Transcriptional and translational analysis of the *ccaR* gene from *Streptomyces clavuligerus*

Liru Wang,[†] Kapil Tahlan, Tracy L. Kaziuk, Dylan C. Alexander[‡] and Susan E. Jensen

Department of Biological Sciences, University of Alberta, Edmonton, Canada T6G 2E9

Correspondence

Susan E. Jensen
susan.jensen@ualberta.ca

Received 16 April 2004
Revised 28 July 2004
Accepted 10 September 2004

CcaR is a positive-acting transcriptional regulator involved in cephamycin C and clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Previous sequence analyses of the *ccaR* gene revealed two possible start codons, an ATG, and a GTG located in-frame 18 bp downstream of the ATG. To determine the true start codon, *ccaR* was expressed, either from the GTG or ATG codon, in *Escherichia coli*. A protein product was only obtained from the ATG construct. Similarly, *ccaR* constructs originating from ATG or GTG and designed for expression from a glycerol-regulated promoter in *Streptomyces* species were prepared and used to complement a *S. clavuligerus ccaR* mutant. Bioassays showed that only the ATG construct could complement the *ccaR* mutant to restore cephamycin C production, and Western analysis confirmed the presence of CcaR in the mutant complemented with the ATG construct only. To ensure that expression of *ccaR* from its native promoter also initiated at the ATG rather than GTG, a conservative point mutation was introduced into *ccaR*, converting the GTG to GTC. The GTC construct still fully complemented a *ccaR* mutant, confirming that ATG is the true start codon. Inspection of the region upstream of *ccaR* by S1 nuclease protection and primer extension analyses indicated the presence of two transcript start points that mapped to residues located 74 and 173 bp upstream of the ATG codon.

INTRODUCTION

The production of cephamycin C, clavulanic acid and four other clavam metabolites in *Streptomyces clavuligerus* is accomplished by biosynthetic enzymes encoded by three separate clusters of genes (Tahlan *et al.*, 2004a, b). The *ccaR* gene, located within the one of these clusters (Alexander & Jensen, 1998; Perez-Llarena *et al.*, 1997), encodes a positive-acting transcriptional regulator that controls the production of both cephamycin C and clavulanic acid. CcaR is a member of the growing class of regulatory proteins referred to as *Streptomyces* Antibiotic Regulatory Proteins or SARPs (Wietzorrek & Bibb, 1997). The importance of CcaR in the production of cephamycin C and clavulanic acid has been established by showing loss of production in *ccaR* mutant strains (Alexander & Jensen, 1998; Perez-Llarena *et al.*, 1997), and increased production when *ccaR* is overexpressed in wild-type *S. clavuligerus* (Perez-Llarena *et al.*, 1997).

However, the detailed mechanism by which *ccaR* regulates clavulanic acid and cephamycin C production has been more difficult to determine at the molecular level, and studies to date have generated conflicting results (Kyung *et al.*, 2001; Santamarta *et al.*, 2002).

While the *ccaR* gene has been located and sequenced, uncertainty remains regarding the translational and transcriptional regulatory signals that control its expression. Initial studies identified an ATG codon as the start codon for *ccaR* (Walters *et al.*, 1994). However, a GTG codon present in-frame with the ATG but 18 bp downstream has also been proposed (Perez-Llarena *et al.*, 1997). Because of the central role that CcaR plays in regulating cephamycin C and clavulanic acid production, there is interest in understanding how CcaR regulates the production of these valuable metabolites. CcaR is present at low levels in wild-type *S. clavuligerus*, and so engineered production of CcaR in *Escherichia coli* or *Streptomyces* species offers the best means to generate large amounts of protein for studies of its regulatory properties. However, this requires knowledge of the transcriptional and translational regulatory signals that control the expression of the *ccaR* gene. In this study we identify the start codon for *ccaR* and use S1 nuclease protection and primer extension studies to determine the location of the transcription start point(s) for the gene.

[†]Present address: CanBioCin, Inc., Suite 1015, 8308-114 St., Edmonton, Canada AB T6G 2E1.

[‡]Present address: Cubist Pharmaceuticals, 65 Hayden Avenue, Lexington, MA 02421, USA.

Abbreviations: SARP, *Streptomyces* antibiotic regulatory protein; TSP, transcription start point.

METHODS

Bacterial strains, plasmids and culture conditions. The bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* strains were grown and maintained on LB medium (Sambrook *et al.*, 1989) at 37 °C. *S. clavuligerus* strains were maintained on ISP medium 4 (Difco) at 28 °C. AS-1 medium supplemented with 10 mM MgCl₂ (Bierman *et al.*, 1992) was used for the isolation of *S. clavuligerus* exconjugants following intergeneric conjugation. Minimal liquid (NMMP) (Kieser *et al.*, 2000) and complex [trypticase soy broth, TSB (Difco)] media, with and without added glycerol, were used to grow *S. clavuligerus* $\Delta ccaR::tsr$ mutants carrying pDA1100, pDA1102 and pDA1103 constructs to assess complementation of the *ccaR* mutation. Spores stored at -70 °C in 20% glycerol (w/v) were inoculated into TSB+1% soluble starch and incubated for 40 h at 28 °C at 250 r.p.m. on a rotary shaker before they were subcultured at 2% (v/v) into TSB or NMMP media for production of cephamycin C. *S. clavuligerus* $\Delta ccaR::tsr$ mutants complemented with pDA1006 and pDA1006*, and *S. clavuligerus* wild-type strains for production of RNA used in SI nuclease

protection and primer extension studies were grown in a complex soy flour-based medium (Mosher *et al.*, 1999) using incubation conditions described above.

For *E. coli*, media were supplemented with ampicillin (100 µg ml⁻¹) or apramycin (Apralan, Provel, 50 µg ml⁻¹), as appropriate. For *S. clavuligerus*, media were supplemented with apramycin (25 µg ml⁻¹) or thiostrepton (5 µg ml⁻¹), as appropriate.

Recombinant DNA techniques. Plasmid isolation, transformation, ligation, end-labelling of DNA using [γ -³²P]dATP and other molecular biology procedures commonly used with *E. coli* were as described by Sambrook *et al.* (1989). Manual sequencing of DNA was performed using the Thermo Sequenase Radiolabelled Terminator Cycle Sequencing Kit (USB) and the reaction mixtures were fractionated by electrophoresis on 6% denaturing polyacrylamide gels. Commonly used molecular biology procedures for *Streptomyces* species were as described by Kieser *et al.* (2000). Plasmids were introduced into *S. clavuligerus* either by protoplast transformation, as described previously (Paradkar & Jensen, 1995),

Table 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant features	Source or reference
Bacterial strains		
<i>S. clavuligerus</i>		
NRRL 3585	Wild-type	NRRL
$\Delta ccaR::tsr$	<i>ccaR</i> mutant	Alexander & Jensen (1998)
<i>E. coli</i>		
DH5 α	<i>SupE44</i> $\Delta lacU169$ ($\phi 80 lacZ\Delta M15$) <i>hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Bethesda Research Laboratories
BL21 (DE3)	F ⁻ <i>ompT hsd S_B (r_B⁻, m_B⁻) gal dcm</i>	Stratagene
ESS	Indicator organism for cephamycin C bioassay	Jensen <i>et al.</i> (1982)
Plasmids		
pBluescript II	pUC18 phagmid with SK+ and KS+, f1 <i>ori</i> , T3 and T7 primer binding sites	Stratagene
pMT3226	pSET152-based plasmid. Apr ^R , carries <i>xylE</i>	Hindle & Smith (1984), Trepanier <i>et al.</i> (2002)
pSL1180	pUC18 phagmid with 'superlinker' MCS	Pharmacia
pSET152	<i>E. coli</i> / <i>Streptomyces</i> shuttle plasmid. Apr ^R , integrates at $\phi C31$ attachment site in <i>Streptomyces</i>	Bierman <i>et al.</i> (1992)
pT7-7	Amp ^R <i>E. coli</i> expression vector	S. Tabor, Harvard Univ.
pDA150	pSL1180 carrying 1.4 kb <i>Bam</i> HI- <i>Bcl</i> II fragment of <i>S. clavuligerus</i> DNA (contains <i>ccaR</i> and part of <i>orf11</i>)	This study
pDA1000	pSET152 carrying the <i>tsr</i> gene	Alexander & Jensen (1998)
pDA1006	pSET152 carrying 1.4 kb <i>Eco</i> RI- <i>Nru</i> I fragment of <i>S. clavuligerus</i> DNA from pDA150	Alexander & Jensen (1998)
pDA1006*	pDA1006 equivalent construct with GTG codon near N-terminus changed to GTC	This study
pDA1100	pMT3226 with <i>Nde</i> I site in MCS and with <i>xylE</i> gene removed	This study
pDA1102	pDA1100 carrying mutant <i>ccaR</i> with <i>Nde</i> I site at putative ATG start codon under the control of <i>gylP1/P2</i>	This study
pDA1103	pDA1100 carrying mutant <i>ccaR</i> with <i>Nde</i> I site at putative GTG start codon under the control of <i>gylP1/P2</i>	This study
pTK-1	pT7-7 carrying mutant <i>ccaR</i> with <i>Nde</i> I site at putative ATG start codon	This study
pTK-2	pT7-7 carrying <i>ccaR</i> with <i>Nde</i> I site at putative GTG start codon	This study
pLW3	pSET152 carrying <i>ccaR</i> with putative GTG start codon mutated to GTC	This study

Table 2. Oligodeoxyribonucleotides used in this study

NdeI sites engineered into primers at the two potential start codons are underlined. Mutations introduced to change the potential GTG start codon to GTC are shown in bold type.

Oligodeoxyribonucleotide	Sequence
BKL96	5'-TCACATCATTCAGGTGTTTCAT-3'
ccaR-UP-Forward	5'-GGATAATGTGCCGGCCAACTGATTGCAGATTCC-3'
ccaR-UP-Reverse	5'-CGGGACCCTCAAGGCATGGACG-3'
DYL30	5'-GCCGCATATGAACACCTGGAATGATGTGACG-3'
DYL34	5'-GCCGGATCCCATATGACGATCCGGCTCCTGGGGC-3'
LWA1	5'-CCC GCGGATGGGTATCGGTACGGAACCTTTCACGAGTGTCA- CCGGCCCCAGGAGCCGGATCGTGACATC-3'
LWA2	5'-ATCGGATCCGCCCAGGTCCG-3'
LWA3	5'-GATGTCACGATCCGGCTCCTGGGGCCG-3'
LWA4	5'-GCCAAGCTTGGGGCCGGTGGGTCTGACGTCTTC-3'
LWA5	5'-GCGCAAGCTTCCCCGGGTTCCTCCGGGTTC-3'
SEJ25	5'-GCCGAATTCGCGGTTTCAGGCCGGGTACCGAC-3'

or by conjugal plasmid transfer. Conjugation was carried out as described by Kieser *et al.* (2000), except that exconjugants were selected on AS-1 agar medium supplemented with 10 mM MgCl₂ (Bierman *et al.*, 1992) and antibiotics as appropriate.

Generation of *ccaR* expression vectors. The oligodeoxyribonucleotide primers used in this study are listed in Table 2. Primers DYL30 (forward) and SEJ25 (reverse) were used in a PCR-mediated process to amplify a mutant form of *ccaR* with an *NdeI* site at the potential ATG start codon at the 5' end of the gene, and an *EcoRI* site just beyond the 3' end of the gene. PCR product was obtained using Vent DNA polymerase (New England Biolabs; 92 °C, 45 s; 57 °C, 60 s; 72 °C, 90 s for 30 cycles) in a reaction mixture containing 2 % DMSO, with pDA150 as template. The resulting PCR product was cloned as a *NdeI*–*BamHI* fragment into pSL1180. To reduce the possibility of unintended PCR-introduced mutations, only the 5' end of the PCR-amplified *ccaR* gene was used for subsequent manipulations. Digestion at a vector-derived *SpeI* site and an *Eco47III* site internal to *ccaR* (position +114 bp relative to the putative ATG start codon) excised the 5' end of the newly amplified gene. It was used to replace the corresponding fragment from a wild-type copy of *ccaR* that had been cloned previously into pBluescript II pSK(+) as a *BamHI*–*NruI* fragment. The reassembled *ccaR* gene was recloned into pBluescript II pSK(+) as a *HindIII*–*XhoI* fragment to pick up appropriate restriction sites, and finally cloned into pT7-7 as a *NdeI*–*BamHI* fragment to yield pTK-1.

pTK-2 was constructed exactly as pTK-1, except that DYL34 (forward) and SEJ25 (reverse) were used to amplify a mutant form of *ccaR* with a *NdeI* site at the potential GTG (now changed to ATG) start codon near the 5' end of the gene.

pDA1100 is an integrative vector based on pMT3226 that was prepared by digesting with *BamHI*/*XbaI* to remove the *xylE* reporter gene, and replacing it with a synthetic DNA fragment that comprises 39 bp of double-stranded sequence flanked on each end by 5' overhanging GATC sequences. The synthetic fragment closely approximates the region upstream of *glpF* (previously called *glpC*) from the start codon to the *gylP1/P2* promoter, but modified to introduce a *BamHI* site at the *gylP1/P2* end, an *NdeI* site at the *glpF* start codon and an *XbaI* site immediately downstream (Fig. 1). In *S. coelicolor*, *glpF* is located just downstream from *gylR*, and its expression is regulated by GylR and glycerol (Hindle & Smith, 1984). pDA1100 maintains this organization and sequence as closely as possible. Mutant versions of *ccaR* with *NdeI*

sites at either the potential ATG or GTG (changed to ATG) start codons were excised from the pBluescript II pSK(+) constructs, which gave rise to pTK-1 and pTK-2, respectively (see above), as *NdeI*–*XbaI* fragments, and cloned into pDA1100 under the control of the *gylP1/P2* promoters, to give pDA1102 and pDA1103 (Fig. 1).

Primers LWA1, LWA2, LWA3 and LWA4 were used in a PCR-mediated process to mutate the GTG codon located near the 5' end of *ccaR* to GTC. With LWA1 (reverse) and LWA2 (forward) as primers and pDA150 as template, a 381 bp PCR product was amplified using the Expand High Fidelity PCR system (Roche; 95 °C, 30 s; 62 °C, 30 s; 72 °C, 1 min for 30 cycles) (Fig. 2a). The product was combined with a 922 bp PCR product amplified using primers LWA3 (forward) and

Fig. 1. Plasmid constructs used for expression of *ccaR* under control of the *gylP1/P2* promoter. Terminator fragments, *tmmr* and *ttd*, are shown as grey boxes; the *gylP1/P2* fragment is shown as an open box; *gylR*, *xylE* and *ccaR* genes are shown as open arrows. B, *BamHI*; Xh, *XhoI*; Xb, *XbaI*; N, *NdeI*.

Fig. 2. Conservative mutation of the potential GTG start codon in *ccaR* to GTC. (a) Primers LWA1, LWA2, LWA3 and LWA4 were used to amplify two overlapping DNA fragments (381 bp and 922 bp) from the *ccaR* gene and upstream region by PCR. The primers LWA1 and LWA3 are mutagenic, resulting in conversion of the putative GTG start codon to a GTC codon. (b) The two amplified fragments were used as template with primers LWA2 and LWA4 in a second PCR reaction. (c) Amplified product (1.23 kb) containing the mutant form of *ccaR*.

LWA4 (reverse) and the same pDA150 template, and then the mixture of the two products was used with primers LWA2 (forward) and LWA4 (reverse) in a PCR reaction with no added template to give a 1.23 kb PCR fragment (Fig. 2b, c). The 1.23 kb fragment was made blunt by treatment with T4 polymerase I, digested with *Bam*HI (LWA2 contains a *Bam*HI site), and then ligated to *Bam*HI/*Eco*RV-digested pSET152, giving pLW3.

pDA1006 carries *ccaR* cloned as a 1.4 kb *Bam*HI–*Nru*I fragment of *S. clavuligerus* DNA into the *Eco*RI–*Eco*RV sites of pSET152 after passage through pSL1180 to pick up compatible restriction sites for cloning. To prepare pDA1006*, a *ccaR* expression construct equivalent to pDA1006, except with the GTG potential start codon changed to GTC, the 300 bp *Kpn*I–*Nru*I fragment of pDA1006 which encompasses the 3' end of *ccaR* and downstream regions was used to replace the corresponding 144 bp *Kpn*I–*Eco*RV fragment of pLW3.

Expression of *ccaR* in *E. coli*. *CcaR* was expressed from either pTK-1 or pTK-2 in *E. coli*. In both plasmids *ccaR* was expressed under the control of the ϕ 10 T7 promoter in the expression vector pT7-7. The plasmids were introduced by electrotransformation into *E. coli* BL21(DE3), which carries the gene for T7 RNA polymerase under the control of the IPTG-inducible *lac* promoter. Cultures were grown to an OD_{600} of 0.6–0.8 before induction with IPTG to a final concentration of 0.4 mM. Samples of cultures were harvested at intervals just prior to, and after induction with IPTG. Cells were resuspended to 1/10 original culture volume in SDS-PAGE sample buffer. After heating for 5 min at 100 °C, 5 and 15 μ l amounts were separated on a 12% SDS polyacrylamide gel and proteins were visualized by staining with Coomassie Brilliant Blue.

Expression of *ccaR* in *S. clavuligerus*. pDA1102, pDA1103 and pDA1100 were transformed into *S. clavuligerus* Δ *ccaR*::*tsr*, a *ccaR* mutant in which the 1.4 kb *Bam*HI–*Nru*I fragment that encompasses *ccaR* is deleted and replaced by a thiostrepton resistance gene (Alexander & Jensen, 1998). Transformants were cultivated in both TSB and NMMP growth media supplemented with various concentrations of glycerol, and assayed for the production of cephamycin C. Cells harvested at each sample time were also washed with 0.85% NaCl, resuspended in 50 mM Tris/HCl, pH 7.2, 0.01 mM EDTA and 0.1 mM DTT, and then disrupted by sonication. After centrifugation for 5 min at 10 000 g, the resulting cell extracts were used for Western analysis.

The pDA1006 and pDA1006* complementation constructs were introduced into the Δ *ccaR*::*tsr* strain by conjugation and with apramycin selection for exconjugants.

Western blot analysis. Forty-microgram amounts of cell extract protein were separated by duplicate SDS-PAGE (12.5%) then electrophoretically transferred onto Immobilon-P PVDF membranes (Millipore) and

analysed as described previously (Alexander & Jensen, 1998). Primary anti-*CcaR* antibodies (Alexander & Jensen, 1998) were used at a dilution of 1:4000. Horseradish peroxidase-conjugated donkey anti-rabbit immunoglobulin G (Amersham Life Sciences) was used as secondary antibody at a dilution of 1:5000.

S1 nuclease protection and primer extension studies. The transcription start points (TSPs) for *ccaR* were determined by S1 nuclease protection analyses using previously described methods (Kieser *et al.*, 2000), and by primer extension analysis. The S1 probe used to locate TSP-1 was amplified by PCR using primers LWA5 (forward) and BKL96 (reverse) with pDA1006 as template, and the Expand High Fidelity PCR system (Roche; 94 °C, 45 s; 65 °C, 45 s; 72 °C, 45 s for 10 cycles followed by 94 °C, 45 s; 70 °C, 45 s; 72 °C, 45 s for 15 cycles). The resulting probe fragment extended from –193 bp to +22 bp relative to the *ccaR* translation start point, and included 10 bp of non-*S. clavuligerus*-derived sequence to enable full-length protection of the probe to be distinguished from probe-probe reannealing. A second S1 protection assay was also performed to investigate the possibility of an additional upstream TSP for *ccaR*. The primers *ccaR*-UP-Forward and *ccaR*-UP-Reverse were used to amplify a probe using pDA1006 as template as described above. This second S1 probe extended from –412 bp to –133 bp relative to the translation start point.

Primer extension analysis was performed using the *C. therm.* Polymerase for reverse transcription in a two-step RT-PCR procedure according to the manufacturer's instructions (Roche), with the following changes. Twenty microlitre reactions were set up using 5 pmol of the end-labelled reverse primer BKL96 and 40 units of RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen) with extension at 55 °C for 60 min and termination at 80 °C for 10 min. For both S1 nuclease protection and primer extension analyses, the reverse primers were used along with the template plasmid pDA1006 to prepare sequencing ladders, which were electrophoresed alongside the reaction products for size estimation.

Cephamycin C bioassay. The production of cephamycin C by *S. clavuligerus* was estimated by an agar diffusion bioassay with *E. coli* ESS as the indicator strain (Jensen *et al.*, 1982).

RESULTS

Expression of *ccaR* in *E. coli*

The *ccaR* gene from *S. clavuligerus* has been cloned and sequenced, but the location of the translational start point is unclear. Two different plasmid constructs, pTK-1 and pTK-2, were prepared to examine the expression of *ccaR* in *E. coli*

Fig. 3. Production of CcaR in *E. coli* cultures carrying either pTK-1 or pTK-2. SDS-PAGE analysis was used to assess CcaR production in cultures induced with 0.4 mM IPTG and grown for 3 h. Lanes 1–5, pTK-1; lanes 6–10, pTK-2. Lanes 1 and 6 are uninduced cultures. Lanes 2, 4, 7 and 9 represent 50 μ l amounts of induced cells; lanes 3, 5, 8 and 10 represent 150 μ l amounts of induced cells. Lanes 2, 3, 7 and 8 represent cells grown at 21 °C after induction; lanes 4, 5, 9 and 10 represent cells grown at 37 °C after induction. Lane 11 is empty; lane 12 contains molecular mass markers.

from the putative ATG and GTG start codons. pTK-1 contains a mutant *ccaR* gene with an *Nde*I site introduced by PCR at the ATG codon. pTK-2 contains a mutant *ccaR* gene with an *Nde*I site introduced at the GTG codon, thereby changing it to an ATG codon. In both cases, the *Nde*I site was used to position the gene precisely behind an optimized *E. coli* ribosome-binding site and ϕ 10 T7 promoter in the plasmid vector pT7-7. The 5' ends of the *ccaR* genes in pTK-1 and pTK-2 were sequenced to ensure that no errors were introduced by the PCR procedure, and none were found.

Cultures of *E. coli* BL21 (DE3) carrying either pTK-1 or pTK-2 were grown in LB medium at 37 °C until mid-exponential phase, then induced by the addition of IPTG to a concentration of 0.4 mM, followed by continued incubation at either 21 or 37 °C. Samples of cultures taken before IPTG induction and at 3 h post-induction were harvested by centrifugation, resuspended to 1/10 of the original culture volume in SDS-PAGE sample buffer, heated for 5 min at 100 °C, and analysed by electrophoresis on a 12 % SDS polyacrylamide gel followed by staining with Coomassie Brilliant Blue. Fig. 3 shows that CcaR accumulated in cultures carrying pTK-1 (lanes 1–5). In contrast, cultures carrying pTK-2 did not show detectable CcaR production (lanes 6–10). To confirm that the absence of CcaR in pTK-2-bearing cells was not due to loss of the expression vector, plasmids were isolated from cell samples taken at each time point. All cultures showed the presence of plasmid, indicating that both pTK-1 and pTK-2 were maintained in the cells (data not shown). Furthermore, to ensure that the failure of pTK-2 to support the production of CcaR was not due to a mutation in the upstream untranslated region of the gene, pTK-2 isolated from a cell sample taken at the end of the expression period was subjected to DNA sequence analysis. No mutations were seen in the upstream region extending well beyond the ϕ 10 T7 promoter region. Expression of *ccaR* from pTK-1 yielded CcaR protein, regardless of whether the cultures were incubated at 37 or 21 °C after induction. However, further analysis showed that the recombinant CcaR protein was located almost exclusively in the insoluble fraction of the cells, regardless of the expression temperature (data not shown).

Expression of *ccaR* in *S. clavuligerus*

The expression of *ccaR* from both the potential ATG and GTG start codons was also investigated in a *ccaR* mutant strain (Δ *ccaR*::*tsr*) of *S. clavuligerus*. Strains of *S. clavuligerus* Δ *ccaR*::*tsr* carrying pDA1100, pDA1102 and pDA1103 were examined for the ability of the plasmid constructs to complement the mutated chromosomal copy of *ccaR* and thereby restore production of cephamycin C. The entire *ccaR* gene has been deleted from *S. clavuligerus* Δ *ccaR*::*tsr* and replaced with a thiostrepton resistance gene, which rules out the possibility that complementation could occur by recombination and repair of the original mutated *ccaR* gene in the chromosome. In each case, the *ccaR* gene was expressed under the control of a glycerol-regulatable promoter, and complementation was tested following growth in both minimal (NMMP) and complex (TSB) media, in the presence and absence of glycerol (Table 3). Cephamycin C production was only seen in wild-type *S. clavuligerus* and in the Δ *ccaR*::*tsr* mutant strain

Table 3. Cephamycin C production by wild-type *S. clavuligerus* and *S. clavuligerus* Δ *ccaR*::*tsr* (pDA1102)

ND, Not done. Data are from single cultures assayed singly, but the entire experiment was repeated three separate times. Glycerol concentrations are % w/v.

Growth medium	Cephamycin C produced (μ g ml ⁻¹)					
	Wild-type at:			Δ <i>ccaR</i> :: <i>tsr</i> (pDA1102) at:		
	24 h	48 h	72 h	24 h	48 h	72 h
NMMP + 0.5 % glycerol	4.8	8.1	6.8	2.8	3.3	2.8
NMMP + 1.0 % glycerol	4.8	5.7	4.0	3.0	2.3	2.3
NMMP + 2.0 % glycerol	3.3	3.3	2.8	2.1	2.0	1.2
TSB, no added glycerol	ND	31	9.7	ND	0	0
TSB + 0.5 % glycerol	0	24	48	0	18	11
TSB + 0.75 % glycerol	0	13	52	0	20	16
TSB + 1.0 % glycerol	0	3.3	40	0	5.2	20
TSB + 1.25 % glycerol	0	1.2	6.8	0	4.4	13

complemented with pDA1102, the *ccaR* expression construct originating at the ATG codon, and then only when grown in the presence of glycerol. While glycerol was essential for production of CcaR from pDA1102, increasing concentrations of glycerol in the growth medium did not necessarily increase the production of cephamycin C in either wild-type *S. clavuligerus* or the $\Delta ccaR::tsr$ mutant strain. At 2.0% glycerol (w/v) in NMMP and at 1.25% glycerol in TSB, cephamycin C production was actually reduced compared to that at lower concentrations of glycerol. No cephamycin C production was detected from the *S. clavuligerus* $\Delta ccaR::tsr$ mutant alone or carrying pDA1103, from the *ccaR* expression construct originating at the GTG codon, or from pDA1100, a control construct lacking a *ccaR* insert, in any of the growth media tested, or at any concentration of glycerol.

CcaR production in complemented cells was also examined by Western analysis using anti-CcaR antibodies. Wild-type *S. clavuligerus*, *S. clavuligerus* $\Delta ccaR::tsr$ (pDA1102) and $\Delta ccaR::tsr$ (pDA1103) were grown in TSB $\pm$ 0.75% glycerol (w/v), and cell extracts from 48 h cultures were analysed. CcaR protein was only detected in wild-type and *S. clavuligerus* $\Delta ccaR::tsr$ (pDA1102), and in the latter case, only when the cells were grown in the presence of glycerol (Fig. 4). *S. clavuligerus* $\Delta ccaR::tsr$ (pDA1102) produced an amount of CcaR protein approximately equal to that seen in the wild-type, whereas the $\Delta ccaR::tsr$ (pDA1103) strain showed no detectable CcaR protein, consistent with the cephamycin C bioassay results.

Complementation of *S. clavuligerus* $\Delta ccaR::tsr$ mutants by a GTG to GTC mutant form of *ccaR*

Cephamycin C production was restored in $\Delta ccaR::tsr$ mutants by complementation with pDA1102. However, expression from this construct is driven by the non-native *gylP1P2* promoter, which might influence the choice of start codon. Furthermore, the ATG expression construct still contains the downstream GTG codon, and so it is possible

that translation could be initiating at the GTG codon within this construct rather than at the ATG codon, as intended. To eliminate these complicating factors, a pSET152-based *ccaR* expression construct (pDA1006*) was prepared, in which *ccaR* is expressed from its native promoter. In this construct, a conservative mutation has been introduced to change the GTG codon near the 5' end of *ccaR* into a GTC codon, thereby precluding its use as a start codon. The corresponding vector, pDA1006, carries the wild-type form of the *ccaR* gene. These two complementation constructs were conjugated into the $\Delta ccaR::tsr$ mutant strain. Exconjugants were cultured in TSBS medium for 48 h and then bioassayed for the production of cephamycin C. *S. clavuligerus* $\Delta ccaR::tsr$ (pDA1006) and $\Delta ccaR::tsr$ (pDA1006*) produced 125% and 96% of the wild-type level of cephamycin C, respectively. In contrast, a control plasmid, pDA1000, consisting of pSET152 carrying only the *tsr* gene, had no ability to restore cephamycin C production in the mutant. These results indicate that *in vivo* expression of *ccaR* in *S. clavuligerus* also initiates from the ATG rather than the GTG codon.

Identification of the TSP for *ccaR*

No promoter was evident upstream of the ATG start codon of *ccaR*. To locate the promoter region for *ccaR*, RNA isolated from wild-type *S. clavuligerus* grown for 24, 48, 72 and 96 h at 28 °C in soy medium was subjected to S1 nuclease protection analysis using a *ccaR*-specific probe. Soy medium supports high-level production of cephamycin C, clavulanic acid and the other clavam metabolites, and so should give correspondingly high levels of *ccaR* expression. A triplet of prominent bands indicating a TSP (TSP-1) at 74–76 bp upstream from the ATG start codon was observed when a DNA fragment extending from –193 to +22 bp relative to the ATG start codon of *ccaR* was used as the probe (Fig. 5a). In addition, a very faint band corresponding to a larger protected probe fragment was also observed, indicating a possible second TSP further upstream (Fig. 5a).

Fig. 4. Western analysis of *S. clavuligerus* wild-type, and $\Delta ccaR::tsr$ mutants carrying the integrative vectors pDA1100, pDA1102 and pDA1103. Forty microgram amounts of cell-extract protein from each strain, harvested after 48 h growth in either TSB or TSB+0.75% glycerol (w/v), were subjected to Western analysis using anti-CcaR polyclonal antisera. The ability of each strain to produce cephamycin C was estimated qualitatively by bioassay. Lanes 1–3, $\Delta ccaR::tsr$ mutants carrying pDA1102, pDA1103 and pDA1100, respectively, and grown in the absence of glycerol; lanes 4–6, the same strains, but grown in the presence of glycerol; lane 7, wild-type culture grown in the absence of glycerol; lane 8, 1 μ g purified CcaR protein. NA, Not applicable.

Fig. 5. Determination of the *ccaR* TSPs. (a) S1 protection analysis of RNA isolated from wild-type *S. clavuligerus* grown in soy medium, using a *ccaR*-specific probe. Lanes 1–4, RNA isolated after 24, 48, 72 and 96 h growth; lane C, probe digested with S1 nuclease; lane P, undigested probe. (b) Primer extension analysis of RNA isolated from wild-type *S. clavuligerus* grown in soy medium. Lane 1, RNA isolated after 72 h growth and subjected to analysis. In both (a) and (b), the TSPs detected by S1 nuclease protection and primer extension analysis are indicated by an asterisk. Lanes G, A, T and C form the sequencing ladder generated using the reverse primer (BKL96). (c) Determination of the *ccaR* TSPs. Diagrammatic representation of the *ccaR* promoter region. The filled arrow represents the *ccaR* coding sequence, including the start codon. TSP-1 and TSP-2 giving rise to the *ccaR* transcripts are shown in bold type and the solid arrows represent the mapped *ccaR* transcripts. The -10 and -35 promoter regions identified are also indicated.

To obtain better resolution, a second probe was designed to protect a region further upstream of *ccaR*, from -412 bp to -133 bp, and overlapping the first probe. When this second probe was employed, very weak protection of multiple bands was observed, and so the second *ccaR* TSP could not be located conclusively (data not shown). Primer extension analysis was also used to confirm results obtained by S1 protection assays. RNA isolated from wild-type *S. clavuligerus* grown in soy medium for 72 h was subjected to analysis using the reverse primer BKL96, the same primer used to prepare the S1 probe to map TSP-1. By primer extension, TSP-1 was mapped to a single nucleotide, 74 bp upstream of the *ccaR* ATG start codon (Fig. 5b). A faint band corresponding to the second TSP (TSP-2) was also observed and was located 173 bp upstream of the *ccaR* ATG start codon (Fig. 5b).

DISCUSSION

Heterologous expression of *ccaR* in *E. coli* from a strong T7-based promoter gave protein product only when constructs starting at the ATG codon were prepared. The ATG-based expression construct gave large amounts of CcaR protein,

visible as a prominent band on SDS-PAGE 3 h post-IPTG induction. This suggests that ATG is the normal start codon for *ccaR* expression. The inability to detect any protein product from the GTG-based expression construct also suggests that, if truncated CcaR protein was produced, it did not persist in the cells long enough to be detected by PAGE analysis. Although CcaR was produced in large amounts from the ATG-based expression construct, essentially all of the protein was in an insoluble form, regardless of the incubation temperature.

Expression of *ccaR* in *S. clavuligerus* gave results consistent with those seen in *E. coli*. CcaR protein was only produced from expression constructs initiating translation at the ATG codon. These results indicate that not only is ATG the true start codon for *ccaR* expression, but also that the extreme amino terminus of the protein must be important for the structural integrity of the protein. Truncation by as few as six amino acid residues results in disappearance of the protein from cell extracts to the limits of detection by Western analysis. CcaR is an example of a class of pathway-specific antibiotic regulatory proteins referred to as SARPs. The amino terminal region has been suggested as the location of the DNA-binding fold that characterizes this

group of regulatory proteins. Truncation may interfere with this important structural element and destabilize the protein.

Expression of *ccaR* in *S. clavuligerus* restored cephamycin C production in the $\Delta ccaR::tsr$ mutant to near wild-type levels. Although the level of expression of *ccaR* in *S. clavuligerus* was modest, the *gylP1P2* promoter was nonetheless seen to be both active, and glycerol-regulated in *S. clavuligerus*. The presence of a TTA codon in *ccaR* may limit expression under the *gylP1P2* promoter, although previous studies have shown that expression of *ccaR* is relatively insensitive to the availability of *bldA* tRNA in *S. clavuligerus* (Trepanier *et al.*, 2002).

While the ATG-based expression construct supported CcaR production in *S. clavuligerus*, whereas the GTG-based construct did not, it was still possible that expression might be regulated differently when driven by the natural *ccaR* promoter. However, a mutant form of *ccaR* under the control of its own promoter, but carrying a conservative mutation that changed the GTG codon to GTC, was still fully able to complement the $\Delta ccaR::tsr$ mutant and restore cephamycin C production. Therefore, these results provide strong evidence to indicate that the ATG codon is the natural start codon for the *ccaR* gene.

The *ccaR* promoter was investigated using SI nuclease protection analysis to locate the transcription start point to a group of three residues -74 to -76 bp upstream from the ATG start codon. A second TSP located further upstream was seen as a faint band in some SI nuclease protection analyses. Primer extension analysis further localized TSP-1 to a G residue at -74 bp. TSP-2 was also located by primer extension analyses to an A residue at -173 bp. The predicted promoter region associated with TSP-1 is TGGAAT-17 bp-AAACAT while that of TSP-2 is TCCCGA-18 bp-GTTCTT. Both of these promoters show only marginal similarity to other promoters associated with cephamycin and clavulanic acid production (Kovacevic *et al.*, 1990; Paradkar *et al.*, 1998; Petrich *et al.*, 1992, 1994).

A feature commonly associated with SARPs is the presence of heptameric sequences repeated at 11 bp (or multiples of 11 bp) intervals in the promoter regions of target genes (Wietzorrek & Bibb, 1997). A consensus sequence for these heptameric repeats has been identified, but no such repeats were evident in the *ccaR* promoter region, despite reports that CcaR regulates its own production (Santamarta *et al.*, 2002). However, if more than one regulatory element is interacting with the promoter, this may result in deviations from this general architecture. In this regard, Folcher *et al.* (2001) have defined an autoregulatory response element (ARE) sequence motif located upstream of a number of SARP-type genes, and associated with expression of the genes under the control of an A factor-like autoregulatory system. As part of that study, a potential ARE motif was identified in the region upstream of *ccaR*, and a recent report

of the cloning of a gene encoding a γ -butyrolactone autoregulator receptor protein from *S. clavuligerus* adds support to the possibility that expression of *ccaR* may be regulated by a γ -butyrolactone-based quorum-sensing system (Kim *et al.*, 2002). However, the ARE-type motif identified by Folcher *et al.* (2001) is located more than 800 bp upstream of *ccaR* TSP-1 and on the opposite strand relative to the promoter, which calls its significance into question. Additional study will be required to substantiate the involvement of an A factor-like autoregulatory system in the production of β -lactam metabolites in *S. clavuligerus*.

ACKNOWLEDGEMENTS

This study was supported by grants from the Natural Sciences and Engineering Research Council of Canada. K. T., T. L. K. and D. C. A. were supported by studentships from the Alberta Heritage Foundation for Medical Research.

REFERENCES

- Alexander, D. C. & Jensen, S. E. (1998). Investigation of the *Streptomyces clavuligerus* cephamycin C gene cluster and its regulation by the CcaR protein. *J Bacteriol* **180**, 4068–4079.
- Bierman, M., Logan, R., O'Brien, E. T., Seno, E. T., Rao, N. & Schoner, B. E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **116**, 43–49.
- Folcher, M., Gaillard, H., Nguyen, L. T., Nguyen, K. T., Lacroix, P., Bamas-Jacques, N., Rinkel, M. & Thompson, C. J. (2001). Pleiotropic functions of a *Streptomyces pristinaespiralis* autoregulator receptor in development, antibiotic biosynthesis, and expression of a superoxide dismutase. *J Biol Chem* **276**, 44297–44306.
- Hindle, Z. & Smith, C. P. (1984). Substrate induction and catabolite repression of the *Streptomyces coelicolor* glycerol operon are mediated through the GylR protein. *Mol Microbiol* **12**, 737–745.
- Jensen, S. E., Westlake, D. W., Bowers, R. J. & Wolfe, S. (1982). Cephalosporin formation by cell-free extracts from *Streptomyces clavuligerus*. *J Antibiot* **35**, 1351–1360.
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich, UK: John Innes Foundation.
- Kim, H. S., Lee, Y. J., Lee, C. K., Hwang, Y. I., Lee, I. S., Yu, T. S. & Nihara, T. (2002). Cloning of the gene encoding γ -butyrolactone autoregulator receptor protein from *Streptomyces clavuligerus*. In *9th International Symposium on the Genetics of Industrial Microorganisms*, p. 79. Gyeongju, Korea.
- Kovacevic, S., Tobin, M. B. & Miller, J. R. (1990). The β -lactam biosynthesis genes for isopenicillin N epimerase and deacetoxycephalosporin C synthetase are expressed from a single transcript in *Streptomyces clavuligerus*. *J Bacteriol* **172**, 3952–3958.
- Kyung, Y. S., Hu, W.-S. & Sherman, D. S. (2001). Analysis of temporal and spatial expression of the CcaR regulatory element in the cephamycin C biosynthetic pathway using green fluorescent protein. *Mol Microbiol* **40**, 530–541.
- Mosher, R. H., Paradkar, A. S., Anders, C., Barton, B. & Jensen, S. E. (1999). Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminic synthase 1 in *Streptomyces clavuligerus*. *Antimicrob Agents Chemother* **43**, 1215–1224.

- Paradkar, A. S. & Jensen, S. E. (1995).** Functional analysis of the gene encoding the clavamate synthase 2 isoenzyme involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *J Bacteriol* **177**, 1307–1314.
- Paradkar, A. S., Aidoo, K. A. & Jensen, S. E. (1998).** A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Mol Microbiol* **27**, 831–843.
- Perez-Llarena, F. J., Liras, P., Rodriguez-Garcia, A. & Martin, J. F. (1997).** A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both β -lactam compounds. *J Bacteriol* **179**, 2053–2059.
- Petrich, A. K., Wu, X., Roy, K. L. & Jensen, S. E. (1992).** Transcriptional analysis of the isopenicillin N synthase-encoding gene of *Streptomyces clavuligerus*. *Gene* **111**, 77–84.
- Petrich, A. K., Leskiw, B. K., Paradkar, A. S. & Jensen, S. E. (1994).** Transcriptional mapping of the genes encoding the early enzymes of the cephamycin biosynthetic pathway of *Streptomyces clavuligerus*. *Gene* **142**, 41–48.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Santamarta, I., Rodriguez-Garcia, A., Perez-Redondo, R., Martin, J. F. & Liras, P. (2002).** CcaR is an autoregulatory protein that binds to the *ccaR* and *cefD-cmcl* promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. *J Bacteriol* **184**, 3106–3113.
- Tahlan, K., Park, H.-U. & Jensen, S. E. (2004a).** Three unlinked gene clusters are involved in clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Can J Microbiol* (in press).
- Tahlan, K., Park, H.-U., Wong, A., Beatty, P. H. & Jensen, S. E. (2004b).** Two sets of paralogous genes encode the enzymes involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Antimicrob Agents Chemother* **48**, 930–939.
- Trepanier, N. K., Jensen, S. E., Alexander, D. C. & Leskiw, B. K. (2002).** The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bldA* mutant. *Microbiology* **148**, 643–656.
- Walters, N. J., Barton, B. & Earl, A. J. (1994).** Novel Compounds. *International patent* WO 94/18326-A 1.
- Wietzorrek, A. & Bibb, M. (1997).** A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25**, 1181–1184.