

PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin

Bertolt Gust*[†], Greg L. Challis[‡], Kay Fowler*[§], Tobias Kieser*, and Keith F. Chater*

*Department of Molecular Microbiology, John Innes Centre, Colney, Norwich NR4 7UH, United Kingdom; and [‡]Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom

Communicated by Richard M. Losick, Harvard University, Cambridge, MA, December 12, 2002 (received for review October 23, 2002)

Streptomyces are high G+C Gram-positive, antibiotic-producing, mycelial soil bacteria. The 8.7-Mb *Streptomyces coelicolor* genome was previously sequenced by using an ordered library of Supercos-1 clones. Here, we describe an efficient procedure for creating precise gene replacements in the cosmid clones by using PCR targeting and λ -Red-mediated recombination. The cloned *Streptomyces* genes are replaced with a cassette containing a selectable antibiotic resistance and *oriT*_{RK2} for efficient transfer to *Streptomyces* by RP4-mediated intergeneric conjugation. Supercos-1 does not replicate in *Streptomyces*, but the clones readily undergo double-crossover recombination, thus creating gene replacements. The antibiotic resistance cassettes are flanked by yeast FLP recombinase target sequences for removal of the antibiotic resistance and *oriT*_{RK2} to generate unmarked, nonpolar mutations. The technique has been used successfully by >20 researchers to mutate around 100 *Streptomyces* genes. As an example, we describe its application to the discovery of a gene involved in the production of geosmin, the ubiquitous odor of soil. The gene, *SCO6073* (*cyc2*), codes for a protein with two sesquiterpene synthase domains, only one of which is required for geosmin biosynthesis, probably via a germacra-1 (10) *E,5E*-dien-11-ol intermediate generated by the sesquiterpene synthase from farnesyl pyrophosphate.

FLP recombinase | λ Red recombinase | *oriT* | SCO5222 | SCO6073

The recently completed 8.7-Mb *Streptomyces coelicolor* A3(2) genome sequence (1) has revealed a wealth of information. Notably, about 400 kb (more than 20 gene clusters) are apparently concerned with secondary metabolites. Predicted gene products include type I modular and type I and type II iterative polyketide synthases, chalcone synthases, nonribosomal peptide synthetases, terpene synthases, and genes with functions that are less well understood. Mutational analysis can provide a significant insight into the functions of these genes.

Here, we describe the adaptation to *S. coelicolor* of a PCR targeting-based gene disruption protocol (2) that permits the deletion of entire gene clusters for secondary metabolism and allows relatively rapid generation of nonpolar, in-frame deletions, thus avoiding polar effects on genes transcriptionally downstream. This procedure is much more efficient than those previously used to disrupt *Streptomyces* genes (3). Our approach was based on the discovery that allelic exchanges on the *Escherichia coli* chromosome can be achieved by recombination with a PCR-generated selectable marker flanked at both ends by extensions of only a few tens of nucleotides homologous to the desired region of the chromosome, when the Red α (*exo*), Red β (*bet*), and Red γ (*gam*) proteins of the phage λ are present in the targeted strain (4–9). We have used λ -Red to promote recombination in *E. coli* between a PCR-amplified antibiotic resistance cassette selectable in *E. coli* and *Streptomyces*, and *S. coelicolor* DNA on a cosmid from the set used to sequence the *S. coelicolor* genome (10). The inclusion of an origin of transfer (*oriT*; RK2) in the disruption cassette allows the conjugal transfer of the PCR-targeted cosmid into *S. coelicolor*, readily yielding excon-

jugants with the desired gene replacement. The strategy also allows the elimination of the disruption cassette by FLP-recombinase-mediated site-specific recombination (11).

In a demonstration of the utility of this PCR-targeting method for the rapid construction of *S. coelicolor* knockout mutants, we set out to identify genes involved in the biosynthesis of the secondary metabolite geosmin. Many actinomycetes (12–14), certain cyanobacteria (15), myxobacteria, liverworts (16), and higher fungi (17) that inhabit aquatic or soil environments produce geosmin, which can impart an undesirable earthy or musty odor to food and water supplies (18, 19). The very ubiquitousness of geosmin, especially among streptomycetes, suggests that it has adaptive importance, but this remains unidentified at present. To the best of our knowledge, no genes that direct geosmin biosynthesis in any of these organisms have been identified. In streptomycetes, geosmin has been postulated to be synthesized from farnesyl pyrophosphate by means of a eudesmanoid or germacranoid sesquiterpene precursor (20). If so, an early step in geosmin biosynthesis should involve a sesquiterpene synthase. Two genes, here designated *cyc1* and *cyc2*, encoding enzymes showing similarity to the sesquiterpene synthase that carries out a cyclization reaction involved in pentalenolactone synthesis in *Streptomyces* UC5319 (21), were found in the *S. coelicolor* genome sequence. Here, we use PCR targeting and transposon mutagenesis to show that one of these two genes and only one of the two protein domains encoded by that gene are essential for geosmin biosynthesis. The accompanying paper by Cane and Watt (22) in this issue of PNAS describes biochemical analysis of the protein encoded by this gene.

Materials and Methods

Strains and Culture Conditions. Media and culture conditions for *E. coli* and *S. coelicolor* were the same as those described (2, 3). Strains are listed in Table 1. *E. coli* BW25113 (2) was used to propagate the recombination plasmid pIJ790 and *S. coelicolor* cosmids (10). *E. coli* DH5 α carrying pCP20 (designated BT340) was used for FLP-mediated site-specific recombination (2). Both strains (CGSC#7636 and CGSC#7629, respectively) were obtained from the *Escherichia coli* Genetic Stock Center (Yale University, New Haven, CN). BW25113 and BT340 were grown in SOB or LB, respectively. *E. coli* DH5 α (Stratagene) was used as a host for plasmid constructions. *E. coli* ET12567/pUZ8002 (23) was the nonmethylating plasmid donor strain for intergeneric conjugation with *S. coelicolor* strain M145 (1). Carbenicillin (Carb, 100 μ g/ml), apramycin (Apra, 50 μ g/ml), chloramphenicol (Cm, 25 μ g/ml),

Abbreviations: FRT, FLP-recombinase recognition target; Carb, carbenicillin; Apra, apramycin; Cm, chloramphenicol; Kan, kanamycin; Vio, viomycin; ^R, resistance; ^S, sensitivity; Strep, streptomycin; Spec, spectinomycin; *oriT*, origin of transfer; GC-MS, gas chromatography-MS.

[†]To whom correspondence should be addressed. E-mail: bertolt.gust@bbsrc.ac.uk.

[§]Present address: Division of Food Safety Science, Institute of Food Research, Norwich NR4 7UA, United Kingdom.

Table 1. Strains and plasmids used in this study

Strain/plasmid	Relevant genotype/comments	Source/ref.
Plasmids		
pKD20	λ -RED (<i>gam</i> , <i>bet</i> , <i>exo</i>), <i>bla</i> , <i>araC</i> , <i>rep101^{ts}</i>	2
pCP20	FLP-recombination plasmid: <i>flp</i> , <i>bla</i> , <i>cat</i> , <i>rep101^{ts}</i>	11
pIJ666	Source of <i>cat</i> gene for pIJ790	24
pOJ436	Source of Apra ^R <i>oriT</i> sequence for pIJ773	25
pIJ790	λ -RED (<i>gam</i> , <i>bet</i> , <i>exo</i>), <i>cat</i> , <i>araC</i> , <i>rep101^{ts}</i>	This study
pIJ773	<i>aac(3)IV</i> (Apra ^R) + <i>oriT</i>	This study
pIJ778	<i>aadA</i> from Ω -fragment (Spec ^R , Strep ^R) + <i>oriT</i>	This study
pIJ779	<i>aadA</i> from Ω -fragment (Spec ^R , Strep ^R)	This study
pIJ780	<i>vph</i> from Ω -fragment (Vio ^R) + <i>oriT</i>	This study
pIJ781	<i>vph</i> from Ω -fragment (Vio ^R)	This study
Supercos-1	<i>bla</i> , <i>neo</i> , <i>cos</i>	Stratagene
pUZ8002	<i>tra</i> , <i>neo</i> , RP4	23
pUB307	<i>neo</i> , RP4	30
pSET Ω	<i>aadA</i> , <i>oriT</i> , <i>attP</i> , <i>int</i>	31
E. coli		
BW25113	K12 derivative: Δ <i>araBAD</i> , Δ <i>rhaBAD</i>	2
ET12567	<i>dam</i> , <i>dcm</i> , <i>hsdS</i> , <i>cat</i> , <i>tet</i>	29
BT340	DH5 α /pCP20	11
S. coelicolor		
M145	SCP1 ⁻ , SCP2 ⁻	1
KF62	<i>cyc1</i> ::Tn4560 insertion at aa position 75	This study
J3001	Δ <i>cyc2</i> : aa 2–725 replaced with pIJ773 cassette	This study
J3002	<i>cyc1</i> ::Tn4560, Δ <i>cyc2</i>	This study
J3003	Δ <i>cyc2scar</i> : aa 2–725 replaced with scar sequence	This study
J3004	<i>cyc2</i> Δ N: aa 6–340 replaced with scar sequence	This study
J3005	<i>cyc2</i> Δ C: aa 380–721 replaced with scar sequence	This study
J3006	<i>cyc1</i> ::Tn4560, <i>cyc2</i> Δ N	This study
J3007	<i>cyc1</i> ::Tn4560, <i>cyc2</i> Δ C	This study

kanamycin (Kan, 50 μ g/ml) or viomycin (Vio, 30 μ g/ml), all from Sigma, were added to growth media when required. Vio is currently available only from Research Diagnostics (Flanders, NJ). L-arabinose (10 mM final concentration; Sigma) was added as indicated to SOB medium from a 1 M sterile filtered stock solution to induce genes under control of the pBAD promoter (2).

Plasmids. The ampicillin resistance marker *bla* on pKD20 (in strain CGSC#7636) was replaced by *cat* (Cm^R) amplified from pIJ666 (24) with primers *catforw* and *catrev* (Table 2), thus generating pIJ790 (Table 1). The PCR product contained *cat*, with its own promoter, flanked by 40 bp of homology to regions adjacent to *bla*. Cm^R colonies were tested for ampicillin or Carb sensitivity. The relevant plasmid structure was confirmed by restriction analysis.

The gene disruption cassette *aac(3)IV* (Apra^R) and the *oriT*

from RK2 were jointly amplified from pOJ436 (25) with primers *apraforw* and *oriTrev* (Table 2). FLP recognition target (FRT) sites and priming sites were then added to the resulting 1,246-bp fragment by amplification of left and right ends of the disruption cassette separately by PCR before joining them by *XhoI* digestion and ligation. The final disruption cassette was further cloned into the *EcoRV* site of pBluescript II SK(+) (Stratagene) to generate pIJ773 (Table 1). *aadA* (spectinomycin-resistance and streptomycin-resistance: Spec^R, Strep^R) and *vph* (Vio^R) were amplified from the omega fragments (accession nos. M60473 and X99314), and *aac(3)IV* (Apra^R) of pIJ773 was replaced with the PCR products by λ -Red-mediated recombination, generating pIJ778 and pIJ780, respectively (Table 1). Disruption cassettes without *oriT* were constructed directly by PCR by using the omega fragments as templates and inserted into the *EcoRV* site of pBluescript II SK(+), generating pIJ779 (Spec^R, Strep^R) and

Table 2. PCR primers used in this study (from Invitrogen)

Primer name	Generation of	Sequence (5'–3')
<i>catforw</i>	pIJ790	ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTt <u>cgggcaccgtaagaggttcc</u>
<i>catrev</i>	pIJ790	TTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGA <u>ttaagggcaccataactgc</u>
<i>apraforw</i>	pIJ773	tcatgagctcagccaatc
<i>oriTrev</i>	pIJ773	cgccagcctcgagagcag
Sc9B1.20forw	<i>cyc2</i> deletion	CGGCGGATGCGGTGCAAGAGCCCTGGGTAGGGCCGGGCC <u>attccggggatccgtcgacc</u>
Sc9B1.20rev	<i>cyc2</i> deletion	CGAGCCACGAAAGAGTGAGACTGAACGTCCGTCCAGCGCGT <u>gttaggctggagctgcttc</u>
Sc9B1.20 Δ Nforw	<i>cyc2</i> Δ N deletion	AGCCCTGGGTAGGGCCGGCC CATG ACGCAACAGCCCTT <u>Cattccggggatccgtcgacc</u>
Sc9B1.20 Δ Nrev	<i>cyc2</i> Δ N deletion	GAGCAGTGCTCCACGTCGCCCGGAGGTGCCGGGCC <u>tgtaggctggagctgcttc</u>
Sc9B1.20 Δ Cforw	<i>cyc2</i> Δ C deletion	GTGCCCTTCCAGAAGGTCCGCCCGTCCGTCCATCCCGAC <u>attccggggatccgtcgacc</u>
Sc9B1.20 Δ Crev	<i>cyc2</i> Δ C deletion	CAGTGAACGTCCGTCCAGCGCT CAGT GCGTCAGTCCGGT <u>gttaggctggagctgcttc</u>

The 3' homology of primers *catforw* and *catrev* to the *cat* gene of pIJ666 and the unique priming sites P1 and P2 of the disruption cassettes are underlined. The oligonucleotides have 39 nt of homology (uppercase) to the relevant gene. Start and stop codons within the 39-nt primer extension are bold.

pIJ781 (Vio^R), respectively (Table 1). All template plasmids contain the same FRT and priming sites, which were verified by sequencing.

Gene Replacement on Cosmids. Electro-competent cells of an *E. coli* BW25113 transformant carrying pIJ790 (Cm^R) and the desired *S. coelicolor* cosmid (Carb^R, Kan^R) containing the gene to be targeted were prepared from a 10-ml SOB culture containing 10 mM L-arabinose to induce the λ -red genes. The culture (30°C, OD₆₀₀ ≈ 0.4) was immediately washed twice with ice-cold 10% glycerol and resuspended in the remaining drop of 10% glycerol after centrifugation, resulting in a 100-fold concentration of cells. Electro-competent cells (50 μ l) were transformed with 100 ng purified PCR product in a 0.2-cm ice-cold electroporation chamber by using a Bio-Rad GenePulser II set to 200 Ω , 25 μ F, and 2.5 kV. Shocked cells were incubated in 1 ml of LB at 37°C for 1 h and plated on LB containing Kan and Carb (to select the cosmid) and Apra (to select the integrated PCR product). Transformants were characterized by restriction analysis and PCR with a set of oligonucleotides priming outside the region of recombination.

PCR products were obtained by amplification of the twice gel-purified 1,383-bp *Eco*RI/*Hind*III pIJ773 disruption cassette with the appropriate primer pairs containing 39-nt extensions for λ -Red-mediated recombination (Table 2). Amplification was performed in a 50- μ l reaction with 100 ng of template DNA, 200 μ M dNTPs, 50 pmol each primer, 5% DMSO, according to the manufacturer's instructions (Expand High Fidelity Polymerase, Roche Molecular Biochemicals): denaturation at 94°C for 2 min, then 10 cycles with denaturation at 94°C for 45 s, annealing at 50°C for 45 s, and extension at 72°C for 1 min 30 s, were followed by 15 cycles with the annealing temperature increased to 55°C. A last elongation step was done at 72°C for 5 min. The PCR product was analyzed by gel electrophoresis and purified by using a PCR Purification kit (Qiagen, Chatsworth, CA).

Allelic Exchange in *S. coelicolor*. The Kan^R, Carb^R, and Apra^R mutagenized cosmids were introduced into *E. coli* ET12567/pUZ8002 by electroporation as above and then transferred to *Streptomyces* by conjugation (3). Apra^R exconjugants were screened for Kan^S, indicating a double-crossover allelic exchange in *S. coelicolor*, and were confirmed by PCR and Southern blot analysis.

In-Frame Deletion Mediated by FLP Recombinase. pCP20 (Carb^R, Cm^R) shows temperature-sensitive replication and thermoinducible expression of the FLP recombinase, which acts on FRT sites to remove the central part of the disruption cassette, leaving behind an 81-bp "scar" sequence (2). Electro-competent *E. coli* DH5 α /pCP20 cells were transformed with the targeted cosmid DNA; Cm^R and Apra^R transformants were selected on LB at 30°C. A few colonies were single colony-purified nonselectively at 42°C. Twenty to fifty percent of the colonies lost Apra^R and pCP20 (Cm^R) simultaneously. Cosmid DNA containing the in-frame deletion was then introduced into the *S. coelicolor* mutant carrying the marked deletion in the gene of interest by polyethylene glycol (PEG)-mediated transformation (3). Selecting for single cross-over events with Kan (the *neo* gene of SuperCos1 confers Kan^R in *Streptomyces*) and subsequent screening for the loss of both Kan^R and the resistance derived from the marked deletion in the *S. coelicolor* mutant indicated the successful replacement of the resistance marker with an unmarked, in-frame deletion. Alternatively to the PEG-mediated transformation procedure, the cosmid containing the in-frame deletion can be made conjugation-proficient by replacing its *neo* gene with the Vio^R*oriT* (pIJ780) or Spec^R, Strep^R*oriT* (pIJ778) cassette by λ -Red-mediated recombination.

Gas Chromatography–Mass Spectrometry (GC-MS) Analysis. Qualitative analysis was performed by coupled GC-MS (Hewlett-Packard gas chromatograph HP 5890 series II coupled with a

Hewlett-Packard mass spectrometer HP 5989 A) (26). Volatiles released from *S. coelicolor* during 2 weeks' growth on MS plates at 30°C were adsorbed onto 100 mg of activated charcoal placed in the lid of each petri dish. The charcoal was extracted with 1 ml of chloroform (Sigma), which was then filtered through cotton wool. Five microliters of each extract was analyzed by GC-MS as follows: capillary column, fused silica (30 m; 0.25-mm inside diameter; 0.25- μ m film thickness); carrier gas, He (0.82 kPa on column injection); temperature program: isothermal for 1 min at 40°C, change from 40 to 210°C at a rate of 10°C per min, and isothermal for 25 min at 210°C; energy of ionization, 70 eV. The presence of geosmin in the samples was confirmed by comparison with synthetic (\pm)-geosmin (Sigma).

Results

Adaptation of PCR Targeting for *S. coelicolor*. The λ -Red expression plasmid pKD20 (2) could not be used with the *S. coelicolor* Supercos-1 clones because both contain the *bla* gene, preventing selection and providing sequence identity for undesirable cointegrate formation by homologous recombination. Thus, the *bla* gene was replaced by *cat* (Cm^R) to give pIJ790. Like pKD20, pIJ790 expressed λ -Red when induced by arabinose and was compatible with the ColE1-derived Supercos-1.

For the disruption of genes cloned in Supercos-1 and the efficient transfer of the mutant derivatives to *S. coelicolor*, a gene replacement cassette (Apra^R*oriT*) containing *aac* (3)*IV* (Apra^R) and *oriT*_{RRK2} was constructed in pBluescript II SK(+) (Fig. 1A). Apra^R and Kan^R (encoded by *neo* of Supercos-1) can be selected independently in both *E. coli* and *S. coelicolor*. The Apra^R*oriT* cassette is flanked at both ends by FRT sequences to allow efficient removal of Apra^R and *oriT* by yeast FLP recombinase (Fig. 1F–H). The template for the cassette amplification was the 1,383-bp *Eco*RI/*Hind*III fragment of pIJ773, which was twice purified by agarose gel electrophoresis to eliminate all traces of the intact plasmid. Without this precaution, Apra^R colonies without the desired gene replacement predominated among the transformants in the next step.

For generation of multiple gene disruptions on either a different or the same *S. coelicolor* cosmid, cassettes similar to the Apra^R*oriT* cassette containing *aadA* (Spec^R/Strep^R) or *vph* (Vio^R) instead of *aac* (3)*IV*, both with and without *oriT*, were also constructed (Table 1). All of these cassettes contain the same priming sequences for PCR amplification and leave the same "scar" sequence after FLP recombinase-mediated excision. Other cassette constructs with different resistance genes (not shown) failed to give the desired gene replacements because they recombined preferentially with Supercos-1 sequences.

In designing primers for PCR amplification, we took into account that λ -Red-mediated recombination frequencies approach their maximum levels with a 40-bp targeting sequence (27). We used 39 bp rather than 40 bp because it comprises an integral number of codons, simplifying primer design. The 39-bp targeting sequence "tails" were supplied as the 5' ends of long PCR primers, whose 3' ends consisted of 19- or 20-bp priming sequences (Fig. 1A). The 39-bp 5' target sequence tails were designed to generate "scar" sequences that were in a reading frame that lacked in-frame stop codons. A Perl program (JIC.GC.BMW) is available to assist in the primer design and in the analysis of the mutants generated (<http://jic-bioinfo.bbsrc.ac.uk/S.coelicolor/>).

Disruptions of *cyc2*. Replacement of *cyc2* in cosmid Sc9B1 (see below) was achieved by introducing a PCR-amplified, tailed gene disruption cassette Apra^R*oriT* by electroporation into *E. coli* BW25113/pIJ790 containing Sc9B1 (Fig. 1B), prepropagated at 30°C in the presence of 10 mM arabinose to induce the λ -Red genes on pIJ790. About 250 Apra^R, Carb^R, Kan^R colonies were obtained (Fig. 1C). These colonies were Cm^S because the temperature-sensitive pIJ790 was lost during cultivation at 37°C.

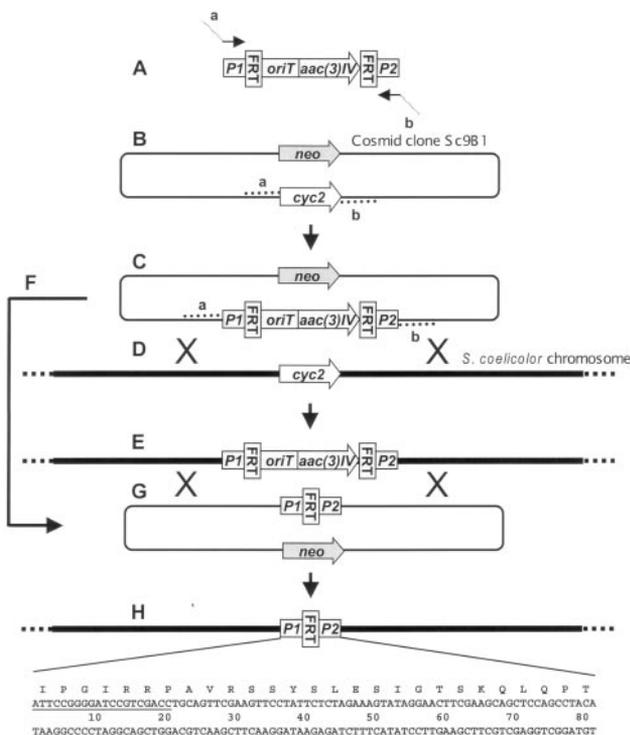


Fig. 1. Strategy for gene replacement in *Streptomyces*. (A) A gene replacement cassette containing *aac(3)IV* (Apra^R), *oriT* of RK2, and two FRT sites was amplified by PCR by using primers containing 39-nt 5' homology extensions corresponding to regions flanking the *S. coelicolor* target gene (dotted lines a and b) and 20-nt 3' homology to the unique priming sites P1 and P2 of the disruption cassette. (B) The PCR product from A was used to transform *E. coli* BW25113/pIJ790 (expressing the λ-Red recombination functions), containing the SupercoS-1 cosmid with the cloned target gene (*cyc2*). (C) Apra^R transformants were selected, and the recombinant cosmid was identified by PCR and restriction analysis. (D) The potent methyl-specific restriction system of *S. coelicolor* was circumvented by introducing the recombinant cosmid into the methylation-deficient *E. coli* host ET12567/pUZ8002. The cosmid was mobilized in *trans* by pUZ8002 to *Streptomyces* by conjugation. (E) Selected exconjugants (Apra^R) were screened for Kan^S (loss of the Kan resistance gene *neo*), and the double cross-over allelic exchange was confirmed by PCR and Southern blot analysis. (F) *E. coli* DH5α cells containing pCP20 were transformed with the mutagenized cosmid DNA, and FLP synthesis was induced. (G) The loss of the disruption resistance marker was tested, and the cosmid was introduced by polyethylene glycol (PEG)-mediated transformation into the *S. coelicolor* mutant carrying a marked deletion within the gene of interest. (H) Kan^R transformants arising from single cross-over events were restreaked nonselectively and screened for the loss of both Kan^R and Apra^R, indicating a successful replacement by the in-frame deletion marked only by a "scar" sequence of 81 bp containing priming sites P1 and P2 (underlined), and one FRT site.

No Apra^R transformants were obtained when arabinose was omitted or when a different cosmid without *cyc2* was substituted for Sc9B1. Agarose gel analysis of digested plasmid DNA from these transformants revealed in most cases mixtures of the original cosmid clone and the expected mutant cosmids. The mutant cosmid constituted between 10% and 90% of the plasmid DNA. [Note that even Apra^R, Carb^R, Kan^R clones containing only a minority of the desired mutant cosmids were still useful for introducing the mutation into *Streptomyces* (see below).]

The high precision of the recombination occurring between 39-bp sequences was more extensively documented in experiments with another gene, *whiI* (28). Cosmid DNA samples from 50 transformants were all shown to contain the desired gene replacement both by restriction analysis and by PCR, using primers annealing 100 bp outside the 39-bp tail sequences. Gene replacement was maximal with 100 ng of the tailed PCR product.

The Apra^R*oriT* cassette which replaced *cyc2* in cosmid Sc9B1 was flanked by 28,085 bp and 3735 bp of *S. coelicolor* DNA. [Note that the EMBL entry for Sc9B1, accession no. AL049727, does not show the entire cosmid sequence; complete (real) cosmid sequences are available from ftp://ftp.sanger.ac.uk/pub/S_coelicolor/cosmid_inserts.]

The next step was to introduce the mutant allele into *Streptomyces*. Because *S. coelicolor* A3(2), like some other streptomycetes, strongly restrict methylated DNA from wild-type *E. coli* K12, the plasmid DNA had to be passed through a nonmethylating host. We used ET12567/pUZ8002 for this purpose, because the plasmid supplies transacting functions for the mobilization of *oriT*-containing cosmids (3, 23, 29). Cosmid DNA from the above Apra^R, Carb^R, Kan^R transformants was introduced by electroporation into ET12567/pUZ8002. By using 100 ng CCC DNA, ≈100 Apra^R, Cm^R, Kan^R transformants were recovered after 24 h of incubation at 37°C. The low transformation frequency and the long time required for visible colonies to develop were typical for this strain (note that the Kan^R selective marker of pUZ8002 becomes ineffectual in the presence of the similarly marked SupercoS-1).

As an alternative to electroporation, mutant cosmids containing *oriT* could also be introduced into ET12567 by RP4 *tra*-mediated mobilization. BW25113 containing the self-transmissible pUB307 instead of the transfer-deficient pUZ8002 was used for this purpose (30). ET12567 contains chromosomal tetracycline resistance and Cm^R markers to counterselect *E. coli* BW25113.

Once in *Streptomyces*, SupercoS-1 does not replicate autonomously, but the long regions of sequence identity in the inserts promote efficient integration by homologous recombination. Thus, intergeneric conjugation between ET12567/pUZ8002/Sc9B1::Apra^R*oriT* and *S. coelicolor* M145 produced ≈10⁴–10⁵ Apra^R colonies per conjugation. Most of these colonies were also Kan^R and presumably contained the entire plasmid integrated into the chromosome by a single crossover event. However, a sizable minority (>10%) were Kan^S, indicating double-crossover gene replacement and loss of the SupercoS-1 vector (Fig. 1 D and E). In the case of the *cyc2* replacement, this finding was verified by PCR and Southern blot analysis for three independent Apra^R Kan^S *S. coelicolor* exconjugants.

The *cyc2* Gene Is Essential for Geosmin Production. One of the best known characteristics of streptomycetes is their production of the earthy odor geosmin, which is thought to be derived by the action of a sesquiterpene synthase on farnesyl pyrophosphate, which in turn is synthesized via the mevalonate pathway. *S. coelicolor* contains homologues of the genes needed for farnesyl pyrophosphate synthesis, as well as two homologues, *cyc1* (SCO5222) and *cyc2* (SCO6073), of the gene for a known sesquiterpene synthase (pentalene synthase, which is involved in pentalenolactone synthesis) from *Streptomyces* UC5319 (21). The *cyc1* product (Cyc1) contains one conserved sesquiterpene synthase domain, and the *cyc2* product (Cyc2) contains two such domains (Fig. 2). *cyc2* was replaced by Apra^R*oriT* in both *S. coelicolor* M145 (generating J3001, Table 1) and an M145 derivative, KF62, containing a Tn4560 insertion in *cyc1* (generating J3002, Table 1). The wild type and the *cyc1* mutant produced amounts of geosmin easily detectable by GC-MS (GC-retention time and MS-fragmentation pattern were identical to those of authentic geosmin). In contrast, no geosmin was detectable in the *cyc2* mutant (Table 3).

It was possible that the putative integral membrane protein gene SCO6074, downstream of *cyc2* (Fig. 2), might have been inactivated by polar effects from the insertion of the Apra^R*oriT* cassette. To eliminate such polar effects, the Apra^R*oriT* sequence was removed by FLP-mediated recombination from the mutant Sc9B1 cosmid. PCR amplification and sequencing confirmed that the expected "scar" sequence without any in-frame stop codons had been generated.

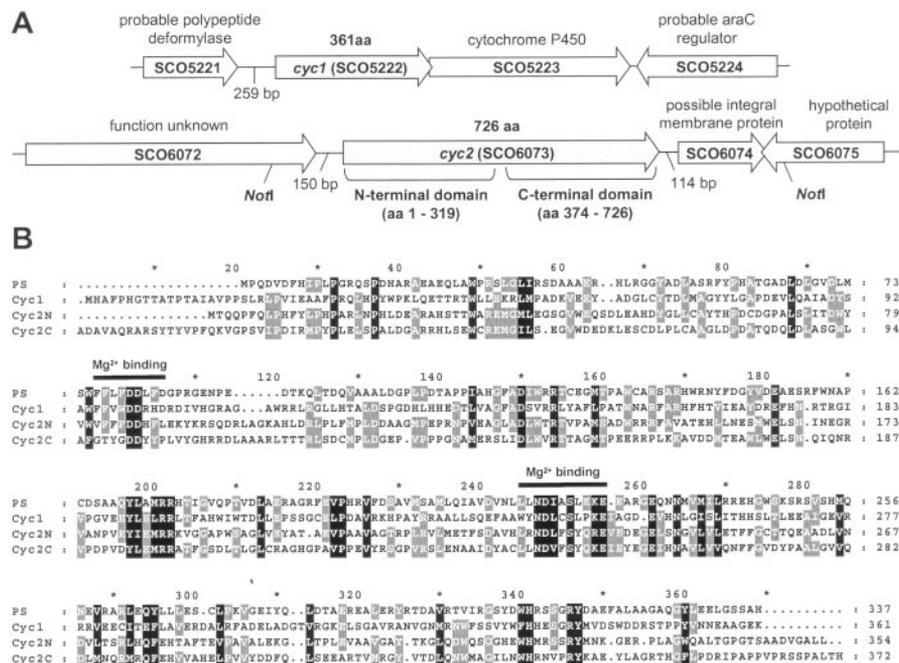


Fig. 2. Sesquiterpene synthase homologues *cyc1* and *cyc2* of *S. coelicolor*. (A) Flanking genes and their proposed functions. The stop codon of *cyc1* overlaps the start of SCO5223, suggesting the possibility of translational coupling of the two genes. SCO5223 encodes a cytochrome P450 enzyme that catalyzes sterol 14 α -demethylation (35). *cyc2* is flanked by SCO6072, encoding a 666-aa protein of unknown function, and SCO6074, encoding a 203-aa protein annotated as a possible integral membrane protein. (B) CLUSTALX alignment of Cyc1 and the N- and C-terminal domains of Cyc2 (Cyc2N, Cyc2C) to the 337-aa sequence of the pentalenene synthase (PS) from *Streptomyces* UC5319 reveals two conserved terpenoid synthase motifs. The residues Asp-80, Asp-81, and Asp-84 within the sequence ⁷⁶FLFLDLD of pentalenene synthase are believed to form an Mg²⁺-binding site (36). Mg²⁺ is required to facilitate pyrophosphate departure in the first step of the cyclization reaction. Cyc1 contains all three conserved aspartate residues, whereas both the N- and the C-terminal domains of Cyc2 contain only the first two aspartate residues. Residues within a second sequence of pentalenene synthase (²¹⁷LLNDIASLEKE) are also thought to be associated with binding of Mg²⁺ ions and are conserved in Cyc1 and both domains of Cyc2.

Sc9B1 containing the scar-marked *cyc2* deletion had only the Kan^R encoded by Supercos-1 for selection in *Streptomyces* and lacked *oriT*. After passage through *E. coli* ET12567, the plasmid was introduced into *S. coelicolor cyc2::Apra^RoriT* (strain J3001) by polyethylene glycol (PEG)-mediated protoplast transformation. About 30 Kan^R transformants were obtained with 100 ng CCC DNA. Nonselective propagation and screening for Apra^S transformants readily yielded an unmarked deletion of *cyc2* (strain J3003) containing a “scar” sequence of 81 bp (27 codons), which was confirmed by PCR and Southern blot analysis.

To confirm that the loss of geosmin production in the constructed *cyc2* mutants was directly caused by the *cyc2* deletions, we were able to restore geosmin production by reintroducing *cyc2* as a 3,574-bp *NotI* fragment in the ϕ C31 *attP* integrating vector pSET Ω (ref. 31; Table 3).

Only the Amino-Terminal Sesquiterpene Synthase Domain of Cyc2 Is Required for Geosmin Biosynthesis. *cyc2* contains two sesquiterpene synthase-like domains. To find out whether both were needed for geosmin synthesis, each domain was deleted separately in *S. coelicolor* M145 by using the FLP system to generate precisely tailored in-frame mutations, as described above. The amino-terminal deletion, *cyc2* Δ N (strain J3004), removed codons 6–340 and retained the ATG start codon, and the carboxyl-terminal deletion, *cyc2* Δ C (strain J3005), removed codons 380–721 and retained the stop codon. Geosmin production was abolished in J3004 (*cyc2* Δ N) but not in J3005 (*cyc2* Δ C), which produced similar amounts to wild-type *S. coelicolor* (Table 3). This result demonstrates that the C-terminal domain was not required for geosmin biosynthesis. The possibility that Cyc1 might complement the loss of the C-terminal domain of Cyc2 was

Table 3. Geosmin production by *S. coelicolor* mutants

Mutant strain	Genotype	Geosmin production	Complemented by	Geosmin production
KF62	<i>cyc1::Tn4560</i>	+		
J3001	Δ <i>cyc2</i>	–	<i>cyc2</i>	+
J3002	<i>cyc1::Tn4560, \Delta</i> <i>cyc2</i>	–	<i>cyc2</i>	+
J3003	Δ <i>cyc2scar</i>	–	<i>cyc2</i>	+
J3004	<i>cyc2</i> Δ N	–	<i>cyc2</i> or <i>cyc2</i> N terminus	+
J3005	<i>cyc2</i> Δ C	+		
J3006	<i>cyc1::Tn4560, cyc2</i> Δ N	–	<i>cyc2</i> or <i>cyc2</i> N terminus	+
J3007	<i>cyc1::Tn4560, cyc2</i> Δ C	+		

Production (+) was indicated by a peak corresponding to the geosmin standard within the total ion current chromatogram (TIC), and a mass spectrum (MS) with the expected $M_r = 182$ and a characteristic fragmentation pattern identical to that observed for the authentic standard. Where a geosmin peak was observed (+), the typical signal to noise ratio was ≈ 100 , although this varied from sample to sample. Where no geosmin peak was observed (–), no ions were detected above background with a mass spectrum similar to authentic geosmin.

eliminated because the constructed *cyc1,cyc2ΔC* double-mutant J3007 still produced geosmin. To verify that the N terminus of Cyc2 without the C terminus of Cyc2 was sufficient for normal geosmin production, the C-terminally deleted gene was cloned into pSETΩ (31) and introduced into the N-terminally deleted *cyc2* mutant. Geosmin production was restored (Table 3).

Discussion

Murphy *et al.* (6) showed that induction of the λ-Red genes (*gam*, *bet*, and *exo*) can significantly increase homologous recombination in *E. coli*. Datsenko and Wanner (2) developed this strategy into a rapid and efficient protocol, which they used to disrupt 13 genes on the *E. coli* chromosome. In other work, genes in a cosmid library have been targeted with PCR products, and the mutagenized cosmids were used subsequently to transform the filamentous fungus *Aspergillus nidulans* (32). Higher recombination frequencies resulted from increasing the length of homologous DNA flanking the transformation marker.

The inclusion of *oriT* from an IncP-group plasmid in the disruption cassette allowed conjugation to be used for intergeneric transfer of cosmid DNA from *E. coli* to *Streptomyces*. In comparison to transformation protocols described for *Streptomyces* (3), conjugation is much more efficient, and it is successful with many actinomycetes (33). It also avoids the processes of protoplasting and regeneration used during transformation procedures, which are often difficult to develop for other strains and which can cause genetic changes.

Typically, tens to hundreds of *E. coli* BW25113 transformants were obtained when using a PCR product with 39 bp of flanking DNA homologous to the targeted *S. coelicolor* sequence; this result reinforces previously published results (6, 27), which showed that 40–50 bp of sequence identity is sufficient to sustain recombination when Redα and Redβ are produced in combination with Redγ (which inhibits the RecBCD exonuclease V).

So far, more than 100 segments of the *S. coelicolor* genome ranging in size from 4 bp to over 7 kb have been replaced by PCR targeting by us alone or in collaboration with others (data not shown). On average, 10–30% of *S. coelicolor* exconjugants showed the desired allelic replacement. Interestingly, the frequency of double crossovers in *Streptomyces* remained higher than 5%, even when only 2–3 kb homology was present on one side and >30 kb on the other side of the gene replaced within the cosmid (results not shown).

In this study, this methodology has been applied both to replace the entire *cyc2* gene and to generate independent in-frame deletions in the regions of *cyc2* coding for the N- and C-terminal sesquiterpene synthase-like domains. In conjunction with a recently established transposon library (34), these mutations have enabled us to identify, for the first time, a gene

involved in the biosynthesis of the well known “earthy odor” metabolite geosmin. We have also shown that only the N-terminal domain of the Cyc2 protein is required for geosmin biosynthesis. Consistent with our findings, Cane and Watt report in the accompanying paper (22) that both the full-length Cyc2 protein and its N-terminal domain catalyze the cyclization of farnesyl pyrophosphate to germacra-1 (10)*E,5E*-dien-11-ol, whereas the recombinant C-terminal domain has no detectable sesquiterpene synthase activity *in vitro*.

These findings reinforce and extend incorporation experiments in *Streptomyces antibioticus*, which first indicated that geosmin is synthesized from farnesyl pyrophosphate via a sesquiterpene precursor (20), and the finding that a geosmin over-producing strain of *Streptomyces citreus* also produced considerable quantities of the sesquiterpene alcohol germacra-1 (10)*E,5E*-dien-11-ol, suggesting that this compound is the sesquiterpene precursor of geosmin (26). Taken together, these results suggest plausible mechanisms for geosmin biosynthesis in *Streptomyces* spp., which are explored in detail in the accompanying paper (22). The Cyc2 synthase is predicted to carry out cyclization of farnesyl pyrophosphate, to give germacra-1 (10)*E,5E*-dien-11-ol after capture of a tertiary carbocation intermediate by water and allylic transposition of the 2,3-double bond. Several further reactions, involving at least three additional enzymes (a cyclase, a reductase, and a hydroxylase) would be required to convert germacra-1 (10)*E,5E*-dien-11-ol into geosmin. Genes for the biosynthesis of *Streptomyces* secondary metabolites are usually clustered on the chromosome. Interestingly, the genes flanking *cyc2* (Fig. 2A) do not appear to code for enzymes likely to be involved in the conversion of germacradienol into geosmin, suggesting that genes for the biosynthesis of the secondary metabolite geosmin may, in fact, be scattered on the *S. coelicolor* chromosome.

In preliminary experiments, we have observed that certain *cyc2* mutations cause reduced sporulation. We anticipate that the further use of PCR-targeted mutagenesis will greatly facilitate further investigations of the role of geosmin in the life of streptomycetes.

We thank David Cane for prepublication information about the biochemical attributes of the SCO6073 (Sc9B1.20) gene product, Nick Le Brun at the University of East Anglia, United Kingdom, for help with GC-MS analysis, and Charlotte E. G. Schoeller at the Technical University of Denmark for access to her Ph.D. thesis, which contained much valuable background information about geosmin. This work was funded by Biotechnological and Biological Research Council (BBSRC) Grant 208/IGF12432 (to K.F.C.), Wellcome Trust Research Fellowship Grant 053086 (to G.L.C.), a BBSRC Research Studentship (to T.K.), a Competitive Strategic Grant from the BBSRC to the John Innes Centre, and the John Innes Foundation.

- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., *et al.* (2002) *Nature* **417**, 141–147.
- Datsenko, K. A. & Wanner, B. L. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645.
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000) *Practical Streptomyces Genetics* (John Innes Foundation, Norwich, U.K.).
- Zhang, Y., Muylers, J. P., Testa, G. & Stewart, A. F. (2000) *Nat. Biotechnol.* **18**, 1314–1317.
- Muylers, J. P., Zhang, Y. & Stewart, A. F. (2000) *Genet. Eng.* **22**, 77–98.
- Murphy, K. C., Campellone, K. G. & Potete, A. R. (2000) *Gene* **246**, 321–330.
- Muylers, J. P., Zhang, Y., Testa, G. & Stewart, A. F. (1999) *Nucleic Acids Res.* **27**, 1555–1557.
- Zhang, Y., Buchholz, F., Muylers, J. P. & Stewart, A. F. (1998) *Nat. Genet.* **20**, 123–128.
- Murphy, K. C. (1998) *J. Bacteriol.* **180**, 2063–2071.
- Redenbach, M., Kieser, H. M., Denapaita, D., Eichner, A., Cullum, J., Kinashi, H. & Hopwood, D. A. (1996) *Mol. Microbiol.* **21**, 77–96.
- Cherepanov, P. P. & Wackernagel, W. (1995) *Gene* **158**, 9–14.
- Gerber, N. N. (1979) *CRC Crit. Rev. Microbiol.* **7**, 191–214.
- Gerber, N. N. & Lechevalier, H. A. (1965) *Appl. Microbiol.* **13**, 935–938.
- Scholler, C. E., Gurtler, H., Pedersen, R., Molin, S. & Wilkins, K. (2002) *J. Agric. Food Chem.* **50**, 2615–2621.
- Saadoun, I., Schrader, K. K. & Blevins, W. T. (2001) *J. Basic Microbiol.* **41**, 51–55.
- Sauerwein, M. & Becker, H. (1990) *Planta Med.* **56**, 364–367.
- Kiviranta, H., Tuomainen, A., Reiman, M., Laitinen, S., Liesivuori, J. & Nevalainen, A. (1998) *Cent. Eur. J. Public Health* **6**, 296–299.
- Dionigi, C. P. (1993) *Food Flavour Safety*, 322–377.
- Wood, S., Williams, S. T. & White, W. R. (1983) *Int. Biodeterior. Bull.* **19**, 83–97.
- Bentley, R. & Meganathan, R. (1981) *FEBS Lett.* **125**, 220–222.
- Cane, D. E., Sohng, J. K., Lamberson, C. R., Rudnicki, S. M., Wu, Z., Lloyd, M. D., Oliver, J. S. & Hubbard, B. R. (1994) *Biochemistry* **33**, 5846–5857.
- Cane, D. E. & Watt, R. M. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 1547–1551.
- Paget, M. S., Chamberlin, L., Atrih, A., Foster, S. J. & Buttner, M. J. (1999) *J. Bacteriol.* **181**, 204–211.
- Kieser, T. & Melton, R. E. (1988) *Gene* **65**, 83–91.
- Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N. & Schoner, B. E. (1992) *Gene* **116**, 43–49.
- Pollak, F. C. & Berger, G. B. (1996) *Appl. Environ. Microbiol.* **62**, 1295–1299.
- Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G. & Court, D. L. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 5978–5983.
- Ainsa, J. A., Parry, H. D. & Chater, K. F. (1999) *Mol. Microbiol.* **34**, 607–619.
- MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H. & MacNeil, T. (1992) *Gene* **111**, 61–68.
- Flett, F., Mersinias, V. & Smith, C. P. (1997) *FEMS Microbiol. Lett.* **155**, 223–229.
- O'Connor, T. J., Kanellis, P. & Nodwell, J. R. (2002) *Mol. Microbiol.* **45**, 45–57.
- Chaveroche, M. K., Ghigo, J. M. & d'Enfert, C. (2000) *Nucleic Acids Res.* **28**, e97.
- Matsushima, P., Broughton, C. G., Turner, J. R. & Baltz, R. H. (1994) *Gene* **146**, 39–45.
- Fowler, K. (2002) Ph.D. thesis (John Innes Centre, Norwich, U.K.).
- Lamb, D. C., Fowler, K., Kieser, T., Manning, N., Podust, L. M., Waterman, M. R., Kelly, D. E. & Kelly, S. L. (2002) *Biochem. J.* **364**, 555–562.
- Lesburg, C. A., Zhai, G., Cane, D. E. & Christianson, D. W. (1997) *Science* **277**, 1820–1824.