The granaticin biosynthetic gene cluster of *Streptomyces violaceoruber* Tü22: sequence analysis and expression in a heterologous host

Koji Ichinose[1,2], David J Bedford[1], Diethild Tornus[3]*, Andreas Bechthold[3†], Maureen J Bibb[1], W Peter Revill[1], Heinz G Floss[3] and David A Hopwood[1]

Introduction: The granaticins are members of the benzoisochromanequinone class of aromatic polyketides, the best known member of which is actinorhodin made by *Streptomyces coelicolor* A3(2). Genetic analysis of this class of compounds has played a major role in the development of hypotheses about the way in which aromatic polyketide synthases (PKSs) control product structure. Although the granaticin nascent polyketide is identical to that of actinorhodin, post-PKS steps involve different pyran-ring stereochemistry and glycosylation. Comparison of the complete gene clusters for the two metabolites is therefore of great interest.

Results: The entire granaticin gene cluster (the *gra* cluster) from *Streptomyces violaceoruber* Tü22 was cloned on either of two overlapping cosmids and expressed in the heterologous host, *Streptomyces coelicolor* A3(2), strain CH999. Chemical analysis of the recombinant strains demonstrated production of granaticin, granaticin B, dihydrogranaticin and dihydrogranaticin B, which are the four known metabolites of *S. violaceoruber*. Analysis of the complete 39,250 base pair sequence of the insert of one of the cosmids, pOJ466-22-24, revealed 37 complete open reading frames (ORFs), 15 of which resemble ORFs from the *act* (actinorhodin) gene cluster of *S. coelicolor* A3(2). Among the rest, nine resemble ORFs potentially involved in deoxysugar metabolism from *Streptomyces* spp. and other bacteria, and six resemble regulatory ORFs.

Conclusions: On the basis of these resemblances, putative functional assignments of the products of most of the newly discovered ORFs were made, including those of genes involved in the PKS and tailoring steps in the biosynthesis of the granaticin aglycone, steps in the deoxy sugar pathway, and putative regulatory and export functions.

Addresses: [1]Department of Genetics, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK. [2]Department of Natural Product Chemistry, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [3]Department of Chemistry, University of Washington, Seattle, WA 98195, USA.

Present addresses: *Institute of Biochemistry, School of Medicine, University of Leipzig, D-04103 Leipzig, Germany. †Institute of Pharmaceutical Biology, Eberhard-Karls-University, D-72076 Tübingen, Germany.

Correspondence: David A Hopwood
E-mail: hopwood@bbsrc.ac.uk

Key words: antibiotic biosynthesis, benzoisochromanequinones, deoxysugar biosynthesis, pathway-specific activator gene, polyketide synthase

Received: **4 August 1998**
Revisions requested: **24 August 1998**
Revisions received: **4 September 1998**
Accepted: **14 September 1998**

Published: **20 October 1998**

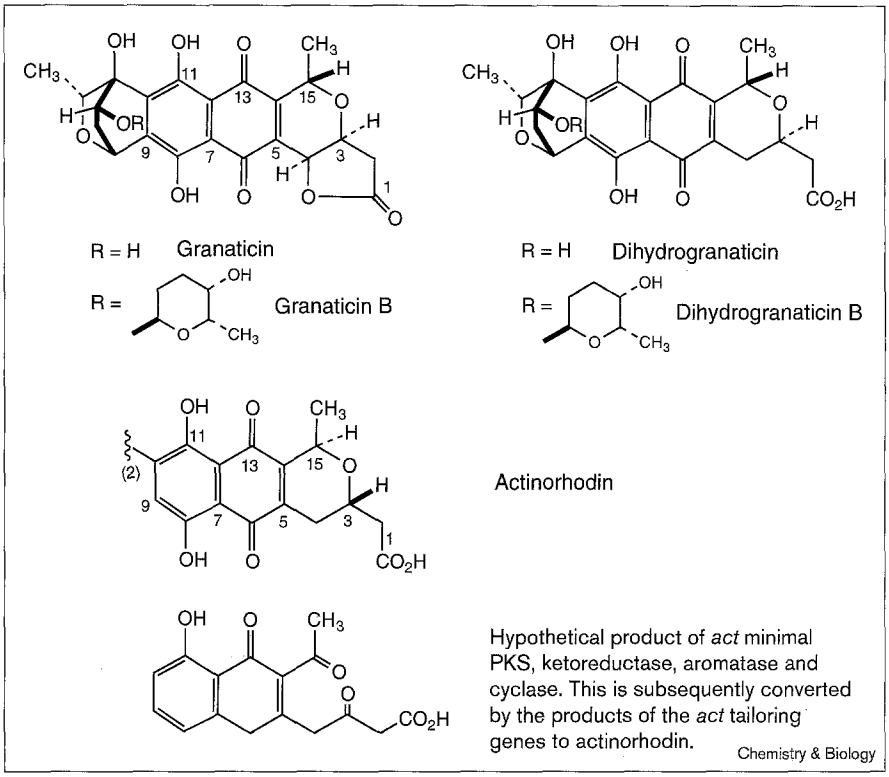
Chemistry & Biology November 1998, **5**:647–659
http://biomednet.com/elecref/1074552100500647

Introduction

Granaticin and the related metabolites dihydrogranaticin, granaticin B and dihydrogranaticin B (Figure 1), made by *Streptomyces violaceoruber* Tü22, are members of a class of *Streptomyces* aromatic polyketide antibiotics known as the benzoisochromanequinones (BIQs), the best known member being actinorhodin (Figure 1) produced by *Streptomyces coelicolor* A3(2). Extensive genetic analysis of actinorhodin biosynthesis revealed a cluster of 22 structural, resistance and regulatory genes (the *act* genes) for antibiotic biosynthesis and export [1–6]. Less extensive genetic studies have been made on other BIQ antibiotics, including frenolicin [7], griseusin [8] and kalafungin [9].

The carbocyclic skeleton of the BIQs is assembled by successive decarboxylative condensations of malonyl residues to give a polyketide carbon chain, which is subsequently reduced and cyclised. These early steps in BIQ biosynthesis are catalysed by the polyketide synthase (PKS) and the corresponding PKS subunit genes are highly conserved throughout the BIQ clusters. A primary stimulus to the study of BIQ genetics has been use of the PKS genes to understand the 'programming' rules that control carbon chain length and the patterns of ketoreduction and cyclisation of such aromatic polyketides. These studies (e.g. [10,11]) led to conclusions about the roles of the different PKS subunits — ketosynthase, chain length factor, acyl carrier protein, ketoreductase, aromatase and cyclase — in the programming choices, and to application of these programming 'rules' to synthesise novel 'unnatural natural products' to order [11]. Isolation and sequencing of the granaticin PKS genes [12] played a significant role in this research [13–15].

Figure 1



Structures of the granaticins, actinorhodin and a hypothetical intermediate in the biosynthesis of both antibiotics.

Control of the post-PKS ('tailoring') steps that give each BIQ its final structure and biological activity has been much less studied. Granaticin provides two particularly interesting dimensions to such research. One question concerns the stereochemistry of the pyran ring of the BIQ antibiotics. The two chiral centres (C-3 and C-15) are always *trans* in respect of the two hydrogen atoms in all natural BIQs, but can be either 3*R*, 15*S*, as in actinorhodin, or 3*S*, 15*R*, as in granaticin (Figure 1). Two putative reductases responsible for establishing the chirality at these two carbon atoms in actinorhodin have been suggested to be the products of two of the *act* genes [6]; what are the roles of the corresponding genes from the granaticin pathway? The second question concerns the 2,6-dideoxyhexose moiety attached to the BIQ chromophore via two carbon–carbon bonds at C-9 and C-10, as well as the second deoxy-sugar (a 2,3,6 trideoxyhexose, L-rhodinose) attached to the first sugar by a conventional glycosidic bond. Carbon origins and some stereochemical features of sugar biosynthesis and attachment were elucidated by feeding ^{13}C-labelled acetate and ^{14}C/^{3}H-labelled glucose [16], and two relevant genes were identified [17]; what further 'sugar genes' exist? Here, we describe the complete DNA sequence of a gene cluster that was proven, by expression in a heterologous host, to be sufficient for biosynthesis of the granaticin metabolites. On the basis of this experimental evidence, we could begin to answer these questions.

Results

Expression of the granaticin gene cluster in *S. coelicolor* A3(2)

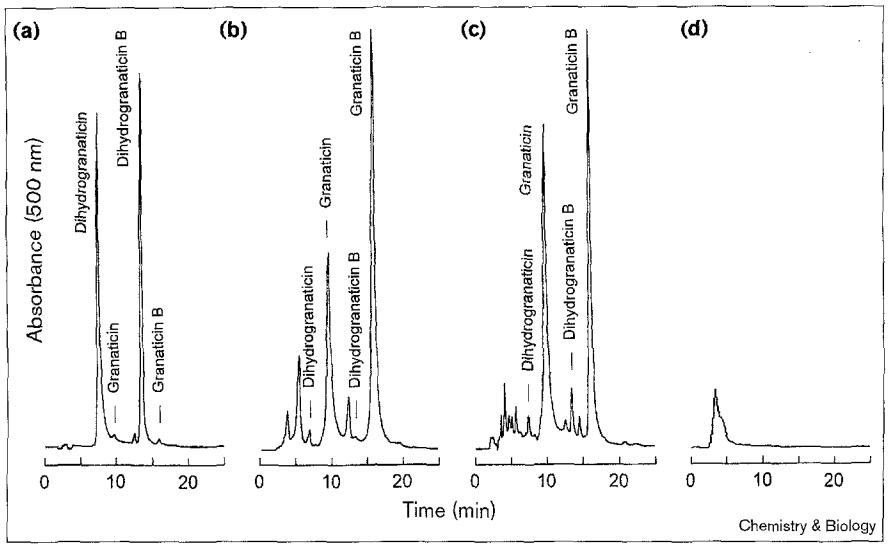
S. coelicolor CH999, which lacks the actinorhodin gene cluster, produces no blue or purple pigments. In contrast, transconjugants of CH999 containing cosmids pOJ446-22-24 or pOJ446-22-16, which carry overlapping inserts of *S. violaceoruber* DNA [17], produced purple and purple-brown pigments, respectively. Analysis using thin layer chromatography (TLC) identified the pigments as dihydrogranaticin, dihydrogranaticin B, and the corresponding lactone forms, granaticin and granaticin B, the four products of *S. violaceoruber* Tü22. High performance liquid chromatography (HPLC) analysis of crude extracts (Figure 2) showed that CH999/pOJ466-22-24 yielded dihydrogranaticin as ~50% of total pigments, with 45% of dihydrogranaticin B and only small quantities of granaticin and granaticin B. In contrast, pOJ466-22-16 yielded ~28% granaticin and 52% granaticin B, with smaller quantities of the dihydro forms; these proportions were similar to those produced by *S. violaceoruber* Tü22 itself (38% granaticin and 42% granaticin B). All of the genetic information required for *S. coelicolor* CH999 to produce the four end products of the granaticin pathway therefore lies on the overlap between pOJ466-22-24 and pOJ466-22-16.

Sequence analysis of cosmid pOJ446-22-24

The complete sequence of the insert in cosmid pOJ446-22-24 (39,250 base pairs) was compiled using the sequence

Figure 2

HPLC traces of granaticins produced by *S. coelicolor* A3(2) strain CH999 carrying **(a)** pOJ446-22-24 or **(b)** pOJ446-22-16. For comparison, **(c)** *S. violaceoruber* Tü22 (the original granaticin producer) and **(d)** CH999/pOJ446 (vector only) are shown. HPLC is described in the Materials and methods section. The retention times are as follows: dihydrogranaticin, 7.25 min; granaticin, 9.58 min; dihydrogranaticin B, 13.26 min; granaticin B, 15.87 min.



of the PKS region (B13 and parts of B8 and B5; Figure 3) and a segment (B32 and B24) containing, among others, two genes involved in early steps in the 2,6-dideoxyhexose pathway, both reported previously [12,17] (a few parts of these sequences were amended), together with the new sequence for the rest of the cosmid insert. Probable protein-coding regions were identified by the FRAME program [18], which makes use of a strong bias towards a G or C in the third position of *Streptomyces* codons, by identification of plausible ribosome-binding sites [19], and by alignment of deduced protein sequences with homologous sequences in the databases. The deduced open reading frames (ORFs) were designated as shown in Figure 3, retaining ORFs 1–6 for the PKS genes [12] and numbering the others consecutively from left to right. We identified 37 putative complete ORFs (Table 1) and a truncated ORF at either end, together with a putative tRNA gene and several inverted repeats in intergenic regions.

Assignment of the PKS and cyclisation ORFs

As described previously [12], *gra* ORFs 1–5, encoding the minimal *gra* PKS (ketosynthase, chain length factor and acyl carrier protein), aromatase and ketoreductase, lie in the same order as in the *act* cluster: downstream of the *gra*-ORF5 ketoreductase gene is a second *act*III homologue (*gra*-ORF6), which might function together with ORF5 to control C-9 ketoreduction as suggested previously [12,13], or, alternatively, might be involved in a tailoring step (see the Discussion section). The final gene expected to be required for assembly of the BIQ carbon skeleton, namely the cyclase (the homologue of the *act*IV gene product) [20], which would catalyse aldol condensation between C-5 and C-14 to close the second ring after the first ring has been dehydrated by the aromatase, was identified in the present study and is clearly *gra*-ORF33.

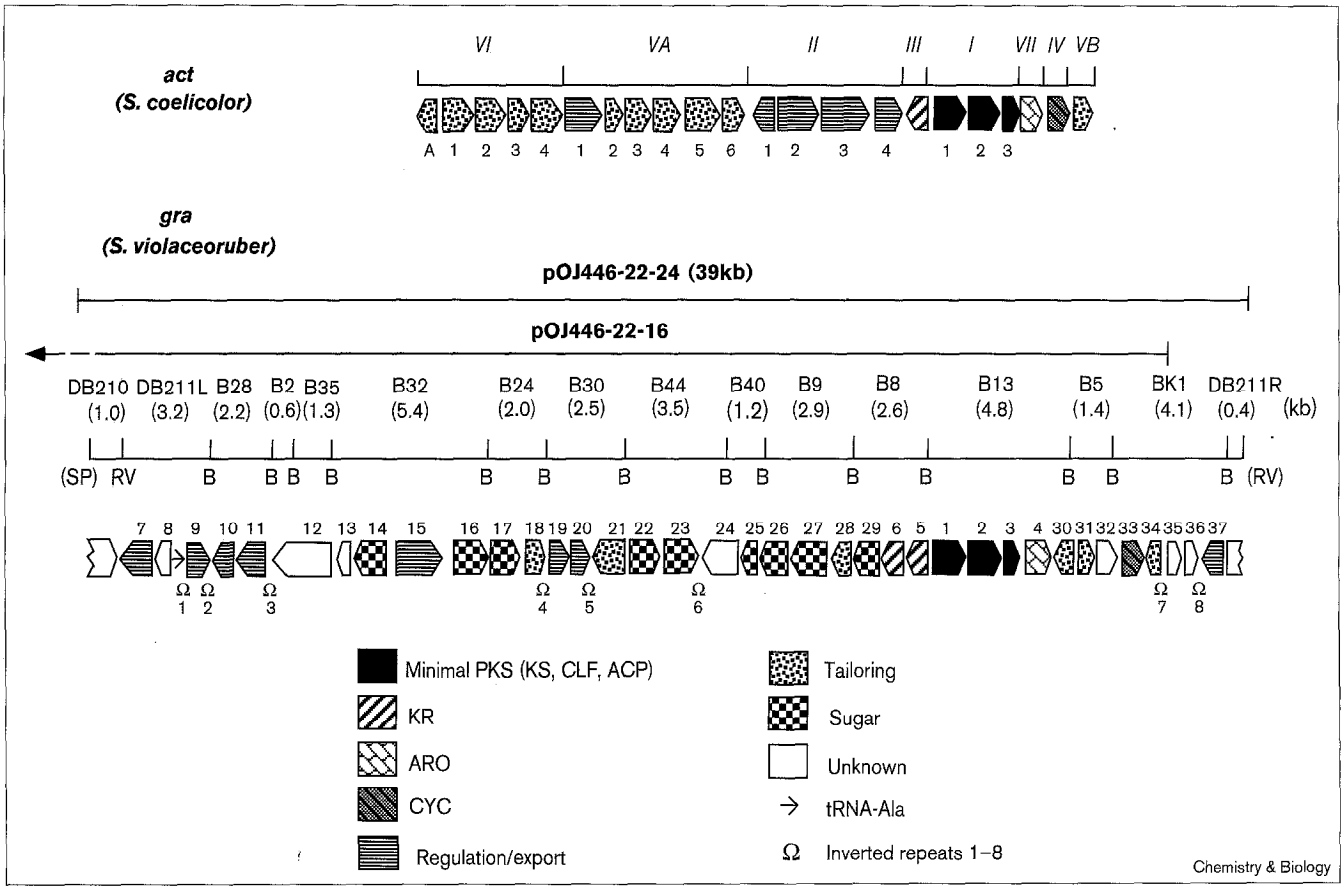
Assignment of ORFs for post-PKS tailoring steps in aglycone biosynthesis

Apart from the stereochemistry of the pyran ring, which is 3*S*, 15*R* for granaticin in contrast to 3*R*, 15*S* for actinorhodin, the structure of the granaticin aglycone is identical to that of the actinorhodin half-molecule. Homologues of many of the 'tailoring' genes identified previously in the *act* cluster of *S. coelicolor* (shown in Figure 3, top), whose products convert the biosynthetic intermediate (Figure 1) generated by the minimal PKS, ketoreductase, aromatase and cyclase to actinorhodin, might therefore be expected to occur in the *gra* cluster. From left to right in Figure 3, the proposed functions of the *act* tailoring genes are as follows.

*act*VI-ORFA is a member of a family of genes found in the gene clusters for *Streptomyces* aromatic polyketides, the founder member being *fren*-ORFX from the frenolicin cluster of *S. roseofulvus* [7] and including *mtmX* from the mithramycin cluster of *S. argillaceus* [21], and *dpsH* from the daunorubicin clusters of *S. peucetius* [22] and *Streptomyces* sp. strain C5 [23]. The function of *act*VI-ORFA is unclear: its disruption reduced, but did not abolish, pigment production [6]. It has recently been suggested that *mtmX* and *dpsH* might encode polyketide cyclases [21,22], but without firm evidence. *gra*-ORF31 is clearly the homologue of *act*VI-ORFA.

The other *act*VI genes — ORFs 1,2,3 and 4 — are candidates for controlling pyran-ring formation [6]: *act*VI-ORF1 and *act*VI-ORF2 were postulated to encode reductases for C-3 and C-15, respectively, *act*VI-ORF3 (disruption of which caused only a delayed blue pigmentation upon ammonia fuming) might encode a dehydratase that assists pyran-ring formation, whereas *act*VI-ORF4 appeared to be a 'silent' homologue of *act*VI-ORF2. In the *gra* cluster,

Figure 3



Organisation of the *gra* cluster of *S. violaceoruber* Tü22 and for comparison, the *act* cluster of *S. coelicolor* A3(2). Restriction sites: B, *Bam*HI; SP, *Spe*I; RV, *Eco*RV. The sites in parentheses are derived from cosmid pOJ446-22-24, and are not present on the chromosome of *S. violaceoruber* Tü22. KS, ketosynthase; CLF, chain length factor; ACP, acyl carrier protein; KR, ketoreductase; ARO, aromatase; CYC, cyclase.

ORF18 is a clear homologue of *act*VI-ORF3, but no homologues of the key genes *act*VI-ORF1 and *act*VI-ORF2/4 are seen, suggesting that the C-3 and C-15 reductions in granaticin biosynthesis are catalysed by enzymes that do not resemble, in sequence, the corresponding actinorhodin enzymes (see the Discussion section).

At least some of the tailoring genes in the *act*VA region have been implicated in introducing oxygen at C-6 and C-8 of the BIQ chromophore. *act*VA-ORF6 was postulated, by homology with *tcmH* in the tetracenomycin cluster of *S. glaucescens* that encodes a monooxygenase [24] and studies of the expressed *act*VA-ORF6 protein [25], to be responsible for oxidation at C-6 in actinorhodin biosynthesis [6]. Evidence that at least some of the other *act*VA tailoring ORFs are involved in C-8 hydroxylation includes the following: 'classical' *act*VA mutants, which accumulate an unhydroxylated intermediate [26], were mapped mostly in *act*VA-ORF5, with some probably in *act*VA-ORF3 and *act*VA-ORF4 [2]; recombinants carrying the *act*VA region converted medermycin (lacking C-8 hydroxylation) to the C-8 hydroxylated mederrhodin [27]; and the *act*VA-ORF5 product resembles a known phenol hydroxylase (Accession number U17960). A clear homologue of *act*VA-ORF5 was found in the *gra* cluster (*gra*-ORF21), but, surprisingly, there was no *act*VA-ORF6 homologue. In addition, both *gra*-ORF28 and *gra*-ORF30 resemble *act*VA-ORF3 to a similar degree, but no homologues of *act*VA-ORF2 or *act*VA-ORF4 were seen. This suggests that oxidation/hydroxylation reactions in granaticin biosynthesis might differ significantly from those in actinorhodin biosynthesis (see the Discussion section).

The final *act* tailoring gene, *act*VB, encodes a flavin mononucleotide (FMN) : NADH oxidoreductase [28], which is likely to be involved either in hydroxylation at C-8 (a blocked mutant accumulates kalafungin, with C-6 but not C-8 hydroxylation [26]), or in dimerisation via C-10 (kalafungin is a monomeric BIQ). *gra*-ORF34 is a clear homologue of *act*VB, at first sight perhaps favouring the idea that its role is in C-8 hydroxylation, because the

Table 1

Deduced functions of the open reading frames.

ORF	Putative functional category	Gene	Deduced role	Homologue SM/ID (%) of product	Origin	Reference	Nucleotide accession number
7	Regulation	*pkaA*	Serine-threonine protein kinase	45/38	*S. coelicolor*	[46]	D86821
8	Unknown	–	–	–	–	–	
9	Regulation	*actII*-4	Pathway-specific transcriptional activator	46/40	*S. coelicolor*	[4]	M64683
10	Regulation	*degU*	Response regulator of two-component system	57/34	*B. subtilis*	[45]	M23558
11	Regulation	*degS*	Sensor kinase of two-component system	45/19	*B. subtilis*	[45]	M23558
12	Unknown	–	–	–	–	–	
13	Unknown	–	–	–	–	–	
14	Sugar	*dnrS*	Glycosyl transferase	52/26	*S. peucetius*	[34]	L47164
15	Export	*actII*-2	Transmembrane protein	57/30	*S. coelicolor*	[4]	M64683
16	Sugar	*strD*	dTDP-1-glucose synthase	70/51	*S. griseus*	[29]	Y00459
17	Sugar	*strE*	dTDP-glucose-4,6-dehydratase	76/61	*S. griseus*	[30]	X62567
18	Tailoring	*actVI*-3	Cyclase-dehydratase ?	52/36	*S. coelicolor*	[6]	X62373
19	Regulation ?	–	Disulphide bond-forming protein ?	–	–	–	
20	Regulation	*soxR*	Transcriptional activator in redox control	72/56	*E. coli*	[42]	M60111
21	Tailoring	*actVA*-5	Hydroxylase	69/53	*S. coelicolor*	[2]	X58833
22	Sugar	*dnmV*	dTDP-4-keto-6-deoxyhexose reductase	51/44	*S. peucetius*	[36]	AF006633
23	Sugar	*rfbH (ascC)*	CDP-4-keto-6-deoxyglucose-3-dehydratase (E1)	71/53	*Y. pseudotuberculosis*	[32]	L01777
24	Unknown	–	–	–	–	–	
25	Sugar	*strM*	dTDP-4-keto-6-deoxyglucose-3, 5-epimerase in streptomycin biosynthesis	52/46	*S. griseus*	[30]	X62567
26	Sugar	*rdmF*	Rhodomycin biosynthesis	65/47	*S. purpurescens*	[37]	U10405
27	Sugar	*dnmT*	dTDP-4-keto-6-deoxyglucose-2, 3-dehydratase	53/46	*S. peucetius*	[39]	U77891
28	Tailoring ?	*actVA*-3	Unknown	46/41	*S. coelicolor*	[2]	X58833
29	Sugar ?	*lmbY*	FMN-dependent monooxygenase in lincomycin biosynthesis ?	48/29	*S. lincolnensis*	[40]	X79146
6	PKS ?	*actIII*	Keto reductase (KR)	47/42	*S. coelicolor*	[3]	M19536
5	PKS	*actIII*	Keto reductase for C-9 (KR)	75/71	*S. coelicolor*	[3]	M19536
1	PKS	*actI*-1	Keto-acyl synthase (KS)	82/76	*S. coelicolor*	[5]	X63449
2	PKS	*actI*-2	Chain length factor (CLF)	68/61	*S. coelicolor*	[5]	X63449
3	PKS	*actI*-3	Acyl carrier protein (ACP)	76/66	*S. coelicolor*	[5]	X63449
4	PKS	*actVII*	First ring aromatisation (ARO)	58/53	*S. coelicolor*	[5]	X63449
30	Tailoring ?	*actVA*-3	Unknown	47/42	*S. coelicolor*	[2]	X58833
31	Tailoring ?	*actVI*-A	Unknown	64/58	*S. coelicolor*	[6]	X62373
32	Unknown	–	–	–	–	–	
33	Tailoring	*actIV*	Second ring cyclisation (CYC)	71/55	*S. coelicolor*	[5]	X63449
34	Tailoring	*actVB*	FMN : NADH oxidoreductase	56/34	*S. coelicolor*	[5]	X63449
35	Unknown	–	–	–	–	–	
36	Unknown	–	–	–	–	–	
37	Regulation	*nshA*	Transcriptional activator	82/70	*S. actuosus*	[47]	U75434

? indicates a very tentative assignment of function. *Y. pseudotuberculosis, Yersinia pseudotuberculosis. B. subtilis, Bacillus subtilis.* SM/ID, % similarity/identity of amino-acid sequences.

granaticin chromophore is not a dimer. It could be that sugar attachment in granaticin biosynthesis requires a similar redox change at C-10 as dimerisation in the biosynthesis of making actinorhodin, however (see the Discussion section).

Assignment of sugar ORFs

Bechthold *et al.* [17] had clearly recognised two *gra* ORFs — now renamed ORF16 and ORF17 — as homologues of the *S. griseus strD* [29] and *strE* [30], respectively. The products of these streptomycin biosynthetic genes [31], dTDP-1-glucose synthase (*strD*) and dTDP-glucose-4,6-dehydratase (*strE*), catalyse the first two steps in deoxyhexose biosynthesis from glucose-1-phosphate. Expression of *gra*-ORF16 in *Escherichia coli* [17] confirmed production of dTDP-4-keto-6-deoxyglucose, the precursor of the two deoxysugars (dTDP-4-keto-2,6-dideoxyglucose and L-rhodinose) that are components of granaticin and its related metabolites (Figure 1).

Database comparisons showed that *gra*-ORF25 would encode a similar protein to the product of *strM* [30] and *rmlC* (Accession number U09876), dTDP-4-keto-6-deoxyglucose-3,5-epimerase.

gra-ORF23 is a homologue of *ascC* (*rfbH*) [32], which encodes the key enzyme CDP-4-keto-6-deoxyglucose-3-dehydratase (E_1) in the 3,6-dideoxyhexose pathway in *Yersinia pseudotuberculosis*. E_1 is a typical pyridoxamine 5′-phosphate (PMP)-dependent enzyme, with a conserved histidine residue (His220) at the active site, which abstracts a C-4′ proton from a PMP-substrate adduct [33]. Another feature of E_1 is a series of cysteine residues commonly found in iron–sulphur-containing enzymes. The ORF23 protein has both features, as well as high end-to-end homology with the *ascC* product, suggesting that the ORF23 product depends on PMP, unlike the other *ascC* (*rfbH*) homologues found in streptomycetes so far, including *dnrJ* (Accession number M80237 from *S. peucetius*, and others [33], which have a conserved lysine residue instead of a histidine residue, indicating their dependence on pyridoxal 5′-phosphate (PLP).

gra-ORF14, reported previously [17], was extended by 29 amino acids at the carboxyl terminus after making some sequence corrections. It resembles sugar transferases from bacteria, including DnrS [34] from *S. peucetius* and others (Accession numbers M74717, U40458, Z22577 and A25110). The homologous regions [34] among these proteins [35] are in the amino terminus (^{2}HIXMXXIAXXG*H*VNPXXXXXRX*L*XAR*GH*R*V*XXA-XXPXXXDXVXXXG^{47}; residues in italics are highly conserved; X indicates a variable residue; residue numbers are those in DnrS; single-letter amino-acid code is used) and carboxyl terminus (^{296}LXXLPXNVXVXXWXPXXA-I*L*XXXXXF*V*XHGAXXXXXLAXXX*P*XIAXPXAX*D*QF-XNADXLXXL*G*XXXXL^{365}). The *gra*-ORF14 product (343 amino acids) showed 32–46% identity with these proteins, with a substantial number of gaps in the alignment. The conserved residues were reasonably matched for the carboxyl terminus, but not for the amino terminus. There is another ATG codon a further 117 base pairs upstream from the putative *gra*-ORF14 start codon (ATG), which would add an extra 39 amino acids to the product (382 amino acids in total). The extended region includes the sequence ^{26}NA*GHEV*IVGA^{35}, which is partially identical to the amino-terminal motifs of the other members of the family, suggesting that the upstream ATG is the true start codon.

The products of *gra*-ORFs 22 and 26 resemble putative biosynthetic gene products involved in the biosynthesis of anthracyclines, including *dnmV* in *S. peucetius* [36], *rdmF* in *S. purpurescens* [37], and a gene from *S. griseus* (Accession number X73148); these genes are assumed to be involved in deoxyhexose biosynthetic pathways. Nucleotide-binding motifs [37,38] were found for both products (ORF22: ^{11}LVL*GGSGFVGRHVCAAFLARGWEV*^{34}; ORF26: ^{12}LR-I*GVAGCADIALRRMLPAFAASP*^{35}; residues in italics are highly conserved). Although the motif from the ORF26 product is less conserved, it also resembles glucose–fructose oxidoreductase from *Zymomonas mobilis* (Accession number M97379). The *gra*-ORF22 and ORF26 proteins are therefore possible oxidoreductases involved in deoxyhexose biosynthesis.

A high degree of similarity (50–53%) was found for the *gra*-ORF27 product with the genes most probably involved in deoxysugar formation for several anthraquinone or macrolide antibiotic biosyntheses. The homologues include *dnmT* from the daunorubicin biosynthetic gene cluster in *S. peucetius* [39], a similar gene from the daunomycin biosynthetic cluster in *Streptomyces* sp. strain C5 (Accession number U43704), *eryBVI* from the erythromycin cluster in *S. erythraea* (Accession numbers U77459, Y11199), *snoH* from the nogalamycin cluster in *S. nogalater* (Accession number AJ224512), and a partially sequenced gene (ORF6) in the tylosin gene cluster in *S. fradiae* (Accession number U08223). Because the final structures of the deoxysugar moieties are different for the different antibiotics, we conclude that the *gra*-ORF27 gene product probably functions at an early step, most likely as a dTDP-4-keto-6-deoxyglucose-2,3-dehydratase (see the Discussion section).

The product of *gra*-ORF29 resembles that of *lmbY* in the lincomycin gene cluster of *Streptomyces lincolnensis*. The *lmbY* product was reported [40] to resemble (21% similarity) FMN-dependent α subunits of alkaline monooxygenase encoded by *luxA* from *Xenorhabdus luminescens* (Accession number M62917), but the possible function of this gene remains unknown.

Assignment of ORFs putatively involved in regulation or antibiotic export

gra-ORF15, described previously [17], is a homologue of *act*II-ORF2, whose product is a putative transmembrane protein [4] implicated, by gene disruption studies, in actinorhodin export and possibly coupled to the conversion of actinorhodin to the lactone form, γ-actinorhodin [1].

Of the newly discovered ORFs, *gra*-ORF9 is a homologue of *act*II-ORF4 [4], the founder member of a family of pathway-specific transcriptional activator genes for antibiotic biosynthetic pathways [41]. The amino-terminal regions of this class of proteins (named SARPs, for *Streptomyces* antibiotic regulatory proteins) contain amino-acid motifs that resemble the DNA-binding fold of the OmpR class of regulatory proteins. These motifs are well conserved in the *gra*-ORF9 product. The SARPs would interact with specific sequences upstream of *Streptomyces* antibiotic biosynthetic genes to ensure that the RNA polymerase engages with the promoter to initiate transcription. Wietzorrek and Bibb [41] also described some characteristic direct repeat sequences in the promoter regions of *Streptomyces* antibiotic biosynthetic genes that are potential binding sites for the SARPs. A search for such sequences in the *gra* cluster revealed a tandem array of trimeric

repeats with the consensus 5′-CNA (where N is any nucleotide), separated by eight nucleotides. This spacing, corresponding to one turn of the DNA helix, places each repeat on the same face of the DNA, perhaps to allow co-operative binding of several SARP molecules. In all the examples in which a putative –10 promoter sequence could be recognised (i.e., excluding ORF8), the promoter-proximal repeat would be on the opposite side of the DNA helix from the conserved T in the –10 region (typically 17 nucleotides upstream of it), allowing simultaneous binding of the SARPs and the RNA polymerase.

The product of *gra*-ORF20 resembles SoxR from *E. coli* [42] which operates, together with the product of the divergently transcribed *soxS*, to bind to specific highly conserved promoter sequences and activate them in response to oxidative stress. The product of the gene immediately upstream of *gra*-ORF20, *gra*-ORF19 transcribed in the same direction, does not resemble SoxS, but contains the motif ^{11}CPWCY^{16} (single-letter amino-acid code), which could be regarded as a CXXC motif (where X is any amino acid), found in members of the glutaredoxin/thioredoxin superfamily of redox sensors [43]. A search of the *gra* sequence for possible SoxR-binding sites revealed two convincing candidates in the promoter regions of ORF19 and ORF15, the latter encoding the putative granaticin-resistance protein. The ORF19 gene product, perhaps together with that of ORF20, might therefore modulate resistance to granaticin in response to changes in the redox potential of the cells; a change in redox state might even be a direct consequence of granaticin biosynthesis.

The products of a further pair of putative regulatory genes, *gra*-ORF11 and *gra*-ORF10, resemble bacterial two-component sensory kinase/response regulator pairs [44]; their closest homologues are *degS* and *degU* from *Bacillus subtilis* [45]. These putative regulators probably play a role in regulating granaticin biosynthesis because they lie to the right of *gra*-ORF9, the very characteristic pathway-specific regulatory gene. On the other hand, *gra*-ORF7, whose product resembles that of *pkaA*, a serine–threonine protein kinase from *S. coelicolor* [46], could well lie outside the *gra* cluster. The same is likely to be true of *gra*-ORF37, the last complete ORF at the right-hand end of pOJ466-22-24, whose product is similar to that of the putative transcriptional activator gene *nshA* ('ORF699') from *S. actuosus* [47].

The 'unknown' ORFs

No meaningful similarities were found with comparisons of the deduced products of ORFs 8, 12, 13, 24, 32, 35 and 36 or the two incomplete ORFs at either end of the sequence.

Identification of a tRNA gene and inverted repeats

A convincing alanine tRNA-like sequence was found between ORFs 8 and 9. Eight obvious inverted repeat (IR) sequences were also identified, many of them in expected positions (see Figure 3). IR1 is located downstream of tRNA–Ala, and four others (IR2, IR5, IR6 and IR8) lie between the stop codons of convergently transcribed ORFs. Two further IRs (IR3 and IR4) lie within a group of ORFs with the same orientation.

Discussion

The *gra* PKS genes were originally cloned by homology with actinorhodin PKS gene probes [48]. Although *S. violaceoruber* Tü22 is recalcitrant to plasmid transformation (our unpublished observations) granaticin nonproducing gene disruptants were made using a φC31 phage vector [48], providing evidence that the cloned genes were indeed those for the granaticin PKS. These genes were later sequenced [12]. Finding genes apparently involved in deoxyhexose biosynthesis within ~7 kilobases of the *gra* PKS genes [17] fitted this assignment. The present results, in which heterologous expression of the putative *gra* genes on either of the cosmids gave rise to the granaticin-related metabolites characteristic of *S. violaceoruber* Tü22, dispel any possible doubts about the correct identification of the *gra* gene cluster.

What are the likely limits of the *gra* cluster? pOJ446-22-16 induced biosynthesis of all four of the granaticin metabolites in similar proportions to those in *S. violaceoruber*, implying that genes lying to the right of the end point of the insert in this cosmid (ORFs 35–37) are outside the *gra* cluster. Sequencing of the right-hand terminus of the insert showed it to end just inside ORF34 (truncating it by 11 amino-acid residues at its amino terminus). Because the ORF34 product resembles that of an essential gene in actinorhodin biosynthesis (*act*VB, encoding an FMN : NADH oxidoreductase) we presume that this ORF forms part of the *gra* cluster.

The first gene from the left end of pOJ466-22-24 that is recognisably a member of the *gra* cluster is ORF9, a member of the SARP family of pathway-specific activators for *Streptomyces* antibiotic gene clusters [41]. This implies that all the genes to the right of *gra*-ORF9, including the putative sensory kinase/response regulator pair encoded by ORFs 11 and 10, and the unassigned genes (ORFs 12 and 13) to their right, play some roles in granaticin biosynthesis, regulation or export. To the left of *gra*-ORF9 lie a $tRNA^{ALA}$ gene, an unassigned gene (ORF8) and a homologue of the serine–threonine protein kinase gene *pkaA* (ORF7). Of these, ORF8 is preceded by a convincing run of SARP-binding sites, implying a functional relationship with the granaticin biosynthetic pathway. There is no reason to believe that the *pkaA* homologue is also involved in the pathway. DNA to the left of the terminus of pOJ446-22-24 is evidently not needed for the production of the four granaticin-related metabolites in *S. coelicolor* A3(2), and so could be regarded as lying outside the *gra*

cluster. Because the recombinant carrying pOJ446-22-16 has a product profile (Figure 2b) more closely resembling that of *S. violaceoruber* Tü22 (Figure 2c) than that carrying pOJ446-22-24 (Figure 2a), however, a gene(s) encoding an enzyme(s) that might aid lactonisation of the dihydro forms into the corresponding granaticins (though not being essential for this) could conceivably lie to the left of the terminus of pOJ446-22-24 in an unsequenced segment of pOJ446-22-16.

The segment from ORF8 to ORF34 consists of 33 genes, in contrast to the 22 for the *act* cluster; this is reasonable because granaticin B contains two deoxysugars that are not part of the actinorhodin molecule, and, in fact, nine of the *gra* genes are candidates for involvement in deoxyhexose metabolism. Based on the direction of transcription of the 33 putative *gra* genes, the genes would be organised in a minimum of 11 transcripts, four of them monocistronic. If we assume that IR3 and IR4 also represent transcription termination points, the number of transcripts would rise to 13, and if, as expected, SARP-binding sites identify further start points for transcription, it would rise to 17, with eight of them monocistronic. Apart from the PKS genes, there is only limited clustering of functionally related genes on the different transcripts (less than in the *act* cluster).

In addition to the SARP gene (*gra*-ORF9) and *gra*-ORF15, the homologue of the actinorhodin export gene (*act*II-ORF2), 13 of the *gra* genes are likely homologues of *act* genes involved in chain building, cyclisation and tailoring steps in the actinorhodin pathway. The earlier finding of *gra* homologues (ORFs 1,2,3,5 and 4, respectively) for the *act* minimal PKS genes (*act*I-ORF1,2,3 for ketosynthase, chain length factor and acyl carrier protein), ketoreductase (*act*III) and aromatase (*act*VII), and of the cyclase (*act*IV) in the present study (ORF33), is completely consistent with the fact that the granaticin biosynthetic pathway would proceed via identical steps to that of actinorhodin as far as the putative biosynthetic intermediate in Figure 1. Although, even beyond formation of this dicyclic intermediate, biosynthesis of the actinorhodin and granaticin chromophores would be expected to proceed by similar pathways (except for the stereochemical course of the C-3 and C-15 reductions for actinorhodin catalysed by many of the *act*VI and *act*VA/B genes), sequence comparisons suggest more profound differences between the two pathways.

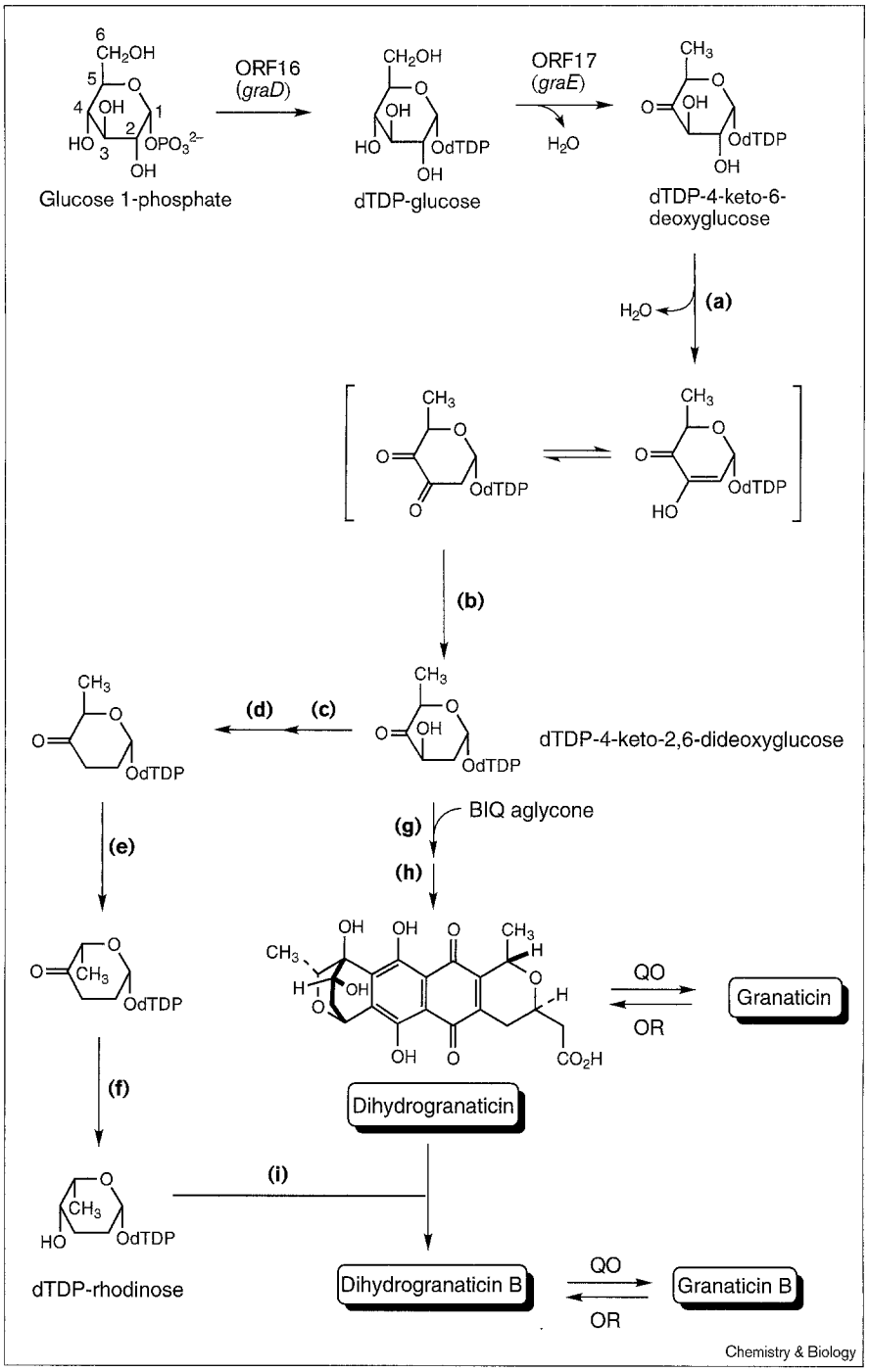
Of particular interest is the lack of *gra* homologues for *act*VI-ORF1 and *act*VI-ORF2/4, postulated to encode stereospecific reductases for C-3 and C-15 [6]. Although there is a precedent for catalytically related enzymes with opposite stereospecificities having similar sequences [49], there are also examples where the two proteins are apparently unrelated [50], and this is presumably the case for the *gra* gene products. We are now using functional tests to seek the equivalents of the enzymes encoded by *act*VI-ORF1 and *act*VI-ORF2/4 (but with opposite stereochemical outcome), focusing initially on unassigned putative gene products that either resemble ketoreductases (*gra*-ORF6), or carry potential NAD(P)H binding motifs (*gra*-ORF26). Oxygenation/hydroxylation at C-6 and C-8 could also be catalysed differently in the two pathways because, although the *gra* cluster contains homologues of *act*VA-ORF5 (*gra*-ORF21) and *act*VA- ORF3 (two homologues: *gra*-ORF28 and *gra*-ORF30), it lacks homologues of *act*VA-ORF2, *act*VA-ORF4 and *act*VA-ORF6. The *gra* cluster has enough unassigned ORFs to provide alternative proteins to catalyse this part of the biosynthetic pathway.

There is currently much interest in biosynthesis of the deoxysugars that decorate many actinomycete secondary metabolites and other products of bacteria such as the O-antigens of the gram-negative cell surface [51,52]. Bechthold *et al.* [17] assigned, plausibly, *gra*-ORF16 (*graD*) and *gra*-ORF17 (*graE*) as the genes for converting glucose-1-phosphate to the generally accepted precursor, dTDP-4-keto-6-deoxyglucose, but little experimental evidence is available for mechanistic aspects of deoxyhexose biosynthesis in granaticin and granaticin B beyond this point [16]. A plausible pathway leading to these sugars is shown in Figure 4 [52]. A suggested candidate gene for the 2,3-dehydration step (a) is ORF27, which has homologues in other 2,6-dideoxyhexose biosynthetic gene clusters. Two ketoreductions, steps (b) and (f), would be catalysed by the oxidoreductases encoded by ORFs 26 and 22. The ORF25 protein is likely to function as an epimerase at step (e), like the *strM* product postulated to epimerise the C-5 methyl group in streptomycin biosynthesis [34]. The discovery of an *ascC* (*rfbH*) homologue [36] in the *gra* cluster (ORF23) is interesting, because all the other *rfbH* homologues reported in streptomycetes [33] are involved in formation of aminosugars, whereas the deoxyhexoses of granaticin lack amino groups. CDP-4-keto-6-deoxyglucose-3-dehydratase (E_1), encoded by *ascC*, requires E_3 (CDP-6-deoxy-^{3,4}-glucoseen 3-reductase) as a catalytic partner in CDP-3,6-deoxyhexose biosynthesis [52]. The lack of an E_3 homologue in the deoxyhexose pathways involved in antibiotic production so far studied suggests the intriguing possibility [34,52] that an *ascC* homologue would encode an aminotransferase depending on coenzyme-B_6 (PLP). The high degree of similarity between the ORF23 product (including the 'PMP' motif and the iron–sulphur cluster motif [33]) and the *ascC* protein might imply that the deoxygenation at steps (c) and (d) would proceed in a similar way to that observed in the 3,6-dideoxyhexose pathway in *Y. pseudotuberculosis*.

Two glycosylation steps are needed for granaticin B formation: (g) + (h), and (i) in Figure 4. Of these, the attachment of rhodinose to granaticin to convert it to granaticin B (i) would be via a conventional glycosidic bond, whereas

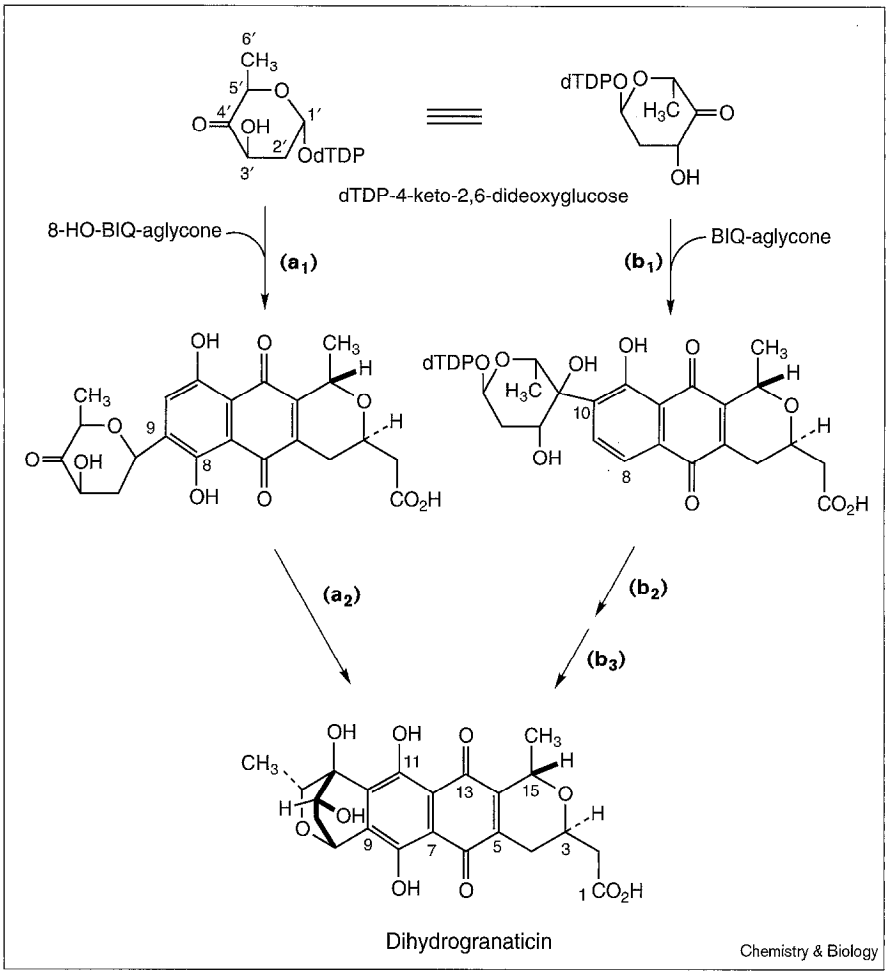
Figure 4

Proposed deoxyhexose pathway involved in the biosynthesis of granaticins. Probable enzymatically catalysed steps are: **(a)** dehydration, **(b)** ketoreduction, **(c)** and **(d)** C-3 deoxygenation, **(e)** epimerisation, **(f)** ketoreduction, **(g)** and **(i)** glycosidations, and **(h)** intramolecular cyclisation. QO (quinone oxidase) and OR (oxidoreductase) are involved in the interconversion of the 'dihydro-' and 'lactone' forms of the granaticins.



attachment of 4-keto-2,6-dideoxyglucose to the aglycone to make the granaticins (g + h) involves the formation of two successive carbon–carbon bonds. Only one clear candidate (*gra*-ORF14) for a glycosyltransferase was identified in the *gra* cluster. We therefore assign this gene to the transfer of rhodinose to granaticin, and postulate that transfer of 4-keto-2,6-dideoxyglucose to the aglycone proceeds by one of the routes shown in Figure 5 [16]. Route (a) involves C-glycosidation of an 8-hydroxylated BIQ aglycone (a_1), which would then undergo cyclisation by an intramolecular aldol condensation (a_2). This mechanism requires recognition of the total structural features of the aglycone by the C-glycosyltransferase in order to account for recognition of C-9 as the point of attachment. Route (b) would involve establishment of the glycosidic bond at C-10 by an intermolecular aldol condensation (b_1), followed by hydroxylation at C-8 (b_2), and then an intramolecular cyclisation (b_3) to establish the carbon–carbon bond at C-9. In

Figure 5



Two possible routes of C-glycosidation in the biosynthesis of granaticin [16] corresponding to steps (g) and (h) in Figure 5. Route a: **(a_1)** glycosyl transfer at C-9 of 8- hydroxy-BIQ; **(a_2)** intramolecular aldol condensation between C-10 and C-4′. Route b: **(b_1)** intermolecular aldol condensation between C-10 and C-4; **(b_2)** hydroxylation at C-8 of BIQ; **(b_3)** intramolecular cyclisation between C-9 and C-1′.

this hypothesis, the enzymes involved might well not have homologues in the databases and could be the products of some of the unassigned *gra* ORFs. Note also that the *act*VB homologue (*gra*-ORF34) could play a role in establishing the appropriate redox charge at C-10 (perhaps analogous to that in actinorhodin biosynthesis to facilitate dimerisation), and that there could well be associated differences in the mechanism of C-8 (and perhaps C-6) hydroxylation, accounting for the observed differences in the assignments of genes for these functions between the actinorhodin and granaticin clusters.

Significance

The granaticin (*gra*) biosynthetic gene cluster is only the second example of a complete cluster of genes for biosynthesis of the benzoisochromanequinone antibiotics, which have played such a significant part in the establishment of the concept of hybrid antibiotic production [27] and its extension into combinatorial biosynthesis [11]. Comparisons of the actinorhodin (*act*) and *gra* gene sets have confirmed or revealed expected similarities in the polyketide synthase (PKS) and immediate post-PKS enzymes, but have suggested significant differences in the tailoring steps involved in pyran-ring formation and the hydroxylation reactions, only some of which might have been predicted. In addition, the *gra* gene cluster has provided further examples of genes for the deoxysugar biosynthetic and transfer reactions that are so crucial for biological activity of a huge range of natural products. A plausible mechanism is shown for the highly unusual C-glycosidation found in the granaticin metabolites.

Materials and methods

Bacterial strains, plasmids, culture conditions and DNA manipulations

pOJ446-22-16 and pOJ446-22-24 [17], derivatives of cosmid pOJ446 [53], harbour inserts of *S. violaceoruber* Tü22 genomic DNA. *S. coelicolor* A3(2) CH999 [15] *(proA1, argA1, redE60, act::ermE* SCP1$^-$, SCP2$^-$) lacks the whole of the actinorhodin (*act*) biosynthetic gene cluster (except for *act*VI-ORFA: [6]). pBluescript SK$^{-/+}$ (pBS-SK$^{-/+}$) was obtained from Stratagene. Genetic manipulations and culture conditions for *Streptomyces* were as described previously [54]. SFM medium for *Streptomyces* was as described previously [55]. General DNA manipulations were according to standard procedures [56].

Cosmids pOJ446-22-16 and pOJ446-22-24 were introduced into strain CH999 by conjugation from *E. coli* [57] on SFM plates; transconjugants were selected by overlaying plates with 1 ml of apramycin sulphate (1 mg/ml in water) and nalidixic acid (0.5 mg/ml).

Purification of granaticins

Transconjugants of CH999 containing pOJ446-22-24 or pOJ446-22-16 were grown as confluent lawns each on ten R5 agar plates [54] containing 50 μg/ml apramycin sulphate. After 6 days of growth at 30°C, the culture medium was cut up and extracted with 1 l ethyl acetate/methanol/acetic acid (80/20/5) in three portions. The combined extracts were evaporated to dryness and the residual acetic acid was stripped with toluene. The crude extract was dissolved in a minimum volume of methanol, centrifuged to remove residual agar, and evaporated to dryness. The extract was then dissolved in 40 ml acetonitrile and filtered, and the pigments were separated on C-18 reverse phase silica gel (YMC gel ODSA, 120 A, I-230/70 mesh) by flash column chromatography with methanol/1% acetic acid (50/50) as eluent. Fractions were pooled, concentrated *in vacuo*, and partitioned between dilute acetic acid and chloroform. The chloroform layer was dried over Na_2SO_4 and evaporated to dryness. Residual acetic acid was stripped with toluene.

Chemical analysis of granaticin

Analytical TLC was performed on 0.2 mm silica gel plates (Kieselgel 60 F_{254}, Merck) precoated with 0.5N oxalic acid as described [58]. The solvent was chloroform/ethyl acetate (60/40). R_F values for granaticin (0.25), dihydrogranaticin (0.20), granaticin B (0.14) and dihydrogranaticin B (0.12) correspond to those of Snipes *et al.* [58].

An Econosil C-18-5U column (5 μm, 250 × 4.6 mm, Alltech) was used for analytical HPLC of purified compounds and crude extracts. Samples were purified over a C-18 adsorption column (Bakerbond) by elution with acetonitrile and applied to the HPLC column (20 μl); the chromatogram was developed (0-8 min, 1% acetic acid/acetonitrile 65:35, 8-30 min 60:40) at room temperature. The flow rate was 1 ml/min and the adsorbance at 500 nm of the effluent was recorded. Retention times were as follows: dihydrogranaticin, 7.25 min; granaticin, 9.58 min; dihydrogranaticin B, 13.26 min; granaticin B, 15.87 min. Purified compounds and granaticin isolated previously [58] were used as standards for coinjection with crude extracts.

Because of the poor solubility of dihydrogranaticin in $CHCl_3$ the methyl ester was prepared in quantitative yield for identification by NMR in $CDCl_3$. Dihydrogranaticin (50 mg) was dissolved in anhydrous methanol (7 ml) and stirred under argon at room temperature; trimethylchlorosilane (0.7 ml) was added slowly from a syringe. After 2 h the mixture was poured into 300 ml ice water and stirred immediately. Dihydrogranaticin methyl ester was extracted with $CHCl_3$ and the organic layer dried over Na_2SO_4, filtered and evaporated to dryness.

NMR data were recorded in $CDCl_3$ on a Bruker AF300 instrument at 300 MHz for ^{1}H-NMR and at 75 MHz for ^{13}C-NMR. Purified compounds gave NMR data consistent with published values for dihydrogranaticin methyl ester [16,59], dihydrogranaticin B [60] and ^{1}H-NMR of granaticin B [60]. ^{13}C-NMR signals (ppm) of granaticin B were assigned by comparison to dihydrogranaticin [59] and dihydrogranaticin B [60] as follows: C1-170.4, C2-37.0, C3-66.2, C4-68.8, C5-139.7, C6-176.7, C7-110.3, C8-166.5, C9-144.6, C10-130.9, C11-167.9, C12-111.5, C13-174.1, C14-145.8, C15-67.0, C16-18.3, C17-61.6, C18-35.1, C19-67.3, C20-78.8, C21-72.7, C22-16.9, L-rhodinose moiety: C1′-95.0, C2′-25.4, C3′-23.5, C4′-75.5*, C5′-66.9*, C6′-17.1 (* can be reversed).

Negative ion electrospray mass spectra for determination of molecular mass (M-1) were recorded with a Fisons VG Quattro II mass spectrometer. Molpeaks were consistent with expected molecular masses for granaticin (M444), dihydrogranaticin (M446), granaticin B (M558) and dihydrogranaticin B (M560).

DNA sequencing

Sequencing was on single-stranded (rescued using M13K07 helper phage) or double-stranded plasmid DNA (pBS-SK$^{-/+}$). Sequencing was performed manually with the TaqTrack® system (Promega) or the Sequenase version 2.0 DNA sequencing kit (United States Biochemicals). Alternatively, sequencing was carried out on an automated DNA sequencer model 4000L (LI-COR Inc., Neb) with the Thermo Sequenase™ cycle sequencing kit (Amersham). Subclones B28, B2, B35, B30, B44, B40 and B9 were sequenced for us by Lark Sequencing Technologies Inc., Houston, Texas and DB211L, DB211R and BK1 by Nigel Hartley at the John Innes Centre, Norwich, UK. DB210, DB211 (part), B32 (part), B30 (part), B9 (part), B40, B8 (part), B13 (part), B5 and BK1 (part) by K.I. at the University of Tokyo; B32 (part) and B24 (part) by A.B. at the University of Tübingen. B40 and B44 were sequenced by D.T. at the University of Washington, Seattle, by sequencing overlapping subclones with the ABI Prism™ Dye Terminator Cycle Sequencing Kit (Perkin Elmer) on an automated DNA sequencer, model ABI Prism 377 (Perkin Elmer).

Computer analysis of DNA and protein sequences

Sequences were analysed with the University of Wisconsin Genetics Computer Group programs and the DNASIS programs (Hitachi Software Engineering Co. Ltd., Japan). A version of the FRAME program [18], MacFRAME (version 1.2 developed by Kevin Kendall (Tulane University, New Orleans, LA), was used to identify potential protein-coding regions.

Accession numbers

The entire sequence reported here, together with revised versions of those previously submitted [12,17], have been deposited in the EMBL database under the accession number AJ011500.

Acknowledgements

We thank J. Niemi and C.R. Hutchinson for useful information on genes relevant to this study and M. Paget for discussions on *SoxR,S* and its homologues. Financial support was from the Biotechnology and Biological Sciences Research Council (Clean Technology Grant no. CT04273) and the John Innes Foundation to D.A.H.; from the National Institutes of Health (AI20264) to H.G.F.; and from the Ministry of Education, Sports and Culture, Japan and the Uehara Memorial Foundation to K.I., D.T. and A.B. were supported at the University of Washington by postdoctoral fellowships from the DAAD and Deutsche Forschungsgeneinschaft, respectively.

References

1. Bystrykh, L.V., Fernández-Moreno, M.A., Herrema, J.K., Malpartida, F., Hopwood, D.A. & Dijkhuizen, L. (1996). Production of actinorhodin-related 'blue pigments' by *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **178**, 2238-2244.
2. Caballero, J.L., Martínez, E., Malpartida, F. & Hopwood, D.A. (1991). Organisation and functions of the *act*VA region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol. Gen. Genet.* **230**, 401-412.
3. Hallam, S.E., Malpartida, F. & Hopwood, D.A. (1988). Nucleotide sequence, transcription and deduced function of a gene involved in polyketide antibiotic synthesis in *Streptomyces coelicolor*. *Gene* **74**, 305-320.
4. Fernández-Moreno, M.A., Caballero, J.L., Hopwood, D.A. & Malpartida, F. (1991). The *act* cluster contains regulatory and antibiotic export genes, direct targets for transcriptional control by the *bldA* tRNA gene of *Streptomyces*. *Cell* **66**, 769-780.
5. Fernández-Moreno, M.A., Martínez, E., Boto, L., Hopwood, D.A. & Malpartida, F. (1992). Nucleotide sequence and deduced functions of a set of co-transcribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J. Biol. Chem.* **267**, 19278-19290.
6. Fernández-Moreno, M.A., Martínez E., Caballero, J.L., Ichinose K., Hopwood, D.A. & Malpartida, F. (1994). DNA sequence and functions of the *actVI* region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor* A3(2). *J. Biol. Chem.* **269**, 24854-24863.
7. Bibb, M.J., Sherman, D.H., Ōmura, S. & Hopwood, D.A. (1994). Cloning, sequencing and deduced functions of a cluster of *Streptomyces* genes probably encoding biosynthesis of the polyketide antibiotic frenolicin. *Gene* **142**, 31-39.

8. Yu, T.-W., Bibb, M.J., Revill, W.P. & Hopwood, D.A. (1994). Cloning, sequencing, and analysis of the griseusin polyketide synthase gene cluster from *Streptomyces griseus*. *J. Bacteriol.* **176**, 2827-2634.
9. Kakinuma, S., Takada, Y., Ikeda, H., Tanaka, H., Ōmura, S. & Hopwood, D.A. (1991). Cloning of large DNA fragments, which hybridize with actinorhodin biosynthesis genes, from kalafungin and nanaomycin A methyl ester producers and identification of genes for kalafungin biosynthesis of the kalafungin producer. *J. Antibiot.* **44**, 995-1005.
10. Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. *Chem. Rev.* **97**, 2465-2497.
11. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1995). Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* **375**, 549-554.
12. Sherman, D.H., Malpartida, F., Bibb, M.J., Kieser, H.M., Bibb, M.J. & 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.
13. Sherman, D.H., Kim, E-S., Bibb, M.J. & Hopwood, D.A. (1992). Functional replacement of genes for individual polyketide synthase components in *Streptomyces coelicolor* A3(2) by heterologous genes from a different polyketide pathway. *J. Bacteriol.* **174**, 6184-6190.
14. Khosla, C., *et al.*, & Hopwood, D.A. (1993). Genetic construction and functional analysis of hybrid polyketide synthases containing heterologous acyl carrier proteins. *J. Bacteriol.* **175**, 2197-2204.
15. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1993). Engineered biosynthesis of novel polyketides. *Science* **262**, 1546-1550.
16. Snipes, C.E., Chang, C.-J. & Floss, H.G. (1979). Biosynthesis of the antibiotic granaticin. *J. Am. Chem. Soc.* **101**, 701-706.
17. Bechthold, A., Sohng, J.K., Smith, T.M., Chu, X. & Floss. H.G. (1995). Identification of *Streptomyces violaceoruber* Tü22 genes involved in the biosynthesis of granaticin. *Mol. Gen. Genet.* **248**, 610-620.
18. Bibb, M.J., Findlay, P.R. & Johnson, M.W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* **30**, 157-166.
19. Strohl, W.R. (1992). Compilation and analysis of DNA sequences associated with apparent *Streptomyces* promoters. *Nucleic Acids Res.* **20**, 961-974.
20. McDaniel, R., Ebert-Khosla, S., Hopwood, D.A. & Khosla, C. (1994). Engineered biosynthesis of novel polyketides: *actVII* and *actIV* genes encode aromatase and cyclase enzymes, respectively. *J. Am. Chem. Soc.* **116**, 10855-10859.
21. Künzel, E., *et al.*, & Rohr, J. (1997). Tetracenomycin M, a novel genetically engineered tetracenomycin resulting from a combination of mithramycin and tetracenomycin biosynthetic genes. *Chem. Eur. J.* **3**, 1675-1678.
22. Gerlitz, M., Meurer, G., Wendt-Pienkowski, E., Madduri, K., & Hutchinson, C.R. (1997). Effects of the daunorubicin *dpsH* gene on the choice of starter unit and cyclisation pattern reveals that type II polyketide synthases can be unfaithful yet intriguing. *J. Am. Chem. Soc.* **119**, 7392-7393.
23. Rajgarhia, V.B. & Strohl, W.R. (1997). Minimal *Streptomyces* sp. strain C5 daunorubicin polyketide biosynthesis genes required for aklanonic acid biosynthesis. *J. Bacteriol.* **179**, 2690-2696.
24. Shen, B. & Hutchinson, C.R. (1993). Tetracenomycin F1 monooxygenase: oxidation of a naphthacenone to a naphthacenequinone in the biosynthesis of tetracenomycin C in *Streptomyces glaucescens*. *Biochemistry* **32**, 6656-6663.
25. Kendrew, S.G., Hopwood, D.A. & Marsh, E.N.G. (1997). Identification of a monooxygenase from *Streptomyces coelicolor* A3(2) involved in biosynthesis of actinorhodin: purification and characterization of the recombinant enzyme. *J. Bacteriol.* **179**, 4305-4310.
26. Cole, S.P., Rudd, B.A.M., Hopwood, D.A., Chang, C.-J. & Floss, H.G. (1987). Biosynthesis of the antibiotic actinorhodin. Analysis of blocked mutants of *Streptomyces coelicolor*. *J. Antibiot.* **40**, 340-347.
27. Hopwood, D.A., *et al.*, & Ōmura, S. (1985). Production of 'hybrid' antibiotics by genetic engineering. *Nature* **314**, 642-644.
28. Kendrew, S.G., Harding, S.E., Hopwood, D.A. & Marsh, E.N.G. (1995). Identification of a flavin: NADH oxidoreductase involved in the biosynthesis of actinorhodin. *J. Biol. Chem.* **270**, 17399-17343.
29. Distler, J., Ebert, A., Mansouri, K., Pissowotzki, K., Stockmann, M. & Piepersberg, W. (1987). Gene cluster for streptomycin biosynthesis in *Streptomyces griseus:* nucleotide sequence of three genes and analysis of transcriptional activity. *Nucleic Acids Res.* **15**, 8041-8056.
30. Pissowotzki, K., Mansouri, K. & Piepersberg, W. (1991). Genetics of streptomycin production in *Streptomyces griseus:* molecular structure and putative function of genes strELMB2N. *Mol. Gen. Genet.* **231**, 113-123.
31. Piepersberg, W. (1997). Molecular biology, biochemistry, and fermentation of aminoglycoside antibiotics. In *Biotechnology of Antibiotics*. pp. 81-161. (Strohl, W.R., ed.) Marcel Dekker Inc, New York.
32. Kessler, A.C., Haase, A. & Reeves, P.R. (1993). Molecular analysis of the 3,6- dideoxyhexose pathway genes of *Yersinia pseudotuberculosis* serogroup IIA. *J. Bacteriol.* **175**, 1412-1422.
33. Lei, Y., Ploux, O. & Liu, H.-W. (1995). Mechanistic studies on CDP-6-deoxy-L-threo- D-glycero-4-hexulose 3-dehydrase: identification of His-220 as the active-site base by chemical modification and site-directed mutagenesis. *Biochemistry* **34**, 4643-4654.
34. Otten, S.L., Liu, X., Ferguson, J. & Hutchinson, C.R. (1995). Cloning and characterization of the *Streptomyces dnrQS* genes encoding a daunosamine biosynthesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. *J. Bacteriol.* **177**, 6688-6692.
35. Strohl, W.R., Dickens, M.L., Rajgarhia, V.B., Woo, A.J. & Priestley, D. (1997). Anthracyclines. In *Biotechnology of Antibiotics* (Strohl, W.R., ed.) pp. 577-657. Marcel Decker, New York.
36. Otten, S.L., Gallo, M.A., Madduri, K., Liu, X. & Hutchinson, C.R. (1997). Cloning and characterization of the *Streptomyces peucetius dnmZUV* genes encoding three enzymes required for biosynthesis of the daunorubicin precursor thymidine diphospho-L-daunosamine. *J. Bacteriol.* **179**, 4446-4450.
37. Rossmann, M.G., Moras, D. & Olsen, K.W. (1974). Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**, 194-199.
38. Persson, B., Rook, M.K. & Jörnvall, H. (1991). Characteristics of short-chain alcohol dehydrogenases and related enzymes. *Eur. J. Biochem.* **200**, 537-543.
39. Scotti, C. & Hutchinson, C.R. (1996). Enhanced antibiotic production by manipulation of the *Streptomyces peucetius dnrH* and *dnmT* genes involved in doxorubicin (adriamycin) biosynthesis. *J. Bacteriol.* **178**, 7316-7321.
40. Peschke, U., Schmidt, H., Zhang, H.Z. & Piepersberg, W. (1995). Molecular characterization of the lincomycin-production gene cluster of *Streptomyces lincolnensis* 78-11. *Mol. Microbiol.* **16**, 1137-1156.
41. Wietzorrek, A. & Bibb, M.J. (1997). A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol. Microbiol.* **25**, 1177-1184.
42. Wu, J. & Weiss, B. (1991). Two divergently transcribed genes, *soxR* and *soxS,* control a superoxide response regulon of *Escherichia coli*. *J. Bacteriol.* **173**, 2864-2871.
43. Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. *J. Biol. Chem.* **264**, 13963-13966.
44. Ninfa, A.J. (1996). Regulation of gene expression by extracellular stimuli. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. (Neidhart, F.C., Curtiss, R. III, Ingraham, J.L., Lin, E.C.C., Lou, K.B., Magosonik, B., Reznikoff, W.S., Riley, M., Schaechter, M. & Umbarger, H.E., eds), pp. 1246-1262. ASM Press, Washington, D.C.
45. Henner D.J., Yang M. & Ferrari E. (1988). Localization of *Bacillus subtilis sacU*(Hy) mutations to two linked genes with similarities to the conserved procaryotic family of two-component signalling systems. *J. Bacteriol.* **170**, 5102-5109.
46. Urabe, H. & Ogawara, H. (1995). Cloning, sequencing and expression of serine/threonine kinase-encoding genes from *Streptomyces coelicolor* A3(2). *Gene* **153**, 99-104.
47. Li, Y., Dosch, D.C., Strohl, W.R. & Floss H.G. (1990). Nucleotide sequence and transcriptional analysis of the nosiheptide-resistance gene from *Streptomyces actuosus*. *Gene* **91**, 9-17.
48. Malpartida, F., *et al.*, & Hopwood, D.A. (1987). Homology between *Streptomyces* genes coding for synthesis of different polyketides used to clone antibiotic synthesis genes. *Nature* **325**, 818-821.
49. Nakajima, K., Hashimoto, T. & Yamada, Y. (1993). Two tropinone reductases with different stereospecificities are short-chain dehydrogenases evolved from a common ancestor. *Proc. Natl Acad. Sci. USA* **90**, 9591-9595.
50. Taguchi, H. & Ohta, T. (1991). D-lactate dehydrogenase is a member of the D-isomer- specific 2-hydroxyacid dehydrogenase family. *J. Biol. Chem.* **266**, 12588-12594.
51. Kirschning, A., Bechthold, A.F.W. & Rohr, J. (1997). Chemical and biochemical aspects of deoxysugars and deoxysugar oligosaccharides. *Topics Curr. Chem.* **188**, 1-84.
52. Liu, H.-W. & Thorson, J.S. (1994). Pathway and mechanisms in the biogenesis of novel sugars by bacteria. *Annu. Rev. Microbiol.* **48**, 223-256.
53. Bierman, M., Logan, R., O'Brien, K., Seno, E.T., Rao, R.N. & Schoner, B.E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**, 43-49.

54. Hopwood, D.A., *et al.*, & Schrempf, H. (1985). *Genetic Manipulation of Streptomyces. A Laboratory Manual.* John Innes Foundation, Norwich, England.
55. Floriano, B. & Bibb, M. (1996). *afsR*, is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.* **21**, 385-396.
56. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual.* 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
57. Flett, F., Mersinias, V. & Smith, C.P. (1997). High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol. Lett.* **155**, 223-229.
58. Snipes, C.E., Chang, C.J. & Floss H.G. (1979). Biosynthesis of the antibiotic granaticin: assignment of the carbon 13 magnetic resonance spectrum. *J. Nat. Prod.* **42**, 627-632.
59. Pyrek, St.J., Achmatowicz, O. & Zamajski, A. (1977). Naphtho- and anthraquinones of *Streptomyces thermoviolaceus* WR-141. Structures and model syntheses. *Tetrahedron* **33**, 673-680.
60. Gilpin, M.L., Box, S.J. & Elson, A.L. (1988) New quinone antibiotics of the granaticin type, isolated from *Streptomyces lateritius*. II. Structure determination. *J. Antibiot.* **41**, 512-518.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.