ARTICLE

OBC
www.rsc.org/obc

The neomycin biosynthetic gene cluster of *Streptomyces fradiae* NCIMB 8233: characterisation of an aminotransferase involved in the formation of 2-deoxystreptamine†

Fanglu Huang,[a] Stephen F. Haydock,[b] Tatiana Mironenko,[b] Dieter Spiteller,[b] Yanyan Li[a] and Jonathan B. Spencer*[a]

[a] *University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge, United Kingdom CB2 1EW. E-mail: jbs20@cam.ac.uk; Fax: +44 1223 336362; Tel: +44 1223 331696*

[b] *Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, United Kingdom CB2 1GA*

Received 25th January 2005, Accepted 15th February 2005
First published as an Advance Article on the web 21st March 2005

The biosynthetic gene cluster of the 2-deoxystreptamine (DOS)-containing aminoglycoside antibiotic neomycin has been cloned for the first time by screening of a cosmid library of *Streptomyces fradiae* NCIMB 8233. Sequence analysis has identified 21 putative open reading frames (ORFs) in the neomycin gene cluster (*neo*) with significant protein sequence similarity to gene products involved in the biosynthesis of other DOS-containing aminoglycosides, namely butirosin (*btr*), gentamycin (*gnt*), tobramycin (*tbm*) and kanamycin (*kan*). Located at the 5′-end of the *neo* gene cluster is the previously-characterised neomycin phosphotransferase gene (*apH*). Three genes unique to the *neo* and *btr* clusters have been revealed by comparison of the *neo* cluster to *btr*, *gnt*, *tbm* and *kan* clusters. This suggests that these three genes may be involved in the transfer of a ribose moiety to the DOS ring during the antibiotic biosynthesis. The product of the *neo-6* gene is characterised here as the L-glutamine : 2-deoxy-*scyllo*-inosose aminotransferase responsible for the first transamination in DOS biosynthesis, which supports the assignment of the gene cluster.

Introduction

2-Deoxystreptamine (DOS)-containing antibiotics constitute the largest subgroup of the aminoglycoside class of antimicrobial agents.[1] This class includes many clinically important antibiotics, of which neomycin and gentamycin are the best known. Neomycin was the first known DOS-containing antibiotic and was isolated in 1949 from *Streptomyces fradiae* and *Streptomyces albogriseus*.[2] Neomycin exhibits a broad spectrum of activity against both Gram positive and Gram negative bacteria. In common with most aminoglycoside antibiotics, neomycin principally interacts with the bacterial ribosomal 16 S rRNA to induce codon misreading during protein synthesis.[3] Experimental evidence shows that the overall antibacterial activity of neomycin may result from such mRNA misreading in addition to inhibition of assembly of the 30 S and 50 S ribosomal subunits[4,5] and impairment of transfer messenger RNA (tmRNA) aminoacylation during *trans*-translation for reprogramming of stalled ribosomes.[6]

Neomycin isolated from bulk fermentation of *Streptomyces fradiae* and *Streptomyces albogriseus* comprises a complex of three related compounds, namely neomycin A, neomycin B and neomycin C (Fig. 1).

Neomycin A (commonly known as neamine) consists of 2,6-diamino-2,6-dideoxy-D-glucose (neosamine C, ring II) attached to the aglycone 2-deoxystreptamine (DOS, ring I). Neomycin B is composed of neamine with 2,6-diamino-2,6-dideoxy-L-idose (neosamine B, ring IV) and ribose (ring III), linked to ring I. Neomycin C differs from neomycin B only with respect to the stereochemistry of the aminomethyl group attached to C5‴ in the ring IV sugar residue. Ring I and the ring II of neomycin are conserved in all DOS-containing antibiotics and are essential for binding to the decoding site of the 16 S rRNA. Aminoglycosides are water-soluble weak bases that are polycations at biological pH. Attraction between the positively charged aminoglycoside and the negatively charged RNA backbone is believed to contribute to the binding of the antibiotics to the rRNA. Targeting of aminoglycosides to other RNAs, mainly ribozymes and the Rev response element (RRE) transcription activation region in human immunodeficient virus (HIV), was also observed.[3,7] As a consequence, many studies have focused on the potential of neomycin and its analogues as a new class of anti-HIV agents.[8–10] Recently neomycin has also been shown to strongly inhibit both translocation of angiogenin and angiogenin-induced cell proliferation and angiogenesis, while other aminoglycosides tested, including the structurally very similar paromomycin, did not show any inhibitory activity. The *in vitro* anti-angiogenic mechanism of action of neomycin has led to its consideration as a potential candidate in anticancer drug development.[11,12]

The emergence and rapid spread of antibiotic resistance among human pathogens poses a serious threat to human health and has provided a potent impetus to the development of new antibiotics that circumvent current resistance mechanisms. One approach to this problem focuses on the generation of new antibiotic agents by modification of existing antibiotic structures. A successful example of such an approach is the semi-synthetic agent amikacin which has been derived from the naturally occurring parent compound kanamycin by attaching a 2-hydroxy-4-amino-butyric acid substituent to the DOS ring.[13] Amikacin has been shown to be active against many strains of bacteria that are resistant to the naturally occurring aminoglycoside antibiotics. However there are limits to potential chemical modifications. A detailed understanding of the biosynthetic pathway of aminoglycoside antibiotics, together with access to the corresponding enzymes, is a prerequisite for a genetic and

† Accession number: The accession number for the neomycin biosynthetic cluster sequence in EMBL Nucleotide Sequence Database is AJ843080.

DOI: 10.1039/b501199j

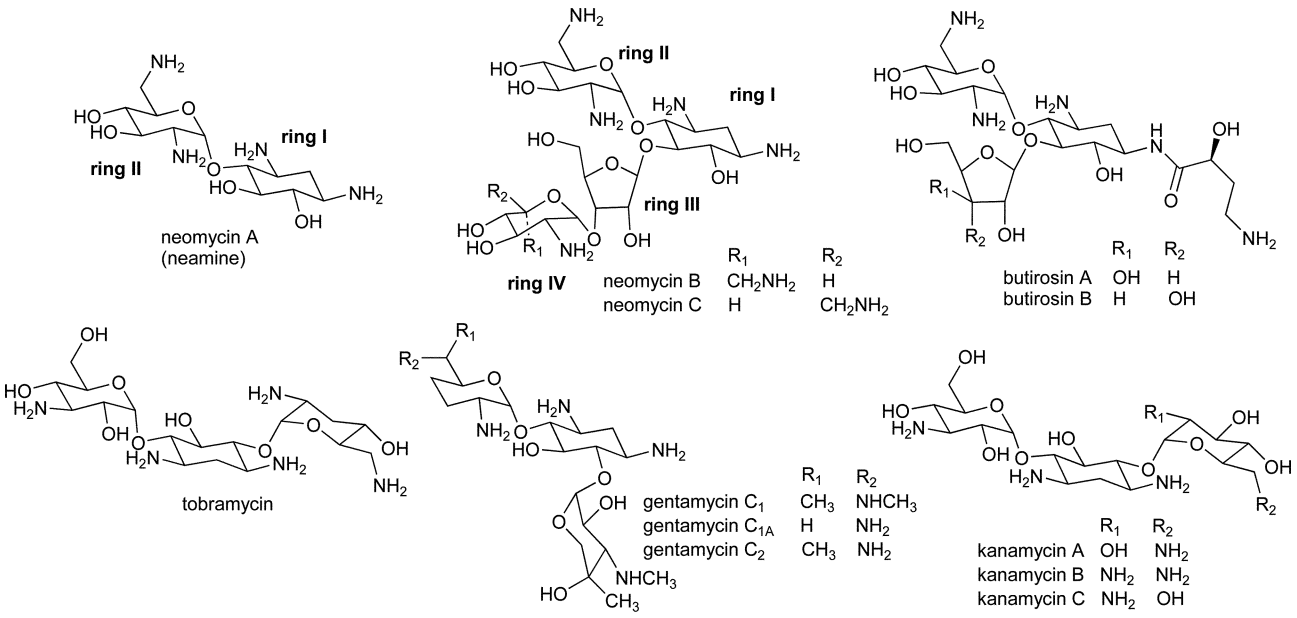


Fig. 1 Structures of neomycin, butirosin, gentamycin, tobramycin and kanamycin.

biological approach to the engineering of antibiotic biosynthesis to produce novel aminoglycoside antibiotics.

Feeding of labelled precursors and studies of intermediates isolated from blocked mutants have identified DOS, paromamine, neamine and ribostamycin as intermediates in neomycin biosynthesis.[14–17] The proposed biosynthetic pathway to ribostamycin has been derived from studies on biosynthesis of butirosin, a related DOS-containing aminoglycoside antibiotic, in *Bacillus circulans.*[18] Neomycin also contains a ribostamycin moiety (ring I + II + III). It is reasonable to speculate that the biosynthesis of this important intermediate is common to both butirosin and neomycin (this common pathway is described in Scheme 1, in the Results and discussion section below). Two enzymes, 2-deoxy-*scyllo*-inosose synthase (BtrC)[19] and L-glutamine : 2-deoxy-*scyllo*-inosamine aminotransferase (BtrR)[20] (also designated BtrS by Tamegai *et al.*[21]), involved in the biosynthesis of DOS have been characterised in the butirosin-producing strain *Bacillus circulans*. Homologues of BtrC have also been characterised recently in biosynthetic gene clusters for tobramycin and kanamycin, respectively.[22,23] In the presence of NAD^+ and Co^{2+}, BtrC and its homologues catalyse the conversion of D-glucose-6-phosphate (G6P) to 2-deoxy-*scyllo*-inosose (DOI), the first step in DOS biosynthesis. BtrR is responsible for the subsequent transamination of DOI to form 2-deoxy-*scyllo*-inosamine (DOIA) using L-glutamine as the amino donor. The BtrR-bound pyridoxal 5′-phosphate (PLP) is the essential cofactor for the amino-transfer activity. Recently, the product of *btrD* from the *btr* gene cluster has been expressed in *E. coli* and characterised as a novel nucleotidyltransferase catalysing the formation of UDP-D-glucosamine from D-glucosamine-1-phosphate and UTP. D-glucose-1-phosphate and dTTP were also good substrates for the enzyme reaction. It has been suggested that the NDP-sugars generated by BtrD reaction are the glycosyl donors for the glycosylation of DOS to give paromamine.[24] Other details of the pathways, and the other enzymes that catalyse the synthesis of neomycin and related aminoglycosides, remain to be elucidated. We describe here the cloning and sequence analysis of the entire gene cluster (*neo*) for the neomycin biosynthesis in *Streptomyces fradiae* NCIMB 8233. The *neo* gene cluster encodes 21 putative open reading frames (ORFs) thought to be involved in the biosynthesis of neomycin. Neo-6, the product of one of the 21 ORFs, has been cloned and expressed in *E. coli.* Enzymatic assays have demonstrated that Neo-6 is the enzyme catalysing the transamination of DOI using L-glutamine as an amino donor to produce DOIA as required for DOS biosynthesis.

Results and discussion

Cloning of the neomycin biosynthetic gene cluster

A genomic library of *S. fradiae* NCIMB 8233 was prepared in the cosmid vector SuperCos 1 (Stratagene) and screened by PCR. Partial DNA sequence of the putative 2-deoxy-*scyllo*-inosose aminotransferase gene in *S. fradiae* was obtained by degenerate PCR. New primers (FRA3, 4, 5, 6) were then synthesised according to this sequence and used for initial screening of the cosmid library. Sequence analysis identified a positive cosmid, FR3F7, with a 37 kbp insert containing 27 complete putative ORFs. 12 of these ORFs show homology to *btr* genes in butirosin biosynthetic gene cluster.[20,25,26] In addition, a partial ORF containing sequence identical to the neomycin phosphotransferase gene[27] is present at one end of the FR3F7 insert. This gene has been shown to confer resistance to neomycin when transferred into a sensitive heterologous host.[28] The clustering of genes encoding proteins with homology to gene products in biosynthetic gene clusters of other aminoglycosides around this characterised resistance gene allows us to assign this locus to the biosynthesis of neomycin in this strain. Four primers, FRA7, FRA8, Primer11 and Primer12, were then designed based on the sequence of the phosphotransferase gene and used to re-screen 672 colonies of the cosmid library. One cosmid was identified (FR20D5) which contained sequence overlap with the insert in FR3F7 and extending further upstream from the 5′-end of FR3F7. Sequence analysis of both FR3F7 and FR20D5 identified a 30 kbp region of DNA containing 21 putative ORFs, including the previously characterised neomycin resistance gene, that are entirely consistent with this locus encoding the neomycin biosynthetic gene cluster (Table 1). The limits of the cluster remain to be determined but the presence of several housekeeping genes both up- and downstream of the locus suggests that the boundaries of the gene cluster lie within the determined sequence. 13 of the ORFs (*neo-5* to *neo-13* and *neo-16* to *neo-19*) have homologues in the butirosin (*btr*) biosynthetic gene cluster (AJ494863, AB066276, AJ847918),[20,25,26] (Fig. 2 and Scheme 1). Many ORFs in the *neo* cluster also have their homologues in gene clusters of other

Table 1 Predicted ORFs in the neomycin biosynthetic gene cluster

Orf	Location[a]	Related gene clusters[b]				Best match[c]		Proposed function[f]
		btr	*gnt*	*tbm*	*kan*	Source[d]	Entry[e]	
1	180–986c	—	*gntI*	—	—	*Streptomyces fradiae*	PKSMR	Phosphotransferase
2	995–2290c	—	—	—	*kacN*	*Streptomyces kanamyceticus*	CAF60526	Unknown
3	2280–2801c	—	—	—	*orf3*	*Streptomyces cinnamoneus DSM 40005*	CAD60535	Unknown
4	3662–4189c	—	—	—	—	*Streptomyces kanamyceticus*	CAF60528	Unknown
5	4401–5423	*btrE*	*gntP*	*tacD*	*kanK*	*Streptomyces tenebrarius*	CAE22477	Alcohol dehydrogenase
6	5502–6776	*btrR*	*gntA* *gntF*	*tbmB*	*kanB*	*Streptomyces tenebrarius*	CAE22472	L-Glutamine : 2-deoxy-*scyllo*-inosose aminotransferase
7	6947–8239	*btrC*	*gntB*	*tbmA*	*kanA*	*Streptomyces kanamyceticus*	CAE46939	2-Deoxy-*scyllo*-inosose synthase
8	8467–9651	*btrM*	*gntZ* *gntD*	*tbmD*	*kanF* *kanE*	*Streptomyces kanamyceticus*	CAE46947	Glycosyltransferase
9	9648–11639	*btrS*	—	—	—	*Streptomyces avermitilis MA-4680*	NP_823966	Exporter
10	11636–13420	*btrT*	—	—	—	*Streptomyces coelicolor A3(2)*	NP_629587	Exporter
11	13629–15254	*btrQ*	*gntX*	*orf1*	*kanI*	*Micromonospora echinospora*	AAR98543	Sugar dehydrogenase
12	15428–16327	*btrN*	—	—	—	*uncultured archaeon GZfos35D7*	AAU84037	Fe–S oxidoreductase
13	16324–17025	*btrP*	—	—	—	*Bacillus circulans*	BAC41220	Phosphomutase
14	17022–17273	—	—	—	—	*Photorhabdus luminescens*	AAL18481	Unknown
15	17266–18366	—	—	—	—	*Kineococcus radiotolerans SRS30216*	ZP_00226821	Unknown
16	18359–19198	*btrD*	—	—	*kacA*	*Streptomyces kanamyceticus*	CAE46942	Nucleotidyltransferase
17	19396–21378c	*btrL*	—	—	—	*Bacillus circulans*	BAC41207	Ribosyltransferase
18	21495–22745	*btrB*	*gntW* *gntL* *gntH* *gntJ*	*orf2*	*kacL*	*Micromonospora echinospora*	AAR98542	Aminomutase
19	22831–26772	*btrA*	—	—	—	*Photorhabdus luminescens*	AAL18480	Unknown
20	26934–27797c	—	—	—	—	*Streptomyces fradiae*	P29809	Aminoglycoside *N*-acetyltransferase
21	28177–30837	—	—	—	—	*Streptomyces vinaceus*	AAP92511	LuxR regulator

[a] Location of genes in the cluster. The first and the last nucleotide of a gene (including a complete stop codon) are indicated by numbers. A letter c next to the numbers denotes a gene that is encoded on the complementary strand. [b] Existence of a homolog in other related aminoglycoside antibiotic gene clusters: *btr* butirosin biosynthesis; *gnt* gentamycin biosynthesis; *tbm* tobramycin biosynthesis; *kan* kanamycin biosynthesis. [c] Best match found by NCBI Protein-Protein BLAST. [d] Organism origin of the best matching sequence. [e] Accession number of the best matching sequence. [f] Proposed function in the neomycin biosynthetic pathway.

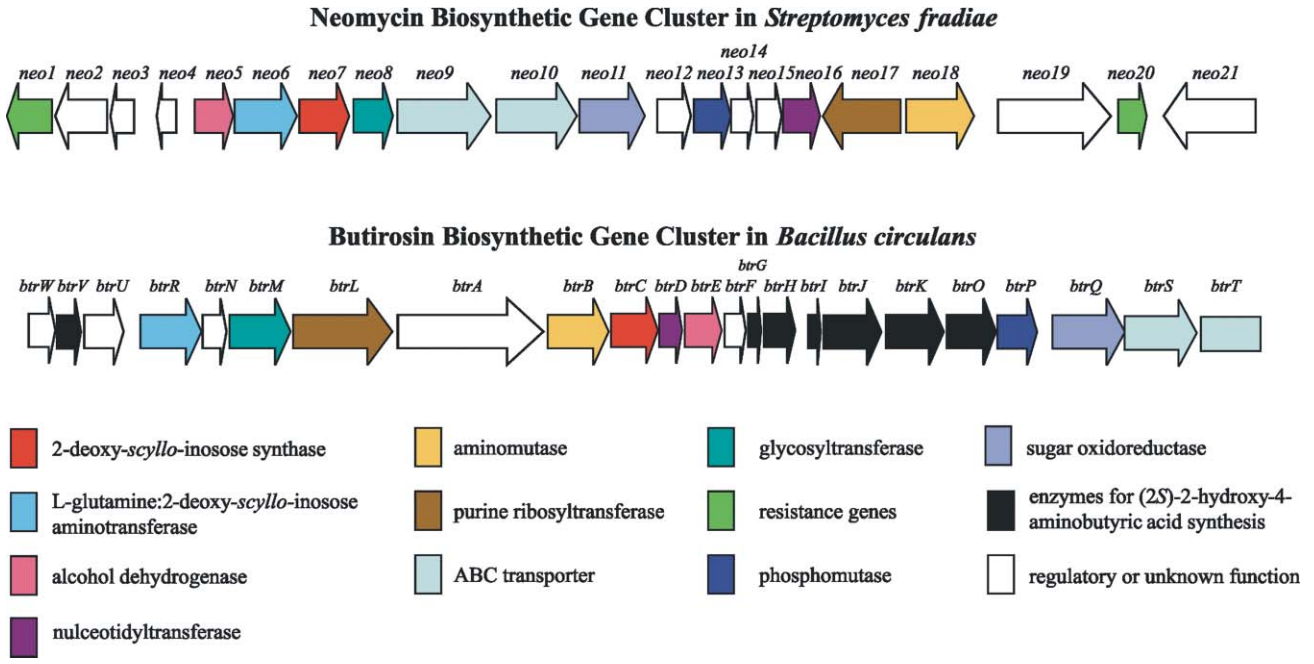


Fig. 2 Comparison of the neomycin (*neo*) and butirosin (*btr*) gene clusters. ORFs with similar predicted functions in both biosynthetic pathways are indicated by the same colour.

related aminoglycosides: *gnt* (gentamycin, AY524043),[29] *tbm* (tobramycin, AJ579650)[22] and *kan* (kanamycin, AJ582817).[23] The genes *neo-13*, *neo-17* and *neo-19* have homologues in the *btr* cluster but no homologues in the *gnt*, *tbm* and *kan* clusters (Table 1).

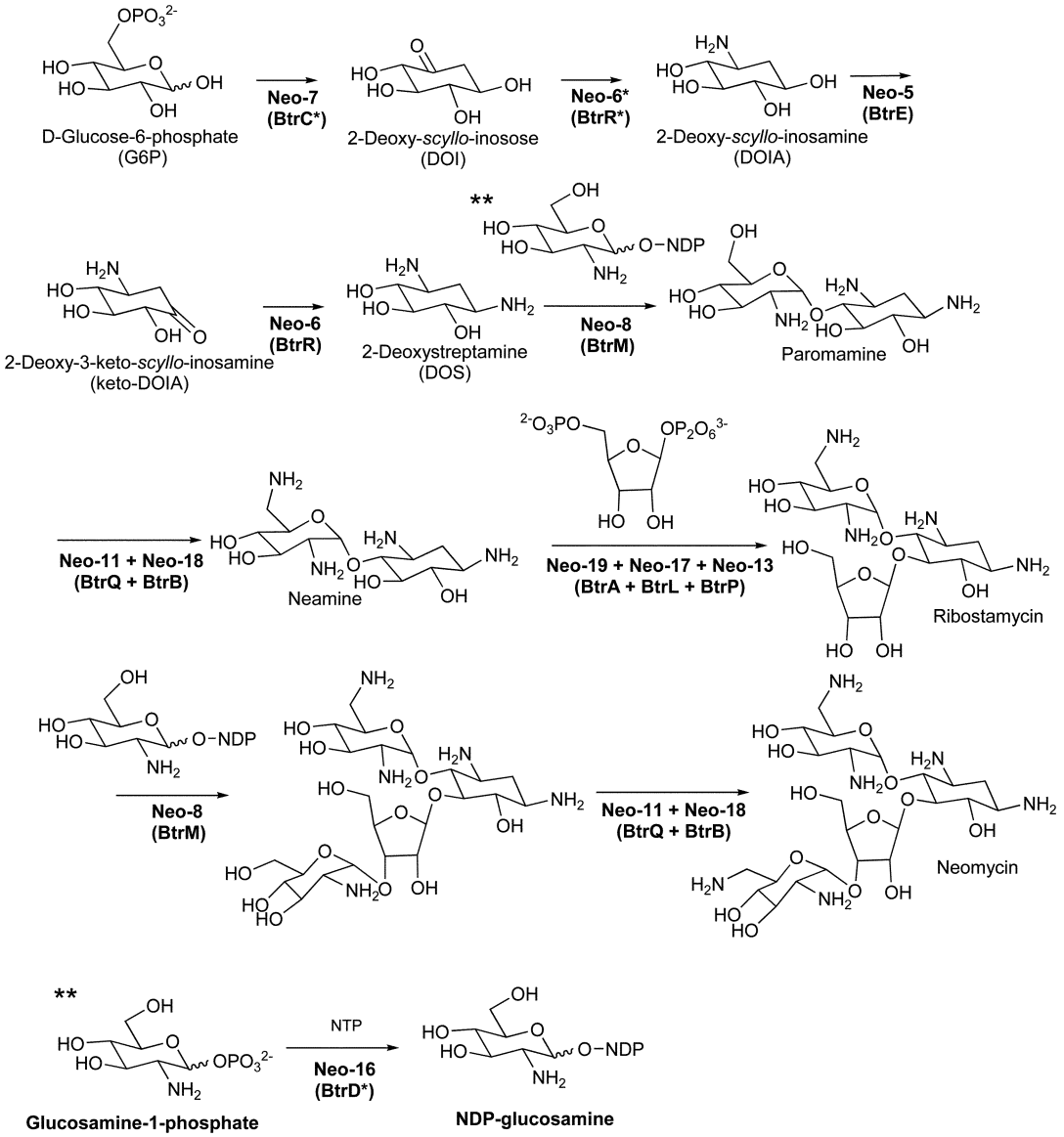
Biosynthesis of 2-deoxystreptamine

2-Deoxystreptamine (DOS) is the aglycone in neomycins. As expected, predicted protein sequences of *neo-7* (430 amino acids) and *neo-6* (424 amino acids) in the *neo* cluster show high homology to the well-characterised BtrC[19] and BtrR[20] in the *btr* gene cluster, respectively, which catalyse the first two reactions in the conversion of G6P to DOIA in the biosynthesis of DOS. In view of our previous observation that BtrR may also participate in the second transamination step in the DOS pathway,[20] a similar role may also be played by Neo-6 in neomycin biosynthesis. Neo-5 is predicted to be a 340 amino acid protein and is similar in sequence to many alcohol dehydrogenases. Neo-5 shows 49% protein sequence identity to BtrE in butirosin biosynthesis,[26] 64% to TacD in tobramycin biosynthesis,[22] 45% to GntP in gentamycin biosynthesis,[29] and 61% to KanK in kanamycin biosynthesis.[23] The possible involvement of the *tacD* gene product in the third step of DOS biosynthesis [oxidation of DOIA at C3 before transamination (Scheme 1)] has been proposed.[22] In preliminary experiments using BtrE, we have observed the NAD^+ dependent oxidation of DOIA (unpublished data), suggesting a similar function for Neo-5.

To prove the function of Neo-6 as the L-glutamine:DOI aminotransferase, *neo-6* has been cloned in pET28a(+) vector and overexpressed in *E. coli* as a His_6-tagged protein. In agreement with the calculated molecular weight (46951 Da), purified Neo-6 appeared on SDS-PAGE gel as a single protein band with a size around 45 kDa and LC-MS analysis of the recombinant Neo-6 further confirmed its molecular weight to be 46801 Da [M-Met]. The Neo-6 protein solution showed characteristic absorptions of protein-bound pyridoxal 5′-phosphate (PLP) at 415 nm and pyridoxamine 5′-phosphate (PMP) at 335 nm (data not shown). This is similar to that of BtrR and StsC, the two homologues of Neo-6 in butirosin and streptamycin biosynthetic pathways.[20,30] Purified Neo-6 has been shown to catalyse amino transfer from L-glutamine to DOI to produce DOIA when coupled with BtrC in the presence of G6P, NAD^+, Co^{2+} and L-glutamine. The production of DOIA was confirmed by TLC (data not shown) and by LC-MS/MS after derivatisation of the product with 9*H*-fluoren-9-yl)methyl 2,5-dioxopyrrolidine-1-carboxylate (FmocOSu)[31] (Fig. 3a). Similar to its counterpart BtrR in the *btr* gene cluster,[20] Neo-6 also catalysed the deamination of DOS to generate keto-DOIA using pyruvate as amino acceptor, suggesting again the possible involvement of both Neo-6 and BtrR in the second transamination reaction in DOS biosynthesis. The keto-DOIA produced in the reverse reaction was derivatised with pentafluorobenzylhydroxylamine hydrochloride[32] and analysed by LC-MS/MS (Fig. 3b). FMOC-inosamine; retention time: 22.3 min; ESI-MS: 386 (100), 342 (3), 208 (4), 195 (3), 179 (8), 164 (7); ESI-MS/MS of 386: 342 (70), 208 (57), 179 (17), 164 (100). PFB-keto-inosamine; retention time: 17.8 min. ESI-MS: 357 (100), 222 (32), 181 (15), 159 (5), 141 (8), 123 (3); ESI-MS/MS of 357: 339 (100), 321 (41), 222 (0.5), 181 (0.5), 159 (7), 141 (37), 124 (22), 113 (7).

Formation and attachment of aminohexoses

The product of *neo-16* is a protein of 279 amino acids and is homologous to BtrD in the *btr* cluster, KacA in the *kan* cluster and to MitC, a putative protein encoded by the mitomycin gene cluster.[33] Recent studies have suggested that BtrD is responsible for making NDP-glucose/glucosamine sugar donors for DOS glycosylation leading to formation of paromamine.[24] Neo-16 shares 25% identity with BtrD and contains all the motifs that were conserved in BtrD and its homologues. A similar function is therefore proposed for Neo-16. BtrM from the *btr* cluster has been proposed to be a glycosyltransferase responsible for the formation of neamine *via* the addition of the aminohexose ring to DOS (Scheme 1). The ORF *neo-8* is a homologue of *btrM* which strongly suggests that this gene also codes a glycosyltransferase with the same function. Neomycin has an additional aminohexose ring to butirosin (Fig. 1), and therefore an additional glycosyltransferase would be predicted to be present in the cluster. A BLAST search against the NCBI protein database showed that Neo-15 has homology (26% identity) to some putative glycosyltransferases, raising the possibility that this enzyme adds the second aminohexose ring. However, since ring II in neomycin B only differs from ring IV in the stereochemistry of the aminomethyl group it would have been predicted that Neo-8 and Neo-15 would have good homology to



Scheme 1 Proposed neomycin biosynthetic pathway. The putative *neo* and *btr* gene products are indicated. Enzymes that have been characterised are indicated with an asterisk.

one another. Since they do not, it raises the possibility that Neo-8 may be responsible for the transfer of both aminohexose rings. A relaxed substrate specificity of many glycosyltransferases that participate in the biosynthesis of secondary metabolites has been indicated by a number of previous studies.[34] It cannot be ruled out that a glycosyltransferase outside the gene cluster participates in *neo* biosynthesis. The predicted translation product of *neo-11* has significant sequence similarity to many dehydrogenases, especially choline and glucose dehydrogenases. The similarity is closest to BtrQ[25] in the *btr* cluster, GntX[29] in the *gnt* cluster, ORF1 in the *tbm* cluster,[22] and KanI in *kan* cluster.[23] We propose that Neo-11 may be the enzyme catalysing the oxidation of the hydroxyl groups at the position 6 of neosamine C and neosamine B prior to transamination. The subsequent transamination could be carried out by the product of *neo-18*, which encodes a putative aminomutase, the only transaminase candidate other than Neo-6 present in the cluster. Genes similar to *neo-18* have been identified in gene clusters for butirosin, gentamycin, tobramycin and kanamycin.

Attachment of ribose

It is noteworthy that homologues of *btrA*, *btrL* and *btrP* in the butirosin gene cluster have been identified in the neomycin gene cluster and designated *neo-19*, *neo-17* and *neo-13*, respectively. No similar genes have been found in the clusters encoding tobramycin,[22] gentamycin[29] and kanamycin[23] biosynthetic genes. Neomycin and butirosin differ from tobramycin, gentamycin and kanamycin in that they contain a ribose moiety that is absent in the latter three antibiotics (Fig. 1). It is therefore possible that Neo-13, Neo-17, Neo-19 and their homologues in the butirosin cluster are involved in ribosyltransfer activities. Supporting this hypothesis, a BLAST search for short, nearly exact matches against protein database shows that (1) a region in both Neo-17 and BtrL (*ca.* 70 amino acids) exhibits some similarity (32% identity) to xanthine guanine phosphoribosyltransferase (EC 2. 4. 22. 22), an enzyme involved in purine metabolism that catalyses ribosyltransfer from 5-phosphoribosyl-1-pyrophosphate (PRPP) to either xanthine or guanine to produce xanthosine-5′-phosphate or guanosine-5′-phosphate, respectively; (2) both Neo-13 and BtrP contain a domain conserved in phosphatase/phosphomutase involved in carbohydrate transport and metabolism (*e.g.* fructose-2,6-bisphosphatase and phosphoglycerate mutase) as revealed by conserved domain search against protein databases; and (3) Neo-19 has a region (*ca.* 77 amino acids) showing 31% identity to the nicotinic acid phosphoribosyltransferase (EC 2.4.2.11) from *Halobacterium salinarum* that carries out reversible transfer of a

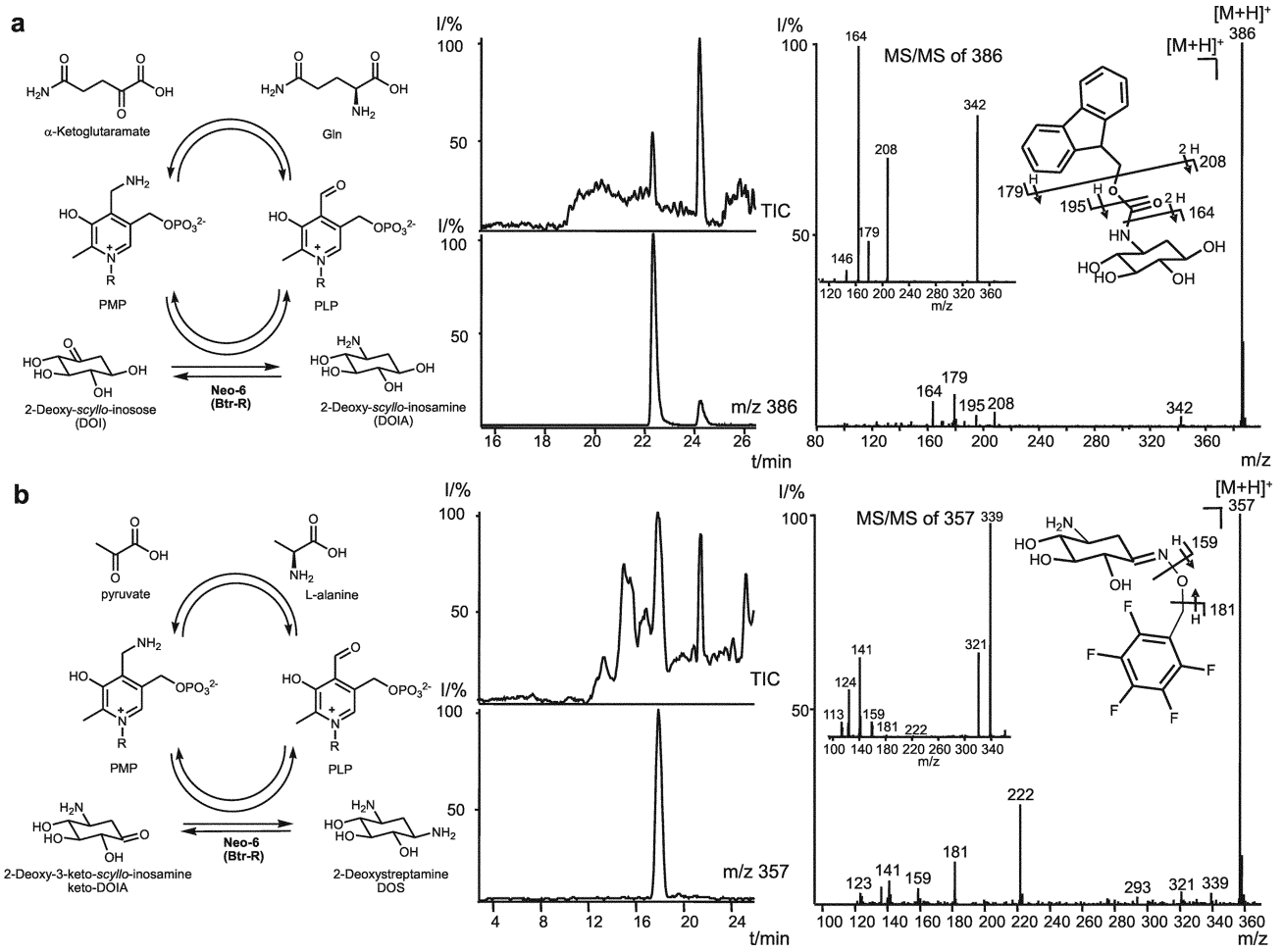


Fig. 3 LC-MS analysis of Neo-6 activity; (a) conversion of DOI to DOIA mediated by Neo-6 using L-glutamine as the amino donor (forward reaction, left panel). The substrate DOI for Neo-6 was generated from glucose-6-phosphate by BtrC and its cofactors (see Scheme 1) present in the reaction mixture. TIC and ion trace *m/z* 386 of the product of the Neo-6 reaction after derivatisation with FMOC-OSu (centre panel). ESI-MS and MS/MS of the $[M + H]^+$ 386 of FMOC-inosamine at 22.3 min (right panel); (b) deamination of DOS catalysed by Neo-6 using pyruvate as the amino acceptor (reverse reaction, left panel). TIC and ion trace *m/z* 357 of the product after derivatisation with PFBHA HCl (centre panel). ESI-MS and MS/MS of the $[M + H]^+$ 357 of PFB-keto-inosamine at 17.8 min (right panel).

ribose-5-phosphate moiety from PRPP to nicotinate. Moreover, the 397 to 415 region in BtrA has significant homology (57%) to GTP pyrophosphokinases (EC 2.7.6.5) found in some bacilli. These enzymes are responsible for reversible pyrophosphate transfer from GTP to the ribose ring of ATP at the C3′ position. It is therefore possible that Neo-17 (or BtrL) may carry out the coupling of a phosphoribose (*e.g.* 5-phospho-α-D-ribose-1-diphosphate) to the DOS ring with the release of the pyrophosphate by Neo-19 (or BtrA) and cleavage of the remaining 5-phosphate may be catalysed by Neo-13 and its homologue BtrP.

Resistance genes

There are two aminoglycoside resistance genes found in the neomycin gene cluster. The gene *neo-1*, located at one end of the cluster, is the previously identified aminoglycoside 3′-phospho-transferase gene *apH*.[27] The product of the *apH* gene inactivates aminoglyoside antibiotics through ATP-dependent phosphorylation of the 3′-hydroxyl of the antibiotics. The protein encoded by *neo-20* is almost identical (with only two differing amino acids) to the aminoglycoside 3-*N*-acetyltransferase previously identified in *S. fradiae* by Salauze *et al.*[35] and belongs to the bacterial aminoglycoside 3-*N*-acetyltransferases family (EC: 2.3.1.81). Members of this family catalyse transfer of an acetyl group from acetyl-CoA to the N3′ position of a 2-deoxystreptamine-containing antibiotic. This inactivates and confers resistance to gentamycin, kanamycin, tobramycin, neomycin and apramycin amongst others.

Exporters and regulators

Proteins encoded by *neo-9* and *neo-10* are similar to a series of ABC-type multidrug efflux/protein/lipid transporters including BtrS and BtrT, the two putative ABC transporters in the *btr* cluster.[25] Neo-9 shows 24% identity to NovA in novobiocin gene cluster.[36] They are likely to be responsible for the ATP dependant export of neomycin from the cell.

Neo-21 is a protein of 886 amino acids with end to end homology to many LuxR class of bacterial regulators, particularly to those encountered in secondary metabolite biosynthetic clusters. Members of the LuxR family bind DNA through a 'helix-turn-helix′ motif located in the C-terminal section of the sequence and are activated when bound to autoinducer molecules such as *N*-(3-oxohexanoyl)-L-homoserine lactone (OHHL).[37]

Neo-6 contains a TTA codon encoding the leucine at position 13. The TTA codon is rarely present in high G + C content DNA of *Streptomyces* and is often found in pathway-specific regulatory genes for biosynthesis of antibiotics.[38] Studies on *S. coelicolor* mutants have linked the deficiency of *bldA* gene, encoding the only tRNA for the rare UUA leucine codon, to the failure in production of aerial mycelium and antibiotic secondary metabolites.[39,40] The well-studied actinorhodin biosynthetic gene cluster[41] contains only two TTA codons. These may facilitate integrated control of differentiation and secondary

metabolism by ensuring that key genes contain such a codon are expressed only at later stages of the life cycle. The presence of a TTA codon in *neo-6* suggests that the expression of Neo-6 is dependent upon expression of *bldA*.

Other ORFs

The deduced amino acid sequences of *neo-2*, *neo-3* and *neo-4* show moderate to significant similarity, to three putative proteins, Cinorf14 (35%), Cinorf13 (66%) and Cinorf12 (70%), respectively, in the cinnamycin biosynthetic gene cluster.[42] Like their counterparts in the *cin* gene cluster, these three *neo* genes are clustered together in an order of *neo-12*, *neo-13* and *neo-14* with the same transcription direction. Neo-2 and Neo-3 also have their homologues, KacI and Orf1, respectively, both in the *kan* cluster. *Neo-14* encodes a protein with homology to several proteins in the databases with no assigned functions.

Neo-12 is predicted to be an iron–sulfur-containing oxidoreductase and also a member of the radical SAM family[43] from sequence comparisons with protein databases. The radical SAM protein superfamily consists of more than 600 members.[43,44] Reactions catalysed by radical SAM proteins include unusual methylations, isomerization, sulfur insertion, ring formation, anaerobic oxidation and protein radical formation during biosynthesis of DNA precursors, vitamins, cofactors, and antibiotics. Neo-12 has a homologue (BtrN) in the butirosin gene cluster but has not been found in the gentamycin, tobramycin or kanamycin clusters. The specific function of Neo-12 and BtrN in the biosynthesis of the antibiotics remains to be clarified.

Conclusion

The biosynthetic gene cluster for neomycin from *Streptomyces fradiae* NCIMB 8233 has been cloned and sequenced. It spans 30 kbp and is proposed to encode 21 ORFs. The first gene encoded by the cluster is the previously-sequenced and characterised neomycin phosphotransferase gene *apH*.[27] Many genes in the *neo* cluster are highly homologous to those involved in the biosynthesis of the related aminoglycosides butirosin,[20,25,26] gentamycin,[29] tobramycin,[22] and kanamycin.[23] The genes *neo-13*, *neo-17* and *neo*-19 have been suggested to be responsible for ribose transfer in the biosynthesis of neomycin. The protein encoded by one of the 21 ORFs, *neo-6*, has been cloned and expressed in *E. coli.* The *in vitro* enzymatic assays have demonstrated that Neo-6, like its homologue in butirosin biosynthesis, is a L-glutamine : 2-deoxy-*scyllo*-inosose aminotransferase responsible for catalyzing the conversion of DOI to DOIA, the first transamination step in DOS biosynthesis.

Experimental

Bacterial strains, culture media and vectors

Streptomyces fradiae MCIMB 8233 was cultured in tryptic soy broth (Difco) at 30 °C. SuperCos 1 (Stratagene) was utilised as the vector for construction of the cosmid library. The cosmid library was transfected into *Escherichia coli* XL1-Blue MR (Stratagene) and selected by resistance to both kanamycin and ampicillin. Recombinant cosmids were subcloned as overlapping partial *Sau3A* I restriction fragments into pSHG397[45] and transformed into *Escherichia coli* DH10B (Gibco BRL) using standard protocols[46] with chloramphenicol selection of recombinants. Plasmid pET28a(+) (Novagen) was used as a vector for gene cloning and expression. *Escherichia coli* XL-2 and BL21(DE3) (Novagen) were used as hosts for recombinant gene construction and gene expression, respectively.

Construction of the cosmid library

Genomic DNA from *Streptomyces fradiae* NCIMB 8233 was isolated by 'procedure B'.[47] A cosmid library of the *S. fradiae* genomic DNA in SuperCos 1 (Stratagene) was constructed in accordance with the manufacturer's instruction: briefly, the purified genomic DNA was partially digested with *BamH*I, dephosphorylated with shrimp intestinal alkaline phosphatase, ligated into *BamH*I-/*Xba*I-cut SuperCos 1, and packaged with Gigapack® III Gold Packaging Extract (Stratagene) without size fractionation. The packaging reaction was transfected into *E. coli* XL1-Blue MR. Individual colonies were inoculated into 1 ml 2xLB medium on 96 well culture plates containing ampicillin and kanamycin and incubated at 37 °C overnight. Small scale cosmid DNA was isolated utilising the High Yield Protocol 1 for Plasmid Minipreparation using MultiScreen (Millipore) as recommended by the manufacturer.

Screening of the cosmid library

Three degenerate PCR primers, FRA1: 5′-T(GC)G G(GC)G T(GC)A A(CT)G C(GC)G T(GC)C C(GC)G T(GC)T T(CT)T GCG T-3′, FRA2: 5′-CCC AT(GC) A(AG)(CT) TC(GC) CC(GC) GT(CT) TC(GC) AC(GC) A(AG)(CT) TCC AT-3′ and FRA1B: T(GC)G G(GC)G T(GC)A A(CT)G C(GC)G T(GC)C C(GC)G T(GC)T T(CT)T GCG A(CT)G T-3′, were designed utilising the partial protein sequence of the putative homologue of BtrR[20] previously identified in *S. fradiae*,[48] taking into account the codon bias of *Streptomyces*. These primers were used in a PCR reaction with *S. fradiae* genomic DNA as template. PCR products of the expected size (*ca.* 400 bp) were obtained using both FRA1/FRA2 and FRA1B/FRA2 primer pairs and cloned in pUC18. Sequencing analysis confirmed that the inserts in the recombinant plasmids contained a sequence encoding a polypeptide with the expected peptide sequence.[48] Four perfectly matched primers, FRA3: 5′-CTC CCC CGA GGC CGT CGA AGC GCT GAT CAC-3′, FRA4: 5′-GGC GGG CCG TCG GAC AGG CAG CGG CCG TCG-3′, FRA5: 5′-TAC TCG GCC GTC GCG GAC ATG GAC GGC CTC-3′ and FRA6: 5′-ACC GCG CCG CCC TCG CCG CTG GTG AGC ACC-3′ were synthesised according to the newly identified sequence and used to screen the cosmid library by PCR. Two cosmids from 768 colonies gave PCR fragments of the expected size under conditions of stringent primer hybridisation. Large scale DNA preparations of these cosmid clones were sequenced using the same PCR primers and one cosmid, FR3F7, which gave the expected sequencing result, was partially digested with *Sau3A*I and the 2–6 kbp fraction of the digest was sub-cloned into pSHG397 plasmid and sequenced with pUC forward and reverse primers. Two primer pairs (FRA7: 5′-GTC GTC GAG CGC GGT GCC GAC GAC ACC-3′/FRA8: CTC CTG CAG GTC GTC GAG GTC CAC CAA-3′, and Primer 11: CCC ACC CCG CAC AAG AAT GTC CGA AAC-3′/Primer 12: CTC GCC GGA CAG GTC GAA GGC GGA GTT-3′) from the end sequences of FR3F7 insert were used to re-screen the library by PCR to identify cosmids that contained sequence overlapping with the FR3F7 insert.

DNA sequencing and analysis

All sequencing experiments were carried out using Big Dye Terminator Kit in an ABI 1373A Sequencer according to manufacturer's protocols (ABI) in the University of Cambridge, Department of Biochemistry DNA Sequencing Facility. Sequence editing was carried out utilising SeqEd v 1.0.3. BLAST algorithm was employed for database search. GAP (Genome Assembly Program) version 4.2 was used for sequence assembly.

Cloning, expression and purification of Neo-6

To construct a plasmid for expression of Neo-6, PCR amplification of *neo-6* was carried out with primers pNR1 (5′-GAAAGGG*CATATG*GTCTCCCCGTTGGC-3′) and pNR2

(5′-ACGGCCGGGGCCC*GGATCC*GTGCCTGATGA-3′) using the cosmid FR3F7 as a template. PCR reactions were performed under the following conditions: 94 °C 7 min, 25 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min at 72 °C, and a final extension at 72 °C for 10 min. The product was digested with *Nde*I and *Hind*III and inserted into expression vector pET28a(+). The sequence of the *neo*-6 gene insert in pET28a(+) was verified by DNA sequencing before transformation into the expression strain *E. coli* BL21 (DE3). A overnight culture of the *neo-6*/pET28a(+) containing BL21 (DE3) was transferred into 1 L fresh LB medium and incubated at 37 °C, 250 rpm until OD_{600} reached 0.5. To induce the expression of Neo-6, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM and the culture was further incubated at 16 °C overnight. The cells were harvested by centrifugation, resuspened in 50 mM Tris–HCl containing 0.5 M NaCl and 1 mM PLP, pH 7.9, and sonicated for 3 min (75% amplitude, 2 sec on, 19 sec off) in a VibraCell™ High intensity ultrasonic processor (Jencons). Cell debris was removed by centrifugation and the recombinant protein in the supernatant was purified by Ni^{2+}-NTA affinity chromatography (Novagen).

Enzyme assays of Neo-6

The ability of Neo-6 to catalyse the transamination of DOI by L-glutamine (forward reaction) was determined in a coupled assay with BtrC as previously described.[20] In a total volume of 100 µl a typical reaction mixture contained 30 mM Tris–HCl, pH 7.5, 5 mM G6P, 5 mM NAD^+, 0.1 mM $CoCl_2$, 10 mM L-glutamine, 100 µg BtrC and 100 µg Neo-6. Control reactions omitted either Neo-6 or G6P. After 1 h at 37 °C, each reaction mixture was mixed with 30 µl of chloroform and the denatured protein was removed by centrifugation. The product DOIA in the supernatant was derivatised to FMOC-DOIA (9*H*-fluoren-9-yl) methyl 2,3,4,5-tetrahydroxycyclohexylcarbamate) by treatment with 20 µl of (9*H*-fluoren-9-yl)methyl 2,5-dioxopyrrolidine-1-carboxylate (Fmoc-OSu)[31] (1 mg ml^{-1} in MeOH) and 20 µl of 0.1 M boric acid buffer, pH 9.0 for 20 min before subjecting to LC-MS/MS analysis. The deamination of DOS (reverse reaction) catalysed by Neo-6 was performed in a 100 µl reaction in the presence of 30 mM Tris–HCl, pH 7.5, 5 mM DOS, 5 mM pyruvate and 100 µg Neo-6. After incubation at 37 °C for 1 h, the enzyme was denatured and removed as described above and 20 µl of pentafluorobenzylhydroxylamine hydrochloride (PFBHA·HCl) solution[32] (1 mg ml^{-1} in MeOH) was added to react with the product keto-DOIA. The oxime derivative generated was then analysed by LC-MS/MS. Synthetic DOIA and DOI[49] were derivatised to FMOC-DOIA and PFBO-DOI, respectively, and used as standards. For the LC-MS/MS experiments a ThermoFinnigan LCQ fitted with an ESI-ion source connected to an Agilent HP 1100 HPLC system was used. 50 to 100 µl samples were injected onto a Phenomenex Synergy polar RP-18 column (125 mm × 2 mm, 5 µm) and separated by gradient elution. HPLC conditions: 100% A for 5 min, gradient to 100% B in 35 min, 100% B for 5 min; A: H_2O 0.1% TFA, B: CH_3CN 0.1% TFA; flow rate 0.3 ml min^{-1}.

Acknowledgements

We are grateful to Professor Dr P. F. Leadlay, Dr H. O'Hare and Dr D. Fowler, University of Cambridge, UK for helpful advice on the preparation of this manuscript and to Professor Dr M. Akhtar, University of Southampton, UK for his donation of 2-deoxystreptamine. This work was supported by project grants from the BBSRC to J.B.S., F.H. and S.F.H.; S.F.H. thanks the University of Cambridge School of Clinical Medicine for additional financial support. D.S. thanks the Deutschen Akademie der Naturforscher Leopoldina, Germany for a postdoctoral fellowship (BMBF-LPD 9901/8-90).

References

1 C. J. R. Pearce, K. L. Rinehart, Jr., in *Antibiotics, Vol. IV*, ed. J. W. Corcoran, Springer-Verlag, Berlin, Heidelberg, New York, 1981, pp. 74–100.
2 S. A. Waksman and H. A. Lechevalier, *Science*, 1949, **109**, 305–307.
3 R. Schroeder, C. Waldsich and H. Wank, *EMBO J.*, 2000, **19**, 1–9.
4 R. Mehta and W. S. Champney, *Antimicrob. Agents Chemother.*, 2002, **46**, 1546–1549.
5 R. Mehta and W. S. Champney, *Curr. Microbiol.*, 2003, **47**, 237–243.
6 S. Corvaisier, V. Bordeau and B. Felden, *J. Biol. Chem.*, 2003, **278**, 14788–14797.
7 M. G. Wallis and R. Schroeder, *Prog. Biophys. Mol. Biol.*, 1997, **67**, 141–154.
8 A. Litovchick, A. G. Evdokimov and A. Lapidot, *FEBS Lett.*, 1999, **445**, 73–79.
9 A. Litovchick, A. Lapidot, M. Eisenstein, A. Kalinkovich and G. Borkow, *Biochemistry*, 2001, **40**, 15612–15623.
10 J. B. Tok, L. J. Dunn and R. C. Des Jean, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 1127–1131.
11 G. F. Hu, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 9791–9795.
12 G. F. Hu, *Biochem. Biophys. Res. Commun.*, 2001, **287**, 870–874.
13 M. J. Cron, J. G. Keil, M. V. Ruggeri and D. Walker, *Chem. Commun.*, 1979, 266–267.
14 H. Baud, A. Betencourt, M. Peyre and L. Penasse, *J. Antibiot.*, 1977, **30**, 720–723.
15 J. R. Fang, C. J. Pearce and K. L. Rinehart, Jr., *J. Antibiot.*, 1984, **37**, 77–79.
16 W. T. Shier, S. Ogawa, M. Hichens and K. L. Rinehart, Jr., *J. Antibiot.*, 1973, **26**, 551–561.
17 K. Suzukake, K. Tokunaga, H. Hayashi, M. Hori, Y. Uehara, D. Ikeda and H. Umezawa, *J. Antibiot.*, 1985, **38**, 1211–1218.
18 T. Furumai, K. Takeda, A. Kinumaki, Y. Ito and T. Okuda, *J. Antibiot.*, 1979, **32**, 891–899.
19 F. Kudo, H. Tamegai, T. Fujiwara, U. Tagami, K. Hirayama and K. Kakinuma, *J. Antibiot.*, 1999, **52**, 559–571.
20 F. Huang, Y. Li, J. Yu and J. B. Spencer, *Chem. Commun.*, 2002, 2860–2861.
21 H. Tamegai, E. Nango, M. Kuwahara, H. Yamamoto, Y. Ota, H. Kuriki, T. Eguchi and K. Kakinuma, *J. Antibiot.*, 2002, **55**, 707–714.
22 M. K. Kharel, D. B. Basnet, H. C. Lee, K. Liou, J. S. Woo, B. G. Kim and J. K. Sohng, *FEMS Microbiol. Lett.*, 2004, **230**, 185–190.
23 M. K. Kharel, B. Subba, D. B. Basnet, J. S. Woo, H. C. Lee, K. Liou and J. K. Sohng, *Arch. Biochem. Biophys.*, 2004, **429**, 204–214.
24 F. Kudo, K. Kawabe, H. Kuriki and T. Eguchi, *J. Am. Chem. Soc.*, 2005, **127**, 1711–1718.
25 Y. Li, PhD thesis, University of Cambridge, Cambridge, 2004.
26 Y. Ota, H. Tamegai, F. Kudo, H. Kuriki, A. Koike-Takeshita, T. Eguchi and K. Kakinuma, *J. Antibiot.*, 2000, **53**, 1158–1167.
27 C. J. Thompson and G. S. Gray, *Proc. Natl. Acad. Sci. U. S. A.*, 1983, **80**, 5190–5194.
28 C. J. Thompson, R. H. Skinner, J. Thompson, J. M. Ward, D. A. Hopwood and E. Cundliffe, *J. Bacteriol.*, 1982, **151**, 678–685.
29 J. Unwin, S. Standage, D. Alexander, T. Hosted, Jr., A. C. Horan and E. M. Wellington, *J. Antibiot.*, 2004, **57**, 436–445.
30 J. Ahlert, J. Distler, K. Mansouri and W. Piepersberg, *Arch. Microbiol.*, 1997, **168**, 102–113.
31 A. Paquet, *Can. J. Chem.*, 1982, **60**, 976–980.
32 F. J. van Kuijk, D. W. Thomas, R. J. Stephens and E. A. Dratz, *Methods Enzymol.*, 1990, **186**, 399–406.
33 Y. Mao, M. Varoglu and D. H. Sherman, *Chem. Biol.*, 1999, **6**, 251–263.
34 X. M. He and H. W. Liu, *Annu. Rev. Biochem.*, 2002, **71**, 701–754.
35 D. Salauze, J. A. Perez-Gonzalez, W. Piepersberg and J. Davies, *Gene*, 1991, **101**, 143–148.
36 M. Steffensky, A. Muhlenweg, Z. X. Wang, S. M. Li and L. Heide, *Antimicrob. Agents Chemother.*, 2000, **44**, 1214–1222.
37 S. Swift, M. K. Winson, P. F. Chan, N. J. Bainton, M. Birdsall, P. J. Reeves, C. E. Rees, S. R. Chhabra, P. J. Hill and J. P. Throup, *Mol. Microbiol.*, 1993, **10**, 511–520.
38 B. K. Leskiw, M. J. Bibb and K. F. Chater, *Mol. Microbiol.*, 1991, **5**, 2861–2867.
39 B. K. Leskiw, E. J. Lawlor, J. M. Fernandez-Abalos and K. F. Charter, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 2461–2465.
40 K. A. Plaskitt and K. F. Chater, *Philos. Trans. R. Soc. London, Ser. B*, 1995, **347**, 105–121.
41 M. A. Fernandez-Moreno, E. Martinez, L. Boto, D. A. Hopwood and F. Malpartida, *J. Biol. Chem.*, 1992, **267**, 19278–19290.

42 D. A. Widdick, H. M. Dodd, P. Barraille, J. White, T. H. Stein, K. F. Chater, M. J. Gasson and M. J. Bibb, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 4316–4321.
43 H. J. Sofia, G. Chen, B. G. Hetzler, J. F. Reyes-Spindola and N. E. Miller, *Nucleic Acids Res.*, 2001, **29**, 1097–1106.
44 E. N. Marsh, A. Patwardhan and M. S. Huhta, *Bioorg. Chem.*, 2004, **32**, 326–340.
45 T. Hashimoto-Gotoh, T. Mizuno, Y. Ogasahara and M. Nakagawa, *Gene*, 1995, **152**, 271–275.
46 J. Sambrook, E. F. Fritsch, T. Maniatis, in *Molecular Cloning, a Laboratory Manual, 2nd ed.*, ed. C. Nolan, Cold Spring Harbor Laboratory Press, 1989.
47 T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical Streptomyces Genetics*, Crowes, Norwich, United Kingdom, 2000, p. 169.
48 H. Tamegai, T. Eguchi and K. Kakinuma, *J. Antibiot.*, 2002, **55**, 1016–1018.
49 J. Q. Yu and J. B. Spencer, *Tetrahedron Lett.*, 2001, **42**, 4219–4221.