Genetic and functional analysis of genes required for the post-modification of the polyketide antibiotic TA of *Myxococcus xanthus*

Yossi Paitan,¹ Elisha Orr,² Eliora Z. Ron¹ and Eugene Rosenberg¹

Author for correspondence: Eliora Z. Ron. Tel: +972 3 6409379. Fax: +972 3 6414138.
e-mail: eliora@ccsg.tau.ac.il

¹ Department of Molecular Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, Israel

² Department of Genetics, University of Leicester, Leicester LE1 7RH, UK

The antibiotic TA of *Myxococcus xanthus* is a complex macrocyclic polyketide, produced through successive condensations of acetate by a type I PKS (polyketide synthase) mechanism. The genes encoding TA biosynthesis are clustered on a 36 kb DNA fragment, which has been cloned and analysed. The chemical structure of TA and the mechanism by which it is synthesized indicate the need for several post-modification steps, which are introduced into the carbon chain of the polyketide to form the final bioactive molecule. These include the addition of several carbon atoms originating from acetate carbonyl, three C-methylations, O-methylation and a specific hydroxylation. This paper reports the analysis of five genes which are involved in the post-modification of TA. Their functional analysis, by specific gene disruption, suggests that they may be essential for the production of the active antibiotic. The characteristics and organization of the genes suggest that they may be involved in the addition of the carbon atoms which arise from acetate.

Keywords: *Myxococcus xanthus*, polyketide synthase, acyl carrier protein, β -ketoacyl ACP synthase III, antibiotic post-modification

INTRODUCTION

Polyketides are a large and highly diverse group of secondary metabolites displaying a broad range of biological activities, some with medical applications. They are synthesized through repetitive condensations of acetate, or propionate, monomers in a similar way to that of fatty acid biosynthesis (Hopwood *et al.*, 1990). In general, polyketide synthase (PKS) systems are classified into two types. Type I consists of large, multifunctional enzymes, containing a separate site for each condensation, or modification step, also referred to as 'modular PKSs'. Type II PKS systems are made up of individual enzymes in which each catalytic site is used repeatedly during the biosynthetic process. Genetic studies have indicated that the gene organization of functional units and motif patterns of various PKSs are similar (Donadio *et al.*, 1991; Donadio & Katz, 1992; Malpartida *et al.*, 1987). Primary structural diversity of polyketides is

achieved through the introduction of different thioester primers with varying chain extensions by PKSs. Additional variations can occur in the stereochemistry and the degree by which intermediates are reduced in each condensation step. A different type of diversity can be achieved by the subsequent processing of polyketide chains, such as alkylations, oxidations, O-methylations, glycosylations and cyclizations.

Genetic studies of prokaryotic PKSs have focused on Gram-positive bacteria, particularly on actinomycetes, and not much is known about PKSs in other bacteria. Myxobacteria are Gram-negative bacteria that produce a large number of secondary metabolites, including polyketides (Reichenbach & Hofle, 1993). They are soil micro-organisms, feeding on proteins and peptides, that undergo a complex life cycle which includes cell-to-cell interactions, signalling, myxospore formation and fruiting body formation (Shimkets, 1990). We have recently studied the biosynthesis of the novel macrolide antibiotic, TA (Fig. 1), produced by *Myxococcus xanthus* (Rosenberg *et al.*, 1973), which inhibits bacterial cell wall synthesis (Zafiri *et al.*, 1981). It is a unique antibiotic in its ability to adhere avidly to a variety of tissues (Rosenberg *et al.*, 1984), a property that has been

Abbreviations: ACP, acyl carrier protein; HMG, hydroxymethylglutaryl; KAS, β -ketoacyl:ACP synthase; PKS, polyketide synthase.

The EMBL accession number for the sequence reported in this paper is AJ132503.

Fig. 1. Chemical structure of the antibiotic TA. TA is a macrocyclic polyketide containing an $\alpha\beta$ unsaturated methoxy group, synthesized from the precursors acetate, methionine and glycine (Fytlovitch *et al.*, 1982; Trowitzsch *et al.*, 1982).

exploited in a successful preliminary clinical trial for treatment of gingivitis (Manor *et al.*, 1989). The production of TA and its chemical properties have been studied extensively (Rosenberg *et al.*, 1973, 1982; Rosenberg & Varon, 1984; Trowitzsch *et al.*, 1982). It is synthesized from the precursors acetate, methionine and glycine (Fytlovitch *et al.*, 1982). Genes required for the biosynthesis of TA have been identified by transposon mutagenesis, and several of them were mapped by P1-transduction to a 36 kb region of the bacterial chromosome (Varon *et al.*, 1992) and shown to be co-regulated (Tolchinsky *et al.*, 1992; Varon & Rosenberg, 1996; Varon *et al.*, 1997). Several genes involved in TA post-modification and regulation, such as a cytochrome P-450 hydroxylase (Paitan *et al.*, 1999b) and a specific transcription anti-terminator (Paitan *et al.*, 1999c), were cloned and analysed. One of the cloned genes in this cluster encodes a type I PKS module, indicating that TA

is produced through a type I PKS mechanism (Paitan *et al.*, 1999a).

The chemical structure of TA, the pattern of precursor incorporation into TA, as studied by NMR, and the type I PKS mechanism by which it is synthesized suggest the existence of several post-modification steps, such as the addition of several carbon atoms that originate from acetate at C-13, C-17 and C-32, a specific hydroxylation at C-20, three C-methylations at C-2, C-4 and C-28, and an O-methylation at C-34 (Fig. 1). The time and mechanism by which these modifications are introduced into the primary polyketide chain, as well as the enzymes catalysing these reactions, are still unknown. In this study, we report the genetic analysis of five ORFs, all believed to be involved in post-modification steps of TA, and discuss their possible functions. The specific gene disruptions of two ORFs (*taB* and *taF*) and the deletion of four ORFs (*taC*–*taF*) suggests that these genes are involved in and are probably essential for the production of a biologically active TA.

METHODS

Bacterial strains and plasmids. The *M. xanthus* strains used in this study were the wild-type strain ER-15 and the TnV mutants ER-1311 and ER-6199, which are blocked in TA production and have been described previously (Varon *et al.*, 1992). *Escherichia coli* TG1 (Bethesda Research Laboratories) and XL-1 Blue MR (Stratagene) were used for cloning and DNA manipulations. A conjugative tagged Tn1000 transposition system was used for the sequencing as described elsewhere (Sedgwick & Morgan, 1994), using the *E. coli* strains MH1578 and MH1599. pUC18, pUC19 (Norrander *et al.*, 1983) and SUPERCOS-1 (Stratagene) are *E. coli* vectors used for both cloning and sequencing.

Media and growth conditions. *E. coli* was grown at 37 °C in Luria broth (LB), or on LB agar, with the appropriate antibiotics (Sambrook *et al.*, 1989). *M. xanthus* was grown at 32 °C in 0.5 CTS, 1 CT or CTK medium, as required, or on media solidified with 1.5% Bacto agar (Difco) as described before (Tolchinsky *et al.*, 1992; Varon *et al.*, 1992).

Fig. 2. Specific gene disruption of *taF* and deletion of all four ORFs by insertion of a kanamycin-resistance gene into pPY10 and pPY11, forming plasmids pPY11S-Kan, pPY10S-Kan and pPY10Nc-Kan. The resulting plasmids were linearized and electroporated into *M. xanthus* ER-15. Integration of the linearized DNA by homologous recombination into the *M. xanthus* ER-15 chromosome blocked TA production. Restriction enzymes: S, *Sall*; X, *XhoI*; H3, *HindIII*; Sp, *SphI*; Nc, *NcoI*.

Fig. 3. Physical map of the TA gene cluster. (a) Chromosomal restriction map of the TA gene cluster; the transcription direction of the Tn5/*lac* mutants is marked by an arrow above the mutant number (Paitan *et al.*, 1999a). (b) The inserts derived from plasmids pPYT1311/p5 and pPYT6199 (Paitan *et al.*, 1999b) were used as TA-specific probes to identify TA-essential genes. (c) Restriction map of the cosmids pPYCC64 and pPYCA111 (Paitan *et al.*, 1999b). (d) The 4.4 kb *Hind*III–*Sph*I fragment and the 0.95 kb *Xho*I–*Sph*I fragment cloned in plasmids pPY10 and pPY11, respectively, used for sequencing, specific gene disruption and the organization of the ORFs in this fragment (grey arrows) and other previously analysed ORFs in the TA cluster (black arrows). Restriction enzymes: S, *Sall*; X, *Xho*I; Bm, *Bam*HI; Ei, *Eco*RI; K, *Kpn*I; H, *Hind*III; Bg, *Bgl*II.

General DNA procedures. Standard genetic techniques, Southern blot analysis, colony hybridization, plasmid preparations and *in vitro* DNA manipulations were performed according to published protocols (Sambrook *et al.*, 1989). Isolation of total DNA from *M. xanthus* is described elsewhere (Paitan *et al.*, 1998). Cosmid DNA and DNA templates for sequencing reactions and electroporation were purified by Qiagen columns. Electroporation of *M. xanthus* was performed as described before (Kashefi & Hartzell, 1995) using a Gene Pulser II (Bio-Rad). Conjugative transposition of Tn1000

for sequencing is described elsewhere (Paitan *et al.*, 1998). Construction of plasmids for specific gene disruption was achieved by inserting a kanamycin-resistance gene to: (i) pPY11, disrupting the *taF* gene by replacing an 11 bp *Sall*–*Sall* fragment (positions 3891–3902), (ii) pPY10, producing a deletion of four ORFs (*taC*–*taF*) by replacing a 2.667 kb *Sall*–*Sall* fragment (positions 1235–3902), and (iii) pPY10, disrupting the *taB* gene by replacing a 249 bp *Nco*I–*Nco*I fragment (positions 429–678; Fig. 2).

Preparation of TA-specific probes. DNA extracted from the

Table 1. Characteristics of the ORFs and their predicted products

ORF	Position*	No. of amino acids	Estimated mol. mass (Da)	Homology			
				Name†	Identity (%)	Similarity (%)	Accession no.
<i>taB</i>	431–667	79	8729	ACP	28	52	gi:2984024
<i>taC</i>	686–1939	420	45496	PksG	66.6	79.7	sp:P40830
				HMG‡	26	40	sp:P23228
				FabH§	24	37	pir:A64493
				Cha	34	50	sp:P48395
<i>taD</i>	1965–2939	325	35483	Several different membrane-associated proteins ¶			
<i>taE</i>	2947–3195	83	9059	ACP	33	54	gi:3603039
<i>taF</i>	3318–4439	374	41090	PksG	45.55	62.6	sp:P40830
				HMG‡	26	45	sp:P54873
				FabH§	24	36	pir:A64493
				Cha	28	46	gi:1944134

* Nucleotide position in the total DNA fragment, accession number AJ132503.

† Protein or putative gene product (in the case of PksG).

‡ HMG-CoA synthetase of *Gallus gallus* (chicken) in the case of *taC* (over a 388 aa stretch) and of *Arabidopsis thaliana* in the case of *taF* (over a 371 aa stretch).

§ KAS III protein (FabH) of *Methanococcus jannaschii* in both *taC* (over a 345 aa stretch) and *taF* (over a 297 aa stretch).

|| Chalcone synthase enzymes (naringenin-chalcone synthase) of *Ipomoea nil* in the case of *taC* (over a 107 aa stretch) and of *Perilla frutescens* in the case of *taF* (over a 297 aa stretch).

¶ See text for homology information.

TnV transposant ER1311 (Varon *et al.*, 1992) was digested with *KpnI* (does not restrict TnV), self-ligated and transformed into *E. coli* TG1, using the transposon-derived kanamycin resistance for selection. The resulting recombinant plasmid, designated pPYT1311/p5, was found to carry a 2.3 kb *M. xanthus* chromosomal fragment (see Fig. 3b) and was used as a TA-specific probe together with an 11.2 kb fragment previously obtained as a TA-specific probe from plasmid pPYT6199 (Paitan *et al.*, 1999b).

DNA sequencing and analysis. Automated DNA sequencing was performed on double-stranded DNA templates by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977), using an Applied Biosystems model 373A sequencer. Several cosmids, originating from chromosomal regions that were found through gene disruption to be essential for TA production, were analysed further. Inserts from two recombinant cosmids, pPYCC64 and pPYCA111 (Paitan *et al.*, 1999b; Fig. 3c), were subcloned into the pUC19 vector.

A 4.5 kb fragment was sequenced. Occasional sequence gaps were filled in by direct sequencing, using synthetic oligonucleotide primers. Analysis and assembly of the primary DNA sequence data were carried out using the MacVector 3.5 (International Biotechnologies) and Sequence Navigator (Applied Biosystems) software. Searches of databases for DNA and protein sequence homology were carried out using the BLAST (Altschul *et al.*, 1990) and FASTA (Pearson, 1990) programs.

TA production assay and developmental analysis. The production of the antibiotic TA was determined by the disc assay described previously (Varon *et al.*, 1992). Fruiting body formation and development were examined on TPM agar medium (10 mM Tris/HCl, 1 mM KH₂PO₄, 8 mM MgSO₄, final pH 7.6, and 1.5% agar), as described by Kroos *et al.* (1986).

RESULTS

Analysis of a *M. xanthus* DNA fragment encoding genes involved in TA biosynthesis

A library of *M. xanthus* ER-15, constructed in the cosmid vector SUPERCOS-1, was screened with two specific TA probes (2.3 kb and 11.2 kb fragments, Fig. 3b) obtained through the cloning of a *M. xanthus* chromosomal DNA fragment adjacent to the TnV transposons which blocked TA production, in ER1311 and ER6199 (Varon *et al.*, 1992; Paitan *et al.*, 1999b; see Methods). Southern blot analysis of *KpnI*-digested chromosomal DNA, probed with these specific TA sequences, indicated the presence of identical genomic 2.3 kb and 11.2 kb fragments in ER-15 (data not shown). In strains transposed with TnV, the size of these two fragments increased to 8.3 kb (ER-1311) and 17.2 kb (ER-6199), in line with the size of TnV (6 kb). The cosmids pPYCC64 and pPYCA111 were found, through colony hybridization, to contain the 2.3 kb and the 11.2 kb *KpnI* fragments, respectively, and were further characterized (Fig. 3c; Paitan *et al.*, 1999b).

The inserts in cosmids pPYCC64 and pPYCA111 were subcloned into pUC19 and two recombinant plasmids – pPY10, carrying a *HindIII*–*SphI* fragment (4.391 kb), and pPY11, carrying a *XhoI*–*SphI* fragment (0.948 kb) – were analysed (Fig. 3d). The nucleotide sequence of the two fragments was determined, while the sequence 3' to pPY10 was obtained from the cosmid pPYCA111, used as a template, and a synthetic oligonucleotide primer complementing pPY10 sequences. The nucleotide se-

Specific gene disruptions

Fig. 4. (a, b) Comparison of the amino acid sequences deduced from (a) *taC* and (b) *taF* (this report) with PksG from *B. subtilis* (Albertini *et al.*, 1995); (c) comparison of *TaC* with *TaF*. The degree of identity between the sequences is indicated as follows: complete identity, vertical bar; conservative replacement, two dots; neutral replacement, one dot; no similarity, no symbol. Numbers at the left and right margins indicate the amino acid position at the beginning and end of each row, respectively.

TaC- <i>M. xanthus</i>	L S R N C R L I E L K S A C Y S G V A G L Q M A V N F I L S G V S P G A K A L V V A S D	144
TaF- <i>M. xanthus</i>	L S R H C R F L E V K Q A C Y A A T G A L Q L A L G Y I A S G V S P G A K A L V I A T D	107
PksG- <i>B. subtilis</i>	L N R N C R L F E V K Q A C Y S G T A G F Q M A V N F I L S Q T S P G A K A L V I A S D	69
HMCS-HUMAN	G N T D I E G I D T T N A C Y G G T A A V F N A V N W I E S S S W D G R Y A L V V A G D	159
HMCS-CHICK	G N T D V E G I D T T N A C Y G G T A A L F N A I N W I E S S S W D G R Y A L V V A G D	159
HMCS-RAT	G N T D I E G I D T T N A C Y G G T A A V F N A V N W I E S S S W D G R Y A L V V A G D	159
FabH - <i>E. coli</i>	G I K G C P A F D V A A A C A G F T Y A L S V A D Q Y V K S G - - A V K Y A L V V G S D	159

Fig. 5. Alignment of amino acid sequences encompassing the presumed acetylation/condensation site, including the conserved cysteine residues of *M. xanthus* TaC, TaF (this report), *B. subtilis* PksG (Albertini *et al.*, 1995), HMG-CoA synthases (HMCS) from human, chicken and rat (Russ *et al.*, 1992; Kattar-Cooley *et al.*, 1990; Ayte *et al.*, 1990) and the active-site cysteine of the condensing site in *E. coli* FabH (Tsay *et al.*, 1992). Boxed letters indicate residues identical in more than four of the seven amino acid sequences. The conserved cysteine residue is arrowed.

(a)

<i>M. xanthus</i> TaB	1	M D K R I I F D I V T S S V R E V V P E L E S H P F E P E D D L	32
<i>M. xanthus</i> TaE	1	M T D E Q I R G V V H Q S I V R V L P R V R S N E I A G H L N L	32
<i>A. aeolicus</i> ACP	1	M S L E E - - R V K E I I A E Q L G V E K E K I T P E A K F V	29
<i>G. theta</i> ACP	1	M N E Q E I F E K V Q T I I S E Q L G V D K S Q V T K D A N F A	32
<i>S. erythraea</i> ACP	1	M D R K E I F E R I E Q V L A E Q L G I P A E Q I T E E A D L R	32
<i>M. xanthus</i> TaB	33	V G L G A N S L D R A E I V N L T L E K L - - A L N I P R V E L	62
<i>M. xanthus</i> TaE	33	R E L G A D S V D R V E I L T S I L D S L - - R L Q K T P L A K	62
<i>A. aeolicus</i> ACP	30	E D L G A D S L D V V E L I M A F E E F G I E I P D E D A E K	61
<i>G. theta</i> ACP	33	N D L G A D S L D T V E L V M A I E E A F N I E I P D D A A E Q	64
<i>S. erythraea</i> ACP	33	E D L G M D S L D L V E L V S A L E D E V G M R V E Q S Q L E G	64
<i>M. xanthus</i> TaB	63	I - D A K T I G G L V D V L H A R L	79
<i>M. xanthus</i> TaE	63	F A D I R N I D A L V A F L A G E V A G G	83
<i>A. aeolicus</i> ACP	62	I Q T V G D V I N Y L K E K V G G	78
<i>G. theta</i> ACP	65	I S N L Q Q A V D F I S Q K V A A	81
<i>S. erythraea</i> ACP	65	I E T V G H V M E L T L D L V A R L A T A S A A D K P E A A S	95

(b)

	<i>M. xanthus</i> TaB	V G L G A N S L D R A E
	<i>M. xanthus</i> TaE	R E L G A D S V D R V E
PKS type I ACPs	<i>Saccharopolyspora erythraea</i> EryA I	A E L G V D S L S A L E
	<i>S. cellulosum</i> Sorphen A	F D L G L D S I M T V E
	<i>B. subtilis</i> PksX-acp2	S E Y G F D S V T F T V
PKS type II ACPs	<i>Streptomyces glaucescens</i> Tcm I-ORF3	Q D L G Y D S I A L L E
	<i>Streptomyces violaceoruber</i> Gra ORF3	E E L G Y D S L A L M E
	<i>Saccharopolyspora erythraea</i>	E D L G M D S L D L V E
FAS ACPs	<i>E. coli</i>	A D L G A D S L D T V E
	<i>Spinacea oleracea</i>	S K L G A D S L D T V E

Fig. 6. (a) Multiple alignment of the deduced amino acid sequences of several ACPs including *M. xanthus* TaB and TaE (this report), *Aquifex aeolicus* (AE000752), *Guillardia theta* (AF041468) and *Saccharopolyspora erythraea* (M64477). Boxed letters indicate residues identical in more than four of the six proteins. (b) Alignment of amino acid sequences encompassing the presumed active-site serine residues (arrowed) in PKS ACPs from *M. xanthus* TaB and TaE (this report), *Sacch. erythraea* EryA I (Donadio & Katz, 1992), *Sorangium cellulosum* Sorphen A (Schupp *et al.*, 1995), *B. subtilis* PksX-Acp2 (Albertini *et al.*, 1995), *Streptomyces glaucescens* Tcm I-ORF3, *Streptomyces violaceoruber* Gra-ORF3 (Bibb *et al.*, 1989), and in fatty acid synthase (FAS) ACPs from *Sacch. erythraea*, *E. coli* and *Spinacea oleracea* (Sherman *et al.*, 1989). Boxed letters indicate residues identical in more than eight of the nine amino acid sequences.

that their activities can be replaced by other bacterial proteins.

Characterization, identification and deduced function of the ORFs

The results, as summarized in Table 1 and displayed in Figs 4 and 5, indicate that the proteins encoded by *taC* and *taF* are acyl enzymes possessing the activities of transacylase and condensing enzymes. This was concluded by the identification of a highly conserved acetylation site containing the conserved cysteine within the active site of condensing enzymes (C-114 in *taC* and C-77 in *taF*; Fig. 5), and the high similarity to proteins such as the *B. subtilis* PksG (Fig. 4a, b) hydroxy-

methylglutaryl-CoA synthase (HMG-CoA synthetase; Fig. 5) and the β -ketoacyl-ACP synthase III proteins (KAS III, FabH, Fig. 5), which are all acyl enzymes possessing these activities. In addition to the highly conserved acetylation site, a second conserved motif, identified in the N-terminal amino acid sequence of both proteins (amino acids 16–24 and 35–43 in TaC and TaF, respectively) is a Myc-type helix-loop-helix, a dimerization domain signature (PDOC00038, PS50075), which is discussed later.

There is also homology between TaC and TaF. The proteins are 47.2 % identical and 55.5 % similar (Fig. 4c), suggesting that they carry out similar functions in TA biosynthesis, as transacylase/condensing enzymes. The

presence of two similar proteins that appear to possess the same activities, together with the gene disruption results (see above), strongly suggests that each protein has a different and specific role in TA biosynthesis.

As summarized in Table 1 and displayed in Fig. 6, the products of the other two homologous genes, *taB* and *taE*, share considerable end-to-end amino acid similarity to many small (~9 kDa) fatty acid synthase ACPs (Wang & Liu, 1991) and PKS type I and type II ACPs (Fig. 6). The search for conserved motifs indicates that both proteins contain a conserved active serine residue within a highly conserved amino acid motif (the 4'-phosphopantetheine binding region) found in many ACPs (Fig. 6b). Like *taC* and *taF*, the *taB* and *taE* products also share sequence similarity (34.3% identity and 40.3% similarity; Fig. 6), implicating them in a similar function as TA-modifying, specific ACPs in the antibiotic biosynthesis; again the presence of two similar proteins strongly suggests that each protein has a different and specific role in TA biosynthesis.

The *taD* gene product is similar in amino acid sequence to several cell-surface, membrane-associated proteins. These include: (i) the PII-type membrane-associated proteinase precursor of *Lactobacillus paracasei* (Q024-70), (ii) the integrin α -E precursor, which is a transmembrane protein involved in cell adhesion and extracellular signalling in mouse (Q60677) and rat (AF020-045), (iii) the histidine protein kinase sensor, HpkY, a transmembrane signal sensor protein of *Pseudomonas syringae* (AF001355), (iv) the CDO cell surface glycoprotein of *Rattus norvegicus* (AF004840), and several other membrane-associated proteins. TaD showed 26–28% identity and 42–45% similarity to these proteins within a 62–100 amino acid stretch. This homology suggests that TaD could be a membrane/transmembrane protein, involved in extracellular signalling.

DISCUSSION

In this work, we report the identification of several genes involved in post-modification steps of TA biosynthesis. We have demonstrated through specific gene disruption that functional inactivation of these genes blocks the production of the biologically active antibiotic.

The results indicate that the proteins encoded by *taC* and *taF* are homologues; they were identified as acyl enzymes, as they contain the highly conserved acetylation site, present in many condensing enzymes. The two proteins may therefore share a similar function in the post-modification process of TA. Both TaC and TaF share sequence homology with several proteins, such as members of the HMG-CoA synthetase family, KAS III proteins, including the *E. coli* 3-oxoacyl-[ACP] synthase III (FabH; Tsay *et al.*, 1992; Fig. 4) and proteins of the chalcone synthase family (naringenin-chalcone synthase). Such homologies imply that TaC and TaF possess, similar to KAS III proteins, the activities of both acetyl transacylase and a condensing enzyme. HMG-CoA synthetases are a defined class of enzymes, involved

in ketogenesis in eukaryotes, condensing acetyl-CoA with acetoacetyl-CoA to form HMG-CoA, which is a substrate of HMG-CoA reductase in cholesterol biosynthesis (Ayte *et al.*, 1990; Kattar-Cooley *et al.*, 1990; Russ *et al.*, 1992). The *E. coli* 3-oxoacyl-[ACP] synthase III protein (FabH; Tsay *et al.*, 1992) is a member of the KAS III family. FabH mediates the decarboxylative condensation between acetyl-CoA and malonyl-ACP to generate acetoacetyl ACP, and unlike KAS I (FabB) and KAS II (FabF), it is the only condensing enzyme able to utilize CoA thioesters rather than acyl-ACP as the primer in the condensation reaction. FabH displays two activities: one is that of acetyl transacylase, which catalyses the transfer of a 4'-phosphopantetheine moiety from a corresponding coenzyme A to a serine residue, present in the conserved active site of ACPs. The second is that of an acetoacetyl-ACP synthase condensing enzyme (Tsay *et al.*, 1992). Chalcone synthase enzymes (naringenin-chalcone synthase) are ACPs which catalyse the formation of naringenin-chalcone from malonyl CoA and 4-coumaryl-CoA, which is presumably formed from polyketide intermediates (Durbin *et al.*, 1995). As shown in Fig. 5, the region responsible for the acetylation/condensation activities, including the active cysteine, is highly conserved in both TaC and TaF.

The effect of disruption of *taF* suggests that it is probably required for the production of an active antibiotic. Moreover, the *taC* gene product does not complement the disrupted *taF* gene, located downstream of it, in the *taF*-disrupted mutant. However, it is possible that the block in TA production is the result of a polar effect, of the insertion in *taF*, on other genes downstream of *taF*. This possibility is difficult to exclude since there are no self-replicating plasmids in *M. xanthus*, which could be used for *trans*-complementation of the disrupted *taF*. This finding, and the presence of two similar homologous proteins, suggests that TaC and TaF have different and specific roles in TA biosynthesis.

Interestingly, while examining the end-to-end similarities between the deduced proteins encoded by *taC*, *taF* and *pksG*, we observed that the two former proteins are longer than the *pksG* product, displaying an extension at their N-termini (75 and 38 amino acids, respectively). In both TaC and TaF, a Myc-type helix-loop-helix, a dimerization domain signature, has been identified within the N-terminal amino acid sequence. This motif is absent from the putative PksG, HMG-CoA synthetase, and KAS III proteins. The Myc-type helix-loop-helix domain is thought to be a dimerization domain of several eukaryotic transcriptional regulators and DNA binding proteins. In the case of the TaC and TaF, this motif could have a physiological importance, either as a dimerization site, or perhaps as a recognition site for an interacting protein or ligand.

The identification of TaB and TaE as TA-modifying specific ACPs is based on their end-to-end homology to many small (~9 kDa) ACPs. Furthermore, in both proteins, a presumed active serine residue is present

within a highly conserved amino acid motif found in other polyketide-ACPs (the 4'-phosphopantetheine binding motif; Fig. 6). A ~9 kDa ACP is a well-characterized component of all fatty acid synthase systems and of several PKS systems. The small ACPs found in the TA cluster could act as specific post-modifying enzymes in TA biosynthesis. The exact role(s) of TaB and TaE is as yet unknown, although gene disruption of *taB* suggests that it may be essential for TA production; but in-frame deletion of *taB* needs to be tested for the production of biologically active TA. In addition, the presence of the two ACPs, the effect of gene disruption of *taB* and the finding that TaE does not complement the disrupted *taB* suggest that each ACP performs a specific step in TA biosynthesis and probably recognizes a different substrate. This suggestion is supported by a recent study (Ritsema *et al.*, 1998) describing the structure-function relationships of the chimeric protein AcpP-NodF. That study suggested that the N-terminal domain of ACP proteins determines the activity, while the C-terminal part carries a specialized domain for protein recognition. Again, it is possible in the *taB taE* case that the lack of complementation is the result of a polar effect (of the insertion in *taB*) on other genes 3' to *taB*.

Considering the structure of TA, the chemical pattern of type I PKS condensation steps and the pattern of precursor incorporation to TA, as studied by NMR (Rosenberg & Varon, 1984; Rosenberg *et al.*, 1982; Trovitzsch *et al.*, 1982), several post-modification steps, including addition of 3 carbon atoms (at C-13, C-17 and C-32) originating from acetate carbon atoms, specific hydroxylation (at C-20), O-methylation (at C-34) and three C-methylations (at C-2, C-4 and C-28) should exist. The identification of *taB*, *taC*, *taE* and *taF* as two sets of genes (*taB-taC* and *taE-taF*) sharing sequence similarity to ACPs and KAS III proteins suggests that they could be involved in the addition of two of the carbon atoms which, as indicated by NMR, originate from acetate (C-32, C-33 and C-34). Their organization further implies a possible coupling of activities, as each ACP synthase (TaC and TaF) probably processes the corresponding ACP (TaB and TaE, respectively). However, further work is required to examine this possibility, as there are no other data from other polyketide systems which display a similar genomic structure. Future analysis of a non-polar disruption of these genes followed by NMR analysis of TA intermediates should resolve the exact functions and the specificities of these proteins.

The most meaningful similarities found between TaD and proteins in the databases were to membrane-associated cell-surface proteins. The inability to assign a function to *taD* from the sequence analysis may reflect a unique feature of TA as a polyketide. TaD could be involved in TA regulation, or in an as yet unknown physiological role of TA. The partial similarities to several membrane/cell surface-associated proteins, involved in signal transduction, suggest that TaD is likely to be a membrane/transmembrane protein involved in

extracellular signalling/sensing related to TA biosynthesis.

In summary, the present study represents the first molecular analysis of genes involved in post-modification of a type I PKS, complex polyketide antibiotic in Gram-negative bacteria. This information should make it possible to compare the Gram-negative systems to the better-characterized polyketide systems of the Gram-positive actinomycetes. In addition, it may reveal novel genetic and biochemical features, which could provide the basis for developing new biologically active polyketides using combinatorial genetics.

ACKNOWLEDGEMENTS

We thank Dr Steven G. Sedgwick for providing the tagged Tn1000 system. This work was supported in part by the Pasha Gol Chair for Applied Microbiology (E.R.), the Morris and Manja Leigh Chair for Biophysics and Biotechnology (E.Z.R.), a British Council Clore Foundation Scholarship awarded to Y.P., and Cancer Research Campaign (no. SP1937/0301) and Wellcome Trust (no. 038060/Z/93/Z) grants to E.O.

REFERENCES

- Albertini, A. M., Caramori, T., Scoffone, F., Scotti, C. & Galizzi, A. (1995). Sequence around the 159° region of the *Bacillus subtilis* genome: the *pksX* locus spans 33.6 kb. *Microbiology* **141**, 299–309.
- Altschul, S. F., Gish, W., Miller, W., Myers, W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **251**, 403–410.
- Ayte, J., Gil-Gomez, G. & Hegardt, F. G. (1990). Nucleotide sequence of a rat liver cDNA encoding the cytosolic 3-hydroxy-3-methylglutaryl coenzyme A synthase. *Nucleic Acids Res* **18**, 3642–3642.
- Bibb, M. J., Biro, S., Motamedi, H., Collins, J. F. & Hutchinson, C. R. (1989). Analysis of the nucleotide sequence of the *Streptomyces glaucescens* *tcml* genes provides key information about the enzymology of polyketide antibiotic biosynthesis. *EMBO J* **8**, 2727–2736.
- Donadio, S. & Katz, L. (1992). Organization of the enzymatic domain in multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*. *Gene* **111**, 51–60.
- Donadio, S., Staver, M. J., McAlpine, J. B., Swanson, S. J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675–679.
- Durbin, M. L., Learn, G. H., Jr, Huttley, G. A. & Clegg, M. T. (1995). Evolution of the chalcone synthase gene family in the genus *Ipomoea*. *Proc Natl Acad Sci USA* **92**, 3338–3342.
- Fytlovitch, S., Nathan, P. D., Zafri, D. & Rosenberg, E. (1982). Amino acid precursors of *Myxococcus xanthus* antibiotic TA. *J Antibiot* **35**, 1525–1530.
- Hopwood, D. A. & Sherman, D. H. (1990). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu Rev Genet* **24**, 37–66.
- Kashefi, A. & Hartzell, P. L. (1995). Genetic suppression and phenotypic masking of a *Myxococcus xanthus* *frzF*-defect. *Mol Microbiol* **15**, 483–494.
- Kattar-Cooley, P. A., Wang, H. H., Mende-Mueller, L. M. & Miziorko, H. M. (1990). Avian liver 3-hydroxy-3-methylglutaryl-

- CoA synthase: distinct genes encode the cholesterogenic and ketogenic isozymes. *Arch Biochem Biophys* **283**, 523–529.
- Kroos, L., Kuspa, A. & Kaiser, D. (1986). A global analysis of developmentally regulated genes in *Myxococcus xanthus*. *Dev Biol* **117**, 252–266.
- Malpartida, F., Hallam, S. E., Kieser, H. M. & 9 other authors (1987). Homology between *Streptomyces* genes coding for synthesis of different polyketides used to clone antibiotic biosynthesis genes. *Nature* **325**, 818–821.
- Manor, A., Eli, I., Varon, M., Judes, H. & Rosenberg, E. (1989). Effect of adhesive antibiotic TA on plaque and gingivitis in man. *J Clin Periodontol* **16**, 621–624.
- Norrander, J., Kempe, T. & Messing, J. (1983). Construction of improved M13 vector. *Gene* **26**, 101–106.
- Paitan, Y., Boulton, N., Ron, E. Z., Rosenberg, E. & Orr, E. (1998). Molecular analysis of the *gyrB* gene from *Myxococcus xanthus*. *Microbiology* **144**, 1641–1647.
- Paitan, Y., Alon, G., Orr, E., Ron, E. Z. & Rosenberg, E. (1999a). The first gene in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus* codes for a unique PKS module coupled to a peptide synthetase. *J Mol Biol* **286**, 465–474.
- Paitan, Y., Orr, E., Ron, E. Z. & Rosenberg, E. (1999b). Cloning and characterization of a *Myxococcus xanthus* cytochrome P-450 hydroxylase required for biosynthesis of the polyketide antibiotic TA. *Gene* **228**, 147–153.
- Paitan, Y., Orr, E., Ron, E. Z. & Rosenberg, E. (1999c). A NusG-like transcription anti-terminator is involved in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus*. *FEMS Microbiol Lett* **170**, 221–227.
- Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* **183**, 63–98.
- Reichenbach, H. & Hofle, G. (1993). Production of bioactive secondary metabolites. In *Myxobacteria II*, pp. 347–398. Edited by M. Dworkin & D. Kaiser. Washington, DC: American Society for Microbiology.
- Ritsema, T., Gehring, A. M., Stuitje, A. R. & 7 other authors (1998). Functional analysis of an interspecies chimera of acyl carrier proteins indicate a specialized domain for protein recognition. *Mol Gen Genet* **257**, 641–648.
- Rosenberg, E. & Varon, M. (1984). Antibiotics and lytic enzymes. In *Myxobacteria, Development and Cell Interaction*, pp. 109–125. Edited by E. Rosenberg. New York: Springer.
- Rosenberg, E., Vaks, B. & Zuckerberg, A. (1973). Bactericidal action of an antibiotic produced by *Myxococcus xanthus*. *Antimicrob Agents Chemother* **4**, 507–513.
- Rosenberg, E., Fytlovitch, S., Carmeli, S. & Kashman, Y. (1982). Chemical properties of *Myxococcus xanthus* antibiotic TA. *J Antibiot* **35**, 788–793.
- Rosenberg, E., Porter, J. M., Nathan, P. N., Manor, A. & Varon, M. (1984). Antibiotic TA: an adherent antibiotic. *Bio/Technology* **2**, 796–799.
- Russ, A. P., Ruzicka, V., Maerz, W., Appelhans, H. & Gross, W. (1992). Amplification and direct sequencing of a cDNA encoding human cytosolic 3-hydroxy-3-methylglutaryl-coenzyme A synthase. *Biochim Biophys Acta* **1132**, 329–31.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Schupp, T., Toupet, C., Cluzel, B., Neff, S., Hill, S., Back, J. J. & Ligon, J. M. (1995). A *Sorangium cellulosum* (*Myxobacterium*) gene cluster for the biosynthesis of the macrolide antibiotic Soraphen A: cloning, characterization, and homology to polyketide synthase genes from Actinomycetes. *J Bacteriol* **177**, 3673–3679.
- Scotti, C., Piatti, M., Cuzzoni, A., Perani, P., Tognoni, A., Grandi, G., Galizzi, A. & Albertini, A. M. (1993). A *Bacillus subtilis* large ORF coding for a polypeptide highly similar to polyketide synthases. *Gene* **130**, 65–71.
- Sedgwick, S. G. & Morgan, B. A. (1994). Locating, DNA sequencing and disrupting yeast genes using tagged Tn1000. *Methods Mol Genet* **3**, 131–140.
- Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M. & Hopwood, D. A. (1989). Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* Tü22. *EMBO J* **8**, 2717–2725.
- Shimkets, L. J. (1990). Social and developmental biology of myxobacteria. *Microbiol Rev* **54**, 473–501.
- Tolchinsky, S., Fuchs, N., Varon, M. & Rosenberg, E. (1992). Use of Tn5lac to study expression of genes required for antibiotic TA production. *Antimicrob Agents Chemother* **36**, 2322–2327.
- Trowitzsch, W., Wray, V., Gerth, K. & Hofle, G. (1982). Structure of Myxovirescin A, a new macrocyclic antibiotic from gliding bacteria. *J Chem Soc Commun* **1982**, 1340–1342.
- Tsay, J. T., Oh, W., Larson, T. J., Jackowski, S. & Rock, C. O. (1992). Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. *J Biol Chem* **267**, 6807–6814.
- Varon, M. & Rosenberg, E. (1996). Transcriptional regulation of genes required for antibiotic TA synthesis in *Myxococcus xanthus*. *FEMS Microbiol Lett* **136**, 203–208.
- Varon, M., Fuchs, N., Monosov, M., Tolchinsky, S. & Rosenberg, E. (1992). Mutation and mapping of genes involved in antibiotic TA production in *Myxococcus xanthus*. *Antimicrob Agents Chemother* **36**, 2316–2321.
- Varon, M., Paitan, Y. & Rosenberg, E. (1997). Trans-acting regulation of antibiotic TA genes in *Myxococcus xanthus*. *FEMS Microbiology Lett* **155**, 141–146.
- Wang, S. L. & Liu, X. Q. (1991). The plastid genome of *Cryptomonas phi* encodes an hsp70-like protein, a histone-like protein, and an acyl carrier protein. *Proc Natl Acad Sci USA* **88**, 10783–10787.
- Zafirri, D., Rosenberg, E. & Mirelman, D. (1981). Mode of action of *Myxococcus xanthus* antibiotic TA. *Antimicrob Agents Chemother* **19**, 349–351.

Received 16 February 1999; revised 10 May 1999; accepted 9 July 1999.