Hybrid compounds generated by the introduction of a nogalamycin-producing plasmid into *Streptomyces argillaceus*†

Tero Kunnari,*[a] Karel D. Klika,[b] Gloria Blanco,[c] Carmen Méndez,[c] Pekka Mäntsälä,[d] Juha Hakala,‡[a] Reijo Sillanpää,[e] Petri Tähtinen,[b] Jose Salas[c] and Kristiina Ylihonko[a]

[a] *Galilaeus Oy, P.O. Box 113, FIN-20781 Kaarina, Finland. E-mail: tero.kunnari@galilaeus.fi; Fax: +358-2-2731460; Tel: +358-2-2741450*
[b] *Structural Chemistry Group, Department of Chemistry, University of Turku, Vatselankatu 2, FIN-20014 Turku, Finland*
[c] *Departamento de Biologia Funcional e Instituto, Universitario de Biotecnologia de Asturias, ESP-33006 Oviedo, Spain*
[d] *Department of Biochemistry, University of Turku, Vatselankatu 2, FIN-20014 Turku, Finland*
[e] *Department of Chemistry, University of Jyväskylä, Survontie 9, FIN-40500 Jyväskylä, Finland*

Received (in Cambridge, UK) 11th January 2002, Accepted 22nd May 2002
First published as an Advance Article on the web 11th July 2002

The combination of genetic material from different antibiotic-producing organisms is a versatile and ever-expanding approach for the production of novel, hybrid compounds possessing bioactivity. The introduction of a plasmid (pSY21b) derived from *Streptomyces nogalater* and encoding PKS for nogalamycin production into the host strain *S. argillaceus* A43 led to the production of three new compounds in addition to the normally produced mithramycin. The new compounds are hybrids in the sense that they share features associated with the genes of both the host and the introduced plasmid. The structural elucidation of the novel compounds relied primarily on NMR spectroscopy, which revealed the three hybrids to be glycosides with the same aglycone common to all. Determination of the relative stereochemistry within the aglycone unit was confirmed by single-crystal X-ray analysis of the aglycone, which also revealed the tautomeric equilibrium to be in a very different position in comparison to that in the solution state. The glycosylation profile was clearly determined by the host, as the typical mithramycin sugars, D-oliose, D-olivose, and D-mycarose, were all expressed. Notable for the mutant was the high titre of the shunt products, 60% of the metabolic output, together with a lack of structural diversity of the hybrid aglycone present in the products, two features which are not normally observed.

Introduction

Since its inception[1] almost two decades ago, the combination of genetic material from different antibiotic-producing organisms has proved to be a versatile, and ever-expanding approach[2] for the production of novel, hybrid compounds possessing latent bioactivity with the potential of being incorporated into the drug discovery process as drug candidates. Considerable attention has been directed towards highlighting the benefits and advantages of both the genetic engineering of antibiotic-producing microbes[3–5] and biotransformation using whole-cell techniques[5,6] and the realisation of the practical application of this approach as a routine, predictable methodology appears imminent.

Mithramycin (**1**) (also known as aureolic acid, plicamycin, mithracin®, *etc.* Fig. 1) is a bis(oligosaccharide) antibiotic produced by *Streptomyces argillaceus* and some other *Streptomyces* strains,[7] whose structural elucidation,[8] since its discovery in the 1950s,[9] has been contraversial. It is a highly toxic antibiotic possessing both antineoplastic and hypocalcemic properties.[10,11] The antibiotics UCH9, the chromocyclomycins, the chromomycins, and the olivomycins are all structurally and biochemically closely related to mithramycin (**1**) and, together, they are classified within the aureolic acid group of compounds.[8] They are potent inhibitors of *in vitro* and *in vivo* RNA synthesis as a consequence of their ability to form stable complexes with DNA.[12] However, unlike intercalating dyes and antibiotics, mithramycin (**1**) does not induce uncoiling of DNA. Despite their demonstrated potency in chemotherapy, the only aureolic acid that is presently in clinical use, albeit limited, is mithramycin (**1**). Within this context mithramycin (**1**) holds promise as a structural template for the application of hybrid techniques to generate novel compounds and thus potential new drug candidates. The structure was initially, and correctly, elucidated using chemical degradation and derivitisation techniques.[13] Subsequently, it was reinvestigated by several workers resulting in the publication of various erroneous structures,[13–15] mainly with respect to the identity, anomeric state, or linkage position of the attendant sugar units. These incorrect structures, needless to say, underpinned a poor appreciation of the mithramycin (**1**) biosynthetic pathway and now that modern NMR methods have firmly established the correct structure of mithramycin[8] (**1**), a better comprehension of the biosynthesis of mithramycin (**1**) is possible.

Nogalamycin (**2**) (Fig. 1) is an antitumour antibiotic belonging to the anthracycline group of compounds. Anthracyclines are useful cytotoxic compounds and some of them, *e.g.* doxorubicin and several analogues, are used clinically for the treatment of various malignancies. Like mithramycin (**1**), nogalamycin (**2**) is also biosynthesised *via* a polyketide pathway and is thus classified as an aromatic type II polyketide on the basis of the molecular genetics of the anthracyclines. The biosynthetic pathways of both mithramycin[16] (**1**) and nogalamycin[17] (**2**) are today well understood regarding the

† Electronic supplementary information (ESI) available: edited GAUSSIAN output of DFT calculations and HyperChem file of the calculated structure. See http://www.rsc.org/suppdata/p1/b2/b200444p/
‡ *Current address*: Lividans Oy, Tykistökatu 4 D, FIN-20520 Turku, Finland.

 DOI: 10.1039/b200444p

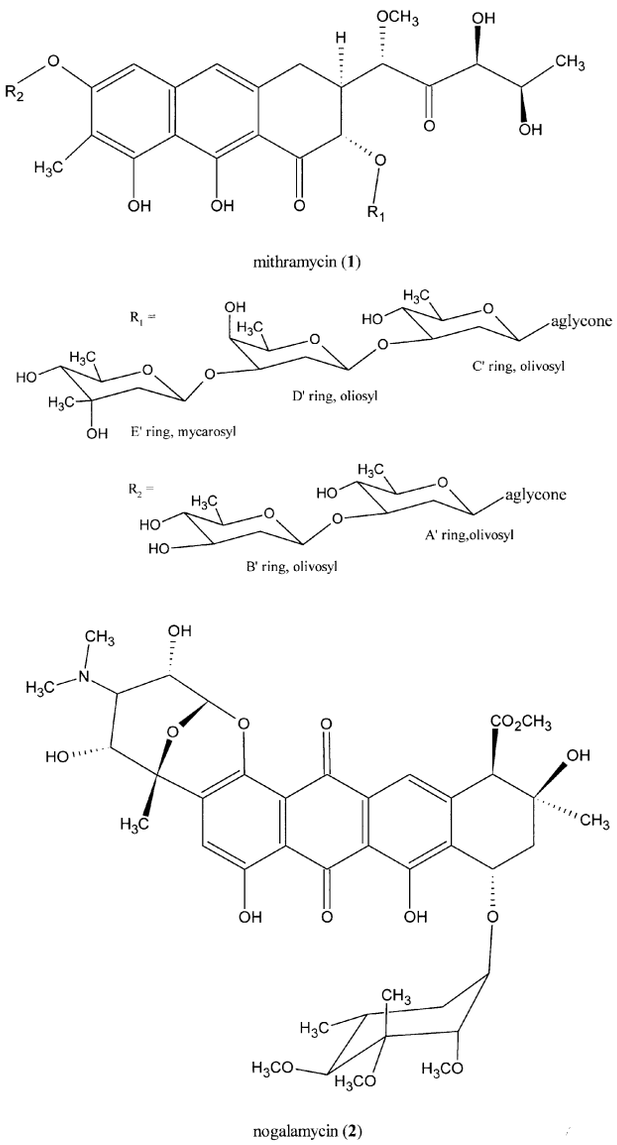


Fig. 1 The structures of mithramycin (**1**) and nogalamycin (**2**).

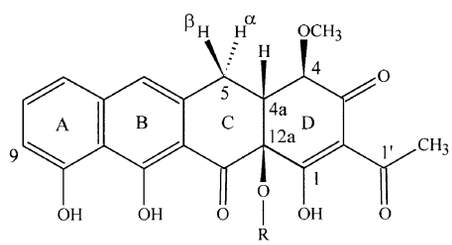
formation of the aglycone moieties. Likewise, the molecular genetics of both these antibiotics is known and the gene clusters for their biosyntheses have been cloned and essentially characterised.[16,18–21] Both mithramycin (**1**) and nogalamycin (**2**) are derived from the condensation of 10 acetates catalysed by minimal polyketide synthase (PKS). However, after formation *in vivo* of the biological equivalent of a polyketide chain, the biosynthetic pathways of these two antibiotics diverge. This divergence thus provides an opportunity to shuffle the biosynthetic routes for the express purpose of generating novel compounds.

Previously, we have described the use of PKS genes for nogalamycin (**2**) (*sno*) derived from *S. nogalater* and introduced them into *S. lividans*,[3] *S. galilaeus*,[4] and *S. steffisburgensis*[22] for the production of new, hybrid compounds. Thus as a continuation of our quest for better drug candidates, herein we report the structures of hybrid compounds resulting from the expression of *sno* genes from *S. nogalater* introduced into a mithramycin (**1**) producer, *S. argillaceus*. Mithramycin (**1**) overproducer *S. argillaceus* A43 was specifically selected as the host for the experiments as it has both a high mithramycin (**1**) titre (300 mg l^{-1}) and a simple production profile. Both of these properties help facilitate the convenient detection of minor metabolites, and the high mithramycin (**1**) titre also partially compensates for the production level decrease which is occasionally observed in hybrid strains. The result was the production of a singular, novel substrate aglycone tethered to a variable, and to some degree controllable, number of sugar units.

Results and discussion

Metabolite production and structural elucidation

A plasmid (pSY21b) derived from *S. nogalater* and encoding PKS for nogalamycin aglycone production[17] was introduced into the host strain *S. argillaceus* A43. During small-scale incubation of the resultant hybrid *S. argillaceus* A43/pSY21b, selection pressure was maintained by the use of thiostrepton and notable differences in the secondary metabolite output were immediately apparent. HPLC analysis of the crude liquid-culture extract indicated that it was composed of four major products (**1** and **3**–**5**, see Figs. 1 and 2) in the ratio 10 : 2 : 5 : 8.



3, R = olivosyl
4, R = olivosyl-oliosyl
5, R = olivosyl-oliosyl-mycarosyl
6, R = H

Fig. 2 The structures of the hybrid compounds **3**–**6** together with the numbering scheme in use [attendant sugar rings are alphabetised similarly to mithramycin (**1**)]. The absolute stereochemistry of positions 4, 4a and 12a is not inferred.

Extraction of the small-scale cultivations showed that the majority of the metabolite mass, *ca.* 90%, was located in the supernatant and consequently the cells were discarded without extraction in subsequent harvests.

The major metabolite was readily identified as mithramycin (**1**) by comparison with an authentic sample of **1** by both UV and HPLC analysis. The similar UV spectra of **3**, **4** and **5** were, however, quite distinct from that of mithramycin (**1**) and the mild acidic hydrolysis of **3**, **4** and **5** resulted in only the one compound, **6** (Fig. 2). Compound **6** was then explicitly shown not to be the aglycone of mithramycin on the basis of HPLC and UV analysis in comparison with an authentic sample of mithramycinone aglycone. The structural elucidation of these novel compounds (**3**–**6**) was then readily facilitated by large-scale fermentation which furnished centigram quantities of each metabolite **3**–**5** (and subsequently **6** by hydrolysis).

The initial structural elucidation focussed on **6**, after which the elucidations of **3**–**5** readily followed on the basis of the structure of **6**. ESI-MS analysis of **6** provided a molecular ion ($398^{+\bullet}$ amu) and degradation pattern consistent with the final structure. The initial expectation was that **6** was a polycyclic structure, and indeed the ^{1}H and ^{13}C spectra of **6** suggested specifically a tetracyclic structure based on their similarity to the spectra of other tetracycles and our prior experience with these types of compound.[3,4,23] The determination of the structural framework and the unambiguous assignment of the carbon shifts were made using a standard combination of DEPT, HSQC, and HMBC experiments. Of immediate note and significance in the ^{1}H NMR was the lack of a hydroxy group at C-8 and saturation transfer NMR experiments indicated the observable presence of only four hydroxy protons upon irradiation of the residual water signal in $CDCl_3$ solution. Extensive keto–enol tautomerism was determined to

be present in the D ring by ^{1}H NMR leading to dramatic changes being observed for H-4 upon variation of the solvent (*e.g.* a sharp doublet in $CDCl_3$, a broad doublet in acetone-d_6, and disappearance of the signal altogether in DMSO-d_6). The size of the vicinal coupling (11 Hz) between H-4 and H-4a clearly indicated a *trans*-diaxial relationship between these two nuclei for the preferred diastereomer in solution (*i.e.* position 4 is a labile stereocentre as a result of the keto–enol tautomerism). However, the relative stereochemistry between the proton at C-4a and the hydroxy group at C-12a could not be unequivocally established by NMR, but the resistance of **6** to chemical dehydration (and consequent aromatisation) and poor loss of H_2O under EI conditions in the MS all implied a *syn* relationship between H-4a and the C-12a hydroxy group; other workers have utilised the same argument.[24] This notion was proved accurate upon obtaining the single-crystal X-ray analysis of **6** which confirmed the *syn* relationship between H-4a and the C-12a hydroxy group (see Fig. 3). Due to the

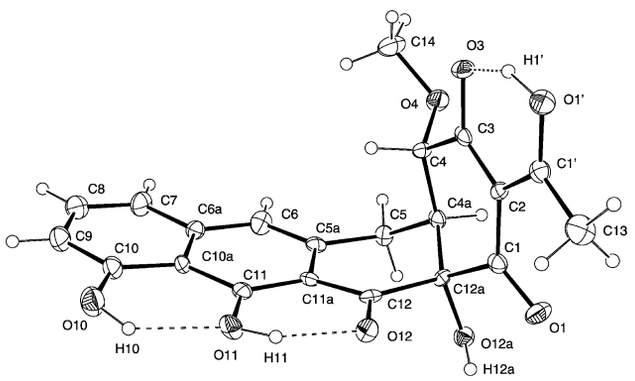


Fig. 3 The single-crystal X-ray structure determination of aglycone **6**. The absolute stereochemistry of positions 4, 4a and 12a is not inferred.

conditions of the measurement (lack of an atom with strong anomalous X-ray scattering and use of Mo-Kα radiation), the absolute stereochemistry of 4, 4a, and 12a could not be ascertained and the chirality of the depiction in Fig. 3 is not inferred. The position of the tautomeric equilibrium in acetone-d_6 for **6** (Fig. 2) was facilitated by observation of the exchangeable proton HO-1 in the NMR and subsequent HMBC correlations supported by the upfield position of C-1 (*ca.* 190 ppm) in comparison to C-3 and C-1′ (*ca.* 196 and 204 ppm, respectively). The similarity of the carbon shifts for all of C-1 to C-3 and C-1′ in compounds **3**–**5** in comparison to **6** implied a similar tautomeric position for these compounds in the same solvent. In contrast to the solution state, in the solid state the position of the tautomeric equilibrium for **6** was decidedly different. Although C-4 was still maintained as a methine carbon and the relative stereochemistry was the same (H-4 *trans* to H-4a), O-1′ is the carbonyl oxygen formally involved in enol formation. Due to strong intramolecular hydrogen bonding, this designation is somewhat arbitrary and, based on all of the bond lengths concerned,§ a better description is that C-2, C-3, O-3, HO-1′, O-1′ and C-1′ form a delocalised, planar enol structure. In contrast to the intramolecular hydrogen bonding of the phenol hydrogens, HO-1′ is located almost equidistant from both O-1′ and O-3 (1.17 and 1.33 Å, respectively), a quite normal occurrence for a β-diketone fragment in its enol form.[25] The planarity, for hybridisation reasons, actually extends out to atoms C-4, C-1 and O-1, with the result that ring D is forced to adopt a half-chair conformation in a similar fashion to the manner in which ring C is forced to adopt a half-chair conformation from being fused to the aromatic ring.

§ Bond lengths: C-2–C-1′, 1.435; C-1′–O-1′, 1.269; O-1′–HO-1′, 1.17; HO-1′–O-3, 1.33; O-3–C-3, 1.294; C-3–C-2, 1.398, Å.

Based on these results it seemed worthwhile to investigate the structure of the molecule in the gas phase for the position of the tautomeric equilibrium derived from the solution state using density functional theory (DFT) calculations. The result is depicted in Fig. 4. Similarities exist between the two structures

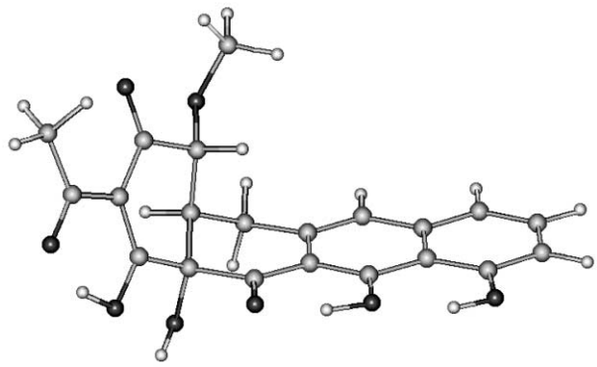


Fig. 4 The calculated structure of aglycone **6** using DFT methodology on the solution-state tautomer in the gas phase. The absolute stereochemistry of poitions 4, 4a and 12a is not inferred and it is the enantiomer of the X-ray structure determination that is depicted here.

of Figs. 3 and 4, namely the intramolecular hydrogen bonding of the phenol hydrogens, the near exact *trans*-periplanar orientation of H-4 and H-4a (179.7°, DFT calc.), and the dihedral angles between H-4a and H-5α and H-4a and H-5β (64.7° and 51.8°, respectively; DFT calc.). However, with the different keto–enol tautomerism it is O-1, C-1, C-2, C-1′, O-1′ and HO-1 that now form a delocalised, planar enol structure based similarly again on all of the bond lengths concerned.¶ As a consequence of the geometric constraints that are imposed by such a structure (*e.g.* co-planarity of C-12a and C-3), subtle changes in the conformation of rings C and D result. Both rings are forced to adopt sofa conformations where in each case it is C-4a that is out of plane. Based on the calculated dihedral angles, it would be expected that H-5β should have a larger observable vicinal coupling constant to H-4a in comparison to H-5α, and experimentally this is indeed the case ($J_{\text{H-5β,H-4a}}$ = 6.7 Hz and $J_{\text{H-5α,H-4a}}$ = 2.5 Hz).

For the compounds **3**–**5**, initial size-exclusion chromatography of the crude extract indicated an approximate mass difference of +100 amu between **3** and **4** and between **4** and **5**. The mass differences were determined more precisely by subsequent ESI-MS analysis (negative mode) on the isolated material which gave the expected $[M - H]^-$ ions for each compound together with diagnostic sequential losses of sugar residues. NMR analysis using standard techniques (COSY, DEPT, HSQC, and HMBC) readily revealed all three compounds (**3**–**5**) to be glycosides with a common aglycone unit (*i.e.* **6**) consisting of one, two, and three sugars, respectively, bound to the C-12a hydroxy group *via* a glycosidic bond. The identity of the sugar units were also readily apparent from NMR analysis and as a series they represent a sequential extension, *i.e.* the sugar in **3** is the first in **4**, and the two sugars in **4** are the first two in **5**. All three sugar units are 2-deoxy sugars and reflect those found in mithramycin (**1**); the trisaccharide sequence of β-D-olivosyl-3-1-β-D-oliosyl-3-1-β-D-mycarosyl for compound **5** is identical to the trisaccharide sequence present in mithramycin (**1**) at the analogous position. The linkage sites in the sugar units in **3**–**5** were established on the basis of the observed interglycosidic HMBC and NOE correlations, particularly the heteronuclear $^3J_{\text{HC}}$ correlations.

¶ Bond lengths: C-1–C-2, 1.394; C-2–C-1′, 1.466; C-1′–O-1′, 1.256; O-1′–HO-1, 1.419; HO-1–O-1, 1.056; O-1–C-1, 1.302, Å. See also http://www.rsc.org/suppdata/p1/b2/b200444p/ for the edited GAUSSIAN output of the DFT calculation and HyperChem file of the calculated structure.

Table 1 ^{1}H NMR data for compounds **3**–**6** in acetone-d_6; δ (ppm), multiplicity, $J_{H,H}$/Hz, area [a]

Comp.	H-4	4-OMe	H-4a	H-5α	H-5β	H-6	H-7	H-8	H-9	H-2′	H–C′1	H–C′2α	H–C′2β	H–C′3	H–C′4	H–C′5
3	3.88, d 11, 1H	3.60, s 3H	2.77, ddd 10.8, 6.5, 2.5, 1H	3.20, dd 16.4, 2.2, 1H	3.57, ddd 16.2, 6.7, 1.8, 1H	7.13, d 1.7, 1H	7.20, dd 8.5, 1.2, 1H	7.55, dd 8.6, 7.2, 1H	6.80, dd 7.2, 1.1, 1H	2.67, s 3H	5.11, dd 9.2, 2.2, 1H	2.66, ddd 12.1, 5.8, 2.0, 1H	1.80, ddd 12.2, 12.0, 9.4, 1H	3.78, ddd 12.1, 9.5, 5.7, 1H	4.55, dd 9.5, 8.8, 1H	3.40, dq 9.0, 6.2, 1H
4	3.92, d 11.2, 1H	3.58, s 3H	2.70, ddd 11.1, 6.2, 2.4, 1H	3.16, dd 16.6, 2.2, 1H	3.50, ddd 16.8, 6.3, 2.0, 1H	7.19, d 1.8, 1H	7.19, dd 8.8, 1.0, 1H	7.50, dd 8.6, 7.5, 1H	6.88, dd 7.5, 1.1, 1H	2.66, s 3H	5.18, dd 9.0, 2.1, 1H	2.70, ddd 12.0, 6.0, 2.2, 1H	1.82, ddd 12.0, 11.6, 9.0, 1H	3.89, ddd 12.0, 9.4, 6.1, 1H	4.50, dd 9.3, 9.0, 1H	3.39, dq 9.0, 6.1, 1H
5	3.91, d 11.0, 1H	3.59, s 3H	2.68, ddd 11.2, 6.4, 2.3, 1H	3.22, dd 16.7, 2.0, 1H	3.52, ddd 16.6, 6.2, 1.5, 1H	7.09, d 1.8, 1H	7.16, dd 8.3, 1.0, 1H	7.55, dd 8.4, 7.4, 1H	6.91, dd 7.3, 1.0, 1H	2.70, s 3H	5.22, dd 9.1, 2.2, 1H	2.62, ddd 12.1, 5.9, 2.1, 1H	1.77, ddd 12.1, 11.8, 9.2, 1H	3.82, ddd 12.1, 9.2, 5.8, 1H	4.49, dd 9.3, 8.9, 1H	3.45, dq 9.1, 6.2, 1H
6	3.95, d 11.1, 1H	3.59, s 3H	2.74, ddd 11.1, 6.6, 2.5, 1H	3.18, dd 16.8, 2.5, 1H	3.54, ddd 16.8, 6.7, 1.6, 1H	7.11, d 1.6, 1H	7.16, dd 8.5, 1.1, 1H	7.52, dd 8.5, 7.4, 1H	6.87, dd 7.4, 1.1, 1H	2.70, s 3H	—	—	—	—	—	—

Comp.	H–C′6	H–D′1	H–D′2α	H–D′2β	H–D′3	H–D′4	H–D′5	H–D′6	H–E′1	H–E′2α	H–E′2β	E′3–Me	H–E′4	H–E′5	H–E′6
3	1.20, d 6.1, 3H	—	—	—	—	—	—	—	—	—	—	—	—	—	—
4	1.22, d 6.1, 3H	4.57, dd 9.3, 2.0, 1H	2.01, ddd 12.4, 4.1, 2.0, 1H	1.77, ddd 12.3, 11.0, 9.1, 1H	3.70, dt 11.1, 4.3, 1H	5.02, dd 4.2, 1.3, 1H	3.51, qd 6.1, 1.2, 1H	1.15, d 6.2, 3H	—	—	—	—	—	—	—
5	1.18, d 6.2, 3H	4.55, dd 9.1, 2.0, 1H	1.98, ddd 12.3, 4.2, 2.1, 1H	1.79, ddd 12.2, 11.1, 9.1, 1H	3.90, dt 11.4, 4.5, 1H	5.00, dd 4.1, 1.0, 1H	3.50, qd 6.2, 1.0, 1H	1.16, d 6.1, 3H	4.88, dd 9.3, 2.0, 1H	1.88, dd 13.0, 2.1, 1H	1.56, dd 13.1, 9.0, 1H	1.10, s 3H	4.49, d 9.3, 1H	3.88, dq 9.2, 6.0, 1H	1.14, d 6.1, 3H
6	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

[a] The exchangeable protons were not observed for compounds **3**–**5**. For compound **6**: 18.23 (s, 1H, 1-OH), 9.44 (s, 1H, 10-OH), 15.00 (s, 1H, 11-OH), 5.00 (s, 1H, 12a-OH).

Table 2 ^{13}C NMR data for compounds **3**–**6** in acetone-d_6; δ (ppm), multiplicity

Comp.	C-1	C-2	C-3	C-4	C-4a	4-OMe	C-5	C-5a	C-6	C-6a	C-7	C-8	C-9	C-10	C-10a	C-11	C-11a	C-12	C-12a	C-1′
(Mult.)	s	s	s	d	d	q	t	s	d	s	d	d	d	s	s	s	s	s	s	s
3	189.2	111.0	195.7	77.2	43.0	51.9	26.2	132.2	118.3	140.7	119.6	134.9	112.0	157.5	113.0	166.3	109.0	197.2	80.1	203.5
4	190.1	112.1	196.2	75.6	44.2	52.2	26.0	130.2	118.0	141.2	119.5	134.2	110.3	157.8	111.1	166.9	108.0	197.7	80.2	203.9
5	189.9	111.2	196.3	76.6	44.0	52.0	26.8	131.8	119.8	141.0	120.7	133.1	110.2	159.0	110.9	166.1	108.1	196.0	79.2	204.0
6	190.5	110.1	196.5	76.8	43.9	51.9	25.9	133.1	118.4	139.9	119.3	133.7	111.7	158.1	112.6	166.8	108.3	196.7	76.9	203.7

	C-2′	C–C′1	C–C′2	C–C′3	C–C′4	C–C′5	C–C′6	C–D′1	C–D′2	C–D′3	C–D′4	C–D′5	C–D′6	C–E′1	C–E′2	C–E′3	E′3–Me	C–E′4	C–E′5	C–E′6
	q	d	t	d	d	d	q	d	t	d	d	d	q	d	t	s	q	d	d	q
3	28.0	100.6	37.0	69.3	75.0	69.2	17.3	—	—	—	—	—	—	—	—	—	—	—	—	—
4	28.1	100.1	36.9	74.2	74.8	69.8	17.5	98.2	32.3	70.2	70.8	68.4	16.9	—	—	—	—	—	—	—
5	28.1	100.3	37.0	73.9	75.1	69.2	17.6	99.0	32.3	74.2	70.2	68.9	16.8	96.2	43.0	71.0	26.8	76.4	67.9	17.7
6	28.1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

This was also evident from the downfield shifts of 4–5 ppm for the relevant carbon in the extended member of the series, *viz.* C–D′3 in **5** in comparison to **4** and C–C′3 in both **4** and **5** in comparison to **3**. Similarly too, C-12a resonated downfield by 2.3–3.3 ppm in compounds **3**–**5** in comparison to **6** due to the attachment of the sugar units to the hydroxy group attached to C-12a. Complete ^{1}H and ^{13}C shift assignments for compounds **3**–**6** are given in Tables 1 and 2, respectively, and are in good accord with reference compounds.[8] The hybrid compounds **3**–**5** and their hydrolysis product **6** were all ascertained to be novel structures.

Biosynthetic pathways

The polyketide backbone of the hybrid compounds **3**–**6**, not surprisingly, originates from the conglomeration of ten acetate units since the number and nature of the building blocks used in the biosynthesis of the aglycones of both mithramycin (**1**) and nogalamycin (**2**) are identical, *viz.* ten acetate units. The structures of the hybrid products **3**–**5**, however, clearly show the interaction of the *S. nogalater* genes with those of the host *S. argillaceus* A43 since the novel metabolites display modifications due to the genes provided by pSY21b. (That the original biogenesis still remains highly viable and presumably operates independently without interference from the introduced enzymes is not without precedence.[26])

The absence of the hydroxy group at C-8 demonstrates the expression of the *S. nogalater* genes for both a ketoreductase and an aromatase. The plasmid pSY21b carries the genes for: minimal PKS (*sno*a1–3), thereby dictating the condensation of ten acetates; a ketoreductase (*sno*aD), thereby reducing the keto group at C-8;[3,4,22] and an aromatase (*sno*aE), which results in the formation of the aromatic A ring. From there, the *sno* genes *S. argillaceus* A43 cyclise the structure further to produce **6**, the C-8 dehydroxylated form of premithramycinone (**7**) and notable is that the same relative stereochemistry between 4, 4a, and 12a is produced. The genes *sno*aD and *sno*aE each separately failed to result in the production of the hybrid compounds as was demonstrated by the introduction of each of the genes in pIJ486 into *S. argillaceus* A43, thereby indicating that reduction by *sno*aD is not occurring as a post-aromatisation process on the decaketide. If such a process were occurring it would actually be operating on a substrate that equates to premithramycinone (**7**), since compounds **6** and **7** differ only by the hydroxy group at C-8. (Note that the different tautomeric positions indicated in Scheme 1 for **6** and **7** merely reflect the different solvents used for their structural elucidation.) In any event, pre-aromatisation reduction has been well documented for *sno*aD.[3,4,22] An inevitable consequence of the reduction of C-8 is, of course, the absence of any sugar units attached to a hydroxy group located at C-8. Interestingly, post-polyketide C-9 methylation is inhibited in the hybrid compounds **3**–**6**, implying that the C-8 hydroxy group plays a pivotal role—most presumably by bonding interactions—in assisting the methylase responsible for the methylation at C-9. In a similar way, oxidative cleavage of **6** akin to the oxidation of premithramycinone (**7**) to yield mithramycin (**1**) is not observed, even though the mechanism remains fully operational as reflected by the high-level production of mithramycin (**1**). Presumably, the hybrid compounds **3**–**6** are not recognised as a target by the enzyme responsible for the last step in the mithramycin (**1**) pathway for formation of the aglycone and perhaps this can similarly be attributed to the missing C-8 hydroxy group. Thus although structurally very similar to premithramycinone (**7**), the aglycone **6** is nevertheless the result of a significantly different biochemical pathway brought about by the introduction of the *S. nogalater* genes. Scheme 1 neatly summarises the biosynthetic pathways for each of the parent strains leading to premithramycinone (**7**) and nogalamycinone (**8**), and the hybrid strain *S. argillaceus* A43/pSY21b leading to the hybrid aglycone **6** and subsequent glycosides **3**–**5**.

The glycosylation profile, together with the premithramycinone (**7**) framework, is, however, determined by the host *S. argillaceus* A43, as the typical mithramycin sugars, D-oliose, D-olivose and D-mycarose, are all expressed. The appearance of derivatives containing from one to three sugar units indicates the sequential addition of sugar residues to the aglycone. This is moreover confirmed by the time profile of the metabolite concentrations which clearly revealed the process of sequential glycosylation as **3** appears first after 48 hours, followed by **4**, and then finally by **5**. Eventually **5** accumulates as the major hybrid product. The aglycone **6** itself was detected only as a minor component in the crude extract (whole broth), reflecting the efficient glycosylation properties of the strain which is nonetheless moderated with respect to glycosylation of the natural aglycone substrate.

Conclusions

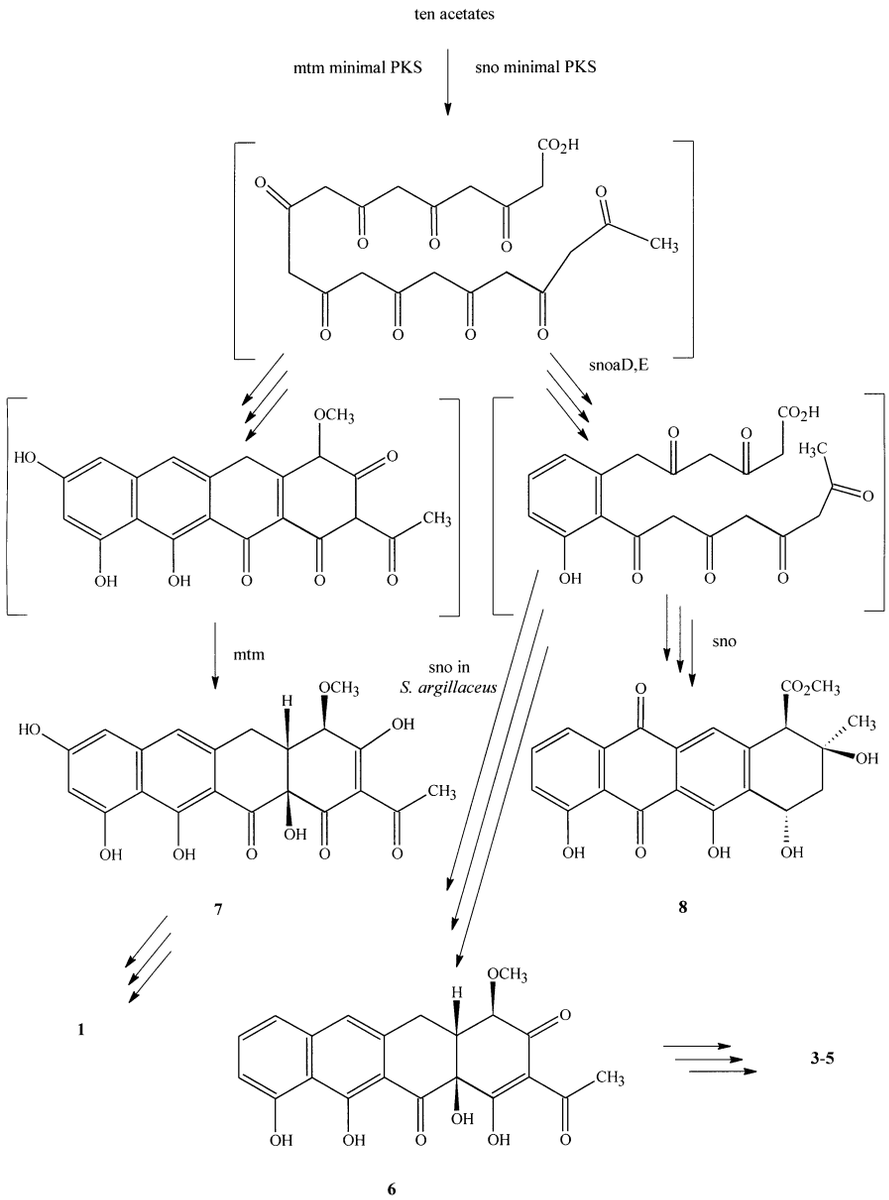
The value of the genetic engineering of antibiotic-producing microbes to produce drug candidates lies in the fact that the approach is more methodical and directed in comparison to the serendipitous nature of random mutagenesis, and in the potential that exists for being able to design multi-step syntheses, particularly when biocompatibility between a substrate and an organism is requisite. These applications are yet to be realised at a practical level but are potentially of immense value upon implementation and this work represents one step closer to that practical realisation. In this particular biosynthesis, three new compounds were produced that shared features of the biosynthetic pathways of each of the "parents" from which the mutant was composed. The synthetic value of the mutant was highlighted by the high titre of the shunt products, 60% of the metabolic output, together with a lack of structural diversity of the hybrid aglycone present in the products, two features which are not normally observed. The specific addition of a sugar residue(s) to a bioactive substrate is important as it may accentuate or mitigate the efficacy of a drug (or drug candidate); the latter response can also be of value for the development of sugar-based prodrug candidates[27] in the case of excessive toxicity where the "magic bullet" approach[28] may be strategically worth pursuing, and this has been explored using ADEPT or GDEPT techniques for particular anthracyclines.[27] In this case here, all four compounds were screened for potential anti-cancer properties but were deemed insufficiently active to warrant further attention.

Experimental

NMR, MS, X-ray structure determination, computational method and optical rotation

NMR spectra were recorded on a JEOL Lambda 400 spectrometer equipped with either a 5 mm normal-configuration ^{13}C{^{1}H} probe or an inverse-configuration ^{1}H{X} probe operating at 399.78 MHz for ^{1}H and 100.54 MHz for ^{13}C. Spectra were recorded at 25 °C and both ^{1}H and ^{13}C spectra were referenced internally to tetramethylsilane (0 ppm for both).

^{1}H spectra were acquired with single-pulse excitation, 45° flip angle, pulse recycle time of 9.5 s and with spectral widths of 7 kHz consisting of 64K data points (digital resolution 0.11 Hz pt^{-1}), zero-filled to 128K prior to Fourier transformation. NOE difference measurements were acquired using saturation times of 6 s and with reduced resolution (3.9 Hz pt^{-1}); 1 Hz of exponential weighting was usually applied prior to Fourier transformation. (NOE difference measurements also provided concomitantly saturation transfer information). DQF COSY spectra were acquired in phase-sensitive mode with spectral



Scheme 1 Depiction of the chemical equivalent of the polyketide chain structure derived from ten acetates and the biosynthetic steps leading to the hybrid aglycone (**6**), premithramycinone (**7**) and nogalamycinone (**8**) by the mutant *S. argillaceus* A43/pSY21b, the host strain *S. argillaceus* A43 and *S. nogalater*, respectively. Note that the position of the tautomeric equilibrium in **6** and **7** is solvent dependent and the depicted position of the equilibrium for **6** is based on the results in this work (using acetone-d_6) whilst the position of the equilibrium for **7** is based on the reports of other workers[24] (using methanol-d_4).

widths and resolution appropriately optimised from the 1D spectra, and processed with zero-filling (×2, ×4) and exponential weighting (1 Hz) applied in both dimensions prior to Fourier transformation.

^{13}C spectra were acquired with single-pulse excitation, 45° flip angle, pulse recycle time of 3.5 s and with spectral widths of 30 kHz consisting of 64K data points (digital resolution 0.46 Hz pt^{-1}), and with 1 Hz exponential weighting applied prior to Fourier transformation. DEPT 135° spectra were acquired with similar spectral windows and with a pulse delay time of 3 s. HSQC PMG[29] and HMBC BIRD experiments were carried out in phase-sensitive mode and magnitude mode, respectively, with spectral widths and resolution appropriately optimised from the 1D spectra and processed with zero-filling (×2, ×4), a 2π/3-shifted sinebell function (for HMBC spectra), and exponential weighting (5 Hz, 25 Hz) applied in both dimensions prior to Fourier transformation. Both HSQC and HMBC spectra utilised a $^1J_{HC}$ coupling of 145 Hz, whilst the HMBC correlations were optimised for a long-range $^nJ_{HC}$ coupling of *ca.* 8 Hz. The length of the purge pulse (typically 0.7 ms) and BIRD relaxation delay (typically 400 ms) were optimised on the incoming FID.

Low-resolution ESI mass spectra in the negative mode were acquired on a PE Sciex API 365 LC/MS/MS instrument using direct infusion as the inlet source (source temperature 300 °C). High-resolution ESI mass spectra were acquired on a VG Analytical ZabSpec instrument under similar conditions. EI mass spectra were acquired on a VG Analytical 7070E instrument using 70 eV for ionisation.

X-Ray crystallographic data were collected on a Nonius Kappa CCD area-detector diffractometer using graphite monochromatised Mo-Kα radiation (λ = 0.71073 Å). Lattice parameters were determined from 10 images recorded with 1° ϕ scans and subsequently refined on all data. The data collection was performed using ϕ and ω scans with 1° steps using an exposure time of 150 s per frame. The crystal-to-detector distance was 30 mm. The data were processed using DENZO-SMN v0.93.0.[30] The structure was solved by direct methods using the SHELXS-97 program[31] and full-matrix least-squares refinements on F^2 were also performed using the SHELXL-97

program.[31] All heavy atoms were refined anisotropically and the hydroxy hydrogen atoms isotropically. The carbon-bound hydrogen atoms were included at calculated distances from their host atoms with fixed displacement parameters. Determination of the absolute configuration was not possible from the data using the Flack parameter, which gave values of −0.7(2.2) and 1.6(2.2) for the two enantiomers. The figure was drawn using the program ORTEP-3 for Windows.[32]

The calculated structure was pre-optimised using the semi-empirical PM3 method[33,34] implemented within the HyperChem version 4.5 program package.[35] A loose convergence criterion was set in the pre-optimisation in order to find a suitable starting geometry for the DFT[36–38] optimisation. DFT geometry optimisation was performed using the GAUSSIAN 98 program[39] using a hybrid B3LYP method[40–43] and 6-31G(d,p) basis set.[44] The optimisation was performed using a spin-restricted method and a tight SCF convergence criterion.

The optical rotation was measured using a Perkin–Elmer model 241 polarimeter and expressed as $[\alpha]_D$ values in 10^{-1} deg cm^2 g^{-1} units.

Bacterial strains, culture conditions and plasmids

S. argillaceus A43 was derived from *S. argillaceus* ATCC 12956 by random chemical mutagenesis using methodology previously described.[45] The plasmid pSY21b containing the genes *sno*a1–3 and *sno*aR,D,E (previously known as *sno*1–3 and *sno*A,D,E, respectively) was constructed using the vector pIJ486 as described earlier.[17] pSY21b was introduced into *S. argillaceus* A43 by protoplast transformation using standard methods.[1] Plasmid carrying strains were maintained on ISP4-agar supplemented by thiostrepton (50 μg ml^{-1}).

HPLC analysis and isolation of compounds 1, 3–5

50 ml of seed culture of *S. argillaceus* A43/pSY21b together with 50 mg of thiostrepton in 1 ml of DMSO were used to inoculate a 13 l Biostat E jar fermentor containing 10 l of media composed of glucose, starch, Pharmamedia (Traders protein), yeast extract, $CaCO_3$, NaCl, $MgSO_4$, and KH_2PO_4 in tap water. Prior to sterilisation, the pH was adjusted to 7.5 with NaOH. After 160 h the broth was cross-flow filtered. The cells were discarded and the supernatant treated with 250 g of Amberlite XAD-7 resin for one hour after which the resin was collected by vacuum filtration and extracted with acetone (3 × 1 l). The acetone solution was vacuum concentrated to *ca.* 5% of its volume and flash chromatographed [RP-18 phase; gradient elution, 1% CH_3CO_2H and CH_3CN (0–100%)]. Fractions were analysed by analytical HPLC (RP-18ec phase); column, 55 × 4 mm; particle size, 3 μm; gradient elution, 0.1% HCO_2H and CH_3CN (0–100%); UV detection at 260 nm. Appropriate fractions were pooled and dried *in vacuo*. The four lots so obtained were then column chromatographed (Sephadex LH-20 phase; column, 30 × 4 cm; elution, 20% H_2O in CH_3OH). Pure fractions were evaporated until precipitation occurred. The yellow precipitates were collected and dried to yield 200 mg of **1** (98% pure by HPLC), 15 mg of **3** (95%), 50 mg of **4** (97%), and 70 mg of **5** (98%). The ^1H and ^{13}C NMR data for compounds **3**–**5** are listed in Tables 1 and 2, respectively. High-resolution MS (−ve mode ESI) analysis yielded the following elemental compositions for the $[M - H]^-$ ions of compounds **3**–**5**: $C_{27}H_{27}O_{11}$ (found: 527.1545; requires: 527.1553), $C_{33}H_{37}O_{14}$ (found: 657.2183; requires: 657.2183), and $C_{40}H_{49}O_{17}$ (found: 801.2979; requires: 801.2970), respectively.

Hydrolysis of compounds 3, 4 and 5 to 6

Fractions containing **2**, **3**, **4**, or mixtures thereof were combined, dried *in vacuo* and taken up in water containing 0.5% of conc. HCl (aq.). The solution was heated to 50 °C and the reaction monitored by TLC ($CHCl_3$–CH_3OH–CH_3CO_2H–H_2O, 834 : 100 : 50 : 16). After 15 min the solution was neutralised with 0.1 M NaOH and washed with dichloromethane. Compound **6** was extracted from the aqueous fraction at pH 3 using chloroform; precipitation of **6** from the chloroform solution was effected by the addition of hexane. The precipitate was dissolved in DMSO and chromatographed semipreparatively [RP-18 phase; column, 250 × 10 cm; particle size, 5 μm; gradient elution (linear), 70% (1% CH_3CO_2H)–CH_3CN to 20% (1% CH_3CO_2H)–CH_3CN]. Fractions of pure **6** were pooled and evaporated to a third of the volume effecting the precipitation of yellow crystals which were dried to yield 50 mg of **6**, 97% pure by HPLC. The ^1H and ^{13}C NMR data for **6** are listed in Tables 1 and 2, respectively. High-resolution MS (+ve mode EI) analysis yielded an elemental composition of $C_{21}H_{18}O_8$ for the $M^{+\bullet}$ ion (found: 398.1003; requires: 398.1002). $[\alpha]_D^{22}$ −33.3° (*c* 0.00096 in acetone).

Crystal structure determination of compound 6‖

Single crystals of compound **6** were grown by the slow evaporation of an acetone solution and obtained as thin plates. Crystal data. $C_{21}H_{18}O_8$, M = 398.35, monoclinic, a = 5.1675(2), b = 10.5383(6), c = 16.2748(9) Å, β = 98.367(2)°, U = 876.84(8) Å^3, T = 173 K, space group $P2_1$ (No. 4), Z = 2, D_c = 1.509 g cm^{-3}, μ(Mo-Kα) = 0.117 mm^{-1}, crystal dimension 0.02 × 0.10 × 0.20 mm, 5231 reflections measured, 3055 unique (R_{int} = 0.088) which were used in all calculations. The final $wR(F^2)$ was 0.132 (all data).

Acknowledgements

The support of the European Commission, grant no. BIO4–CT96–0068, is gratefully acknowledged.

‖ CCDC number 183073. See http://www.rsc.org/suppdata/p1/b2/b200444p/ for the crystallographic file in .cif or other electronic format. The edited GAUSSIAN output of the DFT calculation and the HyperChem file of the calculated structure can also be found at this web address.

References

1 D. A. Hopwood, F. Malpartida, H. M. Kieser, H. Ikeda, J. Duncan, I. Fujii, B. A. M. Rudd, H. G. Floss and S. Omura, *Nature (London)*, 1985, **314**, 642.
2 D. A. Hopwood, *Chem. Rev.*, 1997, **97**, 2465.
3 T. Kunnari, J. Kantola, K. Ylihonko, K. D. Klika, P. Mäntsälä and J. Hakala, *J. Chem. Soc., Perkin Trans. 2*, 1999, 1649.
4 T. J. Kunnari, K. P. J. Ylihonko, K. D. Klika, P. I. Mäntsälä and J. M. L. Hakala, *J. Org. Chem.*, 2000, **65**, 2851.
5 C. E. Anson, M. J. Bibb, K. I. Booker-Milburn, C. Clissold, P. J. Haley, D. A. Hopwood, K. Ichinose, W. P. Revill, G. R. Stephenson and C. M. Surti, *Angew. Chem., Int. Ed.*, 2000, **39**, 224.
6 G. H. Davies, R. H. Green, D. R. Kelly and S. M. Roberts, in *Biotransformations in Preparative Organic Chemistry: The Use of Isolated Enzymes and Whole Cell Systems in Synthesis (Best Synthetic Methods)*, Academic Press, 1989.
7 J. D. Skarbek and M. K. Speedie, in *Antitumor Compounds of Natural Origin: Chemistry and Biochemistry*, ed. A. Aszalos, CRC Press, Boca Raton, FL, 1981, vol. 1.
8 S. E. Wohlert, E. Künzel, C. Méndez, J. A. Salas and J. Rohr, *J. Nat. Prod.*, 1999, **62**, 119.
9 W. E. Grundy, A. W. Goldstein, C. J. Rickher, M. E. Hanes, H. B. Warren Jr. and J. C. Sylvester, *Antibiot. Chemother.*, 1953, **3**, 1215.
10 B. J. Kennedy, J. W. Yarbro, V. Kickertz and M. Sandberg-Wollheim, *Cancer Res.*, 1968, **28**, 91.
11 E. G. Elias and J. T. Evans, *J. Bone Joint Surg.*, 1972, **54A**, 1730.
12 M. Sastry, R. Fiala and D. J. Patel, *J. Mol. Biol.*, 1995, **251**, 674.
13 G. P. Bakhaeva, Yu. A. Berlin, E. F. Boldyreva, O. A. Chuprunova, M. N. Kolosov, V. S. Soifer, T. E. Vasiljeva and I. V. Yartseva, *Tetrahedron Lett.*, 1968, **32**, 3595.
14 J. Thiem and B. Meyer, *Tetrahedron*, 1981, **37**, 551.
15 N. Rama Krishna, D. M. Miller and T. T. Sakai, *J. Antibiot.*, 1990, **43**, 1543.

16 J. Rohr, C. Méndez and J. A. Salas, *Bioorg. Chem.*, 1999, **27**, 41.
17 J. Kantola, T. Kunnari, A. Hautala, J. Hakala, K. Ylihonko and P. Mäntsälä, *Microbiol.*, 2000, **146**, 155.
18 L. Prado, F. Lombó, A. F. Braña, C. Méndez, J. Rohr and J. A. Salas, *Mol. Gen. Genet.*, 1999, **261**, 216.
19 F. Lombó, G. Blanco, E. Fernandez, C. Méndez and J. A. Salas, *Gene*, 1996, **172**, 87.
20 K. Ylihonko, J. Tuikkanen, S. Jussila, L. Cong and P. Mäntsälä, *Mol. Gen. Genet.*, 1996, **251**, 113.
21 S. Torkkell, K. Ylihonko, J. Hakala, M. Skurnik and P. Mäntsälä, *Mol. Gen. Genet.*, 1997, **256**, 203.
22 T. Kunnari, J. Tuikkanen, A. Hautala, J. Hakala, K. Ylihonko and P. Mäntsälä, *J. Antibiot.*, 1997, **50**, 496.
23 T. Kunnari, K. Ylihonko, A. Hautala, K. D. Klika, P. Mäntsälä and J. Hakala, *Bioorg. Med. Chem. Lett.*, 1999, **9**, 2639.
24 J. Rohr, U. Weißbach, C. Beninga, E. Künzel, K. Siems, K. U. Bindseil, F. Lombó, L. Prado, A. F. Braña, C. Méndez and J. A. Salas, *Chem. Commun.*, 1998, 437.
25 G. Gilli, F. Bellucci, V. Ferretti and V. J. Bertolasi, *J. Am. Chem. Soc.*, 1989, **111**, 1023.
26 T. Taguchi, Y. Ebizuka, D. A. Hopwood and K. Ichinose, *Tetrahedron Lett.*, 2000, **41**, 5253.
27 A. K. Ghosh, S. Khan, F. Marini, J. A. Nelson and D. Farquhar, *Tetrahedron Lett.*, 2000, **41**, 4871.
28 C. Rader and B. List, *Chem. Eur. J.*, 2000, **6**, 2091.
29 J.-M. Nuzillard, G. Gasmi and J.-M. Bernassau, *J. Magn. Reson.*, 1993, **104(A)**, 83.
30 Z. Otwinowski and W. Minor, in *Methods in Enzymology Macromolecular Crystallography Part A*, eds C. W. Carter Jr. and R. M. Sweet, Academic Press, New York, 1997, pp. 307–326.
31 G. M. Sheldrick, SHELX-97, University of Göttingen, Germany, 1997.
32 L. J. Farrugia, *J. Appl. Crystallogr.*, 1997, **30**, 565.
33 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209.
34 J. J. P. Stewart, *J. Comput. Chem.*, 1989, **10**, 221.
35 HyperChem Release 4.5 for Windows, Hypercube Inc., 1995.
36 P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864.
37 W. Kohn and L. J. Sham, *Phys. Rev. A*, 1965, **140**, 1133.
38 R. G. Parr and W. Yang, in *Density Functional Theory Of Atoms And Molecules*, Oxford Science Publishers, New York, 1989.
39 GAUSSIAN 98 (Revision A.11), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople, Gaussian Inc., Pittsburgh, PA, 2001.
40 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
41 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
42 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
43 B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200.
44 W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, in *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986.
45 K. Ylihonko, J. Hakala, J. Niemi, J. Lundell and P. Mäntsälä, *Microbiology*, 1994, **140**, 1359.