

An operator associated with autoregulation of the repressor gene in actinophage ϕ C31 is found in highly conserved copies in intergenic regions in the phage genome

Colin J.Ingham, Charles E.Owen⁺, Stuart E.Wilson^{1,§}, Iain S.Hunter and Margaret C.M.Smith^{1,*.§}

Robertson Institute of Biotechnology, Church Street, University of Glasgow, Glasgow G11 5JS and

¹Department of Biological and Molecular Science, University of Stirling, Stirling FK9 4LA, UK

Received November 15, 1993; Revised and Accepted February 17, 1994

EMBL accession no. X77530

ABSTRACT

Previous reports have suggested that the repressor gene, *c*, of ϕ C31 is autoregulated and that likely operators are conserved inverted repeat sequences (CIRs1&2) located just upstream of the promoters, *cp1* and *cp2*. Evidence is now presented that the CIRs 1&2 are indeed binding sites for one of the three inframe, N-terminally different protein isoforms of 42, 54 and 74 kDa produced by the *c* gene. A *cp1* – *aphII* fusion was repressed in a *Streptomyces coelicolor* A3(2) ϕ C31 lysogen and characterisation of an operator-constitutive (*O^c*) mutant showed a single mutation in CIR-1. CIR-1 containing fragments were retarded in electrophoresis gels by the 42 kDa repressor protein isoform and this retardation was inhibited by the addition of competing DNA fragments containing either CIR-1 or CIR-2. Using a combination of Southern blotting and analysis of available DNA sequence we also show that at least 18 copies of the CIRs are present throughout the ϕ C31 genome. Alignment of 9 CIR sequences showed that 8 contained a perfectly conserved 17 bp core whilst the exception had a single mismatch. The core includes a 16 bp inverted repeat (IR), and is usually part of a more extensive and less highly conserved palindrome. When superimposed on a previously derived transcription map of the early region, the CIRs lie in intergenic regions associated with transcription initiation and/or termination.

INTRODUCTION

Streptomyces spp. are mycelial, sporulating, Gram-positive bacteria which produce a wide variety of antibiotics and bioactive molecules (1). Molecular studies on the genes for the biosynthetic

pathways for such secondary metabolites are frequently aimed at increasing productivity and the development of novel pharmaceuticals. The *Streptomyces* temperate phage ϕ C31 has already been used to develop a variety of vectors for genetic manipulation of these commercially important bacteria (1–4) but there still remains scope for the further exploitation of genetic control elements from this bacteriophage in designing regulatable expression systems. Furthermore, as mechanisms of gene control employed by viruses tend to reflect those employed by their hosts, studies on ϕ C31 will result in a greater understanding of gene expression and global regulatory systems in the morphologically and metabolically complex streptomycetes. We are studying the mechanisms which control gene expression in the switch between lytic and lysogenic growth and early events in the lytic cycle of ϕ C31. Clear plaque mutants map to a single locus (*c*) which lies approximately in the middle of the 41.2kbp genome (5,6). The *c* locus is required for the maintenance and probably for the establishment of lysogeny (7,8). Early and late genes are in distinct clusters; the late genes lie to the left of the *c* locus and overlap the *cos* region whereas the early genes lie to the right of *c* and extend to the *attP/int* region (9,10). The *c* gene encodes three inframe proteins of 42, 54 and 74 kDa (8, 11, 12) which are transcribed from at least 2 promoters, *cp1* and *cp2*. Transcription from *cp1* is probably autoregulated as *cp1* – *lacZ* fusions can be repressed in *E. coli* by expression *in trans* of the 54 and/or the 42kDa repressor proteins (13). Although the recognition sequences for binding of the repressor proteins to DNA are not yet known, two completely conserved 20 bp inverted repeat sequences (CIRs) in the *OLcp1* and *cp2* promoter regions are candidate operators (8,12,13). The repressor(s) inhibit expression of the lytic genes, so additional repressor binding sites might be expected to control at least one other ϕ C31 genes which may be required for commitment to lytic growth. Here we present evidence that CIRs 1&2 are indeed operators in the autoregulation

*To whom correspondence should be addressed

Present addresses: ⁺ICRF, 44 Lincoln's Inn Fields, London WC2A 3PX and [§]Department of Genetics, Queens Medical Centre, University Park, Nottingham NG7 2UH, UK

of the repressor gene and we report the mapping of 18 CIR sites within the phage genome including those previously found at *cp1* and *cp2*.

MATERIALS AND METHODS

Basic microbiology and DNA manipulations

Bacterial strains, phages and plasmids used are summarised in Table 1. Growth and transformation of *E. coli* strains, and preparation of plasmid and single-stranded M13 phage DNA was as described by Sambrook *et al* (14). With the exception of chromosomal DNA preparations which were done by the method of Hunter (15), all growth and manipulations of *Streptomyces* strains and phages, including plasmid transformations and preparations of plasmids and phage DNAs, were as described by Hopwood *et al*, (16). Restriction enzyme (Life Technologies) digests, calf intestinal alkaline phosphatase (CIAP; Boehringer-Mannheim) incubations, ligations with T4 ligase (Life Technologies), and agarose gel electrophoresis were performed as described in Sambrook *et al* (14).

Plasmid and phage constructions were performed as follows: For the isolation of the *cp1-O^c* mutant pCJI300 was constructed by isolation of the 300bp *HindIII-PvuI* fragment from pIJ2410 (11), creation of blunt ends by treating with the Klenow fragment of DNA polymerase and deoxyribonucleotide triphosphates and ligation to *HindIII*-cut pUC19 using *HindIII* linkers to form pMS242. The 300bp *HindIII* fragment from pMS242 was then inserted into *HindIII*-cut, phosphatased pIJ486 to form pCJI300. Transformants of *S. lividans* 1326 containing pCJI300 were selected for resistance to thiostrepton and those which also grew on R2YE containing 200µg/ml kanamycin were checked for the presence of the *cp1*-containing *HindIII* fragment in the correct orientation to express the *aphII* gene. The 300bp *HindIII* fragment from pCJI301 was isolated for sequencing as follows: The 1.7kbp *KpnI* fragment from pCJI301 was inserted into pUC19 to form plasmid pCJI302, from which the 300bp *HindIII* fragment was inserted into *HindIII* cut M13mp19 in both orientations to form mCJI303 and mCJI304. For the band shift studies pCO107 was constructed by insertion of the 475 bp *NheI-SmaI* fragment from pIJ2410 into *XbaI/SmaI* cut pUC19. pMS251 was constructed by cutting pMS230 (12) with *NcoI*, treating with S1 nuclease to generate a blunt end and the 300 bp fragment released after ligating on *BamHI* linkers and cleaving with *BamHI* enzyme was then ligated to *BamHI* cut pUC19. CIR sites were mapped using *EcoRI* fragments from a complete digest of ϕC31 (Norwich stock) inserted into pUC19 to form pSE1-4 (Fig. 4). CIR-9 was sequenced using M13 derivatives mMS300 and mMS301 constructed by ligating purified *SphI*-F fragment into M13mp19 in both orientations; mMS300 contained the *SphI*-F fragment with the internal *EcoRV* sites closest to the *HindIII* site.

Purification of 42kDa repressor protein

Purification of the 42kDa repressor protein was carried out essentially as described previously (12) but with some modifications. *E. coli* DS941(pMS221) containing the 2.1 kbp *NcoI-SphI* fragment from the ϕC31 repressor gene was grown overnight in 1 litre of L-broth, washed, lysed in a French Pressure Cell and 50% ammonium sulphate was added to the cleared supernatant. After dialysis against buffer A (20mM Tris-HCl pH8.0, 5mM MgCl₂, 1mM EDTA, 0.1mM DTT, 1mM benzimidazole) the supernatant was loaded onto a 10ml DEAE-Sephacel column and protein was eluted with a linear Buffer

A/NaCl gradient. The 42kDa repressor eluted at approximately 200mM NaCl. Fractions were pooled and loaded straight onto a Heparin-Agarose affinity column. Fractions were eluted using a step gradient of Buffer A/NaCl in 100mM increments to 600mM. The 42kDa repressor protein eluted at approximately 400mM NaCl, the fractions were pooled, glycerol was added to 50% and the protein was stored at -20°C.

DNA binding assays

DNA substrates were prepared using polymerase chain reaction (PCR) and oligonucleotide primers, one of which was labelled at the 5' end. Oligo 1 (12) and reverse sequencing primer were used to generate the binding substrate *cp1/CIR-1:H3* from pMS242 and oligo 1 and reverse sequencing primer were used to make *cp1/CIR-1:Nhe* from pCO107. Oligo 1 was labelled prior to amplification at the 5' end by T4 polynucleotide kinase using 50µCi of γ³²P-ATP (6,000Ci/mmol; 14). PCR was performed as described previously (14); 30 pmoles of the labelled and unlabelled oligos were used in PCR reactions with 2–5ng template DNA, 1.25mM dNTPs; buffer and Taq polymerase. Universal primer and reverse sequencing primer were used to amplify the 300bp *BamHI* insert in pMS251 containing the *cp2* promoter and CIR-2. This amplification product was called *cp2/CIR-2*. 42kDa protein was incubated with 0.65ng of *cp1/CIR-1:H3* or 1.8ng of *cp2/CIR-1:Nhe* in binding buffer (20mM Na phosphate pH6.6, 50mM NaCl, 1 mM EDTA 0.1mM DTT, 5% glycerol, 100µg/ml BSA and 250ng pUC18) for 15 minutes at room temperature, the reaction was quenched on ice and loaded onto a pre-equilibrated (50V, 30 mins), native, 20mM Na phosphate buffer pH6.6, 4% polyacrylamide gel. After electrophoresis (75V for at least 3 hours) the gel was transferred to 3MM paper and subjected to autoradiography. Competition assays were performed by the incubation of unlabelled *cp1/CIR-1:H3* or *cp1/CIR-1:Nhe* or *cp2-CIR-2* substrates with 42kDa protein prior to the addition of the labelled substrate. Monomeric repressor concentrations are used throughout.

Hybridisation mapping of CIR sequences

Blotting and hybridisation procedures were essentially performed as described in Sambrook *et al* (14). ϕC31 DNA (Norwich stock) was digested with restriction enzymes and blotted onto Hybond-N filters (Amersham). Initially a 20mer oligonucleotide (oligo 'A': 5'-CAGCCTACTCACGTGAGTAG) was synthesised (Applied Biosystems) which would hybridise to copies of the 20nt sequence (CIR) common to the *cp1* and *cp2* promoter regions. 200ng of the 20mer oligo was labelled with 100µCi of γ³²P ATP using T4 kinase (Life Technologies) and standard reaction conditions (14). Prehybridisation was at 65°C for 5 hours and hybridisation at 45°C for 2 hours in 6×SSC with 200µg/ml heparin (Sigma). Washing in 5×SSC and 0.1% SDS was at room temperature for 10 mins then either twice at 45°C (non-stringent) or twice at 60°C for 30 mins (stringent). Detection was by autoradiography for 6 days at -70°C with intensifying screens.

The degenerate 21mer oligonucleotide 'B'; (5' CC/TC/AA/TG/ACCTACTCACGTGAGTAG 3') was then used to probe digests of genomic and plasmid DNA blotted onto Hybond-N filters. Prehybridisation and hybridisation were at 50°C in 10 ml Quikhyb solution (Stratagene) supplemented with 2mg sheared but non-denatured salmon sperm DNA (Sigma), and performed according to the Stratagene protocol for use of Quikhyb. The filters were washed (50°C, 3×15 minutes) in 5×SSC, 0.1% SDS and 0.1% sodium pyrophosphate.

DNA sequencing

CIR-9 was sequenced by priming mMS300 ssDNA with oligo 'A' using a Sequenase 2.0 kit (USB) and, from the sequence obtained a 16mer oligo was designed (5'AACTTCTTCATG-ACGG) which was then used to sequence back across the oligo 'A' priming site in mMS301. mCJI303 and mCJI304 were sequenced using a Sequenase 2.0 kit.

RESULTS

Selection and characterisation of a *cp1-O^c* mutant

pCJI300 containing the *cp1-aphII* transcriptional fusion was used to transform a *S. coelicolor* J1501:ϕC31 lysogen and a J1501 non-lysogen. The resulting strains grew on R2YE containing less than 50 µg/ml and more than 200 µg/ml kanamycin respectively suggesting that *cp1* was repressed by the lysogen. Selection for a derepressed promoter was performed by streaking J1501:ϕC31(pCJI300) on R2YE plates containing 200 µg/ml kanamycin. Many of the spradic colonies which grew were non-lysogens. From those that continued to release ϕC31 with a wild type plaque morphology, the plasmid pCJI301 was isolated. The 300bp *HindIII* fragment was inserted into M13mp19 in both orientations and sequenced. The only change to the wild type sequence was a G to T transition at position 14 of the 17 bp core sequence within the CIR (Fig. 1). To verify that this mutation

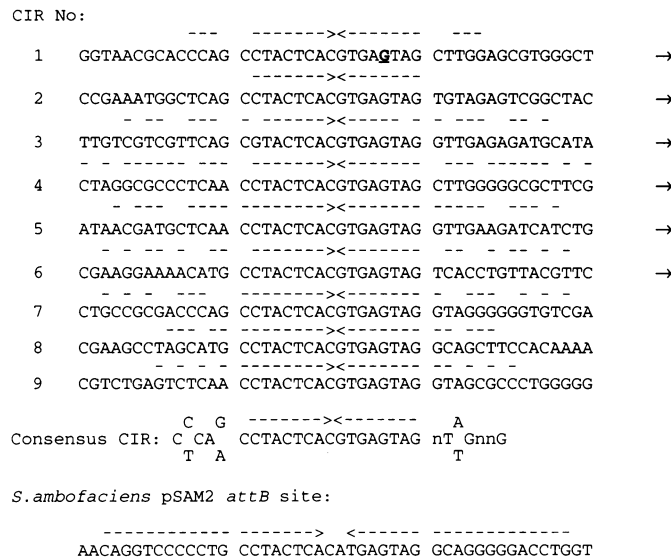


Figure 1. Position of the *cp1O^c* mutation and conservation of CIR sequences. CIRs-1 and 2 start at positions 677 and 1249, respectively in the *c* gene (8); CIRs-3 to 6 start at positions 2426, 5600, 7602 and 7798, respectively, in the sequence of the ϕC31 early region (accession number X76288; N.M. Hartley, C.J. Bruton, G.J.P. Murphy and K.F. Chater, personal communication); CIRs-7 and 8 come from the late region as shown in Fig. 4 (S. Kuhstoss, personal communication; C.W. Howe and M.C.M. Smith., unpublished); CIR-9 lies in the late region to the left of the repressor gene (accession number X77530; this study, Fig. 4). The underlined G in CIR-1 is mutated to T in the *cp1O^c* mutant. CIRs were aligned in the same orientation with respect to their position on the ϕC31 genome reading left to right. The direction of nearby or readthrough transcription (where known) is shown by the arrows. IRs are shown by the dashed lines above the sequences. The consensus CIR sequence was derived when at least 6 CIRs had the same nucleotide or all nine had one of two alternatives at a given position. The core CIR sequence is the almost completely conserved central 17 bp sequence. The *S. ambofaciens* DSM40697 CIR-like sequence (18) is shown.

was indeed responsible for the observed phenotype, the fragment was replaced from mCJI303 into pJ486 and introduced back into *S. coelicolor* J1501:ϕC31 and J1501. As the plasmids continued to show a derepressed phenotype for *cp1*, the mutation in pCJI301 was designated operator-constitutive (*O^c*).

The introduction of multiple copies of a repressor binding site might be expected to affect ϕC31 infection possibly by binding sufficient repressor so that lysogens are induced or so that they cannot establish. In the case of *S. lividans* 1326 containing pCJI300, which contains the *cp1* promoter in a high copy number vector (approx. 50 copies/chromosome), infection with ϕC31 still gave turbid plaques. However the wild type *cp1*-containing fragment did affect an *S. coelicolor* J1501:ϕC31 lysogen in a transformation assay; the frequency of transformation of pCJI300 into protoplasts of J1501:ϕC31 was decreased by approximately 1000-fold when compared with transformation by pCJI301 which contains *cp1-O^c*. In contrast, the transformation frequencies of a non-lysogen of J1501 by pCJI300 and pCJI301 were similar.

The 42kDa repressor protein binds to DNA fragments containing *cp1/CIR-1* and *cp2/CIR-2*

Purified 42kDa repressor protein and DNA substrates containing the *cp1* promoter were used in band shift assays (Fig. 2). Binding of the *cp1/CIR-1:H3* substrate was observed at a repressor concentration of 125nM (Fig. 2a) and the *cp1/CIR-1:Nhe* fragment shifted at a similar repressor concentration (161nM; data not shown). A comparison of DNA concentrations required to inhibit binding in a competition assay indicated specific binding and different relative affinities for the different sites (Fig. 2b).

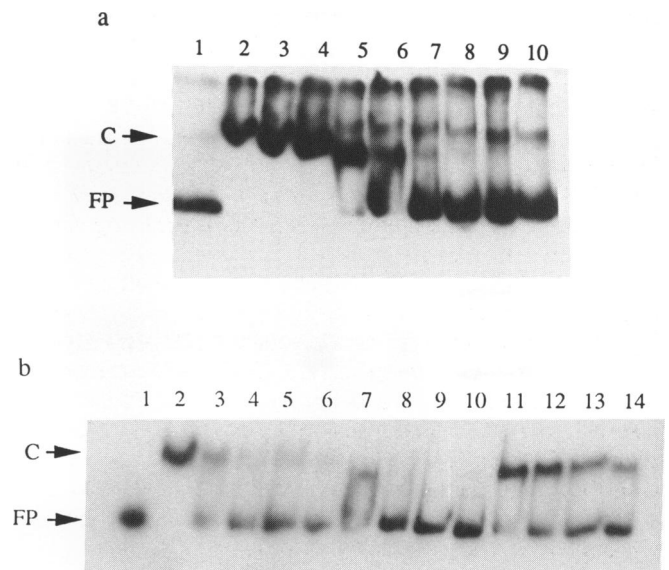


Figure 2. Binding of the 42kDa protein to DNA containing CIRs-1&2. Positions of the free probe (FP) and complex (C) are indicated with arrows. (a) Titration of 42kDa repressor binding to *cp1/CIR-1:H3*. 6.6 fmol of labelled *cp1/CIR-1:H3* was incubated with 0 (lane 1), 684 nM (lane 2), 298 nM (lane 3), 208 nM (lane 4), 149 nM (lane 5), 119 nM (lane 6), 89 nM (lane 7), 60 nM (lane 8), 45 nM (lane 9) and 21 nM (lane 10) of 42kDa repressor. (b) Competition by CIR-containing fragments for 42kDa repressor. 12 fmol of *cp1/CIR-1:Nhe* fragment was added to a binding reaction containing 240nM 42 kDa repressor and unlabelled *cp1/CIR-1:Nhe* (lanes 3-6 containing 7, 37, 74 and 110 ng respectively), *cp2/CIR-2* (lanes 7-10 containing 7, 37, 74 and 110 ng respectively) and *cp1/CIR-1:H3* (lanes 11-14 containing 3, 15, 30 and 45 ng respectively).

A complex formed with 250nM 42kDa repressor and 12 fmoles (1.8ng) of ³²P-labelled *cp1/CIR-1:Nhe* was partially competed by as little as 7.5 ng of unlabelled *cp1/CIR-1:Nhe* and *cp2/CIR-2* fragments and 3 ng of *cp1/CIR-1:H3*; the *cp2/CIR-2* DNA appeared to be the most effective competitor as the complex could be almost completely abolished by adding 36.7 ng competitor DNA compared with greater than 45ng *cp1/CIR-1:H3* and 110ng *cp1/CIR-1:Nhe*. Additions of pUC18 showed no competition even with the addition of 1µg of DNA implying that the interactions between repressor and CIR-containing fragments were specific.

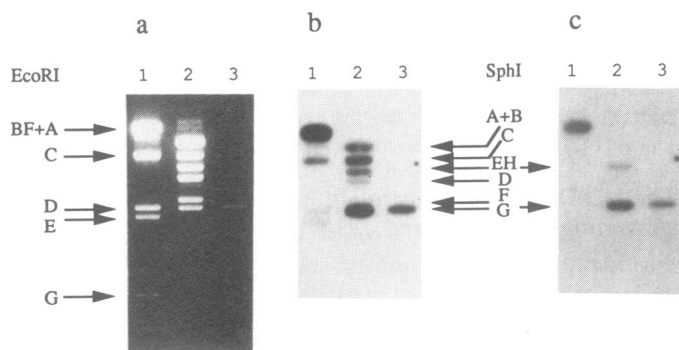


Figure 3. Southern blot of restriction digests of ϕ C31 (Norwich stock) DNA with oligo 'A' as the probe. *EcoRI* (lane 1), and *SphI* (lane 2) digests of ϕ C31 DNA and the purified *SphI*-G fragment (lane 3) were electrophoresed on a 0.8% agarose, 1×TBE gel and ethidium bromide stained (panel a), blotted onto Hybond-N paper, probed with ³²P-labelled oligo 'A' and washed under non-stringent (panel b) and stringent (panel c) conditions.

Mapping of conserved inverted repeats (CIRs) by Southern hybridisation

Southern blots of ϕ C31 DNA digested with *EcoRI* and *SphI* demonstrated an abundance of CIR-like sites when probed with oligo 'A' and washed at 45°C (Fig. 3). Only the *EcoRI*-G fragment and the *SphI*-I fragments lacked a CIR site; the presence or not of CIR sites on the *EcoRI*-A and BF fragments, the *SphI*-A and B and the *SphI*-F fragments was ambiguous due to co-migration (Fig. 3a & b). A 54°C wash gave similar results to the non-stringent wash (data not shown) but after washing at 60°C (Fig. 3c) only the *SphI*-G fragment, the *SphI*-EH fragments and the co-migrating *EcoRI*-BF and *EcoRI*-A fragments were detected.

In order to map CIRs more precisely, digests of ϕ C31 fragments in pUC19 (pSE1-4 containing *EcoRI* fragments) and pBR322 (pKC series containing *SphI* fragments; 9) were probed with oligo 'B' (Fig 4a & b). On the basis of these data a minimum of 16 CIR-like sequences outside the *c* locus were identified. Oligo 'B' hybridised to at least 7 fragments (Fig. 4a lane 1: 0.28, 0.35, 0.50, 0.52, 0.72, 0.82 and 0.99 kbp) from a complete *Sau3AI* digest of pSE1 and 6 fragments (0.35, 0.40, 0.50, 0.63, 0.66 and 0.85 kbp) from a *SalI* digest (lane 2). At least 2 CIRs were located on pKC23 as 2 *SalI* fragments (1.1 and 0.72 kbp) hybridised to oligo 'B'. One of these sites must be the same as that on the 0.18 kbp fragment detected by oligo 'B' in a *Sau3AI* digest of pSE2 (Fig. 4a, lane 4). At least one CIR was located on a 0.98 kbp *Sau3AI* fragment from pKC14 (Fig. 4a, lane 5). pKC13 contains at least 2 CIRs as oligo 'B' hybridised to 3.0 and 1.5 kbp *SalI* fragments. pKC13 contains *SphI*-I in addition to *SphI*-C but probing whole phage digests (Fig. 3) showed that the two CIR sites on this plasmid must be exclusive to *SphI*-C.

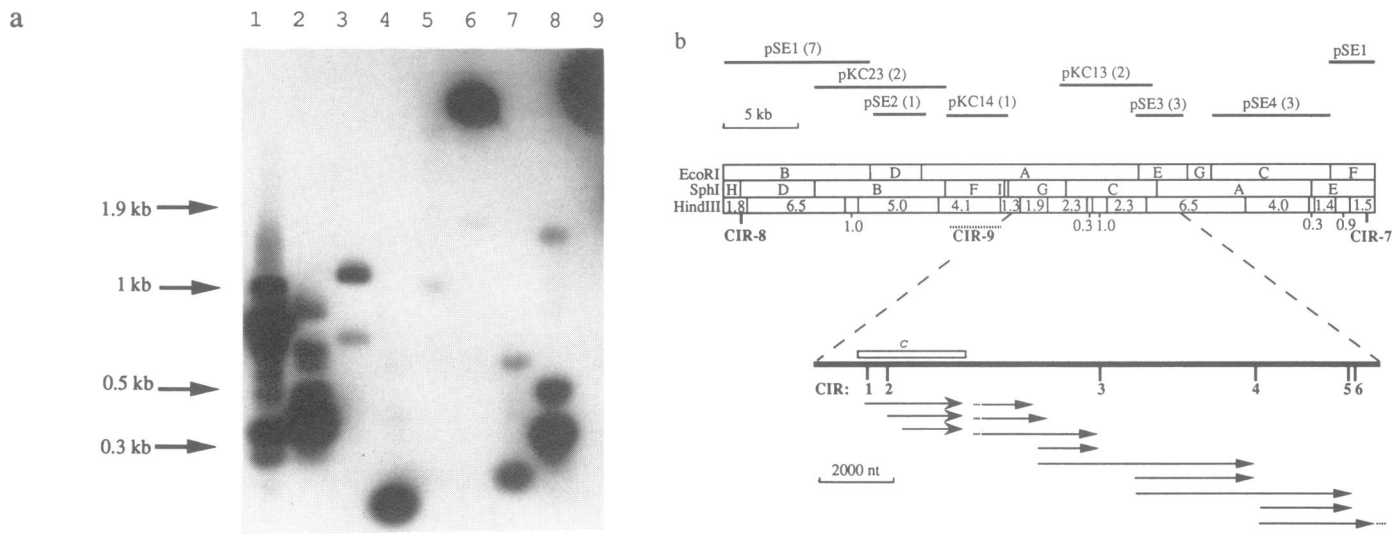


Figure 4. Southern blot of restriction digests of plasmids containing ϕ C31 DNA and probed with oligo 'B'. Panel a; *Sau3AI* (lanes 1, 4, 5, 7 and 8) and *SalI* (lanes 2, 3, 6, and 9) restriction digests of pSE1 (lanes 1 and 2), pKC23 (lane 3), pSE2 (lane 4), pKC14 (lane 5), pKC13 (lane 6), pSE3 (lane 7), pSE4 (lanes 8 and 9). Panel b; Mapping of CIRs onto the ϕ C31 genome and their positions with respect to the transcription map of the early region. *EcoRI*, *SphI* and *HindIII* restriction maps of ϕ C31 genome are shown; the fragments in the *EcoRI* and *SphI* digests are named according to ref 6, whilst the fragments in the *HindIII* map are in kbp. Above the genomic map horizontal lines represent the inserts contained in pSE1-4 and the pKC series plasmids used in panel a. In parenthesis are the numbers of CIR sites on each fragment derived from the experiment in panel a. Below the genomic map are the positions of CIRs-1 to 9 (represented by vertical lines). The expanded region shows the early region and *c* locus. Transcription of this region is represented by arrows, all reading rightwards. The transcripts in the *c* region (shown by the forked arrowheads) terminate at one of two terminators located about 250bp apart (12); for simplicity transcripts ending at only one terminator are shown. CIR-9 has not been mapped precisely and this is represented by the dashed line, see text.

Two fragments (0.22 and 0.63 kbp) hybridised to oligo 'B' after probing the *Sau3AI* digest of pSE3 (Fig. 4a, lane 7) but a search of available ϕ C31 sequence (N.M. Hartley, C.J. Bruton, G.J.P. Murphy and K.F. Chater, personal communication) revealed that the 0.22 kbp band is a doublet and that in fact pSE3 contains at least 3 CIRs (see below). Finally, pSE4 contained at least 3 sites on 0.33, 0.50 and 1.3 kbp *Sau3AI* fragments (Fig. 4a, lane 8), all of which lie within a 3.6 kb *SalI* fragment (lane 9). Probing of a strain of ϕ C31 containing the 1.4 kb Moscow deletion (7) with oligo 'B' suggested that 2 or 3 CIR sites are lost by this deletion (data not shown). In no case did oligos A or B hybridise to vector (pUC19 or pBR322), control (phage λ) or ϕ C31 DNA known to lack CIR sites.

Sequence comparison of CIRs

A search of available ϕ C31 DNA sequence using the FIND program of the University of Wisconsin GCG DNA analysis software package (17) confirmed the mapping data, enabled the precise position of the CIRs to be superimposed onto the transcription map of the early region (Fig. 4b) and further demonstrated the high sequence conservation (Fig. 1). A search of the 10 kbp early region sequence, allowing up to 2 mismatches to oligo 'A', located a total of 4 CIRs (CIRs-3 to 6). Another

CIR (CIR-7) was located downstream of the *int* gene (S. Kuhstoss, personal communication; C.W. Howe and M.C.M.S., unpublished observations) and another (CIR-8) was located within *SphI*-D (C.W.H. and M.C.M.S., unpublished observations). The number and locations of the CIR sequences identified from the DNA sequence agreed with the hybridisation data. FRAME analysis of the available DNA sequence suggests that most CIRs (the exception is CIR-2 which lies within the repressor ORF) lie in untranslated regions (N.M. Hartley, C.J. Bruton, G.J.P. Murphy and K.F. Chater, personal communication; C.W.H. and M.C.M.S., unpublished). Sequencing of CIR-9 was performed using oligo 'A' as a sequencing primer to generate the sequence adjacent to CIR-9 and then a reverse primer was made to sequence back across the CIR site.

The sequence data (Fig. 1) confirmed that the 60°C washes detected perfect matches to oligo 'A' and that sequences containing 1 mismatch or less were detected by the non-stringent wash. Detection of CIR sites was more sensitive with the degenerate oligo 'B' but it is still likely that copies with a larger number of mismatches will not have been detected. Alignment of the known sequences revealed that, with the exception of CIR-3, the 17bp core (Fig. 1) was perfectly conserved. CIR-3 was observed to hybridise particularly poorly to oligo 'B'

Table 1. Strains, plasmids and bacteriophages.

	Relevant genotype	Reference/Source
<i>Streptomyces</i>		
strains		
<i>S.coelicolor</i> A3(2)		
J1501	<i>pgl⁻ hisA1, uraA1</i>	K.F. Chater
J1501 ϕ C31	ϕ C31 (Norwich stock) lysogen	K.F. Chater
<i>S.lividans</i> 1326	Prototroph	(16)
<i>Streptomyces</i>		
Phages		
ϕ C31	Norwich stock	K.F. Chater
ϕ C43	Homoimmune to ϕ C31	K.F.Chater
ϕ CSEA1	Homoimmune to ϕ C31	This work
ϕ CSMA	Homoimmune to ϕ C31	This work
ϕ CA10	Homoimmune to ϕ C31	This work
<i>Streptomyces</i>		
Plasmids		
pIJ486	Vector with promoterless kanamycin resistance gene	(29).
pCJI300	Insert from pMS242 in pIJ486	This study
pCJI301	pCI300 containing <i>cp1O^c</i>	This study
M13 derivatives		(30)
mMS300/1	<i>SphI</i> -F fragment from ϕ C31 in opposite orientations in mp19	This study
mCJI303/4	<i>cp1O^c</i> -containing fragment in opposite orientations in mp18	This study
<i>E. coli</i> plasmids		
pUC19	Vector	(31)
pMS242	300bp <i>HindIII</i> fragment encoding <i>cp1</i> in pUC19	This study
pCO701	400bp <i>NheI-SmaI</i> fragment encoding <i>cp1</i> in pUC19	This study
pMS252	300bp <i>BamHI</i> fragment encoding <i>cp2</i> in pUC19	This study
pMS230	Contains <i>PvuI-SphI</i> fragment from the <i>c</i> gene	(12)
pKC series	Clones of <i>SphI</i> fragment from	
ϕ C31 in pBR322	(9; Fig. 4)	
pSE1-4	Clones of <i>EcoRI</i> fragments from ϕ C31 in pUC19	This study; Fig. 4)
pCJI302	1.7 kb <i>KpnI</i> fragment from pCJI301 in pUC19 including <i>cp1O^c</i>	This study

presumably due to the mismatch in the core sequence (Figs. 1 & 4a). Given that some of the sites are clustered and that doublets may have been unresolved in gels, 18 CIRs on ϕ C31 (Norwich stock) is the minimum estimate.

CIR-like sequences outside of ϕ C31

A search of the GENEMBL databases with the GCG software FIND program (17) using the CIR 17bp core sequence revealed only one closely-related sequence (1 mismatch) and that lies in the *S.ambifaciens* DSM40697 genome. The *S.ambifaciens* CIR-like sequence is within a more extensive IR which is adjacent to the *attB* site for the integrative plasmid pSAM2 (18, 19). This sequence has been suggested to have terminator function (18). Whilst *S.ambifaciens* and *S.lividans* are both hosts for ϕ C31 and pSAM2 (7; CJI, unpublished observations), *S.lividans* does not contain a CIR-like sequence near the pSAM2 *attB* site (18). The significance of the *S.ambifaciens* CIR-like sequence is unclear but its presence does raise the possibility of a specific interaction between ϕ C31 and this host. Furthermore oligo 'B' hybridised to 2 fragments (2.1 and 3.8 kbp) in a chromosomal blot of *SalI*-digested *S.coelicolor* J1501 DNA. Under the same conditions the CIRs in a J1501: ϕ C31 lysogen were easily detected (data not shown).

We analysed a collection of phages homoimmune to ϕ C31 for the presence of CIR sites and all contained them. Probing of *SalI* digests with oligo 'A' of the homoimmune phages ϕ C43, ϕ SEA1, ϕ CSMA and ϕ CA10 revealed at least 10 CIRs with non-stringent washes whilst no hybridisation was detected in 12 actinophages heteroimmune to ϕ C31 including R4 (data not shown).

DISCUSSION

Two completely conserved 20 bp inverted repeat sequences (CIRs) in the *cp1* and *cp2* promoter regions were previously suggested to be operators in the autoregulation of the repressor gene (8,12,13). We have presented evidence here that CIR-1 or a sequence overlapping CIR-1 does indeed act as an operator of *cp1*. A *cp1O^c*-mutation was isolated on the basis of derepressed promoter activity in a ϕ C31 lysogen and was also shown to restore a reduced ability to transform *S.coelicolor* ϕ C31 lysogens. The *cp1O^c* mutant contained a single change in the CIR-1 core sequence. Furthermore purified 42kDa protein bound specifically to fragments containing *cp1*/CIR-1 and *cp2*/CIR-2 *in vitro*. Only homoimmune phages were found to have CIR sites, an observation consistent with a repressor:CIR interaction. Using a combination of Southern blotting and analysis of available DNA sequences we have further shown that at least 18 copies of the CIR are present throughout the ϕ C31 genome. An alignment of 9 CIRs showed that 8 contained a perfectly conserved 17 bp core whilst the exception had a single mismatch. The core includes a 16 bp inverted repeat (IR), and is frequently part of a more extensive and less highly conserved palindrome.

The presence of many sites which interact with repressor is rare amongst the enterobacteriophages. Lytic growth of λ , P22, P2 or Mu is repressed by the action of phage repressors at a few critical promoters (20–23). On the other hand P1 contains 7 regions known to interact with its repressor at eleven operators and as much as 80% of the P1 genome may be expressed in the prophage state (24). The multiple operators in P1 appear to be involved in repression of a few key genes whilst allowing transcription of the majority (24). In a ϕ C31 lysogen, lytic genes and indeed lytic promoters are silent until induced (10) suggesting

complete transcriptional repression more reminiscent of λ , P22, P2 and Mu. At least one CIR in addition to CIRs-1&2 is likely to act as an operator to regulate transcription of a critical gene, possibly an activator of lytic promoters that is required for commitment to lytic growth. An immediate-early promoter, that is one recognised by host RNA polymerase, has recently been identified which overlaps with 2 closely situated CIRs (CIRs-5&6; C J Ingham, I S Hunter and M C M Smith) and early promoters are indeed activated specifically during the lytic cycle (25; C.W. Howe and M.C.M. Smith, unpublished).

Whilst we can assign probable functions as operators for at least CIRs-1&2 (and possibly CIRs-5&6) this now leaves the question of the functions of the remaining CIR sequences located throughout both early and late lytic regions. One possibility is that this number of binding sites is required to form a stable repression complex by looped CIR:repressor interactions such as those reported for AraC:operator interactions (26). Alternatively CIR sites may have roles during lytic growth during which any repressor:CIR interaction cannot be essential as ϕ C31 containing deletions of the *c* locus are viable in the lytic cycle (27). Superimposing the positions of the CIR sites (where known) onto the transcription map of the early region (10) shows that the CIRs lie in regions where transcription initiates (eg CIRs-1&2), terminates (eg CIR-3) or both (CIRs-4 to 6). For all CIRs to function as classical operators would appear redundant because the lytic promoters are apparently under positive control (25). However, as in the action of phage λ Cro protein (20), build up of phage repressor during expression of early genes could eventually turn down transcription from early promoters. If CIR sites perform this role in ϕ C31 it is not clear why so many (9 sites) are present in the late region. The incidence of CIRs with transcriptional terminators may suggest a role in the regulation of termination although preliminary experiments using 42kDa repressor in *in vitro* transcription reactions indicated that this is not the case (C.J. Ingham, I.S. Hunter and M.C.M. Smith, unpublished observations).

The reason why the CIRs are almost completely conserved is not clear. Constraints on the CIR sequence could arise through very stringent requirements for repressor:CIR interactions or via CIR:CIR interactions at the DNA or RNA level. Due to their high sequence conservation, phage-encoded CIRs could present a target through which a host defence mechanism could operate. In fact both the presence of CIR sites and sensitivity to the *pgl* (phage growth limitation) status of *S.coelicolor* A3(2) are specific to phages homoimmune to ϕ C31 (28).

ACKNOWLEDGEMENTS

Thanks to Nigel Hartley, Celia Bruton, George Murphy, Keith Chater, Stuart Kuhstoss and Christine Howe for the communication of results prior to publication. Thanks to Keith Chater and the participants of the 1990 Advanced *Streptomyces* Genetics course at the John Innes Institute for gifts of phages and Rachel Anderson for technical assistance. This work was supported by SERC grant GR/F 53849, MRC grant 9114970, and Wellcome grant 034787/Z/91.

REFERENCES

1. Chater, K.F. In *The Bacteria*, vol. IX. Queener, S.E. and Day, L.E. (eds). Academic Press, Orlando; (1986) pp. 119–158.
2. Kobler, L., Schwertfirm, G., Schmieger, H., Bolotin, A. and Sladkova, I. (1991). *FEMS Microbiol. Letts.* **78**, 347–354.

3. Kuhstoss, S., Richardson, M.A. and Rao, R.N. (1991). *Gene* **97** 143–146.
4. Bruton, C.J., Guthrie, E.P. and Chater K.F. (1991). *Bio/technology* **9**, 652–656.
5. Sladkova, I.A., Vasil'chenko, L.G., Lomovskaya, N.D. and Mkrtumian, N.M. (1980) *Mol. Biol. (Mosk)* **14**, 910–915.
6. Harris, J.E., Chater, K.F., Bruton, C.J. and Piret, J.M. (1983) *Gene* **23**, 167–174.
7. Lomovskaya, N.D., Chater, K.F. and Mkrtumian, N.M. (1980). *Microbiol. Revs.* **44**, 206–229.
8. Sinclair, R.B. and Bibb, M.J. *Mol. Gen. Genet.* (1988). **213**, 269–277.
9. Suarez, J.E., Clayton, T.M., Rodriguez, A., Bibb, M.J. and Chater, K.F. (1992). *J.Gen. Microbiol.* **138**, 2145–2157.
10. Ingham, C.J. and Smith, M.C.M. (1992). *Gene* **122**, 77–84.
11. Sinclair, R.B. and Bibb, M.J. (1989) *Gene* **85**, 275–282.
12. Smith, M.C.M. and Owen, C.E. (1991). *Mol. Micro.* **5**, 2833–2844.
13. Smith, M.C.M., Ingham, C.J., Owen, C.E. and Wood, N.T. (1992). *Gene* **115**, 43–48.
14. Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning, A laboratory manual*, Cold Spring Harbour Laboratory Press, USA (1989).
15. Hunter, I.S. (1989) In Glover, D.M. (ed.) *DNA Cloning - A Practical Approach*. IRL Press, Oxford, Vol. I, pp. 19–44.
16. Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M. Lydiate, D.J., Smith, C.P., Ward, J.M. and Schrepf, H. (1985). *Genetic manipulation of Streptomyces, A laboratory manual*. The John Innes Foundation, Norwich.
17. Devereux, J., Haerberli, P. and Smithies, O. (1984). *Nucl. Acids Res.* **12**, 387–395.
18. Boccard, F., Smokvina, T., Pernodet, J-L., Friedmann, A. and Guérineau, M. (1989). *Plasmid* **21**, 59–70.
19. Boccard, F., Smokvina, T., Pernodet, J-L., Friedmann, A. and Guérineau, M. (1989). *EMBO J.* **8**, 973–980.
20. Ptashne, M. (1991) *A Genetic Switch*, 2nd edition. Cell & Blackwell Scientific Publications, Cambridge, USA.
21. Poteete, A.R. (1988) In Calender, R. (ed.) *The Bacteriophages*, Plenum Press, New York, Vol 2. pp. 647–682.
22. Bertani, E. and Six, E.W. (1988) In Calender, R. (ed.) *The Bacteriophages*, Plenum Press, New York, Vol 2. pp. 73–144.
23. Hershey, R.M. (1988) In Calender, R. (ed.) *The Bacteriophages*, Plenum Press, New York, Vol 1. pp. 193–234.
24. Yarmolinsky, M.B. and Sternberg, N. (1988) In Calender, R. (ed.) *The Bacteriophages*, Plenum Press, New York, Vol 1. pp 291–438.
25. Ingham, C.J., Crombie, H.J., Bruton, C.J. Chater, K.F., Hartley, N.M., Murphy, G.J.P., and Smith, M.C.M. (1993) *Mol. Microbiol.* **9**, 1267–1274.
26. Lobell, R.B. and Schlieff, R.F. (1990) *Science* **250**, 528–532.
27. Vasil'chenko, L.G., Mkrtumian, N.M. and Lomovskaya, N.D. (1981). *Genetika* **17**, 1967–1974.
28. Laity, C., Chater, K.F., Lewis C.G. and Buttner, M.J. (1993). *Mol. Micro.* **7**, 329–336.
29. Ward, J.M., Janssen, G.R., Kieser, T., Buttner, M.J. and Bibb, M.J. (1986). *Mol. Gen. Genet.* **203**, 468–478.
30. Messing, J. (1983) *Meth. Enzymol.* **101**, 20–79.
31. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103–119.