The Journal of Antibiotics (2017) 70, 404–413

www.nature.com/ja

ORIGINAL ARTICLE

Characterization of the biosynthetic gene cluster (*ata*) for the A201A aminonucleoside antibiotic from *Saccharothrix mutabilis* subsp. *capreolus*

Irene Saugar[1], Brian Molloy[1], Eloisa Sanz, María Blanca Sánchez, María Fernández-Lobato and Antonio Jiménez

Antibiotic A201A produced by *Saccharothrix mutabilis* subsp. *capreolus* NRRL3817 contains an aminonucleoside (*N*^{6}, *N*^{6}-dimethyl-3′-amino-3′-deoxyadenosyl), a polyketide (α-methyl-*p*-coumaric acid) and a disaccharide moiety. The heterologous expression in *Streptomyces lividans* and *Streptomyces coelicolor* of a *S. mutabilis* genomic region of ~34 kb results in the production of A201A, which was identified by microbiological, biochemical and physicochemical approaches, and indicating that this region may contain the entire A201A biosynthetic gene cluster (*ata*). The analysis of the nucleotide sequence of the fragment reveals the presence of 32 putative open reading frames (ORF), 28 of which according to boundary gene inactivation experiments are likely to be sufficient for A201A biosynthesis. Most of these ORFs could be assigned to the biosynthesis of the antibiotic three structural moieties. Indeed, five ORFs had been previously implicated in the biosynthesis of the aminonucleoside moiety, at least nine were related to the biosynthesis of the polyketide (*ata-PKS1-ataPKS4, ata18, ata19, ata2, ata4* and *ata7*) and six were associated with the synthesis of the disaccharide (*ata12*, *ata13, ata16, ata17, ata5* and *ata10*) moieties. In addition to AtaP5, three putative methyltransferase genes are also found in the *ata* cluster (Ata6, Ata8 and Ata11), and no regulatory genes were found.
The Journal of Antibiotics (2017) **70,** 404–413; doi:10.1038/ja.2016.123; published online 12 October 2016

INTRODUCTION

A201A is a complex aminoacylnucleoside antibiotic that has recently been totally synthetized chemically,[1] but that was reported much earlier to be produced naturally by *Saccharothrix mutabilis* subsp. *capreolus* (formerly *Streptomyces capreolus*) NRRL3817.[2] Curiously, the same nucleoside antibiotic has been also reported to be isolated from crude extracts of a deep-sea bacterium named *Marinactinospora thermotolerans* SCSIO 00652[3] A201A is very active against Gram-positive aerobic and anaerobic bacteria, as well as most Gram-negative anaerobic species. However, it is much less toxic to aerobic Gram-negative bacteria, some fungi and mammals.[4] Its chemical structure includes a moiety derived from D-rhamnose, the *N*^{6}, *N*^{6}-dimethyl-3′-amino-3′-deoxyadenosyl (aminonucleoside) moiety of puromycin, an α-methyl-*p*-coumaric acid (a polyketide) and an unsaturated furanose moiety closely related to similar structures found in hygromycin A (Figure 1).[5] The structural basis for both Hygromycin A and A201A antibiotics binding and inhibition to ribosomes have been recently showed by Polikanov *et al.*[6] The similarities of A201A structure with puromycin and hygromycin A antibiotics strongly suggest that certain enzymes, and hence the corresponding genes in the A201A biosynthetic pathway, may be related to their counterparts in the puromycin and hygromycin A biosynthetic pathways.[7–11] Homologs of the *ata* open reading frames (ORFs) have indeed been found with for at least 14 ORFs of the hygromycin A biosynthetic cluster in *S. hygroscopicus*[9] and as we showed previously a set of five consecutive genes involved in the biosynthesis of the aminonucleoside moiety of A201A[12] and their deduced products (AtaP3, AtaP4, AtaP5, AtaP7 and AtaP10) are similar to their counterparts from the *pur* cluster of *S. alboniger* (Pur3, Pur4, Pur5, Pur7 and Pur10, respectively), genes implicated in the biosynthesis of the aminonucleoside moiety of puromycin. Not surprisingly, *ataP4*, *ataP5* and *ataP10* can complement corresponding mutations in the *pur* cluster in *S. alboniger*.[12] Here, we describe the cloning, heterologous expression and organization of the A201A biosynthetic *ata* gene cluster from *S. mutabilis* subsp. *capreolus* NRRL3817.

Centro de Biología Molecular Severo Ochoa (CSIC/UAM), Departamento de Biología Molecular, Universidad Autónoma de Madrid, Cantoblanco Madrid, Spain
[1]These authors contributed equally to this work.
This work is dedicated in honor of Professor Julian Davies who has always been a leading scientist in Microbiology and Molecular Biology. Indeed, he reported many important 'firsts' in these fields, thanks to his brilliant insights and ability to act as a mentor to his collaborators. In addition, Julian is an excellent friend, gourmet, head of a delightful family and shares a superb dry sense of humor.
Correspondence: Professor M Fernández-Lobato, Centro de Biología Molecular Severo Ochoa (CSIC/UAM), Departamento de Biología Molecular, Nicolás Cabrera no 1, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain.
E-mail: mfernandez@cbm.uam.es
Received 13 June 2016; revised 16 August 2016; accepted 1 September 2016; published online 12 October 2016

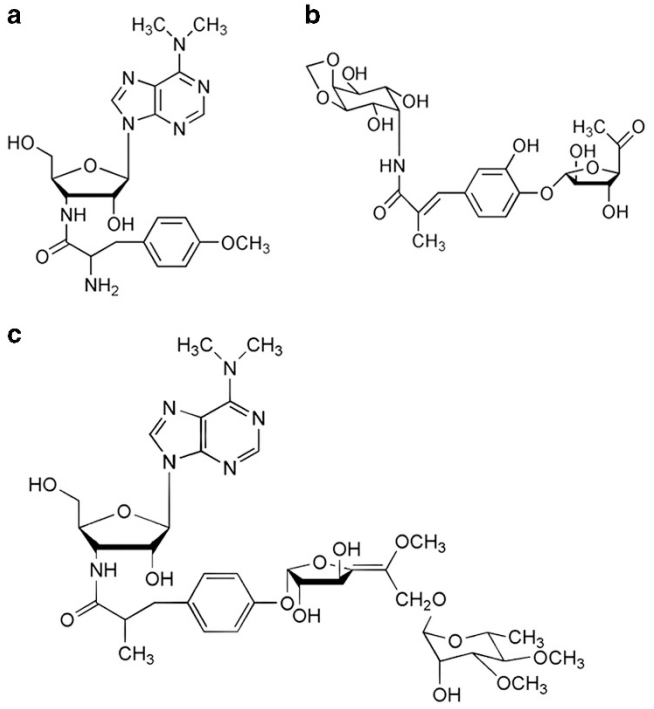

Figure 1 Structures of puromycin (**a**), hygromycin (**b**) and A201A (**c**).

EXPERIMENTAL PROCEDURES

Strains, plasmids and culture conditions

The bacterial strains used in this study are listed in Table 1. *Saccharothrix mutabilis* subsp. *capreolus* NRRL3817 that produces A201A, *Streptomyces lividans* 66–1326, *Streptomyces lividans* 66-TK21 and *Streptomyces colelicolor* M145 have all been described previously (Table 1). The *E. coli* strains DH5α, BW25113(pIJ790) and ET12567(pUZ8002) were employed for DNA amplification, gene disruption by PCR-targeting[13] and *Streptomyces* conjugation, respectively. The pPACS2 *Streptomyces-E. coli* shuttle plasmid, which incorporates the FC31 *attP-int* system and confers resistance to thiostrepton, could be stably maintained in *Streptomyces* by specific integration at the *att*B site of the genome. The plasmid pIJ773, containing the apramycin resistance gene *aac(3)IV*, was used to obtain the disruption cassette by PCR-amplification. pEM4neo, a plasmid derived from pEM4[14] in which the *bla* gene (860 bp) is replaced with *neo* (774 bp) from SuperCos1[15] by λ-Red-mediated recombination, was used for gene expression in *Streptomyces*.

E. coli strains were grown in liquid or agar LB medium. *S. mutabilis* subsp. *capreolus* was grown in NE or MEY medium,[12] whereas *S. lividans* and the other different *Streptomyces* strains generated in this study were grown in either solid SFM[13] or liquid YEME medium.[15] Modified S medium (1% starch, 0.4% Bacto-peptone, 0.4% yeast extract, 50 mg l^{-1} KH_2PO_4, 200 mg l^{-1} K_2HPO_4, 2 mM $MgSO_4$) was employed to study A201A production. When required, antibiotics were added at the following concentrations: 100 μg ml^{-1} ampicillin; 10 and 25 μg ml^{-1} thiostrepton for liquid and solid media, respectively; 50 μg ml^{-1} kanamycin; 25 μg ml^{-1} chloramphenicol; 50 μg ml^{-1} apramycin. A201A was kindly provided by Dr H.A. Kirst (Eli Lilly Research Laboratories). The antimicrobial activity of the *Streptomyces* strains was determined by the agar disc method using *Micrococcus luteus* as the test organism as indicated previously.[16]

Nucleic acid methodology

DNA manipulations, preparation of competent cells, and storage and transformation of *Streptomyces* and *E. coli* were performed as previously described.[15,17] To prepare a gene library from *S. mutabilis* subsp. *capreolus* NRRL3817, total DNA was obtained and partially digested with *Sau*3AI as indicated previously.[12] The resulting 40–80 kb DNA fragments were collected from sucrose gradients and then ligated into the *Bam*HI-digested pPACS2 plasmid.[18] The library was screened by colony hybridization on nitrocellulose Hybond-N membranes (Amersham) using fragments from the ends of the cosmid inserts of pCAR13 and pCAR23 as DNA probes described,[16,19] isolating plasmid pIES. This plasmid contained an insert of ~60 kb that included the A201A biosynthetic (*ata*) gene cluster from *S. mutabilis* subsp. *capreolus*. Appropriate restriction fragments from the aforementioned cosmids were subcloned into the pBS plasmid and sequenced. For heterologous expression of the *ata* cluster in *Streptomyces*, the pIES plasmid was introduced into *S. lividans* 66–1326, *S. lividans* 66-TK21 and *S. coelicolor* (Table 1) using standard integrative methods. Thus, *S. lividans* 66–1326 (pIES) (named *S. lividans ata*), *S. lividans* 66-TK21 (pIES) and *S. coelicolor* M145 (pIES) were generated (Table 1).

To define the *ata* cluster, *orf1*, *ata 11* and *merR* in-frame deletions were generated in pIES via λ-Red-mediated recombination.[20] Disruption cassettes (1.4 kb) containing the *aac(3)IV* apramycin resistance gene were generated by PCR amplification using the vector pIJ773 and specific primers. These primers contained 39 nt regions that match the sequence adjacent to the gene to be inactivated and 19–20 nt regions matching the ends of the disruption cassette. The primers 5′-orf1/3′-orf1, 5′-ata11/3′-ata11 and 5′-merR/3′merR (Table 2) were then used to generate the *orf1* (^{1774}ATG/TGA^{890}), *ata11* (^{30325}GTG/TGA^{31035}) and *merR* (^{32153}GTG/TGA^{31773}) disruption cassettes, respectively. Using this procedure, the plasmids pIESΔ*orf1*, pIESΔ*att11* and pIESΔ*merR* were generated in *E. coli* BW25113, and subsequently introduced by conjugation into *S. lividans* 66–1326. Site-specific integration of apramycin-resistant clones into the genome was checked by PCR. For *ata11*-mutant complementation, *ata11* was amplified with the High fidelity expand PCR system (Roche, Germany) using the primers 5′-ata11ext and 3′-ata11ext (Table 2). To facilitate subcloning, the *Eco*RI and *Xba*I restriction sites were included in 5′-ata11ext and 3′-ata11ext, respectively. The PCR product was inserted into *Eco*RI/*Xba*I linearized pEM4neo, generating pATA11, which was introduced into the *S. lividans* 66–1326(pIESΔ*att11*) strain by protoplast transformation.

A201A determination

A201A or A201A-like material was obtained from culture filtrates (20 ml^{-1}), the pH adjusted to pH 8.5 and the material extracted with chloroform in a rotary shaker as described elsewhere.[21] Organic fractions were evaporated to dryness and dissolved in ethanol (50 μl; 70% (v/v)). A201A was identified and quantified by HPLC with a quaternary pump (Delta 600, Waters). Chromatography was performed through a reverse phase of a C18 Nova-Pack column (4 μm, 3.9×300 mm) with the solvents A (0.1 M ethylacetate, pH 6) and B (0.1% trifluoroacetic acid in acetonitrile). A and B, conditioned with helium, were used as the mobile phases at a flow of 0.5 ml min^{-1} for a total analysis time of 60 min. A linear gradient of 0–100% eluent B was applied and the column temperature was kept constant at 25 °C. A Waters 2487 dual absorbance detector was used at 254 and 280 nm and the data obtained were analyzed using the Waters Empower Software (Milford, MA, USA). The retention time for commercial A201A was used as a reference.

To purify A201A, the filtrate from a culture (200 ml) of *S. lividans* 66–1326 (pIES) was collected and extracted with chloroform as described above. The extract was evaporated to dryness in a vacuum, dissolved in 0.1 ml of EtOH 70% (v/v) and developed by preparative TLC on 2 mm Silica Gel60 F_{254} (Merck, Damstadt, Germany) using ethylacetate/ethanol (6:1, v/v) as the solvent. The plates were examined under UV light (254 nm) and the band migrating as A201A was removed and extracted twice with chloroform (1 ml). This TLC purification process was repeated twice and the final extract was dried under a vacuum and then dissolved in 0.05 ml EtOH 70% (v/v). Samples (25 μl) of this solution were subjected to matrix-assisted laser desorption ionization-time of flight on a Bruker ReflexIII spectrometer (SIDI, Universidad Autónoma Madrid, Spain) and a 2,5 dihydroxybenzoic acid matrix was used. NaI was added to the preparation to improve ionization. The mass spectrometry calculation for $C_{37}H_{50}N_6O_{14}$+Na was 825.5.

In addition, A201A was identified using the Ard2 phosphotransferase enzymatic assay described previously.[19]

Table 1 Strains and plasmids used in this study

Strain/plasmid	*Relevant properties*	*A201A production*	*Reference/source*
E.coli			
DH5α	*supE*44, Δ*lacU*169 (ϕ80*lacZ*ΔM15), *hsdR*17, *rec*A1, *endA*1, *gyrA*96, *thi*1, *relA*1	NA	13
DH5α (pIJ773)	*aac(3)IV, oriT*(RK2)	NA	14
DH5α (pPACS2)	*attp-int-tsr* cassette	NA	15
DH5α (pIES)	pPACS2 including *ata* cluster	NA	This work
DH5α (pEM4neo)	pEM4 Δ*bla* e including *neo*	NA	This work
DH5α (pATA11)	pEM4neo including *Ata11*	NA	This work
BW25113 (pIJ790)	λ-RED (*bet,exo,gam*), repA101^{ts}, *araC, cat*	NA	14
BW25113 (pIJ790)(pIES)	*ata* cluster	NA	This work
BW25113 (pIJ790)(pIESΔ*ata11*)	*ata* cluster, *ata11* deleted	NA	This work
BW25113 (pIJ790)(pIESΔ*orf1)*	*ata* cluster, *orf1* deleted	NA	This work
BW25113 (pIJ790)(pIESΔ*merR*)	*ata* cluster, *merR* deleted	NA	This work
ET12567 (pUZ8002)	*tra, neo*, RP4	NA	14
ET12567 (pUZ8002)(pIES)	Including *ata* cluster	NA	This work
ET12567 (pUZ8002)(pIESΔ*ata11*)	*ata* cluster, *ata11* deleted	NA	This work
ET12567 (pUZ8002)(pIESΔ*orf1*)	*ata* cluster, *orf1* deleted	NA	This work
ET12567 (pUZ8002)(pIESΔ*merR*)	*ata* cluster, *merR* deleted	NA	This work
Saccharothrix mutabilis subsp.*capreolus* NRRL3817	Wild type	+	5
Streptomyces			
lividans 66–1326	Wild type	–	16
lividans 66–1326 (pPACS2)	*attp-int-tsr* cassette	–	This work
lividans 66–1326 (pIES)	pPACS2 including *ata* cluster, clon 16c	+	This work
lividans 66–1326 (pIESΔ*ata11*)	*ata* cluster, *ata11* deleted-	–	This work
lividans 66–1326 (pIESΔ*ata11*/ pATA11)	*ata* cluster, Δ*ata11*, *Ata11* complemented	+	This work
lividans 66–1326 (pIESΔ*orf1*)	*ata* cluster, *orf1* deleted	+	This work
lividans 66–1326 (pIESΔ*merR*)	*ata* cluster, *merR* deleted	+	This work
lividans 66-TK21	SLP2$^-$, SLP3$^-$	–	16
lividans 66-TK21 (pPACS2)	*attp-int-tsr* cassette	–	This work
lividans 66-TK21 (pIES)	pPACS2 including *ata* cluster	+	This work
coelicolor M145	Wild type	–	16
coelicolor M145 (pPACS2)	*attp-int-tsr* cassette	–	This work
coelicolor M145 (pIES)	pPACS2 including *ata* cluster	+	This work

Abbreviations: NA, not applicable. *S. lividans* 66–1326 (pIES) was named *S. lividans ata.*

For A201A-resistance assays, suspensions of *Streptomyces* spores were streaked on agar plates containing different antibiotic concentrations (0–1000 μg ml^{-1}). Growth was scored after a 5-day incubation at 30 ºC.

Computer analysis

Currently available methods were employed to analyze the nucleotide and amino acid sequences.[22–24] The sequence analyzed here was submitted to the EMBL database (accession number X84374). The amino acid sequences were analyzed using BLAST against the Swiss-Prot/TrEMBL protein database (http://www.expasy.org/tools/blast/).

Table 2 Oligonucleotides used in this work

5′-orf1:	ctgaggactccgctcaccggcggggaaggagatccgatg**attccggggatccgtcgacc**
3′-orf1:	ccgggccgtcgcggtgggtgccggtcgggaaggccttca**tgtaggctggagctgcttc**
5′-ata11:	gcggtcgccgaggacgcactggacgatgtcctgtgacgg**attccggggatccgtcgacc**
3′-ata11:	actccatggtcgaacacgctcaccggcgtcggcgcgtca**tgtaggctggagctgcttc**
5′-merR:	ggtcaagtccggctgcatcggacccggacacgaagggtg**attccggggatccgtcgacc**
3′-merR:	gcatcggggccgctcaagacgtgctggagcgtcccatca**tgtaggctggagctgcttc**
5′-ata11ext:	gctctagatggccatggcggtcgccga
3′-ata11ext:	cggaattcatggtcgaacacgctcac

Sequences matching the ends of the disruption cassette are in bold. Underlined sequences indicate restriction sites.

RESULTS AND DISCUSSION

Identification and expression of the full *ata* cluster in heterologous hosts

In actinomycetes, antibiotic biosynthetic gene clusters are generally situated in a single stretch of DNA that includes the genes for self-resistance, export, enzymatic activities and usually, but not always, regulation.[25] A profound examination of antibiotic biosynthetic gene clusters requires the use of a wide variety of recombinant DNA techniques, including the genetic modification of relevant strains that produce them, a requisite that can often not be fulfilled. A reasonable approach to overcome this problem is to express the antibiotic gene clusters in a heterologous host amenable to gene manipulation, like *Streptomyces lividans* or *S. coelicolor*.[26–28] In the case of A201A, many years ago we successfully identified two resistance determinants, *ard1*[16] and *ard2*,[19] as well as seven ORFs involved in the biosynthesis of the antibiotic.[12] Moreover, we next isolated a plasmid named pIES from a *S. mutabilis* subsp. *capreolus* gene library, where we sequenced the full putative *ata* gene cluster responsible for A201A antibiotic biosynthesis (GeneBank acc. number X84374, unpublished data). After many unsuccessful attempts to introduce exogenous DNA into the producer

strain *S. mutabilis* subsp. *capreolus* by a variety of strategies, we introduced this putative *ata* gene cluster containing plasmid pIES into *S. lividans* 66–1326 to test whether its heterologous expression was possible. The *S. lividans ata* strain was resistant to 500 μg ml^{-1} of A201A on agar plates, and as indicated by a disc diffusion test on agar plates for biological activity, was in addition able to inhibit *Micrococcus luteus* growth, an A201A-sensitive organism (data not shown). These data indicated that A201A producing *S. lividans ata* strain could well be used as a bona fide tool to study the *S. mutabilis* subsp. *capreolus* biosynthetic A201A cluster.

To detect the heterologous production of A201A in *S. lividans ata* we first use an *in vitro* phosphorylation assay with Ard2, the ATP/GTP-binding/phosphorylating protein (an A201A-inactivating activity) encoded by the *ard2* gene from *ata*.[19] Accordingly, chloroform extracts from culture filtrates from *S. lividans ata* and those from *S. mutabilis* subsp. *capreolus*, *S. lividans* 66–1326 and *S. lividans* 66–1326 (pPACS2; Table 1), as well as standard A201A, were used as substrates for Ard2 assays in the presence of [γ-^{32}P]ATP. Cell-free extracts from *S. lividans ata* and *S. lividans* 66–1326 (pPACS2) were used as a source of the Ard2 enzyme and as a negative control, respectively. Reaction products were developed by TLC and radioactive spots were detected by exposure to X-ray films, where a radioactive spot corresponding to P-A201A served as positive control. Only the presumed A201A-containing samples (*S. mutabilis* subsp. *capreolus* and *S. lividans ata*) appeared to incorporate [γ-^{32}P]ATP into a P-A201A spot, suggesting that this antibiotic was expressed heterologously in *S. lividans ata* (data not shown). To confirm the production of A201A in the heterologous host, *S. lividans ata* was cultured and the chloroform extracts of culture filtrates were examined by HPLC (Figure 2a). Two compounds were identified in the HPLC profiles of the presumed A201A-containing samples. Their retention times corresponded to those of A201A and a putative A201A molecule lacking the disaccharide moiety (probably that considered as hydrolysis product 1 by Kirst *et al*.[5]). The identity of these substances was confirmed by ESI-MS (Figure 2b), yielding the characteristic molecular ion (*m/e* for $[M+H]^+$) of 803 for A201A and 455 for that lacking the disaccharide (Figure 2c) that correspond to their molecular formulae ($C_{37}H_{50}N_6O_{14}$ and $C_{22}H_{26}N_6O_5$, respectively).

Taken together these results indicate that the ~34 kb-long *S. mutabilis* subsp. *capreolus* DNA fragment in *S. lividans ata* strain is sufficient for the heterologous production of A201A as well as to confer resistance to it, strongly suggesting that it corresponds to the full *ata* biosynthetic cluster.

Time course of A201A production in *S. mutabilis subsp. capreolus* and heterologous hosts

Cultures of *S. mutabilis* subsp. *capreolus* and *S. lividans ata* were grown to compare the levels of antibiotic production over relevant time courses (Figure 3). In both cases, A201A production appeared to be strictly controlled, since it commenced during late-exponential growth to reach maximum values at the late stationary phase, close to 96 and 144 h, respectively. Interesting, A201A biosynthesis commences much earlier in *S. lividans* than in *S. mutabilis* subsp. *capreolus*. In addition, it appears that the production by the recombinant organism was approximately twofold lower than that in the wild-type strain. A similar result was found when the puromycin biosynthetic (*pur*) gene cluster was cloned.[7,29] The fact that antibiotic production starts in the late logarithmic phase indicates that expression of the *ata* cluster could be under the control of some regulatory mechanism. The *ata* cluster included in the pIES plasmid was also introduced in *S. lividans* 66-TK21 and *S. coelicolor* (Table 1). Similar profiles of A201A production were observed in these organisms derivatives containing the integrated *ata* sequence (data not shown).

Targeted gene disruption to identify the boundaries of the *ata* gene cluster

A201A production in both *S. lividans* and *S. colelicolor* indicates that the *S. mutabilis* subsp. *capreolus* DNA fragment inserted in their genomes contains the genes for the enzymes required and sufficient for the antibiotic synthesis. An examination of the deduced amino acid sequences from this 34 kb DNA stretch suggested that the ends of the putative *ata* cluster included *ard1* and *ata11* (Figure 4a and Table 3). These genes encode a putative ABC transporter that confers resistance to A201A (*ard1*)[16] and a putative O-methyltransferase (*ata11*), respectively. To determine whether this was the case, a variety of deletions of the 34 kb sequence from *S. lividans ata* were subjected to relevant phenotypic assays (see Experimental Section and Figure 4a). At the right-hand end, a *merR* deletion did not alter the production of A201A in the heterologous host, whereas destruction of *ata11* completely abolished it. Moreover, the *ata11* gene from pATA11 complemented a specific A201A non-producing Δ*ata11* mutant of *S. lividans ata* (Table 1). Despite the fact that in *S. lividans* the regulation of the *ata* cluster might differ from that in *S. mutabilis* subs. *capreolus*, or even that an homologous to *merR* in the host genome could functionally complement this transcriptional regulator absence, our data indicated that *ata11* will be the last gene on the right-hand end of the *ata* gene cluster. With regards the left-hand end, while a deletion of *orf1* did not alter A201A production, all attempts to delete *ard1* were by now unsuccessful. This could be explained by the determinant function of this A201A-resistance gene, since its disappearance would cause lethality. Because such genes are usually members of antibiotic biosynthetic clusters,[28] these findings indicate that *ard1* is the terminal gene at the left-hand end of *ata*. Finally, it is noteworthy that a region close to 750 nucleotides with no apparent coding capacity flanks the 3′ region of the coding sequence of these two terminal genes, which contrasts with the compact organization of the DNA coding regions of streptomycetes.

Bioinformatics analysis of *ata* biosynthetic cluster

The *ata* biosynthetic cluster bioinformatics analysis from *ard1* to *ata11* has revealed the presence of 28 putative ORFs (Figure 4a and Table 3), including *ard1* and *ata11*, most of which can be aligned with homologous sequences in GeneBank using BLAST P programs. The proposed functions of these *ata* gene products based on their homologies are listed in Table 3, together with the corresponding homologs from puromycin and hygromycin A antibiotics biosynthetic pathways from *S. hygroscopicus* and *S. alboniger*, respectively (see also Figure 4b). On the other hand, the organization of the A201A biosynthetic cluster is mostly conserved between *S. mutabilis* subs. *capreolus* and the other A201A producer *M. thermotolerans*, with the exception of the resistance gene *ard1*, which is placed between *ata14* and *ata15* in *M. thermotolerans* and the *ataP7-P10-P4-P5* ORFs, which are in the opposite orientation in *M. thermotolerans*.[3]

Genes related to the biosynthesis of the polyketide moiety

We suggested previously that the central A201A α-methyl-*p*-coumaric acid moiety was derived from the condensation of methyl-malonyl CoA with *p*-hydroxybenzoyl CoA, which must be mediated by a type II polyketide synthase (PKS).[30] The minimal PKS involves a ketoacyl synthase activity (KS), a chain length factor and an acyl carrier protein (ACP).[30] Analysis of the *ata* cluster sequence indicated the presence of nine putative genes that might be involved in the biosynthesis of the

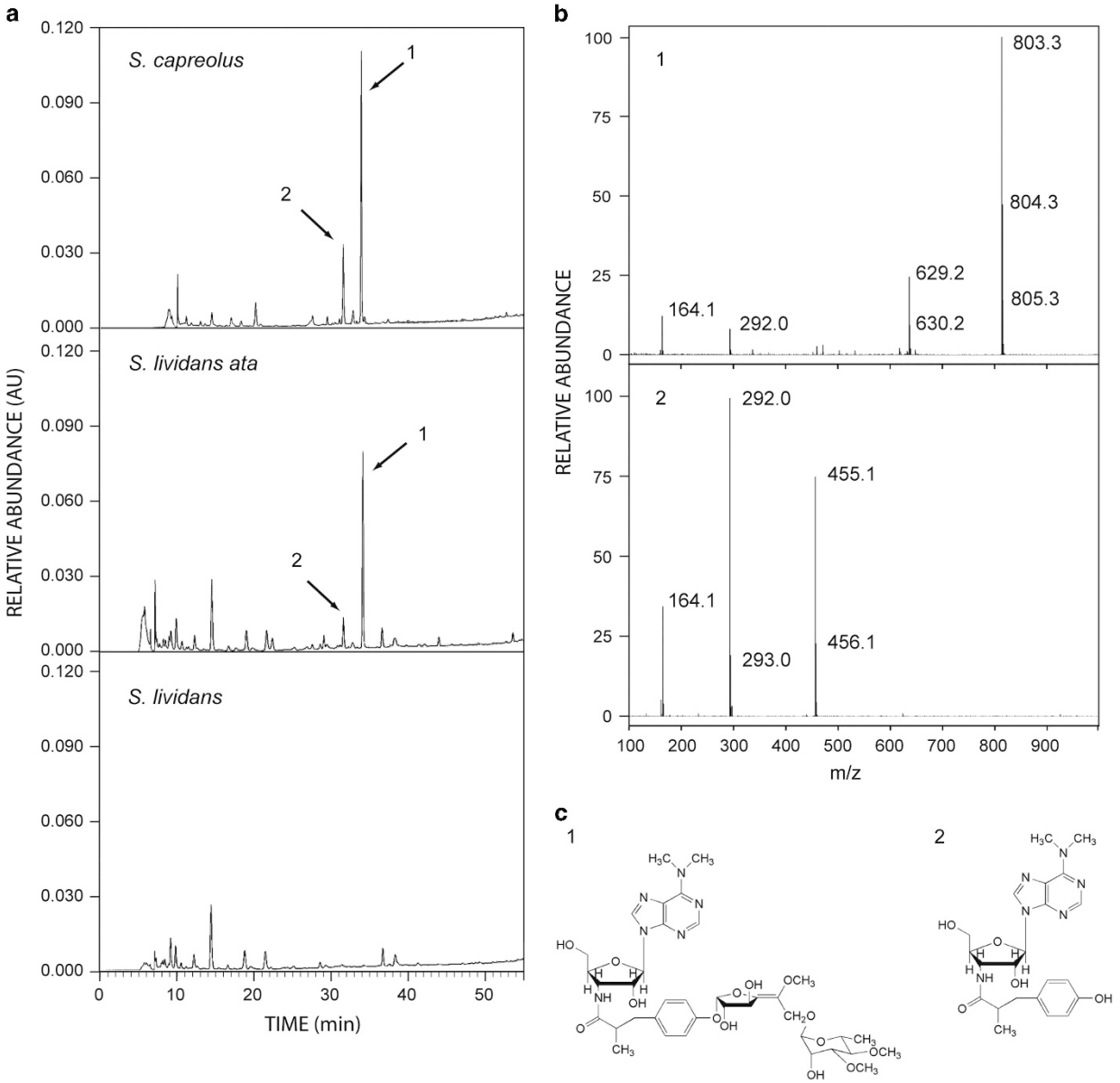

Figure 2 HPLC and MS analyses of the A201A produced in *S. lividans ata*. (**a**) HPLC chromatograms of chloroform extracts from culture filtrates of several *Streptomyces* strains. (**b**) Samples were developed by TLC, extracted with chloroform and subjected to matrix-assisted laser desorption ionization-time of flight. (**c**) Chemical structure of compound corresponding to 1: A201A and 2: the putative A201A lacking the disaccharide moiety.

polyketide moiety (Ata-PKS1-AtaPKS4, Ata18, Ata19, Ata2, Ata4 and Ata7). Ata–PKS1 (436 residues) aligned with PKS acyltransferases and includes the substrate-binding motif ^{196}GHSMG^{200}.[31] Likewise, Ata-PKS2 (108 residues) and Ata19 (95 residues) showed similarity to ACP and they contain the serine 4′-phosphopantetheine-binding sites ^{47}DX$_6$GXDSX$_2$LX$_7$E^{68} and ^{35}DX$_6$GXDSX$_2$LX$_7$E^{56}, respectively.[32] Ata-PKS3 (388 residues) and Ata-PKS4 (265 residues) showed an overall similarity with KS, and their N terminal domains contain most of the structures involved in dimer formation. In contrast, Ata-PKS4 lacks the highly conserved cysteine residue in the proposed active site ^{120}G-M^{135} required to link acyl groups.[33] However, the Ata-PKS3 ^{5}G-V^{23} sequence includes a cysteine, ^{16}C and could act as the protein active site. The presence or absence of key cysteine residue is a distinctive characteristic of chain length factor and KS activities.[34] Ata18 (486 residues) showed similarity to AMP-dependent synthetase and ligases, including the CoA ligases that are implicated in fatty acid and polyketide biosynthesis.[35] These proteins share an AMP-binding motif that includes a Ser/Gly/Thr-rich domain followed by a conserved Pro-Lys,[36] which is also presents in Ata18 (^{142}SGTQGRPK^{149}). Ata7 (219 residues) aligned with several bacterial chorismate pyruvate-lyases (*p*-hydroxybenzoic acid synthases), including that of *Mycobacterium tuberculosis* (42% similarity over a 160-amino acid overlap). The chorismate pathway is only present in bacteria, fungi and plants, and it provides a wealth of compounds with diverse biological functions, including aromatic amino acids, folate cofactors, pigments, mycobactins and glycosylated *p*-hydroxybenzoic acid methyl esters. The chorismate pyruvate-lyase from *M. tuberculosis* is responsible for the direct conversion of chorismate to *p*-hydroxybenzoate, and it has been identified as the sole enzymatic source of *p*-hydroxybenzoic acid in this organism.[37]

In this context, eight of the nine gene products putatively required for the biosynthesis of A201A α-methyl-*p*-coumaric acid moiety (disregarding Ata7) showed clear homology to the same number of genes identified in *S. hygroscopicus* NRRL2388 from the hygromycin A biosynthetic gene cluster (Hyg9-15 and Hyg22).[9] Hygromycin A and A201A share a similar polyketide structural moiety and therefore, it seems plausible that some analogous enzymes must be required for the biosynthesis of these two antibiotics. However, the organization of the homologous genes in both these clusters is not completely coincident (Figure 4b). This contrasts with the organization of the genes in the *ata* cluster that are involved in the biosynthesis of the aminonucleoside moiety when compared with those of the *pur* cluster from *S. alboniger*[29] (Figure 4b).

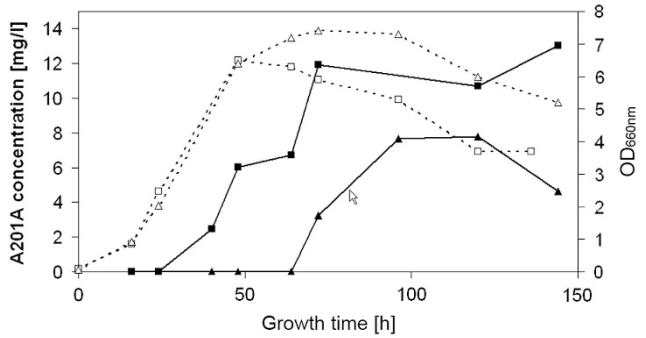

Figure 3 Time courses of A201A production by *S. mutabilis* subsp. *capreolus* and *S. lividans ata*. *S. mutabilis* subsp. *capreolus* (squares) and *S. lividans ata* (triangles) were grown at 30 °C. At the indicated times, samples were taken to determine the OD (broken lines) and then filtered. Chloroform extracts from the filtrates were used to quantify A201A by HPLC (solid lines).

It has been proposed that the biosynthetic pathway of the 3,4-dihydroxy α-methyl-*p*-coumaric acid moiety in hygromycin A commences with 4-hydroxybenzoic (or 3,4-dihydroxybenzoic) acid, which is derived from chorismate via the putative chorismate lyase Hyg4.[9] Although there was little similarity between Hyg4 and Ata7, it seems likely that both putative proteins could be responsible for the biosynthesis of 4-hydroxybenzoic acid. Presumably, this acid must be activated by conversion to a thioester, and Ata18 (homologous to Hyg12) could catalyze this reaction based on its homology to AMP-dependent synthetases and CoA ligases (Figure 5). Ata-PKS2 and Ata19 (which aligned with Hyg9 and Hyg13, respectively) are homologous to ACP and either of them could a priori be required to produce methyl-malonyl ACP for condensation with the activated hydroxybenzoyl thioester (Figure 5). The putative acyltransferase Ata-PKS1 (as proposed due to its similarity to with Hyg22) may affect the methyl-malonyl CoA-methyl-malonyl ACP interconversion and the proposed KS, Ata-PKS3 (homologous to Hyg10), might be required for the decarboxylative condensation with methyl-malonyl ACP. In addition, Ata-PKS4 (homologous to Hyg11) could be involved in KS dimer formation and/or, be related to chain length factor, thereby controlling the shutdown in chain elongation (Figure 5). Once the polyketide backbone is formed, different enzymes act to modify the final structure, including ketoreductases, oxygenases, dehydratases or aromatases. Accordingly, Ata4 (a putative 3-ketoacyl ACP-reductase homologous to Hyg15) and Ata2 (a 3-hydroxylacyl ACP dehydratase homologous to Hyg14) could be responsible for additional changes in the polyketide moiety (double bonding; Figure 5).

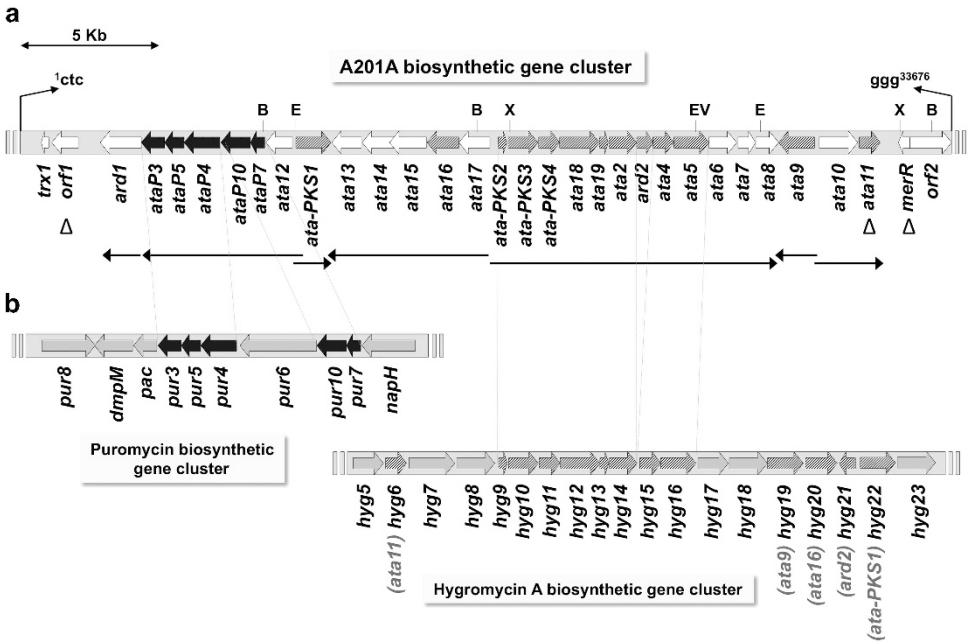

Figure 4 A201A biosynthetic gene cluster. (**a**) Scheme of the 32 open reading frame (ORFs) sequenced from *S. mutabilis* subsp. *capreolus* including the A201A (*ata*) gene cluster. Boxes indicates the direction of transcription of each predicted (ORFs). The horizontal arrows indicate the putative transcription units. Deleted genes are marked by open triangles. The ORFs with homology to putative genes in the hygromycin A biosynthetic gene cluster from *S. hygroscopicus* NRRL2388 or to puromycin biosynthetic genes from *S. alboniger* are hatched and in dark boxes, respectively. Note that *trx1* and *orf1* were previously named *orfB* and *orfA*, respectively.[12] ORFs from *trx1* to *ata-PKS1* and from *ata2* to *ata5* were sequenced previously. Ev, *Eco*RV; E, *Eco*RI; X, *Xho*I; B, *Bgl*II. (**b**) Puromycin biosynthetic cluster from *S. alboniger* (left) and the region in the hygromycin A biosynthetic gene cluster from *S. hygroscopicus* NRRL2388 showing homology to A201A genes (right) are shown for comparison. Dotted lines indicate those sets of 2 or more genes where orientation and organization is conserved between clusters.

Table 3 Open reading frames of the ata cluster and their proposed functions.

Protein	*Start/stop codon*	*Proposed function*	*Pur/HygA homolog*
Ard1	^{4170}ATG/TGA^{2494}	ABC transporter	–
AtaP3	^{5097}GTG/TGA^{4291}	Monophosphatase	Q53743 (Pur3)
AtaP5	^{5780}ATG/TGA^{5094}	*N*-methyl-transferase	Q53742 (Pur5)
AtaP4	^{7083}ATG/TGA^{5800}	Aminotransferase	Q53741 (Pur4)
AtaP10	^{8165}GTG/TGA^{7080}	Oxidoreductase	Q53739 (Pur10)
AtaP7	^{8680}ATG/TGA^{8162}	Pyrophosphohydrolase	Q53738 (Pur7)
Ata12	^{9704}CTG/TGA^{8682}	GDP-mannose 4,6-dehydratase	
Ata-PKS1	^{9790}ATG/TGA^{11098}	Acyltransferase	Q27YQ0 (Hyg22)
Ata13	^{12207}GTG/TGA^{11083}	Mannosyltransferase	–
Ata14	^{13208}GTG/TGA^{12204}	Unkown	–
Ata15	^{14575}ATG/TGA^{13205}	Unkown	–
Ata16	^{15750}GTG/TGA^{14572}	Transglycosylase	Q27YQ2 (Hyg20)
Ata17	^{16897}GTG/TGA^{15743}	NAD-dependent epimerase/dehydratase	–
Ata-PKS2	^{17174}GTG/TGA^{17500}	Acyl carrier protein	Q27YR3 (Hyg9)
Ata-PKS3	^{17497}TTG/TGA^{18663}	Ketoacyl synthase	Q27YR2 (Hyg10)
Ata-PKS4	^{18656}ATG/TGA^{19453}	Ketoacyl synthase	Q27YR1 (Hyg11)
Ata18	^{19440}ATG/TGA^{20900}	CoA-ligase	Q27YR0 (Hyg12)
Ata19	^{20897}ATG/TGA^{21184}	Acyl carrier protein	Q27YQ9 (Hyg13)
Ata2	^{21181}GTG/TGA^{22212}	3-hydroxyacyl ACP dehydratase	Q27YQ8 (Hyg14)
Ard2	^{22226}ATG/TGA^{22816}	Phosphotransferase	Q27YQ1 (Hyg21)
Ata4	^{22813}GTG/TGA^{23583}	3-oxoacyl-ACP-reductase	Q27YQ7 (Hyg15)
Ata5	^{23597}ATG/TGA^{24874}	Glycosyltransferase	Q27YQ6 (Hyg16)
Ata6	^{24875}ATG/TGA^{25927}	Methyltransferase	–
Ata7	^{25920}ATG/TGA^{26579}	Chorismate-pyruvate lyase	–
Ata8	^{26580}GTG/TGA^{27410}	Methyltransferase	–
Ata9	^{28733}ATG/TAG^{27387}	MFS transporter	Q27YQ3 (Hyg19)
Ata10	^{28840}ATG/TGA^{30328}	Oxidoreductase	–
Ata11	^{30325}GTG/TGA^{31035}	Methyltransferase	Q2KC54 (Hyg6)
MerR	^{32153}GTG/TGA^{31773}	MerR-family transcriptional regulator	–

Abbreviations: ABC, ATP-binding cassete; ACP, acyl carrier protein; MFS, major facitilator superfamily; ORF, open reading frame; PFS, polyketide synthase.

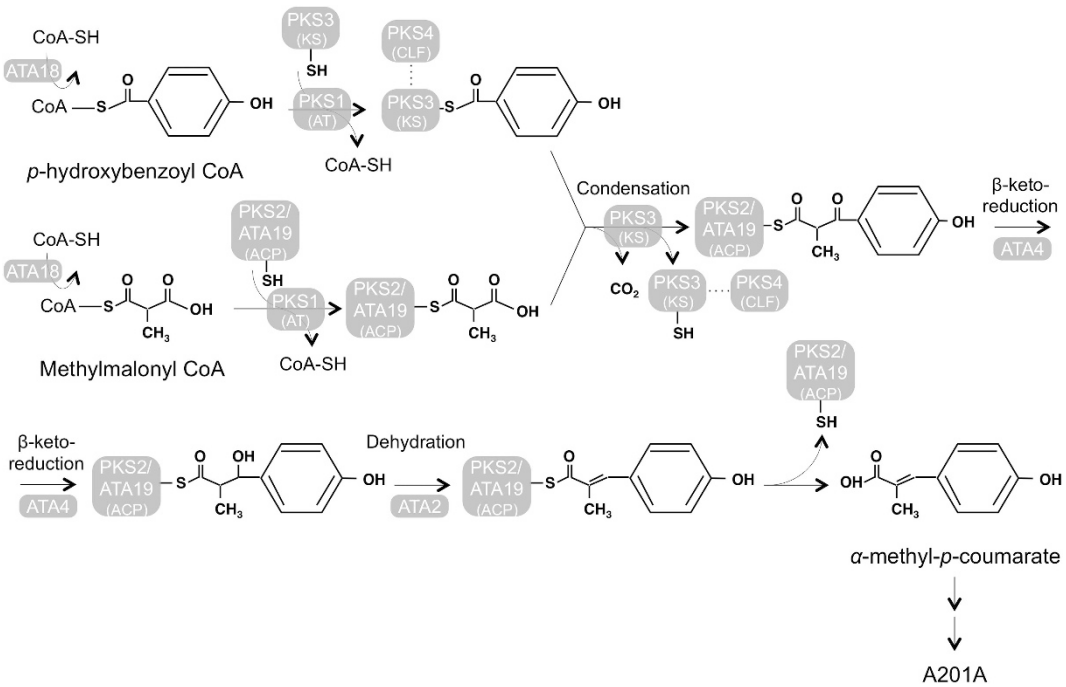

Figure 5 Proposed working hypothesis of A201A polyketide moiety biosynthesis. Putative *S. mutabilis* subsp. *capreolus* enzymes involved in the biosynthesis of the A201A antibiotic polyketide moiety are shown as gray boxes. See text for details.

Genes related with the biosynthesis of the disaccharide moiety

The chemical structure of A201A includes two sugar moieties, an unsaturated hexofuranose associated with the α-methyl-*p*-coumaric acid moiety and a 3,4-dimethyl-D-rhamnose. Analysis of the *ata* cluster seemed to indicate that at least six gene products might be involved in disaccharide biosynthesis (Ata12, Ata13, Ata16, Ata17, Ata5 and Ata10). Ata13 (374 residues) aligned with bacterial glycosyltransferases and mannosyltransferases that catalyze the transfer of sugar moieties from activated donor molecules to specific acceptors, forming glycosidic bonds. Prokaryotic mannosyltransferases generally use GDP-mannose as the sugar donor in lipopolysaccharide and polysaccharide biosynthesis. These proteins are larger than 350 amino acids and while they do not display an overall similarity, they do share the $EXFGX_4E$ motif in the C-terminal half of the protein that is involved in the catalytic activity of the enzyme.[38,39] This motif is also found in sucrose phosphate synthase, which also shows some homology to Ata13 (Q0I5Z9; 30% identity; 42% similarity). However, although Ata13 contains the ^{264}EAFGLSILE^{272} sequence, it also contains the ^{78}P, ^{84}H, ^{127}S and ^{186}K residues that are highly conserved in mannosyltransferases but that are not present in sucrose phosphate synthases.[38] Ata16 (392 residues) and Ata5 (425 residues) were homologous to some bacterial and plant transglucosylases and glycosyltransferases, respectively. Interestingly, these two putative proteins aligned with Hyg20 and Hyg16 from the hygromycin A biosynthetic gene cluster,[9] and this antibiotic also contains a pyranose–furanose moiety closely related to those found in A201A. Ata17 (384 residues) displayed similarity to NAD-dependent epimerase/dehydratases that contain a N-terminal NAD(P) binding motif[40] (^{56}GAGGFIG^{62} in Ata17), including epimerase-ketoreductase, epimerase-dehydratase and dehydrogenase, which utilize nucleotide-sugar substrates in a variety of chemical reactions. Finally, Ata10 (496 residues) showed overall homology to glucose-methanol-choline oxidoreductases that act on CH–OH groups of different donors (glucose, methanol and choline),[41] and which share a FAD–ADP-binding domain in the N-terminal portion of the protein (^{11}GAGSAG^{16} in Ata10).

The formation of activated-sugar precursors by nucleotidyltransferases must be a crucial branch point in secondary carbohydrate metabolic pathways from primary metabolism. However, there did not appear to be any genes involved in the formation of activated sugars within the *ata* cluster (Table 3). Similar results were recently found in the moenomycin A and the butirosin gene clusters from *Streptomyces ghanaensis*[42] and *Bacillus circulans*,[43] respectively. Biosynthesis of these carbohydrate-containing antibiotics requires a sugar-nucleotide building block directly derived from primary metabolism. In this context, either GDP-D-rhamnose or GDP-L-fucose is the 6-deoxyhexose produced from GDP-D-mannose. Biosynthesis of GDP-L-fucose is mediated by mannose 4,6-dehydratase, 3,5 epimerase and 4-ketoreductase activity in both eukaryotic and prokaryotic organism.[44,45] Accordingly, Ata12 (mannose 4,6-dehydratase) and Ata17 (epimerase/dehydratase) may be involved in the synthesis of GDP-D-mannose. Furthermore, the unsaturated hexofuranose moiety could be synthesized from GDP-D-rhamnose through Ata10 (oxidoreductase), and Ata13 (mannosyltransferase) could link it to that D-rhamnose. If so, Ata16 (transglucosylase) might well be involved in the glycosyl ring biosynthetic pathway and Ata5 (glycosyltransferase) could be responsible for the formation of the glycosidic link between the hexofuranose and the polyketide moieties.

Genes encoding methyltransferases

Besides AtaP5, a putative SAM-dependent methyltransferase that could dimethylate N^6 of the A201A aminonucleoside moiety, three other putative methyltransferases are also found in the *ata* cluster: Ata6 (350 residues), Ata8 (276 residues) and Ata11 (236 residues). As Ata11 exhibited some sequence similarity to Hyg6 of the hygromycin A gene cluster (38% identity and 50% similarity over a 140-amino acid overlap), it would presumably be related with *p*-coumaric acid methylation which occurs in both these antibiotic structures. In addition, and given that the two glycoside residues in A201A are *O*-methylated, Ata 6 and Ata8 may be the enzymes responsible for these methylations. Nevertheless, despite these sequence homologies, further molecular and biochemical analysis is necessary to determine the role of these putative proteins in the A201A biosynthetic pathway and the specific order of events that take place.

Genes encoding resistance, regulatory or proteins of unknown function

Antibiotic-producing actinomycetes, usually posses more than one mechanism that confer resistance to the relevant antibiotic. Besides *ard1* and *ard2*, the fragment sequenced from *S. mutabilis* subsp. *capreolus* NRRL3817 contained a gene whose product could be a priori involved in providing resistance to A201A, Ata9 (488 residues). This is a putative transmembrane proteins, which aligns with permeases of the major facilitator superfamily (MSF transporters) and thus, they may be involved in the efflux of A201A.

The regulation of *ata* cluster expression remains unclear since the cluster appears to lack dedicated regulatory genes. A putative regulatory protein MerR (125 residues) with similarity to DNA binding proteins of the MerR family of transcriptional regulators (Table 3) was present in the sequence analyzed, but removal of this putative protein did not affect either negatively or positively the A201A production (see further and Table 1). Nevertheless a misfunction of the regulation mechanism in the heterologous organism cannot be fully excluded. In addition, a higher level of regulation by the regulatory *bldA* gene is unlikely given that none of the genes analyzed contain the rare TTA leucine codons.[46] This is not surprising considering that $<0.2\%$ of genes in the *S. coelicolor* genome contain TTA codons.[47] A number of pleiotropic regulatory genes have been isolated from *S. coelicolor* that affect multiple antibiotic pathways but that are not pathway-associated.[48] The isolation and characterization of these genes may reveal novel regulatory mechanisms in the biosynthetic pathways of secondary metabolites within this important group of bacteria.

Finally, the role of two genes of the cluster in A201A biosynthesis could not be defined, since the putative proteins they encode did not clearly align with any protein sequence in databases. Still, some similarities were encountered. Ata14 (334 residues) includes the nucleotide (FAD, NAD(P)) binding site ^{236}GPGFDG^{24},[49] and not surprisingly, its corresponding homolog in *M. thermotolerans* A201A biosynthetic cluster (MtdJ) has been to function as a oxidoreductase.[3] For Ata15 (456 residues), the sequence from ^{137}V to R^{250} aligns with bacterial glucokinases (Q2RRF9; 30% identity and 45% similarity). This 113 amino acid stretch includes the highly conserved ^{206}GTGL^{209} and residues ^{165}N and ^{166}D that might be associated with glucose binding (all of which can also be found in its corresponding homolog in *M. thermotolerans* cluster MtdK[3]), as well as the aspartic acid functioning as the general base in the enzyme catalytic mechanism.[50] However, Ata15 did not display overall similarity to glucokinases and neither were other highly conserved residues in these enzymes present, such as those associated with ATP-binding.

In conclusion, here we have studied the gene cluster involved in the biosynthesis and resistance of A201A from *S. capreolus*, expressed it in various heterologous systems, and the predictions drawn from the bioinformatics analysis were largely consistent with the expected enzymology. A201A is an interesting compound, which could be viewed as a natural 'hybrid molecule', with moieties similar to those found in other antibiotics such as puromycin and hygromycin A. Besides a moiety derived from D-rhamnose, its chemical structure includes the aminonucleoside moiety of puromycin, as well as a polyketide and an unsaturated furanose moiety closely related to similar structures found in hygromycin A. Both our group and the Reynolds group have done a remarkable work in understanding the puromycin and hygromycin A biosynthetic pathway steps and some of these are likely to occur identically for A201A biosynthesis in *S. mutabilis*. In this regard, the aminonucleoside moiety of A201A synthesis is likely to be initiated from ATP by the ATAP10 oxidoreductase and followed sequentially by five enzymatic steps catalyzed by ATAP4, ATAP7 and ATAP5/ATAP3, as described for puromycin.[12] Likewise, we have placed at least eight enzymatic activities homologous to hygromycin A ones in those steps leading to the A201A polyketide moiety biosynthesis, all of which needs to be confirmed with further studies. The remaining enzymes sharing homology with hygromycin A ones are likely to be involved in the synthesis and incorporation of the disaccharide moieties, methylation of the antibiotic and resistance mechanisms and will require to be studied further too. Thus, the biosynthetic pathway of A201A consists in a combination of an interesting variety of known (from puromycin and hygromycin A biosynthesis) and novel enzymatic steps for the synthesis and modification of each moiety together with those that correctly link them together. It will be interesting to see how our knowledge of these complex antibiotic biosynthetic pathways is expanded with new studies and hopefully completed in the next years.

Characterization of all the genes involved in A201A biosynthesis will be useful to detect homologous genes from other actinomycetes, to identify biosynthetic clusters for other antibiotics that include nucleoside, polyketide and/or sugar moieties, and also to produce novel hybrid antibiotics. Future combinatorial biosynthetic strategies, and cross-complementation between selected segments from *ata*, *pur* and *hyg* clusters, could potentially provide a tool for the generation of new compounds with different biotechnological applications.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank A. Ceballos for preliminary experiments with *Streptomyces* growth and mutant strain preparation, and A. Martín for technical assistance. This work was supported by grants BIO2003-03314 and BIO2013-48779-C4-4 from the Spanish CICYT. We also thank the Fundación Ramón Areces for an institutional grant to the CBMSO (CSIC-UAM).

1 Nie, S., Li, W. & Yu, B. Total synthesis of nucleotide antibiotic A201A. *J. Am. Chem. Soc.* **136**, 4157–4160 (2014).
2 Hamill, R. L. & Hoehn, M. M. Antibiotics A201A and A201B and process for the production thereof. US Patent, Eli Lilly & Co. 3,843,784. Oct. 23.
3 Zhu, Q. *et al.* Discovery and engineered overproduction of antimicrobial nucleoside antibiotic A201A from the deep-sea marine actinomycete *Marinactinospora thermotolerans* SCSIO 00652. *Antimicrob. Agents Chemother.* **56**, 110–114 (2012).
4 Ensminger, P. V. & Wright, W. E. A201A, a new antibiotic produced by *Streptomyces capreolus*. II. Biological studies. *Abst.16th Intersci. Conference Antiomicrob. Agents Chemother.* **63** (1976).
5 Kirst, H. A. *et al.* The structure of A201A, a novel nucleoside antibiotic. *J. Antibiot. (Tokyo)* **38**, 575–586 (1985).
6 Polikanov, Y. S. *et al.* Distinct tRNA accommodation intermediates observed on the ribosome with the antibiotics hygromycin A and A201A. *Mol. Cell* **58**, 832–844 (2015).
7 Lacalle, R. A., Tercero, J. A. & Jiménez, A. Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *EMBO J.* **11**, 785–792 (1992)
8 Habib, S. E., Scarsdale, J. N. & Reynolds, K. A. Biosynthetic origin of hygromycin A. *Antimicrob. Agents Chemother.* **47**, 2065–2071 (2003).
9 Palaniappan, N., Ayers, S., Gupta, S., Habib, S. E. & Reynolds, K. A. Production of hygromycin A analogs in *Streptomyces hygroscopicus* NRRL 2388 through identification and manipulation of the biosynthetic gene cluster. *Chem. Biol.* **13**, 753–764 (2006).
10 Palaniappan, N. *et al.* Biosynthesis of the aminocyclitol subunit of the hygromycin A in *Streptomyces hygroscopicus* NRRL 2388. *Chem. Biol.* **16**, 1180–1189 (2009).
11 Dhote, V., Starosta, A. L., Wilson, D. N. & Reynolds, K. A. The final step of hygromycin A biosynthesis, oxidation of C-5′-dihydrohygromycin A, is linked to a putative proton gradient-dependent efflux. *Antimicrob. Agents Chemother.* **52**, 3580–3588 (2009).
12 Saugar, I., Sánz, E., Rubio, M. A., Espinosa, J. C. & Jiménez, A. Identification of a set of genes involved in the biosynthesis of the aminonucleoside moiety of antibiotic A201A from *Streptomyces capreolus*. *Eur. J. Biochem.* **269**, 5527–5535 (2002).
13 Gust, B., Challis, G. L., Fowler, K., Kieser, T. & Chater, K. F. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl Acad. Sci. USA* **100**, 1541–1546 (2003).
14 Quirós, L. M., Aguirrezabalaga, I., Olano, C., Méndez, C. & Salas, J. A. Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by Streptomyces antibioticus. *Mol. Microbiol.* **28**, 1177–1185 (1998).
15 Hopwood, D. A. *et al. Genetic manipulation of Streptomyces. A laboratory manual* (The John Innes Foundation, Norwich, England, 1985).
16 Barrasa, M. I., Tercero, J. A., Lacalle, R. A. & Jiménez, A. The *ard1* gene from *Streptomyces capreolus* encodes a polypeptide of the ABC-transporters superfamily, which confers resistance to the aminonucleoside antibiotic A201A. *Eur. J. Biochem.* **228**, 562–569 (1995).
17 Sambrook, K. J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning. Molecular Cloning. A laboratory Manual* (Cold Spring Harbor Laboratory Press, NY, USA, 1989).
18 Sosio, M. *et al.* Artificial chromosome for antibiotic producing actinomycetes. *Nat. Biotechnol.* **18**, 343–345 (2000).
19 Barrasa, M. I., Tercero, J. A. & Jiménez, A. The aminonucleoside antibiotic A201A is inactivated by a phosphotransferase activity from *Streptomyces capreolus* NRRL 3817, the producing organism. Isolation and molecular characterization of the relevant encoding gene and its DNA flanking regions. *Eur. J. Biochem.* **245**, 54–63 (1997).
20 Datsenko, K. A. & Warner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
21 Vara, J., Marlpartida, F., Hopwood, D. A. & Jiménez, A. Cloning and expression of a puromycin N-acetyl transferase gene from *Streptomyces alboniger* in *Streptomyces lividans* and *Escherichia coli*. *Gene* **33**, 197–206 (1985).
22 Wright, F. & Bibb, M. J. Codon usage in the G+C-rich *Streptomyces* genome. *Gene* **113**, 55–65 (1992).
23 Devereux, J., Haeberli, P. & Smithies, O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387–395 (1984).
24 Marck, C. *Strider* (Centre d'Estudes de Saclay, Gif-Sur-Yvette, Cedex, France, 1991).
25 Cundliffe, E. Antibiotic production by actinomycetes: the Janus faces of regulation. *J. Ind. Microbiol. Biotechnol.* **33**, 500–506 (2006).
26 Wendt-Pienkowski, E. *et al.* Cloning, sequencing, analysis, and heterologous expression of the fredericamycin biosynthetic gene cluster from *Streptomyces griseus*. *J. Am. Chem. Soc.* **127**, 16442–16452 (2005).
27 Hu, Z., Reid, R. & Gramajo, H. The leptomycin gene cluster and its heterologous expression in *Streptomyces lividans*. *J. Antibiot. (Tokyo)* **58**, 625–633 (2005).
28 Eustaquio, A. S. *et al.* Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl. Environ. Microbiol.* **71**, 2452–2459 (2005).
29 Tercero, J. A., Espinosa, J. C., Lacalle, R. A. & Jiménez, A. The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the pur cluster of *Streptomyces alboniger*. *J. Biol. Chem.* **271**, 1579–1590 (1996).
30 Hopwood, D. A. Genetic contributions to understanding polyketide synthases. *Chem. Rev.* **97**, 2465–2498 (1997).
31 Grimm, A., Madduri, K., Ali, A. & Hutchinson, C. R. Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase. *Gene* **151**, 1–10 (1994).
32 Hopwood, D. A. & Sherman, D. H. Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet* **24**, 37–66 Review (1990).
33 Kim, E. S., Bibb, M. J., Butler, M. J., Hopwood, D. A. & Sherman, D. H. Sequences of the oxytetracycline polyketide synthase-encoding otc genes from *Streptomyces rimosus*. *Gene* **141**, 141–142 (1994).
34 McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. Engineered biosynthesis of novel polyketides. *Science* **262**, 1546–1550 (1993).
35 Dieckmann, R., Lee, Y. O., van Liempt, H., von Dohren, H. & Kleinkauf, H. Expression of an active adenylate-forming domain of peptide synthetases corresponding to acyl-CoA-synthetases. *FEBS Lett.* **357**, 212–216 (1995).

36 Kleinkauf, H. & Von Dohren, H. A nonribosomal system of peptide biosynthesis. *Eur. J. Biochem.* **236**, 335–351 (1996).
37 Stadthagen, G. *et al.* *p*-Hydroxybenzoic acid synthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* **280**, 40699–40706 (2005).
38 Geremia, R. A., Petroni, E. A., Ielpi, L. & Henrissat, B. Towards a classification of glycosyltransferases based on amino acid sequence similarities: prokaryotic alpha-mannosyltransferases. *Biochem. J.* **318**, 133–138 (1996).
39 Abdian, P. L., Lellouch, A. C., Gautier, C., Ielpi, L. & Geremia, R. A. Identification of essential amino acids in the bacterial alpha –mannosyltransferase AceA. *J. Biol. Chem.* **275**, 40568–40575 (2000).
40 Wierenga, R. K., Terpstra, P. & Hol, W. G. Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint. *J. Mol. Biol.* **187**, 101–107 (1986).
41 Cavener, D. R. GMC oxidoreductases. A newly defined family of homologous proteins with diverse catalytic activities. *J. Mol. Biol.* **223**, 811–814 (1992).
42 Ostash, B., Saghatelian, A. & Walker, S. A streamlined metabolic pathway for the biosynthesis of moenomycin A. *Chem. Biol.* **14**, 257–267 (2007).
43 Truman, A. W., Huang, F., Llewellyn, N. M. & Spencer, J. B. Characterization of the enzyme BtrD from *Bacillus circulans* and revision of its functional assignment in the biosynthesis of butirosin. *Angew. Chem. Int. Ed.* **46**, 1462–1464 (2007).
44 Andrianopoulos, K., Wang, L. & Reeves, P. R. Identification of the fucose synthetase gene in the colanic acid gene cluster of *Escherichia coli* K-12. *J. Bacteriol.* **180**, 998–1001 (1998).
45 Sullivan, F. X. *et al.* Molecular cloning of human GDP-mannose 4,6-dehydratase and reconstitution of GDP-fucose biosynthesis *in vitro*. *J. Biol. Chem.* **273**, 8193–8202 (1998).
46 Leskiw, B. K., Lawlor, E. J., Fernández-Abalos, J. M. & Chater, K. F. TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. *Proc. Natl Acad. Sci. USA* **88**, 2461–2465 (1991)
47 Chater, K. F. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* **361**, 761–768 (2006).
48 Fernández-Moreno, M. A. *et al. abaA*, a new pleiotropic regulatory locus for antibiotic production in *Streptomyces coelicolor*. *J. Bacteriol.* **174**, 2958–2967 (1992).
49 Wierenga, R. K., Drenth, J. & Schulz, G. E. Comparison of the three-dimensional protein and nucleotide structure of the FAD-binding domain of p-hydroxybenzoate hydroxylase with the FAD- as well as NADPH-binding domains of glutathione reductase. *J. Mol. Biol.* **167**, 725–739 (1983).
50 Lunin, V. V. *et al.* Crystal structures of *Escherichia coli* ATP-dependent glucokinase and its complex with glucose. *J. Bacteriol.* **186**, 6915–6927 (2004).