

NovG, a DNA-binding protein acting as a positive regulator of novobiocin biosynthesis

Alessandra S. Eustáquio,¹ Shu-Ming Li¹ and Lutz Heide¹

Correspondence

Lutz Heide

heide@uni-tuebingen.de

¹Pharmazeutische Biologie, Pharmazeutisches Institut, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

The biosynthetic gene cluster of the aminocoumarin antibiotic novobiocin contains two putative regulatory genes, i.e. *novE* and *novG*. The predicted gene product of *novG* shows a putative helix–turn–helix DNA-binding motif and shares sequence similarity with StrR, a well-studied pathway-specific transcriptional activator of streptomycin biosynthesis. Here functional proof is provided, by genetic and biochemical approaches, for the role of NovG as a positive regulator of novobiocin biosynthesis. The entire novobiocin cluster of the producer organism *Streptomyces spheroides* was expressed in the heterologous host *Streptomyces coelicolor* M512, and additional strains were produced which lacked the *novG* gene within the heterologously expressed cluster. These Δ *novG* strains produced only 2% of the novobiocin formed by the *S. coelicolor* M512 strains carrying the intact novobiocin cluster. The production could be restored by introducing an intact copy of *novG* into the mutant. The presence of *novG* on a multicopy plasmid in the strain containing the intact cluster led to almost threefold overproduction of the antibiotic, suggesting that novobiocin biosynthesis is limited by the availability of NovG protein. Furthermore, purified N-terminal His₆-tagged NovG showed specific DNA-binding activity for the *novG*–*novH* and the *cloG*–*cloY* intergenic regions of the novobiocin and clorobiocin biosynthetic gene clusters, respectively. By comparing the DNA sequences of the fragments binding NovG, conserved inverted repeats were identified in both fragments, similar to those identified as the binding sites for StrR. The consensus sequence for the StrR and the putative NovG binding sites was GTTCTACTG(N)₁₁CRGTGTAAC. Therefore, NovG and StrR apparently belong to the same family of DNA-binding regulatory proteins.

Received 30 September 2004

Revised 24 January 2005

Accepted 24 February 2005

INTRODUCTION

Streptomycetes produce two-thirds of the clinically useful antibiotics of natural origin (Kieser *et al.*, 2000). Therefore, considerable efforts are directed towards understanding how these filamentous soil bacteria synthesize antibiotics, in order to manipulate the biosynthetic gene clusters and to produce novel compounds with improved properties. In addition, the identification and characterization of regulatory genes is crucial for the rational engineering of strains with enhanced antibiotic production.

Biosynthesis of secondary metabolites by *Streptomyces* species is a complex process involving several levels of regulation. In this respect, two phylogenetically distant species have received the most extensive attention so far: *Streptomyces coelicolor* A3(2) and *Streptomyces griseus*. The

former because of its early development as an excellent genetic system, which allowed the analysis of morphological differentiation and secondary metabolite formation, and the latter because it provided the first well-studied bacterial example of extracellular signalling by a hormone-like acylated lactone (the γ -butyrolactone A-factor) (Chater & Horinouchi, 2003). In both cases, pleiotropic regulatory genes affect antibiotic biosynthesis by influencing the expression of pathway-specific regulatory genes, which are usually clustered with the structural biosynthetic genes (Chater & Bibb, 1997).

In *S. griseus*, StrR has been characterized as a pathway-specific regulatory protein. It activates the expression of streptomycin biosynthetic genes by binding to DNA loci which are characterized by an inverted repeat with the consensus sequence GTTCGActG(N)₁₁CagTcGAAC, and located upstream of their respective promoter regions (Retzlaff & Distler, 1995). StrR has a putative helix–turn–helix (HTH) motif in the central region of its primary structure (Retzlaff & Distler, 1995), which is typical for a family of bacterial and phage DNA-binding proteins (Pabo & Sauer, 1992).

Abbreviations: DIG, HTH, helix–turn–helix; PHD, Profile network prediction Heidelberg; SARP, *Streptomyces* antibiotic regulatory protein.

The GenBank/EMBL/DDBJ accession numbers for the genes and DNA regions used in this study are AF170880 (novobiocin cluster) and AF329398 (clorobiocin cluster).

Analysis of ActII-ORF4 from *S. coelicolor* A3(2) (Arias *et al.*, 1999; Wietzorrek & Bibb, 1997) and of DnrI from *Streptomyces peucetius* (Sheldon *et al.*, 2002; Wietzorrek & Bibb, 1997) revealed another family of *Streptomyces* antibiotic regulatory proteins (SARPs). They show in their predicted secondary structure an OmpR-like DNA-binding domain with a different structure than the typical HTH motif (Sheldon *et al.*, 2002; Wietzorrek & Bibb, 1997). These proteins activate transcription of target genes by binding to DNA loci characterized by direct (rather than inverted) heptameric repeats with the consensus sequence TCGAG-CG/C located close to the transcriptional start sites (Arias *et al.*, 1999; Sheldon *et al.*, 2002; Wietzorrek & Bibb, 1997).

Furthermore, other types of pathway-specific regulatory genes have been identified from different *Streptomyces* strains, such as *srmR* of the spiramycin cluster in *Streptomyces ambofaciens*, the predicted product of which shows no significant sequence similarity to any other known regulatory protein; *mmvR* of the methylenomycin cluster in *S. coelicolor* A3(2), representing the first identified negative pathway-specific regulator of antibiotic production; and response regulator genes of two-component systems like *dnrN* of the daunorubicin cluster in *S. peucetius* and *redZ* of the undecylprodigiosin cluster in *S. coelicolor* A3(2) (Chater & Bibb, 1997).

The aminocoumarin antibiotics novobiocin, clorobiocin and coumermycin A₁ are very potent inhibitors of DNA gyrase produced by different *Streptomyces* strains (Maxwell, 1993). The biosynthetic gene clusters for these antibiotics have been cloned and sequenced (Pojer *et al.*, 2002; Steffensky *et al.*, 2000; Wang *et al.*, 2000), which allowed detailed investigations of the biosynthetic pathways (reviewed by Li & Heide, 2004) as well as the generation of novel antibiotics by metabolic engineering (Eustáquio *et al.*, 2003a, 2004), chemo-enzymic synthesis (Xu *et al.*, 2004) and precursor-directed biosynthesis (Galm *et al.*, 2004).

In contrast, little is known about how aminocoumarin antibiotic production is regulated. *novE* is likely to have a positive regulatory function in novobiocin biosynthesis, as shown by an inactivation experiment (Eustáquio *et al.*, 2003b). The predicted gene product of *novE* shows sequence similarity only to LmbU, which may be involved in the regulation of lincomycin biosynthesis (Peschke *et al.*, 1995). The present study concentrated on another putative regulatory gene of novobiocin biosynthesis, i.e. *novG*. NovG shows sequence similarity to StrR, the well-established positive regulator of streptomycin biosynthesis (Retzlaff & Distler, 1995).

The aims of the present study were to provide functional proof, by genetic and biochemical approaches, for the role of *novG* as a regulator of novobiocin biosynthesis, and to investigate whether overexpression of *novG* can be used to upregulate novobiocin production.

METHODS

Bacterial strains, plasmids, cosmids and culture conditions.

Bacterial strains, plasmids and cosmids used in this study are listed in Table 1. The REDIRECT technology kit for PCR targeting (Gust *et al.*, 2003) was obtained from Plant Bioscience (Norwich, UK).

S. coelicolor M512 strains were routinely cultured in baffled Erlenmeyer flasks containing a stainless steel spring. YMG medium containing 1 % malt extract, 0.4 % yeast extract and 0.4 % glucose (pH 7.3) was used routinely for cultivation, carried out at 30 °C and 200 r.p.m. for 2 days. For preparation of protoplasts or isolation of genomic DNA, strains were grown in YEME medium (Kieser *et al.*, 2000). For analysis of secondary metabolites, cells from a YMG pre-culture (1 ml) were inoculated in 50 ml CDM production medium (Kominek, 1972) and cultured at 30 °C and 200 r.p.m. for 7 days. The cultivation in production medium was carried out without addition of antibiotics. *Escherichia coli* strains were cultured in LB medium as described by Sambrook & Russell (2001).

Kanamycin (15 µg ml⁻¹ in liquid medium and 50 µg ml⁻¹ in solid medium for *Streptomyces* species; 50 µg ml⁻¹ for *E. coli*), chloramphenicol (25–50 µg ml⁻¹), apramycin (50 µg ml⁻¹), carbenicillin (50–100 µg ml⁻¹) and thiostrepton (15 µg ml⁻¹ in liquid medium and 50 µg ml⁻¹ in solid medium) were used for selection of recombinant strains.

Before transformation of *S. coelicolor* M512 strains, the recombinant plasmids and cosmids were amplified in *E. coli* ET12567 to avoid methyl-sensing restriction (MacNeil *et al.*, 1992).

Plasmid construction. All the plasmid constructions are summarized in Table 1. For the construction of the NovG expression plasmid pAE-G5, *novG* was amplified by PCR using cosmid nov-BG1 as template and the primer pair PnovG_f (5'-TGG GGA TCC CAT GAC CAA CAG-3') and PnovG_r (5'-GAT TCA AGC TTT TGA ACG TCA GG-3'); bold letters represent mutations inserted in comparison to the original sequence to give the underlined restriction sites *Bam*HI and *Hind*III, respectively. The PCR reaction was carried out in 50 µl volume with 100 ng template, 0.2 mM dNTPs, 50 pmol each primer, 5 % (v/v) DMSO, using the Expand High Fidelity PCR System (Roche Molecular Biochemicals): denaturation at 94 °C for 2 min, then 30 cycles with denaturation at 94 °C for 45 s, annealing at 55 °C for 45 s, extension at 72 °C for 90 s, and a final elongation step at 72 °C for 5 min. After purification, the PCR product was ligated into pGEM-T to give pAE-G4. The insert of pAE-G4 was checked by nucleotide sequencing, and the *Bam*HI–*Hind*III fragment was ligated into the same sites of pRSET B to give pAE-G5.

DNA isolation, manipulation and cloning. Standard procedures for DNA isolation and manipulation were performed as described by Sambrook & Russell (2001) and Kieser *et al.* (2000). Isolation of DNA fragments from agarose gel and purification of PCR products were carried out with the NucleoSpin 2 in 1 Extract Kit (Macherey–Nagel). Isolation of cosmids and plasmids from *E. coli* was carried out with ion-exchange columns (Nucleobond AX kits, Macherey–Nagel) according to the manufacturer's protocol. Isolation of plasmid DNA from *Streptomyces* strains was carried out by alkaline lysis and potassium acetate precipitation, procedure D (Kieser *et al.*, 2000). If required, the plasmid DNA isolated from *Streptomyces* species was amplified in *E. coli* XL-1 Blue MRF' before restriction analysis.

Genomic DNA was isolated from *S. coelicolor* M512 strains using the Kirby mix procedure (Kieser *et al.*, 2000). Southern blot analysis was performed on Hybond-N nylon membrane (Amersham Biosciences) with a digoxigenin-labelled probe by using the DIG high prime DNA labelling and detection starter kit II (Roche Molecular Biochemicals).

Table 1. Bacterial strains, plasmids and cosmids

Strain, plasmid or cosmid	Description*	Reference or source
Strain		
<i>S. coelicolor</i> M512	$\Delta redD \Delta actII-ORF4$ SCP1 [−] SCP2 [−]	Florianio & Bibb (1996)
<i>S. coelicolor</i> (nov-BG1)	<i>S. coelicolor</i> M512 containing the novobiocin cluster, Km ^r	Eustáquio <i>et al.</i> (2004)
<i>S. coelicolor</i> (nov-BG1)pAE-G2_1	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-BG1)pAE-G2_2	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-BG1)pAE8	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-BG1)pWHM3	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-AE10)	<i>S. coelicolor</i> M512 containing a <i>novG</i> -defective novobiocin cluster, Km ^r	This work
<i>S. coelicolor</i> (nov-AE10) pAE-G2_1	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-AE10) pAE-G2_2	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-AE10) pAE8	Km ^r Thio ^r	This work
<i>S. coelicolor</i> (nov-AE10) pWHM3	Km ^r Thio ^r	This work
<i>E. coli</i> XL-1 Blue MRF'	General cloning host, Tet ^r	Stratagene
<i>E. coli</i> ET12567	Strain triply defective in DNA methylation (<i>dam dcm hsdM</i>), Cm ^r	MacNeil <i>et al.</i> (1992)
<i>E. coli</i> BL21(DE3)pLysS	Host for the heterologous expression of His ₆ -tagged NovG, Cm ^r	Novagen
Plasmid or cosmid		
pGEM-T	Cloning vector for PCR products, Amp ^r	Promega
pBluescript SK(−)	Cloning vector, Amp ^r	Stratagene
pLitmus28	Cloning vector, Amp ^r	New England Biolabs
pSPORT1	Cloning vector, Amp ^r	Invitrogen
pRSET B	Expression vector, T7 promoter, N-terminal His ₆ tag, Amp ^r	Invitrogen
pWHM3	<i>E. coli</i> – <i>Streptomyces</i> shuttle vector, Amp ^r Thio ^r	Vara <i>et al.</i> (1989)
nov-BG1	Cosmid containing the novobiocin biosynthetic gene cluster in a modified SuperCos1 vector, Km ^r	Eustáquio <i>et al.</i> (2004)
nov-AE10	nov-BG1($\Delta novG$) Km ^r	This work
pMS33	1.43kb <i>Bam</i> HI– <i>Eco</i> RI fragment (position 6393–7821, AF170880) of cosmid 9-6G in pBluescript SK(−), Amp ^r	Steffensky <i>et al.</i> (2000)
pAE-G1	1.19kb <i>Aat</i> II– <i>Eco</i> RI fragment (<i>novG</i> , position 6628–7821, AF170880) of pMS33 in the same sites of pLitmus28, containing 103bp before the putative start codon of <i>novG</i> , Amp ^r	This work
pAE-G2_1	1.20kb <i>Xba</i> I– <i>Spe</i> I fragment of pAE-G1 (<i>novG</i>) in the <i>Xba</i> I site of pWHM3 (against <i>lacZ</i> orientation), Amp ^r Thio ^r	This work
pAE-G2_2	1.20kb <i>Xba</i> I– <i>Spe</i> I fragment of pAE-G1 (<i>novG</i>) in the <i>Xba</i> I site of pWHM3 (same orientation as <i>lacZ</i>), Amp ^r Thio ^r	This work
pAE8	1.43kb <i>Bam</i> HI– <i>Eco</i> RI fragment of pMS33 (<i>novG</i> , position 6393–7821, AF170880) in the same sites of pWHM3, containing 336bp before the putative start codon of <i>novG</i> , Amp ^r Thio ^r	This work
pAE7	209bp <i>Eco</i> RI– <i>Sal</i> I fragment of pMS33 (position 7613–7821, AF170880) in the same sites of pBluescript SK(−), containing the NovG binding site (<i>novH</i>), Amp ^r	This work
Cosmid K1F2	Clorobiocin cluster in SuperCos1, Amp ^r Km ^r	Pojer <i>et al.</i> (2002)
pAE9	235bp <i>Not</i> I– <i>Mlu</i> I fragment of cosmid K1F2 (position 12 654–12 888, AF329398) in the same sites of pSPORT1, containing the NovG binding site (<i>cloY</i>) Amp ^r	This work
pAE-G4	PCR fragment (<i>novG</i>) in pGEM-T (see Methods)	This work
pAE-G5	PCR fragment (<i>novG</i>) in pRSET B (see Methods)	This work

*Km^r, kanamycin-; Thio^r, thiostrepton-; Tet^r, tetracycline-; Cm^r, chloramphenicol-; Amp^r, ampicillin-resistant.

Heterologous expression of the novobiocin biosynthetic gene cluster. In cosmid 10-9C, containing the complete novobiocin biosynthetic gene cluster, the ampicillin resistance gene *bla* of the SuperCos1 backbone was replaced, using λ -Red-mediated

recombination, with a cassette from pIJ787 containing the integrase gene, *int*, and attachment site, *attP*, of phage ϕ C31, as well as a tetracycline resistance gene, *tet*, giving nov-BG1. Cosmid nov-BG1, still carrying the kanamycin resistance gene *neo*, was then introduced

into *S. coelicolor* M512 via PEG-mediated protoplast transformation (Kieser *et al.*, 2000). Kanamycin-resistant clones were checked for specific genomic integration of the cosmid into the ϕ C31 attachment site by Southern blot analysis. For details, see Eustáquio *et al.* (2004, 2005).

Inactivation of *novG* in cosmid nov-BG1, and heterologous expression of the Δ *novG* cosmid. In cosmid nov-BG1, *novG* was replaced, via λ -Red-mediated recombination (Gust *et al.*, 2003), by an apramycin resistance (*aac(3)IV*) cassette from pUG019 (Eustáquio *et al.*, 2004, 2005), which is flanked by *Xba*I and *Spe*I recognition sites. The cassette for replacement of *novG* was generated by PCR using the primer pair P1-*novG* (5'-GAT CCG GAC CAG ACC ATT AAG TCC TAT GGG GGT TAC ATG ATT CCG GGG ATC TCT AGA TC-3') and P2-*novG* (5'-CAA CCG AAT GAT TCG AGC AGT TGA ACG TCA GGC GGT GTC ACT AGT CTG GAG CTG CTT C-3'). Underlined letters represent 39 nt homologous extensions to the DNA regions immediately upstream and downstream of *novG*, including the putative translational start and stop codons of *novG*, respectively; the *Xba*I and *Spe*I restriction sites are presented in bold letters. The PCR reaction was carried out in 50 μ l volume with 100 ng template (pUG019 digested with *Eco*RI, *Hind*III and *Dra*I), 0.2 mM dNTPs, 50 pmol each primer, 5% (v/v) DMSO, using the Expand High Fidelity PCR System (Roche Molecular Biochemicals): denaturation at 94 °C for 2 min; 10 cycles with denaturation at 94 °C for 45 s, annealing at 45 °C for 45 s and extension at 72 °C for 90 s; 15 cycles with annealing at 48 °C, and a final elongation step at 72 °C for 5 min.

After isolation from the non-methylating *E. coli* ET12567, cosmid DNA was digested with *Xba*I and *Spe*I and ligated overnight at 4 °C. *E. coli* XL-1 Blue MRF' cells were transformed with 100 ng DNA. Apramycin-sensitive, kanamycin-resistant clones were analysed using restriction enzymes and gel electrophoresis. The generated *novG*⁻ cosmid nov-AE10, carrying the kanamycin resistance gene *neo*, was introduced into *S. coelicolor* M512 by PEG-mediated protoplast transformation (Kieser *et al.*, 2000). Kanamycin-resistant clones were checked for specific genomic integration of cosmid nov-AE10 into ϕ C31 *attB* by Southern blot analysis.

Complementation with plasmids pAE-G2_1, pAE-G2_2, pAE8 or pWHM3. Introduction of plasmids pAE-G2_1, pAE-G2_2, pAE8 or pWHM3 into the Δ *novG* mutant [*S. coelicolor*(nov-AE10)], or into *S. coelicolor*(nov-BG1) was carried out by PEG-mediated protoplast transformation (Kieser *et al.*, 2000).

Production and analysis of secondary metabolites. Integration mutants, transformants and the parental strains of *S. coelicolor* M512 were cultured as described above, and assayed for novobiocin production by HPLC as described by Eustáquio *et al.* (2003b).

Negative-ion FAB mass spectra were recorded on a TSQ70 spectrometer (Finnigan), using diethanolamine as matrix. The substance isolated from the Δ *novG* mutants gave the following signals: *m/z* (relative intensity in %): 611 (4, [M-H]⁻), 568 (2, [M-CONH]⁻), 416 (7), 394 (14), 255 (100), 209 (74), corresponding to novobiocin (C₃₁H₃₆N₂O₁₁; molecular mass, 612).

Expression and purification of His₆-tagged NovG. Five litres of LB medium (Sambrook & Russell, 2001) containing 50 μ g carbenicillin ml⁻¹ and 35 μ g chloramphenicol ml⁻¹ was inoculated with *E. coli* BL21(DE3)pLysS(pAE-G5) cells; sufficient inoculum was used to reach an initial OD₆₀₀ of 0.1. Cells were then cultured at 30 °C. When the OD₆₀₀ reached 0.6, IPTG was added to a final concentration of 0.5 mM. After further growth for 1.5 h at 30 °C, cells were harvested by centrifugation and frozen at -70 °C. All subsequent steps were carried out at 4 °C. After thawing on ice, cells (16 g) were suspended in 20 ml lysis buffer (50 mM NaH₂PO₄,

pH 8.0, 300 mM NaCl, 10 mM imidazole, 2 mg lysozyme ml⁻¹) and incubated on ice for 30 min. The cell suspension was sonicated for 7 \times 30 s with 30 s intervals between each treatment (Branson Sonifier 250). Cellular debris was removed by centrifugation (17 500 g for 30 min). Three millilitres of Ni-NTA-agarose slurry [50% (w/v) nickel-nitrilotriacetic acid agarose resin suspension in 30% (v/v) ethanol, precharged with Ni²⁺] (Qiagen) was added to 3 ml lysis buffer and stirred gently for about 15 min. The supernatant was added to the Ni-NTA-agarose mixture described above and stirred gently for 60 min. Ten millilitres of washing buffer (50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 20 mM imidazole) were added and the protein-Ni-NTA-agarose mixture was harvested by centrifugation (5000 g for 10 min at 4 °C) and washed twice with 30 ml washing buffer (5000 g for 10 min at 4 °C) to remove unbound proteins. The pellet was suspended in washing buffer and loaded into a column. After further washing with 3 \times 4 ml washing buffer containing 50 mM imidazole, unspecifically bound proteins were eluted using a stepwise imidazole gradient: 2 \times 1 ml 100 mM, 1 \times 1 ml 150 mM and 1 \times 1 ml 200 mM imidazole in washing buffer. The NovG fusion protein was subsequently eluted with 4 ml elution buffer (50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 250 mM imidazole). Aliquots (10 μ l) were analysed by SDS-PAGE, carried out according to the method of Laemmli (1970), and protein bands were stained with Coomassie brilliant blue R-250. Eluate fractions (1 ml) containing the NovG fusion protein (38.1 kDa) were applied onto a NAP 10 column (Amersham Biosciences) and eluted with 1.2 ml storage buffer [25 mM Tris/HCl (pH 7.5), 10% (w/v) glycerol, 2 mM DTT]. Aliquots of NovG fusion protein in storage buffer were either frozen in liquid nitrogen and stored at -70 °C, or used immediately in DNA-binding assays.

Preparation of 3'-end DIG-labelled DNA fragments. The following DNA fragments were used in gel mobility-shift assays: (1) the 325 bp *Ava*I fragment obtained from pMS32 (position 4463–4787 in AF170880, i.e. upstream of *novE*); (2) the 341 bp *Pvu*I–*Sal*I fragment obtained from pMS32 (position 4231–4571 in AF170880, i.e. upstream of *novE*); (3) the 272 bp PCR product obtained using pMS63 as template and the primer pair Pnov\$F_f (5'-AGG ACC ACT GGC TCG ATT TCG-3') and Pnov\$F_r (5'-GTC ACG CGC GAA GCC GTG AG-3') (position 5238–5509 in AF170880, i.e. upstream of *novF*); (4) the 265 bp PCR product obtained using pMS33 as template and the primer pair Pnov\$G_f (5'-GAG CTG GCC CGC CTC TTC GA-3') and Pnov\$G_r (5'-ACT TAA TGG TCT GGT CCG GAT CG-3') (position 6451–6715 in AF170880, i.e. upstream of *novG*); (5) the 180 bp *Van*9II–*Hind*III fragment obtained from pMS63 (position 7725–7904 in AF170880, i.e. upstream of *novH*); (6) 209 bp *Sal*I–*Eco*RI fragment obtained from pMS33 (position 7613–7821 in AF170880, i.e. upstream of *novH*); (7) the 244 bp *Bam*HI fragment obtained from pMS61 (position 17 264–17 507 in AF170880, i.e. upstream of *novO*); (8) the 252 bp PCR product obtained using pMS61 as template and the primer pair Pnov\$O_f (5'-TGT ACG AGC TGC TCA CCC ACG-3') and Pnov\$O_r (5'-TGA ATT GAG CCT ACA CGG ACA C-3') (position 17 101–17 352 in AF170880, i.e. upstream of *novO*); (9) the 231 bp PCR product obtained using pMS62 as template and the primer pair Pgyr\$f (5'-GCG CAG AGG TGC TCT CGT TCC-3') and Pgyr\$r (5'-TGC GGG TGT CGT AAG AAG TCA C-3') (position 168–398 in AF205854, i.e. upstream of *gyrB*^R); (10) the 277 bp *Sma*I fragment obtained from pMS62 (position 25 608–25 617 in AF170880 and 4–270 in AF205854, i.e. upstream of *gyrB*^R); (11) the 235 bp *Not*I–*Mlu*I fragment obtained from pAE9 (Table 1) (position 12 654–12 888 in AF329398, i.e. upstream of *cloY*); (12) the 255 bp PCR product obtained using cosmid K1F2 (Table 1) as template and the primer pair Pclo\$H₁f (5'-GAA CGG CTC CTA TCT GGT CC-3') and Pclo\$H₁r (5'-CGG CCT TCG AAC AAC CTT CG-3') (position 12 946–13 200 in AF329398, i.e. upstream of *cloH*); and (13) the 260 bp PCR product obtained using cosmid K1F2 as

template and the primer pair Pclo\$H_2f\$ (5'-TGG GTG GCG AGT AGC ATC TG-3') and Pclo\$H_2r\$ (5'-CTT AAG TCT CCA TGC CAT TGG-3') (position 13 122–13 381 in AF329398, i.e. upstream of *cloH*). PCR products were obtained using the Expand High Fidelity PCR System (Roche Molecular Biochemicals) as described above ('Plasmid construction'). pMS plasmids are subclones of cosmid 10-9C or 9-6G (Steffensky *et al.*, 2000).

After purification by NuSieve GTG-agarose (FMC BioProducts) gel electrophoresis, DNA fragments were 3'-end-labelled with DIG-11-ddUTP using the DIG Gel Shift Kit, 2nd Generation (Roche Molecular Biochemicals) according to the manufacturer's instructions.

Gel mobility-shift assays. These assays were performed using the DIG Gel Shift Kit, 2nd Generation (Roche Molecular Biochemicals) according to the manufacturer's instructions. The DNA-binding reaction conditions were adapted from Retzlaff & Distler (1995). The reaction was carried out at 25 °C in 20 µl 12.5 mM Tris/HCl (pH 7.5), 5 % (w/v) glycerol, 62.5 mM KCl, 1 mM DTT, 5 mM MgCl₂, 50 ng poly [d(I-C)] µl⁻¹, 5 ng poly L-lysine µl⁻¹. About 4 ng DIG-labelled DNA fragment and approximately 0.5 µg of purified His₆-tagged NovG were used for each assay. For testing the specificity of binding, competitor plasmid DNA or the respective empty vector were added in approximately 125-fold molar excess in comparison to the labelled fragment. After 15 min incubation, the reaction mixture was applied to a pre-run (20 min at 35 V) 6 % (w/v) native polyacrylamide gel (85 × 75 × 0.75 mm) with 0.5 × TBE (Sambrook & Russell, 2001) as running buffer. The gel was run at 35 V for about 3 h and transferred to a positively charged Hybond-N+ nylon membrane (Amersham Biosciences) by contact blotting. Cross-linking and detection were carried out following the manufacturer's instructions.

Computer-assisted sequence analysis. The DNASIS software package (version 2.1; Hitachi Software Engineering, San Bruno, California) and the BLAST program were used for sequence analysis and for homology searches in the GeneBank database, respectively. The secondary structure of NovG was predicted using the PHD (Profile network prediction Heidelberg) method (Rost, 1996) (available on the web at <http://www.embl-heidelberg.de/predictprotein>).

RESULTS

Sequence analysis of *novG*

novG encodes a protein which comprises 318 aa and shows 41 % identity at the amino acid level to StrR, the pathway-specific transcriptional activator of streptomycin biosynthesis in *Streptomyces glaucescens* and *S. griseus* (Retzlaff &

Distler, 1995; Thamm & Distler, 1997). The predicted amino acid sequence of NovG shows a putative HTH motif in the central region of the protein, which is typical for a family of bacterial and phage DNA-binding proteins (Pabo & Sauer, 1992). This motif is also found in StrR (Retzlaff & Distler, 1995; Thamm & Distler, 1997). Prediction of the secondary structure of NovG using the PHD method confirmed the presence of an HTH motif (Fig. 1).

novG contains the rare TTA leucine codon (codon 182), which indicates its dependence on *blaD*, the structural gene for tRNA^{UUA}. This may represent a further indication for a regulatory role of NovG in novobiocin biosynthesis, since most of the known TTA-containing genes specify regulatory or resistance proteins associated with biosynthetic gene clusters for antibiotics (Leskiw *et al.*, 1991). Orthologues of *novG* are found in the biosynthetic gene clusters of the aminocoumarin antibiotics clorobiocin and coumermycin (Pojer *et al.*, 2002). They are termed *cloG* and *couG* and both code for proteins of 319 aa which share a mean 78 % sequence identity with *novG*. Just as for *novG*, both *cloG* and *couG* show the putative HTH motif in their central region. *cloG* contains three TTA codons (codons 28, 83 and 149), whereas *couG* contains none.

Inactivation of *novG*

In order to investigate the function of *novG*, we first examined the effects of its inactivation and its over-expression. Since the natural novobiocin producer *Streptomyces spheroides* is difficult to manipulate genetically (Hussain & Ritchie, 1991), we expressed the biosynthetic gene cluster of novobiocin in *S. coelicolor* M512 by previously described methods (Eustáquio *et al.*, 2004, 2005). As shown in Fig. 2(a), cosmid nov-BG1, containing the entire novobiocin cluster as well as *int* and *attP* of phage ϕ C31, was used for this purpose.

Cosmid nov-BG1 was introduced into *S. coelicolor* M512 (Floriano & Bibb, 1996) by PEG-mediated protoplast transformation and integrated into the *attB* site of the chromosome. Southern blot analysis confirmed the site-specific integration of the entire cosmid into the genome of the heterologous host (Fig. 2b, d).

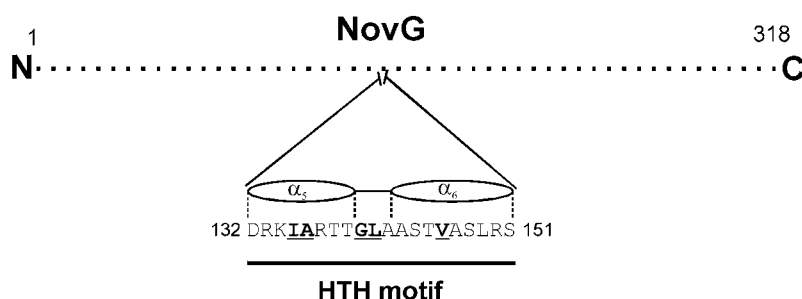


Fig. 1. Putative helix–turn–helix (HTH) motif of NovG. The most highly conserved residues in the HTH motif are underlined and include a glycine (G) in the turn and several hydrophobic residues (I, A, L, V) (Pabo & Sauer, 1992). The secondary structure was predicted using the PHD method (Rost, 1996; <http://www.embl-heidelberg.de/predictprotein>).

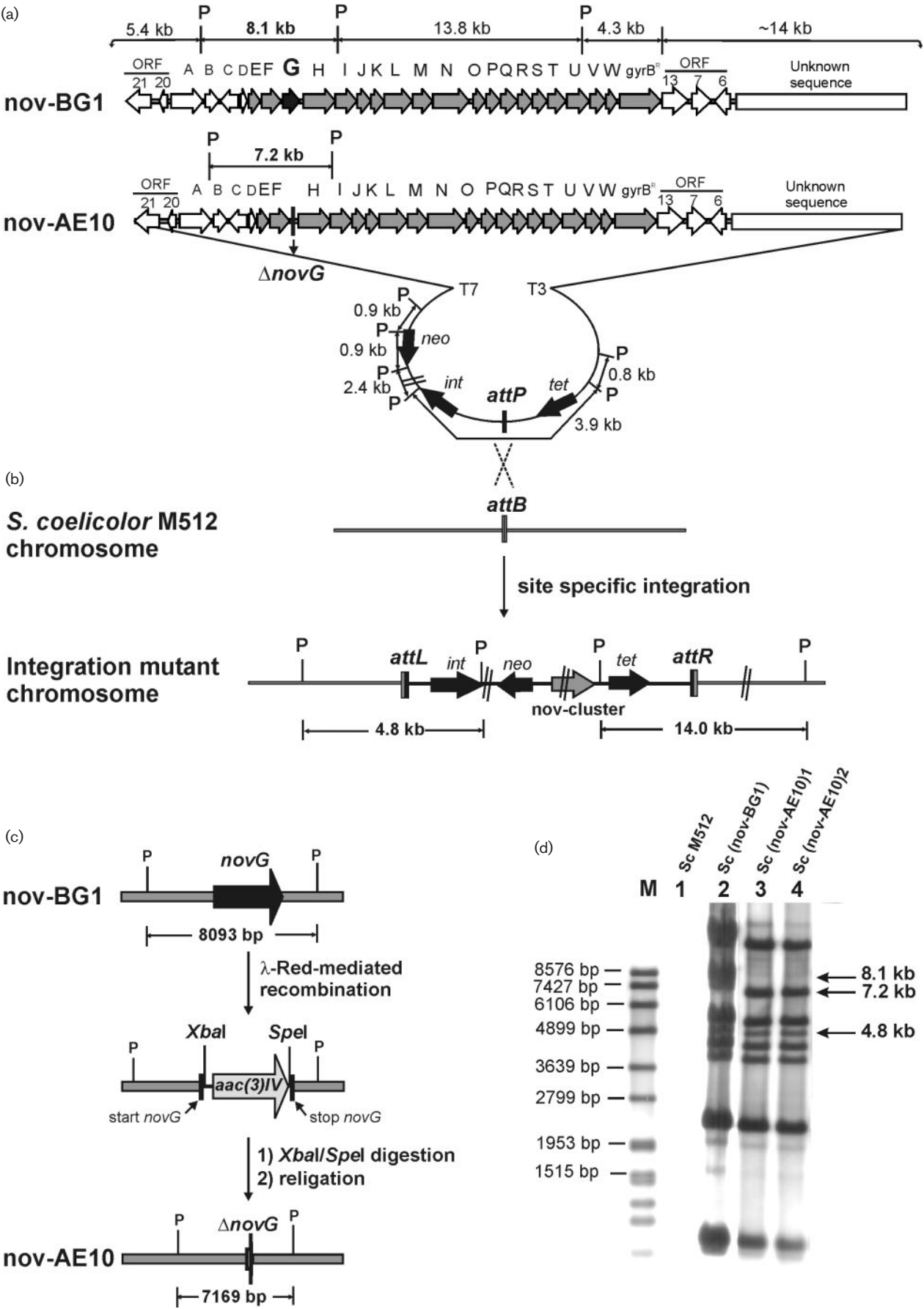


Fig. 2. Inactivation of *novG*. (a) Schematic representation of cosmids nov-BG1, containing the intact novobiocin biosynthetic gene cluster, and nov-AE10, in which *novG* was deleted. The novobiocin cluster is presented in grey, the flanking regions in white and *novG* in black. P, *Pst*I restriction site; T3, T7, promoters of the SuperCos1 vector; *tet*, tetracycline resistance gene; *neo*, neomycin/kanamycin resistance gene; *int*, *attP*, integrase and attachment sites of phage ϕ C31. Cosmid backbone not to scale. (b) Schematic representation of site-specific integration. The integrase, *int*, derived from *Streptomyces* phage ϕ C31, catalyses integration via recombination between *attP* (from phage or vector) and *attB* (from *Streptomyces* genome) sites, generating the hybrid sites *attL* and *attR* (Thorpe *et al.*, 2000). The sizes of junction fragments which allow proof of site-specific integration into the ϕ C31 attachment site of the *S. coelicolor* genome are indicated (4.8 and 14.0 kb). Not to scale. (c) Schematic representation of *novG* inactivation. *novG* (957 bp) was replaced by an apramycin resistance (*aac(3)IV*) cassette. The cassette was then excised by digestion with *Xba*I and *Spe*I and religated, leaving an in-frame 'scar' of 18 nt between the start and stop codons of *novG*. Not to scale. (d) Southern blot analysis of *S. coelicolor* M512 (lane 1) and M512-derived integration mutants bearing cosmid nov-BG1 (lane 2) or the *novG*[−] cosmid nov-AE10 (lanes 3 and 4). M, DIG-labelled DNA molecular mass Marker VII (Roche). Genomic DNA was digested with *Pst*I. The DIG-labelled cosmid nov-BG1 was used as a probe. The *novG* inactivation is confirmed by the 7.2 kb band (lane 3 and 4) instead of an 8.1 kb band (lane 2); otherwise the band pattern is expected to be identical.

The analysis of secondary metabolites by HPLC showed that the integration mutants, in contrast to the parental strain *S. coelicolor* M512, accumulated novobiocin. The identity of the produced novobiocin was confirmed by ¹H-NMR and negative-ion FAB MS analysis (molecular ion [*M*−*H*][−] at *m/z* 611; novobiocin, C₃₁H₃₆N₂O₁₁, has a molecular mass of 612). The obtained spectroscopic data were identical to those from authentic novobiocin.

Functional investigations of *novG* could now be carried out by modification of the cosmid nov-BG1 prior to its introduction into *S. coelicolor* M512. Thus, we generated a Δ *novG* strain by deletion of *novG* in cosmid nov-BG1 and introduction of this modified cosmid into the genome of *S. coelicolor* M512. For this purpose, *novG* was replaced by an apramycin resistance cassette flanked by *Xba*I and *Spe*I recognition sites via λ -Red-mediated recombination (Datsenko & Wanner, 2000; Gust *et al.*, 2003). The cassette was then removed by digestion with *Xba*I and *Spe*I, enzymes which create compatible ends, allowing re-ligation of the outer ends and consequent excision of the cassette (Fig. 2c). The modified cosmid (named nov-AE10) was introduced into *S. coelicolor* M512 by protoplast transformation.

Southern blot analysis confirmed the site-specific integration into the genome, and the deletion of *novG* was clearly shown by the size of the relevant *Pst*I restriction fragment in comparison to nov-BG1 strains (Fig. 2d, lanes 2–4).

Analysis of secondary metabolites by HPLC showed that the resulting Δ *novG* strains still produced novobiocin, which was identified by negative-ion FAB MS analysis in comparison to an authentic standard (molecular ion [*M*−*H*][−] at *m/z* 611; novobiocin, C₃₁H₃₆N₂O₁₁, has a molecular mass of 612). However, the amount of novobiocin produced by the Δ *novG* mutants was reduced by 98 % in comparison to *S. coelicolor* M512 strains carrying the intact novobiocin cluster (Table 2).

Complementation of the *novG* mutation

To prove that the inactivation of *novG* was indeed responsible for the low productivity of the Δ *novG* strains, we attempted to complement one of these strains by expressing *novG* under the control of its own promoter. For this purpose, we used a derivative of the promoter-less, replicative shuttle vector pWHM3 (Vara *et al.*, 1989), into

Table 2. Effect of *novG* inactivation and expression on novobiocin production by recombinant *S. coelicolor* strains carrying the novobiocin cluster

Strain	Expression plasmid	Novobiocin production*	
		(mg l ^{−1})	(%)
<i>S. coelicolor</i> M512	–	0	0
<i>S. coelicolor</i> (nov-BG1), complete cluster	pWHM3 (empty vector)	20	100
<i>S. coelicolor</i> (nov-AE10), Δ <i>novG</i>	pWHM3 (empty vector)	0.4	2
<i>S. coelicolor</i> (nov-AE10), Δ <i>novG</i>	pAE-G2_1 (<i>novG</i> with 100 bp upstream sequence, opposite orientation to <i>lacZ</i>)	0.4	2
<i>S. coelicolor</i> (nov-AE10), Δ <i>novG</i>	pAE-G2_2 (<i>novG</i> with 100 bp upstream sequence)	3	15
<i>S. coelicolor</i> (nov-AE10), Δ <i>novG</i>	pAE8 (<i>novG</i> with 336 bp upstream sequence)	16	79
<i>S. coelicolor</i> (nov-BG1), complete cluster	pAE-G2_2 (<i>novG</i> with 100 bp upstream sequence)	38	190
<i>S. coelicolor</i> (nov-BG1), complete cluster	pAE8 (<i>novG</i> with 336 bp upstream sequence)	54	270

*Values are means from two to four independent mutants with two determinations each.

which the coding region of *novG* and 336 bp of the DNA region upstream of the putative translational start point was cloned, resulting in pAE8.

Transformants of the Δ *novG* strain carrying pAE8 produced on average 80 % of the novobiocin amount accumulated by *S. coelicolor* M512 strains bearing the intact novobiocin cluster (Table 2). Therefore, complementation was successful, establishing that the low productivity of Δ *novG* strains was due to lack of *novG* and not to possible polar effects of the deletion on downstream genes.

Notably, a pWHM3 construct containing *novG* with only 100 bp of the upstream DNA region (termed pAE-G2_2) was much less effective in restoration of novobiocin productivity (Table 2). If the orientation of its insert was reversed (i.e. placed against the *lacZ* orientation), the resulting plasmid pAE-G2_1 was completely unable to enhance novobiocin production. Apparently, these constructs did not contain a functional promoter upstream of *novG*, and the low increase of productivity observed in pAE-G2_2 transformants may be due to low activity of the *lacZ* promoter in the *S. coelicolor* host.

Overexpression of *novG* in *S. coelicolor* (*nov-BG1*) leads to overproduction of novobiocin

The results presented above indicate that NovG could act as a positive regulator in novobiocin biosynthesis. It has been reported that overexpression of pathway-specific activators can lead to overproduction of the respective antibiotic (Gramajo *et al.*, 1993; Stutzman-Engwall *et al.*, 1992). Therefore, we introduced the multicopy plasmids pAE-G2_2 and pAE8 into *S. coelicolor* (*nov-BG1*) by protoplast transformation. As presented in Table 2, pAE-G2_2 led to a 1.9-fold and pAE8 a 2.7-fold increase in novobiocin biosynthesis in comparison to strains carrying only the empty vector pWHM3.

Overexpression and purification of NovG as a His₆ fusion protein

Retzlaff & Distler (1995) successfully expressed the regulator protein StrR both in *E. coli* and in *Streptomyces lividans*, and found no difference in the DNA binding property between the proteins from either expression system. For further investigation of the function of *novG*, its gene product was therefore expressed as an N-terminal His₆ fusion protein in *E. coli* (see Methods). Upon cultivation at 30 °C and induction with 1 mM IPTG for 5 h, most of the resulting protein was insoluble. The amount of soluble protein did not improve significantly when the growth temperature was reduced to 15 °C and the IPTG concentration to 0.25 mM. However, useful amounts of soluble protein could be obtained reproducibly when the induction period was shortened to only 1.5 h, using 0.5 mM IPTG at 30 °C (see Methods). The His₆-tagged NovG protein was purified from the soluble fraction by nickel affinity chromatography. SDS-PAGE analysis showed a band of 40 kDa (calculated mass

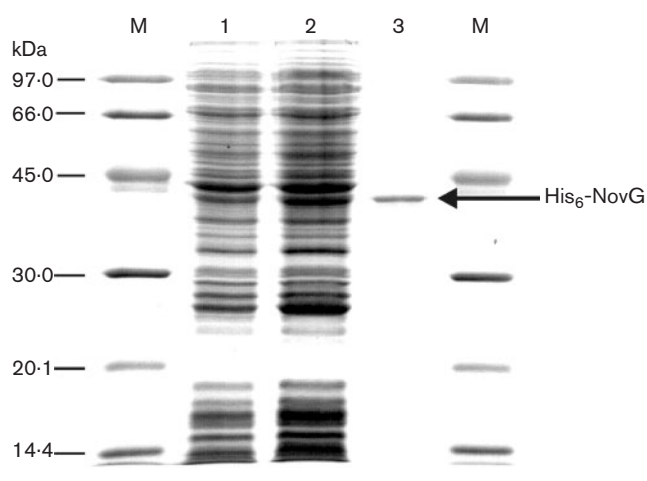


Fig. 3. Purification by SDS/12% PAGE of NovG after overexpression as a fusion protein with an N-terminal His₆ tag in *E. coli* BL21(DE3)pLysS. Lane 1, soluble protein before induction; lane 2, soluble protein after induction with 0.5 mM IPTG for 1.5 h; lane 3, eluate from nickel affinity chromatography; M, molecular mass marker. The molecular masses (kDa) of the marker proteins are indicated.

38.1 kDa) in the eluate (Fig. 3). The protein yield of purified NovG was about 15 µg per g cells (fresh weight), determined by SDS-PAGE.

DNA-binding activity

As described above, we speculated that NovG may act as a positive regulator of novobiocin biosynthesis by binding at or near the promoter region of biosynthetic genes and thereby activating their transcription. In order to verify whether NovG does indeed have DNA-binding activity, gel mobility-shift assays were carried out using the purified His₆-tagged NovG and selected DNA fragments from the novobiocin biosynthetic gene cluster. In the novobiocin cluster, the intergenic regions between most genes are very short, and often the coding regions even overlap, indicating a possible translational coupling (Steffensky *et al.*, 2000). Only six of the intergenic regions are larger than 70 bp (Fig. 4a) and those were chosen for gel shift experiments. As shown in Fig. 4(b), DNA fragments extending into the coding sequences upstream of these intergenic regions were also investigated. The NovG fusion protein showed very clear binding activity for one of the two fragments obtained from the *novG*–*novH* intergenic region (Fig. 4), but not for any other of the ten DNA fragments investigated.

The novobiocin and the clorobiocin clusters show very high similarity between each other. For most of the *nov* genes, orthologues exist in the clorobiocin cluster, which show 75–80 % sequence identity to the genes of the novobiocin cluster at the amino acid level (Pojer *et al.*, 2002). However, in the novobiocin cluster, the gene *novH*, which encodes a biosynthetic enzyme (Chen & Walsh, 2001), is situated

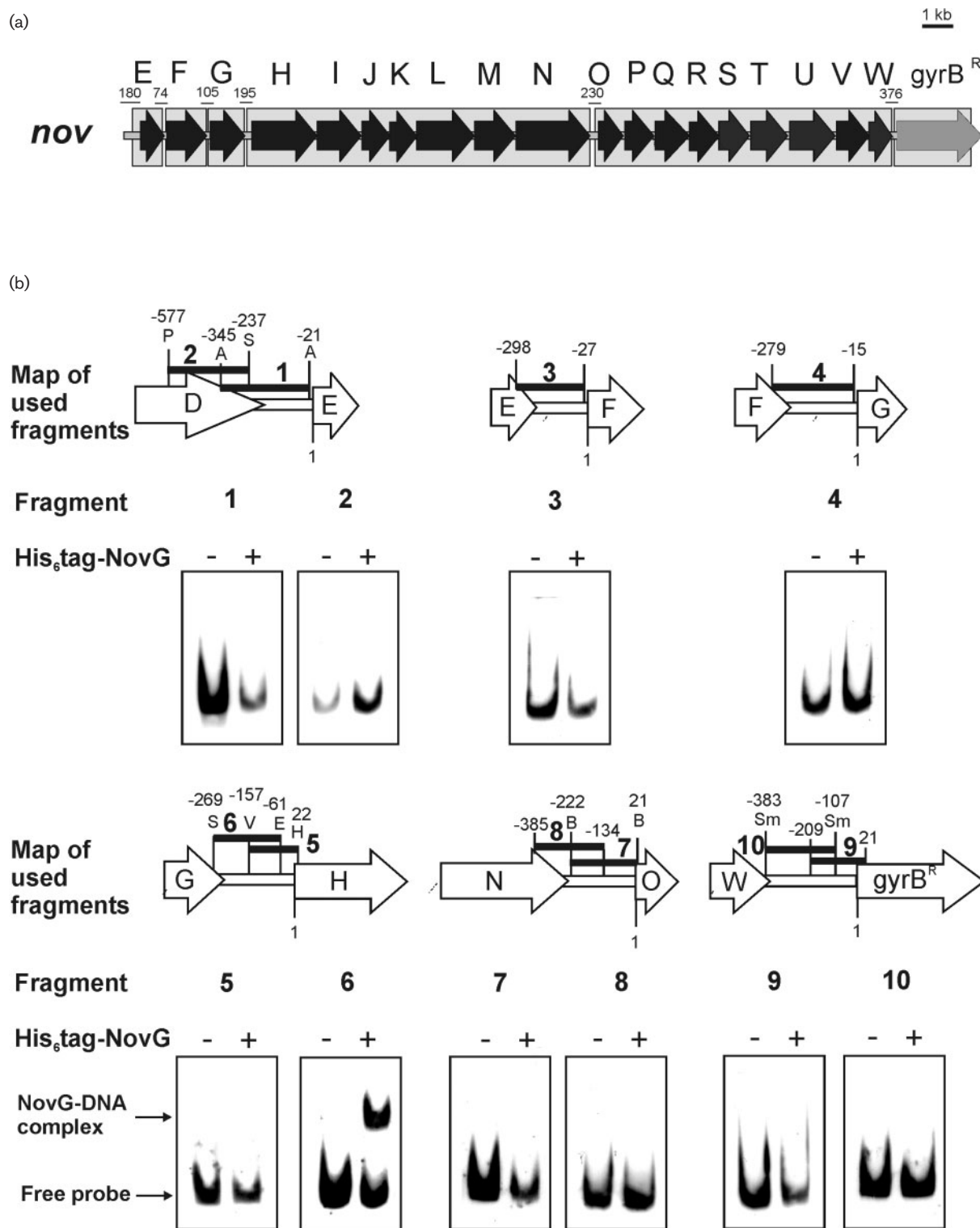


Fig. 4. Gel mobility-shift assays with DNA fragments from the novobiocin biosynthetic gene cluster. (a) Schematic representation of the novobiocin biosynthetic gene cluster. All intergenic regions larger than 70 bp are shown and their exact size is given. (b) Autoradiogram of gel mobility-shift assays using purified His₆-tagged NovG. The DNA-binding activity was analysed as described in Methods, using about 4 ng of the indicated DIG-end-labelled fragment and no protein (–) or 0.5 μg of purified His₆-tagged NovG protein (+). A map showing the location of the DNA fragments used is given above each autoradiogram (not to scale; the intergenic regions are oversized in comparison to the ORFs). The fragments were obtained by digestion of appropriate plasmids or by PCR amplification. For fragments obtained by digestion, restriction sites are indicated: P, *Pvu*I; S, *Sal*I; A, *Ava*I; V, *Van*91I; E, *Eco*RI; H, *Hind*III; B, *Bam*HI; Sm, *Sma*I.

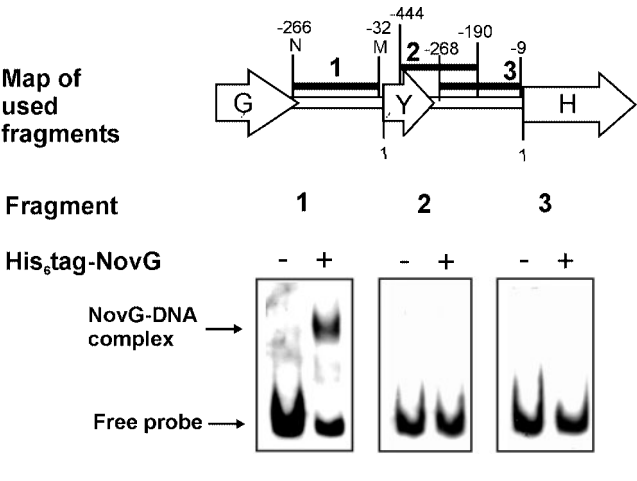


Fig. 5. Gel mobility-shift assays with DNA fragments from the clorobiocin biosynthetic gene cluster. The fragments were obtained by digestion of pAE9 (fragment 1) or by PCR amplification from cosmid K1F2 (fragments 2 and 3). N, *NotI*; M *MluI*. See legend of Fig. 4 for further explanations.

directly downstream of the regulatory gene *novG*, whereas in the clorobiocin cluster their orthologues *cloG* and *cloH* are separated by a small ORF of unknown function, i.e. *cloY* (Fig. 5). As shown in Fig. 5, we tested whether NovG could also bind to the DNA regions upstream of *cloY* and *cloH*. Very clearly, NovG did bind to the DNA region upstream of *cloY*, but not to those upstream of *cloH*.

NovG binds specifically to the *novG*–*novH* and *cloG*–*cloY* intergenic regions

To provide proof that binding of NovG to the *novG*–*novH* and *cloG*–*cloY* intergenic regions is specific, a gel mobility-shift assay was carried out in which a competitor plasmid containing the respective binding site was added in approximately 125-fold molar excess. As a negative control, the same amount of empty vector was used. As shown in Fig. 6, the presence of the specific competitor plasmid strongly reduced the intensity of the band representing the NovG–DNA complex, while the empty vector had no significant effect.

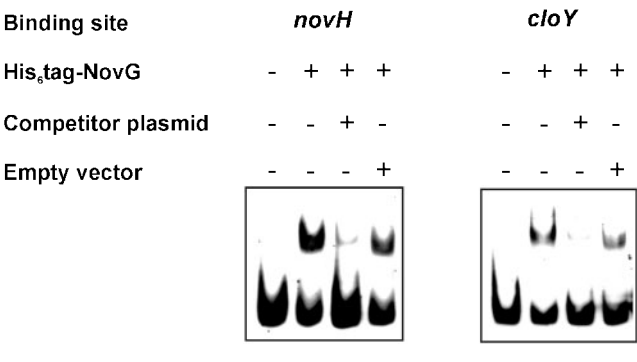


Fig. 6. Specificity of His₆-tagged NovG binding to DNA. Approximately 4 ng DIG-end-labelled fragment (*novH*, fragment 6 of Fig. 4; *cloY*, fragment 1 of Fig. 5) and no protein (–) or 0.5 µg purified His₆-tagged NovG protein (+) were used in each lane for a standard DNA-binding assay. When indicated (+), an additional 125-fold molar excess of specific competitor plasmid DNA (pAE7 or pAE9, respectively) or non-specific empty vector DNA (pBluescript SK(–) or pSPORT1, respectively) was added.

***In silico* analysis of the NovG binding DNA fragments**

When the DNA sequences of the *novG*–*novH* and *cloG*–*cloY* intergenic regions were compared, similar palindromic structures could be found in both fragments (Fig. 7). Notably, similar inverted repeats were identified as the binding site for StrR by DNase I footprinting assay (Retzlaff & Distler, 1995). The putative NovG-binding sites and the StrR-binding sites all possess a conserved palindromic structure with the consensus sequence GTTCTACTG(N)₁₁CRGTYGAAC.

DISCUSSION

The present work provides experimental proof for the function of NovG as a DNA-binding protein which acts as a positive regulator of novobiocin biosynthesis. Inactivation of *novG* led to a 98 % reduction of the novobiocin productivity of the heterologous producer strain. Since Δ *novG* strains still produced some novobiocin, an essential catalytic

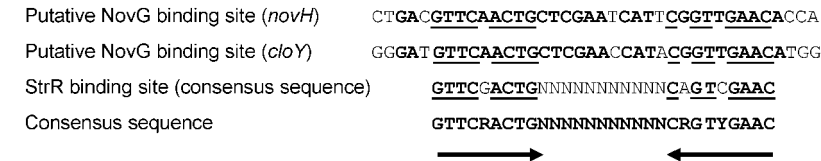


Fig. 7. Sequences of putative NovG-binding sites of *S. spheroides* (upstream of *novH*) and *Streptomyces roseochromogenes* (upstream of *cloY*) in comparison to the consensus sequence of StrR-binding sites of *S. griseus* and *S. glaucescens* (see Retzlaff & Distler, 1995). Nucleotides conserved in at least two of the sequences are in bold. Nucleotides conserved in all three sequences are underlined. The inverted repeats are indicated with arrows. A putative consensus sequence for NovG- and StrR-binding sites is indicated.

role for *novG* can be ruled out. By complementation with *novG* under control of its own promoter, novobiocin production was restored nearly to the same level as observed before *novG* inactivation, indicating that the observed phenotype of Δ *novG* strains was indeed caused by lack of *novG*.

The inactivation of positive regulators in the biosynthetic gene clusters of various antibiotics has been reported to result in a complete loss of antibiotic production (Antón *et al.*, 2004; Lombó *et al.*, 1999; Otten *et al.*, 1995, 2000; Pérez-Llarena *et al.*, 1997; Wilson *et al.*, 2001). The Δ *novG* strains constructed in this study, however, still produced some novobiocin (see above). This indicates a low level transcription of the biosynthetic genes in the absence of NovG. Similarly, *in vivo* analysis of regulatory genes of the nystatin biosynthetic gene cluster showed that nystatin production in the deletion mutants was reduced by 91–99.5 % compared to that in the wild-type strain; i.e. some nystatin production could still be detected (Sekurova *et al.*, 2004).

Expression of *novG* from a multicopy plasmid in a *S. coelicolor* M512 strain carrying the intact novobiocin cluster led to almost threefold overproduction of the antibiotic, suggesting that novobiocin biosynthesis in the heterologous expression host investigated is limited by availability of the activator protein, as has been shown for other pathway-specific activators in their respective natural antibiotic producers (Gramajo *et al.*, 1993; Stutzman-Engwall *et al.*, 1992).

Furthermore, it was demonstrated that His₆-tagged NovG protein binds specifically to the *novG*–*novH* intergenic region and, in so doing, probably activates the transcription of *novH*, just as described for StrR, which activates transcription of streptomycin biosynthetic genes by binding to their promoter regions (Retzlaff & Distler, 1995). The putative binding site of NovG is located directly downstream of the translational stop codon of *novG*, i.e. between –165 and –194 bp upstream of the putative translational start codon of *novH*. Since nickel affinity purified protein, generated in *E. coli*, was used for the DNA-binding assays, it is very likely that NovG can bind to DNA with no further macromolecular factor involved. However, our results do not completely rule out the possibility that other proteins may be required for activating the transcription of *novH*.

The novobiocin and clorobiocin clusters show very high similarity between each other. Indeed, NovG could also bind to a DNA region of the clorobiocin cluster, i.e. the *cloG*–*cloY* intergenic region; the putative binding site is located between positions –160 and –189 upstream of the putative translational start of *cloY*. The function of the small ORF *cloY* remains to be elucidated. Genes with obvious similarity to *cloY* have been found in many other clusters, such as those for coumermycin A₁ (Wang *et al.*, 2000), teicoplanin (Sosio *et al.*, 2004), complestatin (Chiu *et al.*, 2001), CDA (Hojati

et al., 2002) and balhimycin (Pelzer *et al.*, 1999), but not in the novobiocin cluster.

The *in silico* analysis of the DNA fragments from the novobiocin and clorobiocin clusters which bind NovG showed the presence of a perfectly conserved 9 bp inverted repeat, separated by a somewhat less-conserved (two mismatches) 11 bp spacer sequence (Fig. 7). Notably, the previously identified StrR binding sites in *S. griseus* and *S. glaucescens* contain a similar palindromic structure, i.e. conserved inverted repeats of 9 bp each, separated by a non-conserved 11 bp spacer (Retzlaff & Distler, 1995). The same putative NovG binding site, with exactly the same inverted repeat and spacer sequences as found upstream of *cloY*, is also present in the coumermycin A₁ cluster, between genes *couG* and *couY*. The close similarity of the putative NovG/CloG/CouG binding sites in the novobiocin, clorobiocin and coumermycin A₁ clusters further indicates a common evolutionary origin for these clusters (Eustáquio *et al.*, 2003b).

The consensus sequence GTTCTACTG(N)₁₁CRGTYGAAC or similar motifs were not found anywhere else in the gene clusters of novobiocin or clorobiocin, except in the mentioned regions upstream of *novH* and *cloY*, respectively. This is in accordance with the results of the gel shift assays with the novobiocin cluster depicted in Fig. 4. In contrast, four StrR binding sites have been identified in the streptomycin biosynthetic gene cluster in *S. griseus* and three in the 5'-OH-streptomycin cluster in *S. glaucescens* (Retzlaff & Distler, 1995). It cannot be completely excluded that NovG may also bind to additional sequences which are different from the motif shown above. However, it is quite possible that only a single transcription unit is expressed under control of *novG*. In the novobiocin and clorobiocin clusters, all genes are orientated in the same direction and may, in principle, be transcribed as a single mRNA from *novH* to *novW* or from *cloY* to *cloZ*, respectively. In contrast, previous results suggest that the resistance gene *gyrB^R*, at the right border of the cluster depicted in Fig. 4, is under control of its own promoter and may be regulated by changes in DNA superhelical density (Thiara & Cundliffe, 1989). NovG was not required for expression of *gyrB^R* in *S. lividans* (Thiara & Cundliffe, 1988) and, correspondingly, we did not observe binding of NovG to the promoter sequence of *gyrB^R* in our gel-shift assays.

Analysis of different gene clusters encoding biosynthetic pathways for antibiotics has revealed the existence of distinct families of pathway-specific regulatory proteins. One of these families comprises the *strR*-like genes and includes, besides *novG*, *cloG* and *couG*, also *dbv4* of the glycopeptide antibiotic A40926 (Sosio *et al.*, 2003), *bbr* of the balhimycin (Pelzer *et al.*, 1999) and *tcp28* of the teicoplanin (Sosio *et al.*, 2004) clusters. Another family comprises the SARPs, i.e. ActII-ORF4 and DnrI (see Introduction) and SnoA of the nogalamycin, RedD of the undecylprodigiosin and CcaR of the cephamycin clusters, respectively (Wietzorrek & Bibb, 1997). However, no genes with sequence similarity to this

family have been found in the gene clusters of the aminocoumarin antibiotics.

An understanding of the complex regulatory mechanisms that determine the onset of antibiotic biosynthesis in actinomycetes is only just beginning to emerge (Wietzorrek & Bibb, 1997). Activation of any particular pathway in any particular organism might be expected to require its own combination of signals (Chater & Bibb, 1997). Therefore, analysis of regulatory genes from different species is crucial to build a comprehensive picture of these processes.

The aminocoumarin antibiotics may provide useful model systems for such studies, due to their very stringent genetic organization (Pojer *et al.*, 2002) and the availability of detailed data on the function of most genes contained therein (Li & Heide, 2004). The fact that they can be successfully expressed in *S. coelicolor* M512, and the genome sequence of its wild-type strain *S. coelicolor* A3(2) is available (Bentley *et al.*, 2002), makes them even more attractive for investigating the regulation cascades for improved antibiotic production in streptomycetes.

ACKNOWLEDGEMENTS

We thank E. Takano, R. Shawky and W. Wohlleben for valuable discussions. We are also grateful to E. Takano and Janet White for kindly providing *S. coelicolor* M512. This work was supported by a grant from the European Community (No. 503466 to L.H. and S.-M.L.).

REFERENCES

- Antón, N., Mendes, M. V., Martín, J. F. & Aparicio, J. F. (2004). Identification of PimR as a positive regulator of pimarin biosynthesis in *Streptomyces natalensis*. *J Bacteriol* **186**, 2567–2575.
- Arias, P., Fernández-Moreno, M. A. & Malpartida, F. (1999). Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) as a DNA-binding protein. *J Bacteriol* **181**, 6958–6968.
- Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A. M. & 40 other authors (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147.
- Chater, K. F. & Bibb, M. J. (1997). Regulation of bacterial antibiotic production. In *Biotechnology*, vol. 7. *Products of Secondary Metabolism*, pp. 57–105. Edited by H. Kleinkauf & H. von Döhren. Weinheim: VCH.
- Chater, K. F. & Horinouchi, S. (2003). Signalling early developmental events in two highly diverged *Streptomyces* species. *Mol Microbiol* **48**, 9–15.
- Chen, H. & Walsh, C. T. (2001). Coumarin formation in novobiocin biosynthesis: β -hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chem Biol* **8**, 301–312.
- Chiu, H.-T., Hubbard, B. K., Shah, A. N., Eide, J., Fredenburg, R. A., Walsh, C. T. & Khosla, C. (2001). Molecular cloning and sequence analysis of the complestatin biosynthetic gene cluster. *Proc Natl Acad Sci U S A* **98**, 8548–8553.
- Datsenko, K. A. & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640–6645.
- Eustáquio, A. S., Gust, B., Luft, T., Li, S.-M., Chater, K. F. & Heide, L. (2003a). Clorobiocin biosynthesis in *Streptomyces*. Identification of the halogenase and generation of structural analogs. *Chem Biol* **10**, 279–288.
- Eustáquio, A. S., Luft, T., Wang, Z.-X., Gust, B., Chater, K. F., Li, S.-M. & Heide, L. (2003b). Novobiocin biosynthesis: inactivation of the putative regulatory gene *novE* and heterologous expression of genes involved in aminocoumarin ring formation. *Arch Microbiol* **180**, 25–32.
- Eustáquio, A. S., Gust, B., Li, S.-M., Pelzer, S., Wohlleben, W., Chater, K. F. & Heide, L. (2004). Production of 8'-halogenated and 8'-unsubstituted novobiocin derivatives in genetically engineered *Streptomyces coelicolor* strains. *Chem Biol* **11**, 1561–1572.
- Eustáquio, A. S., Gust, B., Galm, U., Li, S.-M., Chater, K. F. & Heide, L. (2005). Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl Environ Microbiol* **71** (in press).
- 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.
- Galm, U., Dessoy, M. A., Schmidt, J., Wessjohann, L. A. & Heide, L. (2004). *In vitro* and *in vivo* production of new aminocoumarins by a combined biochemical, genetic and synthetic approach. *Chem Biol* **11**, 173–183.
- Gramajo, H. C., Takano, E. & Bibb, M. J. (1993). Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. *Mol Microbiol* **7**, 837–845.
- Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. (2003). PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* **100**, 1541–1546.
- Hojati, Z., Milne, C., Harvey, B. & 9 other authors (2002). Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*. *Chem Biol* **9**, 1175–1187.
- Hussain, H. A. & Ritchie, D. A. (1991). High frequency transformation of *Streptomyces niveus* protoplasts by plasmid DNA. *J Appl Bacteriol* **71**, 422–427.
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*, 2nd edn. Norwich: John Innes Foundation.
- Kominek, L. A. (1972). Biosynthesis of novobiocin by *Streptomyces niveus*. *Antimicrob Agents Chemother* **1**, 123–134.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Leski, B. K., Bibb, M. J. & Chater, K. F. (1991). The use of a rare codon specifically during development? *Mol Microbiol* **5**, 2861–2867.
- Li, S.-M. & Heide, L. (2004). Functional analysis of biosynthetic genes of aminocoumarins and production of hybrid antibiotics. *Curr Med Chem Anti-Infective Agents* **3**, 279–295.
- Lombó, F., Braña, A. F., Méndez, C. & Salas, J. A. (1999). The mithramycin gene cluster of *Streptomyces argillaceus* contains a positive regulatory gene and two repeated DNA sequences that are located at both ends of the cluster. *J Bacteriol* **181**, 642–647.
- MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H. & MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**, 61–68.

- Maxwell, A. (1993). The interaction between coumarin drugs and DNA gyrase. *Mol Microbiol* **9**, 681–686.
- Otten, S. L., Ferguson, J. & Hutchinson, C. R. (1995). Regulation of daunorubicin production in *Streptomyces peuceitii* by the *dnrR*₂ locus. *J Bacteriol* **177**, 1216–1224.
- Otten, S. L., Olano, C. & Hutchinson, C. R. (2000). The *dnrO* gene encodes a DNA-binding protein that regulates daunorubicin production in *Streptomyces peuceitii* by controlling expression of the *dnrN* pseudo response regulator gene. *Microbiology* **146**, 1457–1468.
- Pabo, C. O. & Sauer, R. T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem* **61**, 1053–1095.
- Pelzer, S., Süßmuth, R., Heckmann, D., Recktenwald, J., Huber, P., Jung, G. & Wohlleben, W. (1999). Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908. *Antimicrob Agents Chemother* **43**, 1565–1573.
- Pérez-Llarena, F. J., Liras, P., Rodríguez-García, A. & Martín, J. F. (1997). A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both β -lactam compounds. *J Bacteriol* **179**, 2053–2059.
- 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.
- Pojer, F., Li, S.-M. & Heide, L. (2002). Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics. *Microbiology* **148**, 3901–3911.
- Retzlaff, L. & Distler, J. (1995). The regulator of streptomycin gene expression, StrR, of *Streptomyces griseus* is a DNA binding activator protein with multiple recognition sites. *Mol Microbiol* **18**, 151–162.
- Rost, B. (1996). PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol* **266**, 525–539.
- Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd edn. New York: Cold Spring Harbor Laboratory.
- Sekurova, O. N., Brautaset, T., Sletta, H., Borgos, S. E., Jakobsen, M. O., Ellingsen, T. E., Strom, A. R., Valla, S. & Zotchev, S. B. (2004). In vivo analysis of the regulatory genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455 reveals their differential control over antibiotic biosynthesis. *J Bacteriol* **186**, 1345–1354.
- Sheldon, P. J., Busarow, S. B. & Hutchinson, C. R. (2002). Mapping the DNA-binding domain and target sequences of the *Streptomyces peuceitii* daunorubicin biosynthesis regulatory protein, DnrI. *Mol Microbiol* **44**, 449–460.
- Sosio, M., Stinchi, S., Beltrametti, F., Lazzarini, A. & Donadio, S. (2003). The gene cluster for the biosynthesis of the glycopeptide antibiotic A40926 by *Nonomuraea* species. *Chem Biol* **10**, 541–549.
- Sosio, M., Kloosterman, H., Bianchi, A., de Vreugd, P., Dijkhuizen, L. & Donadio, S. (2004). Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. *Microbiology* **150**, 95–102.
- Steffensky, M., Mühlenweg, A., Wang, Z.-X., Li, S.-M. & Heide, L. (2000). Identification of the novobiocin biosynthetic gene cluster of *Streptomyces sphaeroides* NCIB 11891. *Antimicrob Agents Chemother* **44**, 1214–1222.
- Stutzman-Engwall, K. J., Otten, S. L. & Hutchinson, C. R. (1992). Regulation of secondary metabolism in *Streptomyces* spp. and overproduction of daunorubicin in *Streptomyces peuceitii*. *J Bacteriol* **174**, 144–154.
- Thamm, S. & Distler, J. (1997). Properties of C-terminal truncated derivatives of the activator, StrR, of the streptomycin biosynthesis in *Streptomyces griseus*. *FEMS Microbiol Lett* **149**, 265–272.
- Thiara, A. S. & Cundliffe, E. (1988). Cloning and characterization of a DNA gyrase B gene from *Streptomyces sphaeroides* that confers resistance to novobiocin. *EMBO J* **7**, 2255–2259.
- Thiara, A. S. & Cundliffe, E. (1989). Interplay of novobiocin-resistant and -sensitive DNA gyrase activities in self-protection of the novobiocin producer, *Streptomyces sphaeroides*. *Gene* **81**, 65–72.
- Thorpe, H. M., Wilson, S. E. & Smith, M. C. (2000). Control of directionality in the site-specific recombination system of the *Streptomyces* phage ϕ C31. *Mol Microbiol* **38**, 232–241.
- Vara, J., Lewandowska-Skarbek, M., Wang, Y. G., Donadio, S. & Hutchinson, C. R. (1989). Cloning of genes governing the deoxysugar portion of the erythromycin biosynthesis pathway in *Saccharopolyspora erythraea* (*Streptomyces erythreus*). *J Bacteriol* **171**, 5872–5881.
- Wang, Z.-X., Li, S.-M. & Heide, L. (2000). Identification of the coumermycin A₁ biosynthetic gene cluster of *Streptomyces rishiriensis* DSM 40489. *Antimicrob Agents Chemother* **44**, 3040–3048.
- Wietzorrek, A. & Bibb, M. (1997). A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25**, 1181–1184.
- Wilson, D. J., Xue, Y., Reynolds, K. A. & Sherman, D. H. (2001). Characterization and analysis of the PikD regulatory factor in the pikromycin biosynthetic pathway of *Streptomyces venezuelae*. *J Bacteriol* **183**, 3468–3475.
- Xu, H., Heide, L. & Li, S. M. (2004). New aminocoumarin antibiotics formed by a combined mutational and chemoenzymatic approach utilizing the carbamoyltransferase NovN. *Chem Biol* **11**, 655–662.