The Journal of Antibiotics (2010) 63, 135–138
© 2010 Japan Antibiotics Research Association All rights reserved 0021-8820/10 $32.00
www.nature.com/ja

NOTE

Transcriptional organization of ThnI-regulated thienamycin biosynthetic genes in *Streptomyces cattleya*

Miriam Rodríguez, Carmen Méndez, José A Salas and Gloria Blanco

The Journal of Antibiotics (2010) **63,** 135–138; doi:10.1038/ja.2009.133; published online 22 January 2010

Keywords: antibiotic; β-lactam; carbapenem; LysR-type transcriptional activator; regulator

Streptomyces are soil mycelial bacteria with a complex life cycle, which produce an extraordinary diversity of antibiotics and other biologically active secondary metabolites of medical and industrial significance. Antibiotic production in these species is coordinated with morphological differentiation and regulated in a process operating at different levels.[1,2] Expression of antibiotic gene clusters is controlled by different families of regulatory proteins. Complex regulatory cascades link environmental and physiological signals with pleiotropic and pathway-specific regulatory proteins that control the expression of individual antibiotic gene clusters. Carbapenems constitute a potent class of non-conventional β-lactam antibiotics with important clinical application as antimicrobial agents in infections mediated by multidrug-resistant bacteria. Carbapenem biosynthesis and its regulation in bacteria have been reviewed.[3]

Thienamycin, the first carbapenem described,[4] is considered one of the most potent, most broad spectrum of all natural antibiotics known so far and has an important clinical role in the treatment of severe nosocomial infectious diseases.[5] The gene cluster for thienamycin biosynthesis (*thn*) has been cloned from the producer *S. cattleya* NRRL 8057,[6] opening up the possibility to study its regulation. We recently showed that ThnI, an LysR-type pathway-specific transcriptional activator, constitutes a key factor in controlling thienamycin biosynthesis. ThnI is required for activation of the expression of a subset of nine genes in the *thn* cluster involved in thienamycin bosynthesis and export: *thnH*, *thnJ*, *thnK*, *thnL*, *thnM*, *thnN*, *thnO*, *thnP* and *thnQ*,[7] whereas the regulation of the rest of the *thn* genes, including *thnI*, remained unknown.

The deduced products of the ThnI-dependent genes are predicted to be involved in the assembly of thienamycin molecule and its export. From left to right in Figure 1a, *thnH* encodes a hydrolase of the haloacid dehalogenase superfamily proposed to hydrolyze 4-phosphopantetheine to render pantetheine in the generation of cysteaminyl side chain in thienamycin biosynthesis.[8] The predicted product of *thnJ* belongs to a group of efflux systems of the major facilitator superfamily exporters[9] and might have a role in thienamycin secretion through the cell membrane. Three genes, *thnK*, *thnL* and *thnP*, encode methyltransferases proposed to be responsible of the methylation steps that give rise to the hydroxyethyl side chain of thienamycin. Insertional inactivation of *thnL*[6] and lately *thnP* (Nuñez LE, unpublished results) has shown their involvement in thienamycin biosynthesis. The *thnM* gene encodes a β-lactam synthetase proposed to have a role in carbapenem ring formation in thienamycin biosynthesis through an alternative mechanism to that of classic β-lactam antibiotics.[6] Both *thnN* and *thnO* are required for thienamycin biosynthesis as has been shown in independent insertional inactivation experiments.[6] The *thnN-thnO* genes are homologues to *griC-griD,* which encode a carboxylic acid reductase complex essential for grixazone biosynthesis in *S. griseus*.[10] These authors proposed that ThnO-ThnN may be responsible for the reduction of a carboxylic acid to the corresponding aldehyde in thienamycin biosynthesis. Finally, *thnQ* encodes one of the two potential 2-oxoglutarate-dependent dioxygenases that might catalyze the desaturation reaction between C2 and C3 in thienamycin biosynthesis, in a similar way that CarC does in carbapenem 2-em-3-carboxylic acid biosynthesis.[11] However, other *thn* genes involved in thienamycin biosynthesis have been found to be differentially regulated. Particularly striking is the case of *thnE*, which encodes a carboxymethylproline synthase recently demonstrated to be responsible of the pyrrolidine ring formation in the first step of thienamycin biosynthesis;[12] that is not regulated by ThnI.[7]

The aim of this work is to analyze the transcriptional organization of the ThnI-dependent genes by reverse transcription polymerase chain reaction (RT–PCR) and promoter probe assays, and get some insights into the expression of *thnI* itself to explain its observed differential regulation with respect to the ThnI-regulated genes. From the nucleotide sequence observation of the *thn* gene cluster (EMBL Accession number AJ421798) it is deduced that, among the

Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias, Universidad de Oviedo, Oviedo, Spain
Correspondence: Dr G Blanco, Departamento de Biologia Funcional, Universidad de Oviedo, Julian Claveria, Oviedo, Asturias 33006, Spain.
E-mail: gbb@uniovi.es
Received 30 September 2009; revised 3 December 2009; accepted 22 December 2009; published online 22 January 2010

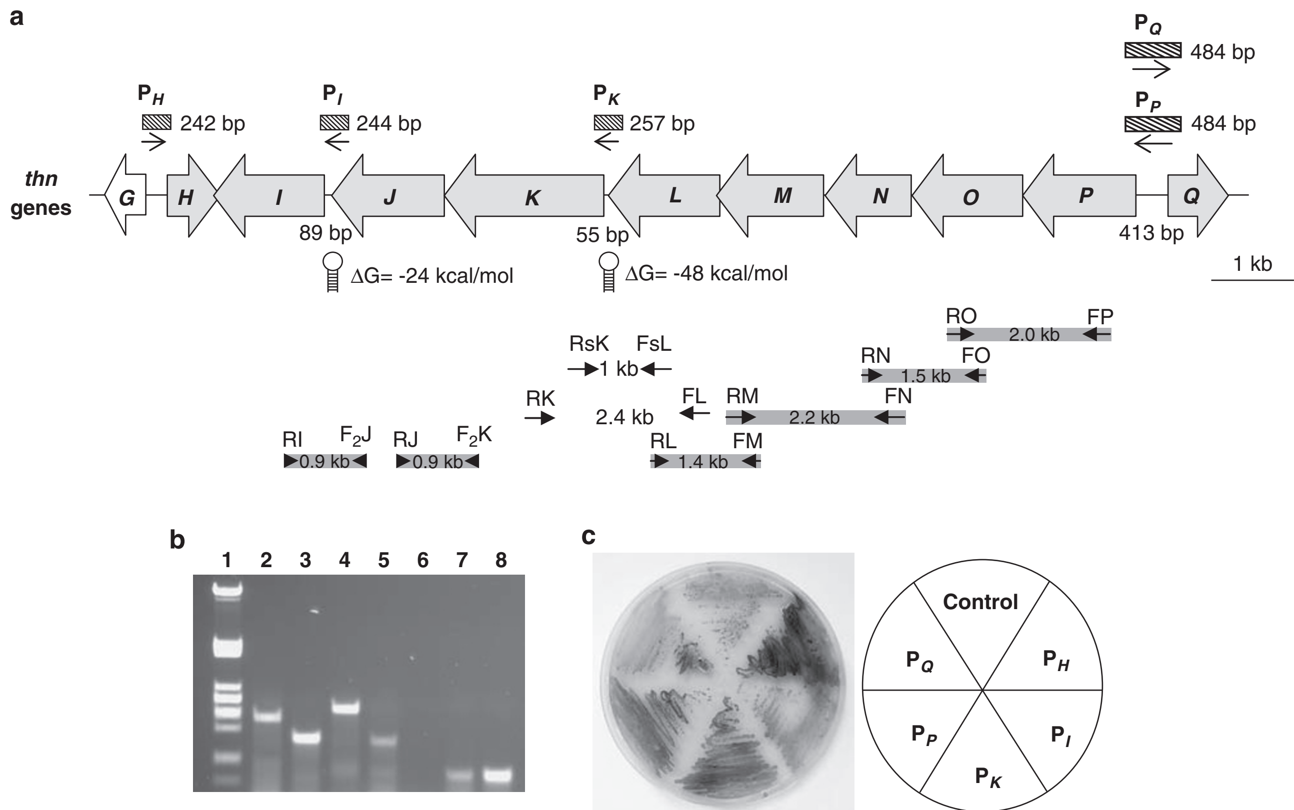

Figure 1 Transcriptional organization of ThnI-activated thienamycin biosynthesis genes. (**a**) Physical map of the studied *thn* genes, RT–PCR and promoter probe analysis from *S. cattleya* wild type. The primers used for RT–PCR and transcriptional fusions are indicated in Supplementary Table S1. Big arrows show the order and direction of transcription of the genes. Gray arrows indicate the ThnI-regulated genes as shown in this paper. The gray bars at the bottom represent the expected fragments amplified by RT–PCR with each pair of primers. Each fragment, whose size is also indicated, partially spans two adjacent genes. Putative stem-loop secondary structures are represented by a hairpin. The hatched boxes on the upper part of the map indicate the amplified fragments used in promoter probe experiments, with the small arrow below pointing to the direction of transcription and indicating the orientation of fragments in the transcriptional fusions. The size and designation of amplified regions with promoter activity are also indicated. Promoter probe assays with the different transcriptional *xylE* fusions are shown in Figure 2. (**b**) The gel shows the expected amplification cDNA bands obtained after RT–PCR analysis for the various primer pairs separated in agarose gel and visualized by ethidium bromide staining. Lane 1, λ DNA digested with *Pst*I as molecular marker; lanes 2–8, RT–PCR products. The following primers were used: FP and RO for *thnP-thnO* intergenic region (lane 2); FO and RN for *thnO-thnN* intergenic region (lane 3); FN and RM for *thnN-thnM* intergenic region (lane 4); FM and RL for *thnM-thnL* intergenic region (lane 5); FL and RK for *thnL-thnK* intergenic region (lane 6) (although not presented in the gel, R_SK and F_SL were used as well with the same result); F_2K and RJ for *thnK-thnJ* intergenic region (lane 7); and F_2J and RI for *thnJ-thnI* intergenic region (lane 8). The amount of RNA used was optimized for each pair of primers and 40 ng were used for most cases, with the exception of lanes 7 (100 ng) and 8 (60 ng). The analysis was performed at least three times for each pair of primers. Negative controls to determine whether RNA was free of contaminant DNA are not represented in the figure. (**c**) Promoter probe assays. *S. lividans* TK21 strains carrying the different transcriptional *xylE* fusions (in pIJ4083) were tested for catechol 2, 3-dioxygenase activity using *S. lividans* carrying pIJ4083 as a control strain. Transformants containing the different constructs turned yellow ($xylE^+$ phenotype) when sprayed with catechol, whereas the one containing pIJ4083 did not. The assays were carried out at least three times. The picture was taken after 5 min exposure to catechol. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

previously described ThnI-dependent genes, *thnH* and *thnQ* are organized into two monocistronic transcriptional units (Figure 1a). The rest of genes (from right to left), *thnP*, *thnO*, *thnN*, *thnM*, *thnL*, *thnK*, *thnJ* and *thnI* (previously described as ThnI independent), are adjacent and transcribed in the same direction, and it has been suggested that most of them might be co-transcribed.[7] DNA sequence analysis, carried out with the FRAME program[13] suggests a form of translational coupling among some of them, with the stop codon of one gene overlapping the start codon of the following one. Such overlapping occurs between the genes *thnP-thnO*, *thnO-thnN* and *thnK-thnJ*, whereas a 29 bp overlap was found between *thnM-thnL*. No such overlap was found, however, between *thnN-thnM* (1 bp intergenic distance), *thnL* and *thnK* (55 bp intergenic region) and *thnJ* and *thnI* (89 bp intergenic region). These last two intergenic regions could form stem-loop secondary structures with a free energy of −48.1 and $-24\,\mathrm{kcal\,mol^{-1}}$, respectively, as determined with the MFOLD program[14] that might act as transcriptional terminators (Figure 1a).

To investigate whether these adjacent genes are co-transcribed, we conducted transcriptional analysis of ThnI-dependent thienamycin genes by RT–PCR. Total mycelial RNA was prepared from *S. cattleya* wild-type strain after growth in R5A liquid medium (minus sucrose) for 95 h, when thienamycin is being actively produced, and RT–PCR was performed as described earlier.[7] Primers for the co-amplification of adjacent genes in this region were designed from the pairs *thnP-thnO*, *thnO-thnN*, *thnN-thnM*, *thnM-thnL*, *thnL-thnK*, *thnK-thnJ* and *thnJ-thnI* (Supplementary Table S1). Amplification products corresponding to overlapping transcripts were obtained with all the

pairs assayed (Figure 1b, lanes 2–5, 7 and 8) with the exception of *thnL-thnK*, where no amplification band was obtained even using two different pairs of primers (Figure 1b, lane 6). These pair of primers gave the expected amplified fragments when used in PCR experiments on DNA, thus proving that the predicted secondary structure did not prevent amplification of this region. The absence of contaminating DNA in the RNA samples was verified in PCR-negative controls, containing DNA polymerase but lacking reverse transcriptase, in which amplified products were not detected with any of the primers. These results showed that in *S. cattleya*, *thnP*, *thnO*, *thnN*, *thnM*, *thnL* seem to be expressed as a pentacistronic operon, named *thnPONML*, whereas *thnK*, *thnJ* and *thnI* might be transcribed independently as a tricistronic operon, termed *thnKJI*. This is in agreement with the presence of the stem-loop secondary structure between *thnL* and *thnK* acting as a transcriptional terminator. In contrast, and despite the presence of another putative hairpin loop structure, transcription spanning the *thnJ-thnI* intergenic region was detected.

To identify promoter regions, transcriptional fusions were made between selected intergenic DNA sequences of the *S. cattleya thn* cluster and the promoter-less *xylE* reporter gene, contained in the *Streptomyces* vector pIJ4083.[15] Primers were designed (Supplementary Table S1) to amplify the regions upstream of *thnQ*, *thnP*, *thnK*, *thnI* and *thnH* (Figure 1a). Amplification products for the corresponding fragments, obtained with *Pfx* polymerase (Invitrogen S.A. El Prat De Llobregat, Barcelona, Spain), were cloned into pIJ4083. To analyze the *in vivo* functionality of these putative promoter regions, DNA constructs were obtained after transformation of *S. lividans* TK21 protoplasts. Cloning of a promoter upstream of the *xylE* structural gene allows the expression of this gene that encodes a catechol 2,3-dioxygenase activity, which converts the colorless cathechol into a yellow derivative. Qualitative assays were performed by spraying with 0.5 M cathechol colonies grown on agar plates as described earlier,[16] using *S. lividans* carrying pIJ4083 as a control. As shown in Figure 1c, significant promoter activity (yellow phenotype) was detected in all five regions assayed, termed as P_Q, P_P, P_K, P_I and P_H. Bidirectional gene expression in the *thnQ-thnP* intergenic region seems to be carried out by the divergently oriented promoters P_Q and P_P, which drive divergent transcription.

The data presented here, together with previous RT–PCR reports on the *thnI* deletion mutant,[7] are consistent with a putative model for the regulation of ThnI-dependent thienamycin genes, which provides some clues to how *thnI* expression itself occurs (Figure 2). The model proposes that ThnI activates transcription of 10 *thn* genes, the ThnI regulon, that might be organized in four transcriptional units associated to regions with promoter activity. P_Q and P_H drive expression of the downstream genes as two monocistronic operons, termed *thnQ* and *thnH*, respectively. P_P drives *thnP*, *thnO*, *thnN*, *thnM* and *thnL* transcription as the pentacistronic operon *thnPONML*, whereas P_K drives transcription of *thnK*, *thnJ* and *thnI* as the tricistronic operon *thnKJI*. A caveat must be mentioned, however, as only intergenic regions were analyzed, we cannot rule out the presence of additional promoters lying within protein coding sequences.

Concerning *thnI* itself, according to the RT–PCR results reported here, it seems to be co-transcribed along with the ThnI-dependent genes *thnK* and *thnJ*, in a ThnI-dependent way. It was surprising to find transcription spanning the *thnJ-thnI* intergenic region, as it was previously shown that the main *thnI* expression is ThnI independent. No significant differences were detected in the transcript levels between a *thnI* deletion mutant and the wild-type strain, thus suggesting that *thnI* is transcribed from a ThnI-independent promoter putatively located upstream in the *thnI* coding region.[7] In this study, promoter probe analysis has shown the presence of a promoter activity upstream of the *thnI* coding region, termed P_I, which could be the promoter driving ThnI-independent *thnI* transcription. The results provided here support the idea that *thnI* expression might be driven by two promoters, P_I and P_K. Most *thnI* transcription might be carried out from the ThnI-independent P_I promoter as a monocistronic transcript, thus explaining the observed differential expression of *thnI* in comparison with the ThnI-regulated genes.[7] In addition, *thnI* transcription might be driven from the P_K promoter as the ThnI-dependent tricistronic transcript *thnKJI*. This finding strongly suggests that ThnI is also involved in activating its transcription, which means that its expression is in part autoregulated. By activating transcription from the P_K promoter, ThnI seems to establish a positive autoregulatory circuit. The presence of a putative stem-loop secondary structure between *thnJ* and *thnI*, rather than a transcriptional terminator, might represent a constant level transcriptional attenuator, compensating for the additional transcription of *thnI* from its own promoter. The situation is reminiscent of the one found in the closest ThnI homologue, ClaR, which activates transcription of late clavulanic acid biosynthetic genes in *S. clavuligerus*.[17,18] ClaR is expressed as a monocistronic transcript from its own promoter, as well as co-transcribed with a biosynthetic-regulated gene as a bicistronic transcript, where a similar read through of transcription spanning a stem-loop structure was observed.[17]

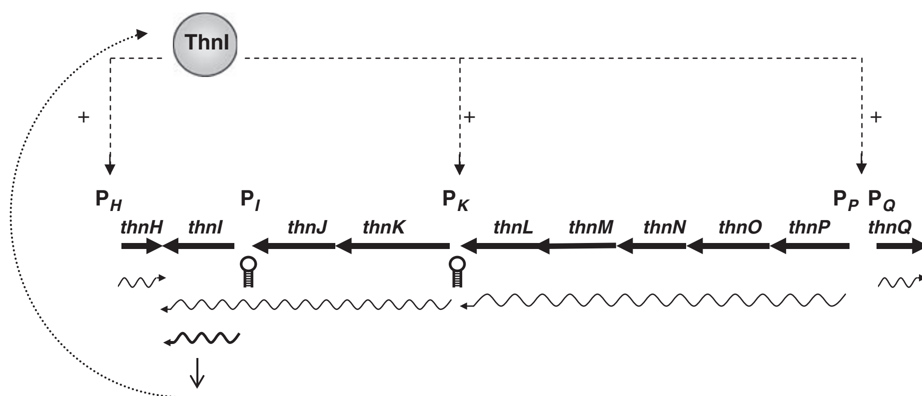

Figure 2 Proposed model for the ThnI regulon in *S. cattleya*. Big arrows indicate *thn* genes. The thin wavy arrows at the bottom of the figure represent the putative mRNAs transcribed after ThnI activation and the highlighted one the ThnI-indepentent *thnI* transcription. Activation of transcription is indicated as +. Putative stem-loop secondary structures are represented by a hairpin.

The LysR family is the most abundant type of transcriptional regulators among prokaryotes. Most studied LysR-type transcriptional activators interact with DNA boxes that contain the T-N_{11}-A conserved motif sequence within an inverted repeat, usually in regions of overlapping divergent promoters that allow simultaneous bidirectional control of transcription.[19,20] Although several putative LysR T-N_{11}-A boxes were found in the nucleotide sequence upstream of *thnQ*, *thnH*, *thnP* and *thnK* coding regions, additional studies are required to identify the binding sites for ThnI to exert its regulatory role. Regarding the autoregulatory nature of most characterized LysR-type transcriptional activators it should be mentioned that, although the 'classical' type of regulation is by transcriptional activation and negative autoregulation,[20] there is an increasing body of knowledge indicating that an additional class of LysR-type regulators, acting as transcriptional activators or repressors, is subjected to positive autoregulation.[19] The data presented here suggest that ThnI belongs to this class of positively autoregulatory activators of the LysR family.

In conclusion, the results presented here showed that 10 genes constitute the ThnI regulon, including *thnI*, the regulator itself. The ThnI-regulated genes seem to be organized in four operons driven by four ThnI-dependent promoters: a pentacistronic (*thnPONML*), a tricistronic (*thnKJI*) and two monocistronic units (*thnQ* and *thnH*). We propose a putative model for the mechanism of ThnI regulation in which, interestingly, ThnI seems to positively autoregulate in part its own expression.

ACKNOWLEDGEMENTS

We are grateful to Alfredo F Braña for his continuous support and critical reading of the manuscript and to José A Guijarro for fruitful discussions. We also thank Javier Méndez for providing us with Imipenem and José L Caso for valuable help.

1 Bibb, M. J. Regulation of secondary metabolism in streptomycetes. *Curr. Opin. Microbiol.* **8,** 208–215 (2005).
2 Chater, K. F. & Bibb, M. J. Regulation of bacterial antibiotic production. in *Products of Secondary Metabolism,* Vol.7 (eds Kleinkauf, H. & Von Dohren, H.) 57–105 (VCH, Weinheim, 1997).
3 Coulthurst, S. J., Barnard, A. M. L. & Salmond, G. P. C. Regulation and biosynthesis of carbapenem antibiotics in bacteria. *Nat. Rev. Microbiol.* **3,** 295–306 (2005).
4 Kahan, J. S. *et al.* Thienamycin, a new β-lactam antibiotic.I. Discovery, taxonomy, isolation and physical properties. *J. Antibiot.* **23,** 1255–1265 (1979).
5 Rodloff, A. C., Goldstein, E. J. C. & Torres, A. Two decades of imipenem therapy. *J. Antimicrob. Chemother.* **58,** 916–929 (2006).
6 Núñez, L. E., Méndez, C., Braña, A. F., Blanco, G. & Salas, J. A. The biosynthetic gene cluster for the β-lactam carbapenem thienamycin in *Streptomyces cattleya*. *Chem. Biol.* **10,** 1–20 (2003).
7 Rodríguez, M. *et al.* Identification of transcriptional activators for thienamycin and cephamycin C biosynthetic genes within the thienamycin gene cluster from *Streptomyces cattleya*. *Mol. Microbiol.* **69,** 633–645 (2008).
8 Freeman, M. F., Moshos, K. A., Bodner, M. J., Li, R. & Townsed, C. A. Four enzymes define the incorporation of coenzyme A in thienamycin biosynthesis. *Proc. Natl Acad. Sci. USA* **105,** 11128–11133 (2008).
9 Martin, J. F., Casqueiro, J. & Liras, P. Secretion systems for secondary metabolites: how producer cells send out messages of intercellular communication. *Curr. Opin. Microbiol.* **8,** 282–293 (2005).
10 Suzuki, H., Ohnishi, Y. & Horinouchi, S. GriC and GriD constitute a carboxylic acid reductase involved in grixazone biosynthesis in *Streptomyces griseus*. *J. Antibiot.* **60,** 380–387 (2007).
11 Sleeman, M. C. & Schofield, C. J. Carboxymethylproline synthase (CarB), an unusual carbon-carbon bond-forming enzyme of the crotonase superfamily involved in carbapenem biosynthesis. *J. Biol. Chem.* **279,** 6730–6736 (2004).
12 Hamed, R. B., Batchelar, E. T., Mecinovic, J., Claridge, T. D. & Shofield, C. J. Evidence that thienamycin biosynthesis proceeds via C-5 epimerization: ThnE catalyzes the formation of (2S, 5S)-trans-carboxymethylproline. *Chem. Bio. Chem.* **10,** 246–250 (2009).
13 Bibb, J. J., Findlay, P. R. & Johnson, M. W. The relationship between base composition and codon usage in bacterial genes and its use in the simple and reliable identification of protein-coding sequences. *Gene* **30,** 157–166 (1984).
14 Zucker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31,** 3406–3415 (2003).
15 Clayton, T. M. & Bibb, M. J. *Streptomyces* promoter-probe plasmids that utilise the xylE gene of *Pseudomonas putida*. *Nucleic Acids Res.* **18,** 1077 (1989).
16 Ingram, C., Brawner, M., Youngman, P. & Westpheling, J. *xylE* functions as an efficient reporter gene in *Streptomyces* spp.: use for the study of galP1, a catabolite-controlled promoter. *J. Bacteriol.* **171,** 6617–6624 (1989).
17 Paradkar, A. S., Aidoo, K. A. & Jensen, S. E. A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Mol. Microbiol.* **27,** 831–843 (1998).
18 Pérez-Redondo, R., Rodríguez-García, A., Martín, J. F. & Liras, P. The claR gene of *Streptomyces clavuligerus*, encoding a lysR-type regulatory protein controlling clavulanic acid biosíntesis, is linked to the clavulanate-9-aldehyde reductase (car) gene. *Gene* **211,** 311–321 (1998).
19 Maddocks, S. E. & Oyston, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154,** 3609–3623 (2008).
20 Schell, M. A. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47,** 597–626 (1993).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (http://www.nature.com/ja)