones.[47]

Extended C-termini of mACPs A and D might facilitate protein–protein interactions

Amino acid alignments of the tailoring mAcps with other type II ACPs identified extended C-termini on mAcps A and D (Figure 2). The C terminus of mAcpD is a mixture of hydrophobic and hydrophilic residues, whereas the mAcpA extension predominantly consists of alanine residues, interspersed with isoleucine, arginine and serine. To investigate the importance of these regions the *macp*s with their C-termini deleted were cloned into pJH10, and their ability to complement a complete in-frame deletion in the chromosomal *macp* when expressed *in trans* was compared to that of the wild type (WT) *macp*s. Nine residues were deleted from mAcpA and 21 from mAcpD. For Δ*macpD* containing pJH10 only, mupirocin production was just 6% of WT levels. Expression of WT *macpD* restored antibiotic biosynthesis to 54% of WT levels without induction, increasing to 93% in the presence of 0.1 mM IPTG, while for the truncated *macpD* the equivalent figures were 20% of WT levels without induction and 75% in the presence of IPTG. For Δ*macpA* expression of WT mAcpA restored PA production to 90% of WT levels without induction, whereas the mutant *macp* was unable to complement Δ*macpA* even in the presence of IPTG. Some or all of the C-terminal nine residues of mAcpA therefore appear to be essential for function.

To investigate further which of the nine C-terminal mAcpA residues are most important to the function of mAcpA, we created mutants in which three or six residues were deleted in the C terminal, and Ser100 was substituted by alanine. These mutants were able to restore PA-A biosynthesis to between 40 and 50% of WT levels (Table 2). It was therefore hypothesised that the N-terminal three amino acids of the nine-residue section (IAA) might be essential for mAcpA activity. The mAcpA I93A mutant was unable to complement the *macpA* deletion without induction (in contrast to the WT gene), but in the presence of 0.1 mM IPTG mupirocin production was restored to 91% of WT levels (Table 2).

Although we cannot discount the possibility that these mutations reduce the level of *macp* expression in the absence of the inducer IPTG, this does not seem very likely in view of the nature and position of the mutations. We are therefore inclined to attribute the low activity of the mutant genes in the absence of IPTG to a reduced ability of the mutant proteins to participate in polyketide biosynthesis. This decrease in activity might be due to a number of possible factors. First, the protein might show decreased affinity for its partners. Second, the mutation might disrupt the protein structure and result in a protein fold that is less stable or competent than a substrate of the enzymes that modify the attached intermediate. Third, the ACP might be less able to perform a required biochemical function (for instance the protection of biosynthetic intermediates through chain binding).

Others have also suggested that the ACP C terminal might be involved in protein–protein interactions[45,46] but given the diversity of probable partners it is not surprising that a consensus is not found in these C-terminal sequences. In the case of mAcpD we certainly expect protein–protein interactions to be critical because it appears to be associated with MupQ, which is a putative Acyl-CoA synthase that is hypothesised to be involved in substrate loading. In the case of mAcpA the "extra" amino acids might be an integral part of an atypical C-terminal domain, although interaction with protein neighbours might also be indicated by the phenotype of the *macpA* I93A mutation.

Conclusions

A key finding of this study is that representatives of both the modular type I and the putative type II ACPs of the tailoring region were modified by MupN with addition of a mass equivalent to phosphopantetheine; this supports the conclusion that these polypeptides are indeed the apo-forms of ACPs and

that MupN is a phosphopantetheinyl transferase that activates them to the holo-form. While many PKS/NRPS systems lack a PPTase, there are examples of some with a dedicated PPTase, such as EntD (Sfp-type) in the enterobactin cluster,[16] FdmW (AcpS-type) in the fredericamycin synthase[48] and SePptI (integrated) in an uncharacterised PKS in *S. erythraea*.[25]

The general lack of phosphopantetheinylation of mupirocin ACPs in *E. coli* in the absence of MupN, together with the absolute requirement of MupN for mupirocin biosynthesis,[30] might suggest that the PPTase of *P. fluorescens* NCIMB 10586 that activates carrier proteins of primary metabolism does not have the broad substrate specificity of the PPTases previously identified in *P. aeruginosa*, *P. syringae* and *P. putida*.[26–28,49] It would be of interest to know which ACPs of the *mup* cluster are not activated by this host PPTase.

Interestingly, only mAcpA was observed to be a substrate for the *E. coli* AcpS, which has previously been found to be capable of modifying heterologous ACPs,[18–21,50] although some *Streptomyces* PKS ACPs were observed to be better substrates than others in vivo.[51,52] A number of studies have suggested the influence of ACP surface charge in recognition by PPTases,[18] particularly in the region of helix 2 and the DSL motif containing the active site serine.[53,54,50] In an extensive mutational study of *E. coli* ACP, De Lay and Cronan[50] showed that the Asp preceding the active site serine is critical for activation by AcpS of *E. coli*. Because mAcpA is the only type II ACP tested that has Asp in this position it might provide an explanation for why mAcpA but not mAcpC or mAcpD is activated by *E. coli* AcpS. However, because the type I ACPs ACP3, ACP5 and ACP6 all have Asp at this position and are not modified by *E. coli* AcpS there must be additional specificity determinants. The crystal structure of AcpS from *Bacillus subtilis* on its own and in complex with a cognate ACP actually implicates that helix 2 of the conserved ACP structure interacts with helix 1 of AcpS.[22] A critical interaction appears to be between the equivalent of Arg14 in *B. subtilis* AcpS and the Asp next but one along from the active site serine in the ACP. These residues are conserved in MupN and each of the ACPs studied here, respectively, so there is clearly scope for further mutational analysis to establish specificity rules.

The results presented thus lay the basis for a detailed analysis of the subset of tailoring enzymes that work with each specific mAcp and exploration of the proposed protein–protein interactions that create the multiprotein complexes that retain substrates while they are being processed. The genetic recruitment of such diverse ACPs along with their associated enzymes indicates that combinatorial genetics has been proceeding for many years before the idea of creating new pathways by in vitro recombinant DNA technology. Defining the functional units, understanding their role in mupirocin biosynthesis and the constraints on how they fit together should help us to perform further genetic shuffling ourselves in future.

Experimental Section

Bacterial strains and plasmids: *Pseudomonas fluorescens* NCIMB 10586 was used as the wild-type strain for mupirocin production. *P. fluorescens* mutants with in-frame deletions of each tailoring *macp* (named 10586Δ*macpA-E*) were used in cross-complementation studies. *E. coli* strain DH5α[55] was used for genetic manipulation, strain S17–1[56] was used for mobilisation of plasmids into *P. fluorescens* by conjugative transfer and *E. coli* BL21(DE3) was used for overexpression of ACPs. *Bacillus subtilis* 1064 was used as the mupirocin-sensitive organism in bioassays to test levels of antibiotic production by mutant strains of *P. fluorescens*. PCR fragments were initially cloned into pGEM-TEasy (Promega) for sequencing. Other plasmids used or constructed in this study are listed in Table S1.

Growth and culture conditions: *P. fluorescens* was grown at 30 °C in L-broth and L-agar (L-broth supplemented with agar (1.5%, w/v)). *E. coli* strains and *B. subtilis* were grown at 37 °C unless stated otherwise in L-broth and L-agar. For plasmid selection media was supplemented with ampicillin (50 $\mu g\,mL^{-1}$), tetracycline (15 $\mu g\,mL^{-1}$) or kanamycin (50 $\mu g\,mL^{-1}$) as appropriate.

DNA isolation and manipulation: Plasmid DNA was purified by using the alkaline SDS method[57] or by using the Wizard Plus SV Mini Preps DNA Purification Systems (Promega). Restriction endonuclease digestion of DNA was carried out by using enzymes purchased from New England Biolabs or MBI Fermentas (Burlington, Ontario, Canada). PCR amplification of genes was carried out by using the BIO-X-ACT™ Long DNA Polymerase kit (Bioline, Taunton, Massachusetts, USA). DNA fragments excised from agarose gels were purified using GeneClean Kit (Bio101, La Jolla, California, USA), and ligated by using T4 DNA ligase (Invitrogen).

DNA sequencing: Sequencing reactions were performed by using the ABI PRISM® Big Dye™ V3 Terminator kit. Sequencing reactions were separated on an ABI 3700 DNA Analyzer (Functional Genomics Laboratory, University of Birmingham)

Construction of expression plasmids: PCR fragments were initially cloned and sequenced in pGEM-T Easy (Promega). WT *macps* were excised by restriction digest and ligated into pJH10 as EcoRI-XbaI fragments to give plasmids pASRA11, pASRB11, pASRC11, pASRD11 and pASRE11. The hybrid *macp*A-B was constructed by cloning and sequencing the N terminus of *macpA* and the C terminus of *macpB* in pGEM-T-Easy, followed by ligation into pJH10 as a single EcoRI-XbaI fragment with an internal ClaI site joining the two halves, giving pASRAB11. The opposite hybrid (N terminus of *macpB* and C terminus of *macpA*) was constructed in the same way to give pJS321. C-terminal deletions of mAcps A and D were carried out by using mutagenic primers designed so that the genes could be cloned into pJH10 as EcoRI-XbaI/KpnI fragments (primers are listed in Table S2). For expression of his-tagged mACPs in *E. coli*, *macps* A-E were cloned into pGTB340 as EcoRI-HindIII fragments to give plasmids pJS551–555.

Bioassay for mupirocin production and HPLC analysis were carried out as previously described.[30]

Expression and purification of ACPs: E. *coli* seed cultures grown overnight at 37 °C were used to inoculate Luria–Bertani medium (LB; 400 mL; 2% inoculum). The production culture was incubated at 37 °C until the OD_{600} reached 0.4–0.6 followed by induction with IPTG (0.5 mM) and incubation at 30 °C for 4 h. The cells were harvested by centrifugation (15000 *g*) and stored at −20 °C. Pellets were thawed on ice and resuspended in Bugbuster mastermix (5 $mL\,g^{-1}$ of cell paste, Novagen) plus half a protease inhibitor cocktail tablet (Roche). The suspension was incubated at RT for >20 min on a rocking platform, followed by centrifugation (15000 g) for 20 min. Ni-NTA (Qiagen) was added to cell lysate

(1 mL per 4 mL of cell lysate) and rotated for 1 h at 4 °C. The slurry was transferred to a disposable polypropylene column (Qiagen), washed with wash buffer (4×8 mL; NaH_2PO_4 (50 mM), NaCl (300 mM), imidazole(20 mM)) and the proteins eluted with elution buffer
(6×1 mL) with increasing concentration of imidazole (NaH_2PO_4 (50 mM), NaCl (300 mM), imidazole(20 to 250 mM)).

Mass spectrometry: Proteins were desalted and concentrated as reported elsewhere.[58] Proteins were analysed in positive-ion mode by electrospray ionisation mass spectrometry (ES-MS) on an Applied Biosystems Qstar XL qTOF mass spectrometer with an Advion Nanomate™ source. Source gas pressure was 0.3 psi and voltage 1.4 kV. Curtain gas voltage was set to 20 V, declustering potential was 75 V and focussing potential was 280 V.

Acknowledgements

This work was funded by BBSRC grants P15257, P07071 and E021611 supporting work in Birmingham as well as EPSRC grant S78124 supporting work in Bristol. J.A.S. was funded by a BBSRC DTA studentship. A.S.R. was the recipient of a studentship from the Darwin Trust of Edinburgh.

Keywords: antibiotics · polyketides · PPTase · *Pseudomonas fluorescens* · pseudomonic acid

[1] A. C. Mercer, M. D. Burkart, *Nat. Prod. Rep.* **2007**, *24*, 750 – 773.
[2] M. P. Crump, J. Crosby, C. E. Dempsey, M. Murray, D. A. Hopwood, T. J. Simpson, *FEBS Lett.* **1996**, *391*, 302 – 306.
[3] M. P. Crump, J. Crosby, C. E. Dempsey, J. A. Parkinson, M. Murray, D. A. Hopwood, T. J. Simpson, *Biochemistry* **1997**, *36*, 6000 – 6008.
[4] S. E. Evans, C. Williams, C. J. Arthur, S. G. Burston, T. J. Simpson, J. Crosby, M. P. Crump, *ChemBioChem* **2008**, *9*, 2424 – 2432.
[5] M. C. Oswood, Y. Kim, J. B. Ohlrogge, J. H. Prestegard, *Proteins* **1997**, *27*, 131 – 143.
[6] X. Y. Qiu, C. A. Janson, *Acta Crystallogr. Sect. D Acta* **2004**, *60*, 1545 – 1554.
[7] A. Roujeinikova, C. Baldock, W. J. Simon, J. Gilroy, P. J. Baker, A. R. Stuitje, D. W. Rice, A. R. Slabas, J. B. Rafferty, *Structure* **2002**, *10*, 825 – 835.
[8] E. Ploskoń, C. J. Arthur, S. E. Evans, C. Williams, J. Crosby, T. J. Simpson, M. P. Crump, *J. Biol. Chem.* **2008**, *283*, 518 – 528.
[9] S. E. Evans, C. Williams, C. J. Arthur, E. Ploskoń, P. Wattana-Amorn, R. J. Cox, J. Crosby, C. L. Willis, T. J. Simpson, M. P. Crump, *J. Mol. Biol.* **2009**, *389*, 511 – 528.
[10] W. P. Revill, M. J. Bibb, D. A. Hopwood, *J. Bacteriol.* **1996**, *178*, 5660 – 5667.
[11] G. Florova, G. Kazanina, K. A. Reynolds, *Biochemistry* **2002**, *41*, 10462 – 10471.
[12] C. Khosla, R. McDaniel, S. Ebertkhosla, R. Torres, D. H. Sherman, M. J. Bibb, D. A. Hopwood, *J. Bacteriol.* **1993**, *175*, 2197 – 2204.
[13] Y. Ma, L. H. Smith, R. J. Cox, P. Beltran-Alvarez, C. J. Arthur, T. J. Simpson, *ChemBioChem* **2006**, *7*, 1951 – 1958.
[14] C. Y. Kim, V. Y. Alekseyev, A. Y. Chen, Y. Y. Tang, D. E. Cane, C. Khosla, *Biochemistry* **2004**, *43*, 13892 – 13898.
[15] K. J. Weissman, R. Müller, *ChemBioChem* **2008**, *9*, 826 – 848.
[16] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923 – 936.
[17] H. D. Mootz, R. Finking, M. A. Marahiel, *J. Biol. Chem.* **2001**, *276*, 37289 – 37298.
[18] A. M. Gehring, R. H. Lambalot, K. W. Vogel, D. G. Drueckhammer, C. T. Walsh, *Chem. Biol.* **1997**, *4*, 17 – 24.
[19] R. J. Cox, T. S. Hitchman, K. J. Byrom, I. S. C. Findlow, J. A. Tanner, J. Crosby, T. J. Simpson, *FEBS Lett.* **1997**, *405*, 267 – 272.
[20] L. Kremer, K. M. Nampoothiri, S. Lesjean, L. G. Dover, S. Graham, J. Betts, P. J. Brennan, D. E. Minnikin, C. Locht, G. S. Besra, *J. Biol. Chem.* **2001**, *276*, 27967 – 27974.
[21] Z. C. Suo, C. C. Tseng, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 99 – 104.
[22] K. D. Parris, L. Lin, A. Tam, R. Mathew, J. Hixon, M. Stahl, C. C. Fritz, J. Seehra, W. S. Somers, *Struct. Fold. Des.* **2000**, *8*, 883 – 895.
[23] L. E. N. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585 – 1595.
[24] E. Zazopoulos, K. X. Huang, A. Staffa, W. Liu, B. O. Bachmann, K. Nonaka, J. Ahlert, J. S. Thorson, B. Shen, C. M. Farnet, *Nat. Biotechnol.* **2003**, *21*, 187 – 190.
[25] K. J. Weissman, H. Hong, M. Oliynyk, A. P. Siskos, P. F. Leadlay, *ChemBioChem* **2004**, *5*, 116 – 125.
[26] R. Finking, J. Solsbacher, D. Konz, M. Schobert, A. Schafer, D. Jahn, M. A. Marahiel, *J. Biol. Chem.* **2002**, *277*, 50293 – 50302.
[27] N. Barekzi, S. Joshi, S. Irwin, T. Ontl, H. P. Schweizer, *Microbiology-SGM* **2004**, *150*, 795 – 803.
[28] H. F. Seidle, R. D. Couch, R. J. Parry, *Arch. Biochem. Biophys.* **2006**, *446*, 167 – 174.
[29] A. K. El-Sayed, J. Hothersall, S. M. Cooper, E. Stephens, T. J. Simpson, C. M. Thomas, *Chem. Biol.* **2003**, *10*, 419 – 430.
[30] J. Hothersall, J. Wu, A. S. Rahman, J. A. Shields, J. Haddock, N. Johnson, S. M. Cooper, E. R. Stephens, R. J. Cox, J. Crosby, C. L. Willis, T. J. Simpson, C. M. Thomas, *J. Biol. Chem.* **2007**, *282*, 15451 – 15461.
[31] V. Simunovic, R. Müller, *ChemBioChem* **2007**, *8*, 1273 – 1280.
[32] C. T. Calderone, W. E. Kowtoniuk, N. L. Kelleher, C. T. Walsh, P. C. Dorrestein, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8977 – 8982.
[33] D. J. Edwards, B. L. Marquez, L. M. Nogle, K. McPhail, D. E. Goeger, M. A. Roberts, W. H. Gerwick, *Chem. Biol.* **2004**, *11*, 817 – 833.
[34] L. C. Gu, B. Wang, A. Kulkarni, T. W. Geders, R. V. Grindberg, L. Gerwick, K. Hakansson, P. Wipf, J. L. Smith, W. H. Gerwick, D. H. Sherman, *Nature* **2009**, *459*, 731 – 735.
[35] X. H. Chen, J. Vater, J. Piel, P. Franke, R. Scholz, K. Schneider, A. Koumoutsi, G. Hitzeroth, N. Grammel, A. W. Strittmatter, G. Gottschalk, R. D. Sussmuth, R. Borriss, *J. Bacteriol.* **2006**, *188*, 4024 – 4036.
[36] C. Sanchez, L. C. Du, D. J. Edwards, M. D. Toney, B. Shen, *Chem. Biol.* **2001**, *8*, 725 – 738.
[37] J. N. Copp, B. A. Neilan, *Appl. Environ. Microbiol.* **2006**, *72*, 2298 – 2305.
[38] K. F. Geoghegan, H. B. F. Dixon, P. J. Rosner, L. R. Hoth, A. J. Lanzetti, K. A. Borzilleri, E. S. Marr, L. H. Pezzullo, L. B. Martin, P. K. LeMotte, A. S. McColl, A. V. Kamath, J. G. Stroh, *Anal. Biochem.* **1999**, *267*, 169 – 184.
[39] S. M. Cooper, W. Laosripaiboon, A. S. Rahman, J. Hothersall, A. K. El-Sayed, C. Winfield, J. Crosby, R. J. Cox, T. J. Simpson, C. M. Thomas, *Chem. Biol.* **2005**, *12*, 825 – 833.
[40] J. Wu, S. M. Cooper, R. J. Cox, J. Crosby, M. P. Crump, J. Hothersall, T. J. Simpson, C. M. Thomas, C. L. Willis, *Chem. Commun.* **2007**, 2040 – 2042.
[41] C. T. Calderone, D. F. Iwig, P. C. Dorrestein, N. L. Kelleher, C. T. Walsh, *Chem. Biol.* **2007**, *14*, 835 – 846.
[42] A. Roujeinikova, C. Baldock, W. J. Simon, J. Gilroy, P. J. Baker, A. R. Stuitje, D. W. Rice, J. B. Rafferty, A. R. Slabas, *Acta Crystallogr. Sect. D* **2002**, *58*, 330 – 332.
[43] G. Y. Xu, A. Tam, L. Lin, J. Hixon, C. C. Fritz, R. Powers, *Structure* **2001**, *9*, 277 – 287.
[44] Q. Li, C. Khosla, J. D. Puglisi, C. W. Liu, *Biochemistry* **2003**, *42*, 4648 – 4657.
[45] S. C. Findlow, C. Winsor, T. J. Simpson, J. Crosby, M. P. Crump, *Biochemistry* **2003**, *42*, 8423 – 8433.
[46] H. C. Wong, G. H. Liu, Y. M. Zhang, C. O. Rock, J. Zheng, *J. Biol. Chem.* **2002**, *277*, 15874 – 15880.
[47] A. Ranganathan, M. Timoney, M. Bycroft, J. Cortes, I. P. Thomas, B. Wilkinson, L. Kellenberger, U. Hanefeld, I. S. Galloway, J. Staunton, P. F. Leadlay, *Chem. Biol.* **1999**, *6*, 731 – 741.
[48] Y. Huang, E. Wendt-Pienkowski, B. Shen, *J. Biol. Chem.* **2006**, *281*, 29660 – 29668.
[49] F. Gross, D. Gottschalk, R. Müller, *Appl. Microbiol. Biotechnol.* **2005**, *68*, 66 – 74.
[50] N. R. De Lay, J. E. Cronan, *J. Biol. Chem.* **2007**, *282*, 20319 – 20328.
[51] B. Shen, R. G. Summers, H. Gramajo, M. J. Bibb, C. R. Hutchinson, *J. Bacteriol.* **1992**, *174*, 3818 – 3821.

[52] J. Crosby, D. H. Sherman, M. J. Bibb, W. P. Revill, D. A. Hopwood, T. J. Simpson, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1995**, *1251*, 32 – 42.
[53] M. R. Mofid, R. Finking, M. A. Marahiel, *J. Biol. Chem.* **2002**, *277*, 17023 – 17031.
[54] K. J. Weissman, H. Hong, B. Popovic, F. Meersman, *Chem. Biol.* **2006**, *13*, 625 – 636.
[55] D. Hanahan, *J. Mol. Biol.* **1983**, *166*, 557 – 580.
[56] R. Simon, U. Priefer, A. Pühler, *Bio-Technology* **1983**, *1*, 784 – 791.
[57] H. C. Birnboim, J. Doly, *Nucleic Acids Res.* **1979**, *7*, 1513 – 1523.
[58] R. L. Winston, M. C. Fitzgerald, *Anal. Biochem.* **1998**, *262*, 83 – 85.

Received: September 9, 2009
Published online on December 10, 2009

