Analysis of the ColE1 stability determinant Rcd

Michaela E. Sharpe,† Heather M. Chatwin, Cindy Macpherson, Helen L. Withers‡ and David K. Summers

Author for correspondence: David K. Summers. Tel: +44 1223 333991. Fax: +44 1223 333992. e-mail: d.summers@gen.cam.ac.uk

Department of Genetics, Downing Street, Cambridge CB2 3EH, UK

Multimer formation is an important cause of instability for many multicopy plasmids. Plasmid ColE1 is maintained stably because multimers are converted to monomers by Xer-mediated site-specific recombination at the cer site. However, multimer resolution is not the whole story; inactivation of a promoter (Pcer) within cer causes plasmid instability even though recombination is unaffected. The promoter directs the synthesis of a short transcript (Rcd) which is proposed to delay the division of multimer-containing cells. Mapping of the 5' terminus of Rcd confirms that transcription initiates from Pcer. The 3' terminus shows considerable heterogeneity, consistent with a primary transcript of 95 nt being degraded via intermediates of 79 and 70 nt. Secondary structure predictions for Rcd are presented. Of four mutations which abolish Rcd-mediated growth inhibition, one reduces the activity of Pcer while the other three map to the rcd coding sequence and reduce the steady-state level of the transcript. RNA folding analysis suggests that these three mutant transcripts adopt a common secondary structure in which the major stem-loop differs from that of wild-type Rcd. A survey of 24 cer-like multimer resolution sites revealed six which contain Pcer-like sequences. The putative transcripts from these sites have similar predicted secondary structures to Rcd and contain a highly conserved 15 base sequence. To test the hypothesis that Rcd acts as an anti-sense RNA, interacting with its target gene(s) through the 15 nt sequence, we used DNA hybridization and sequence analysis to find matches to this sequence in the Escherichia coli chromosome. Our failure to find plausible anti-sense targets has led to the suggestion that Rcd may interact directly with a protein target.

Keywords: ColE1, dimer resolution, cell cycle, anti-sense RNA, plasmid stability

INTRODUCTION

Multimer formation and associated copy number depression are acknowledged causes of cloning vector instability (Summers et al., 1993; Summers, 1998). Multimers of the naturally occurring plasmid ColE1 are resolved to monomers by site-specific recombination between directly repeated cer sites (Summers & Sherratt, 1984). In addition to the recombination site, ColE1 multimer resolution requires four chromosome-encoded proteins: ArgR (Stirling et al., 1988), PepA (Stirling et al., 1989), XerC (Colloms et al., 1990) and XerD (Blakely et al., 1993). However, multimer resolution alone is insufficient to ensure stable maintenance of ColE1-like plasmids. A promoter (Pcer) which is located centrally within the cer site (Summers & Sherratt, 1988) is also required and mutational inactivation of this promoter results in plasmid instability, even though recombination at cer is unaffected (Patient & Summers, 1993).

Pcer directs the synthesis of Rcd (regulator of cell division); a short, untranslated transcript. The growth of cells overexpressing Rcd from an up-promoter derivative of Pcer or from aPpe is severely retarded on solid media and microscopic examination suggests that there is a ‘bottle-neck’ in their cell cycle between nucleoid segregation and septation (Patient & Summers, 1993). To explain the role of Rcd in plasmid stability it has been proposed that its expression from Pcer in response to plasmid multimerization delays cell division until multimer resolution is complete (Patient &
Summers, 1993). The need to delay cell division implies that Xer–cer recombination is relatively slow and an increasing amount of circumstantial evidence supports this view (Summers, 1998).

The mechanism of action of Rcd is unknown but some clues have come recently from studies of Rcd overexpression in broth culture. Rcd overexpression slows growth of wild-type cells in broth but the culture eventually reaches stationary phase at a normal cell density. In contrast, Rcd induction in hns mutant cells blocks cell division and growth ceases completely after 2–3 h (Summers & Rowe, 1997; Rowe & Summers, 1999). The nucleoid in these so-called quiescent cells is highly condensed, but they remain capable of transcription and translation and are capable of expressing high levels of plasmid-encoded protein. The continuing metabolic activity of quiescent cells in the absence of growth suggests that Rcd inhibits cell division directly, rather than via a block on RNA or protein synthesis. We have pursued the idea that it might act as an anti-sense RNA, achieving its effect on growth by inhibiting mRNA translation. Two examples of trans-acting anti-sense RNAs in Escherichia coli are the 53 nt DicF (Tétart & Bouche, 1992), which regulates ftsZ mRNA, and the 93 nt MicF (Andersen & Delilas, 1990), which inhibits translation of ompF mRNA. In both cases part of the anti-sense transcript displays complementarity to the region of the target mRNA containing the ribosome-binding site. In DicF, a region of 31 nt shows 81% complementarity to ftsZ mRNA immediately upstream of the start codon. In MicF a region of 54 nt shows 80% complementarity to the corresponding region of ompF mRNA.

Searching for the target of Rcd has been difficult in the absence of precise knowledge of the 5’ and 3’ termini, or any idea of which part of the transcript is involved in the target interaction. In this report we describe the mapping of the 5’ and 3’ ends of Rcd which has facilitated prediction of its secondary structure. We have employed a bioinformatic approach to identify the active site of Rcd and have looked for complementarity to the presumed active site by hybridization and by searching the E. coli genome sequence. Our failure to find a target by either of these approaches has led us to speculate that Rcd might interact directly with a protein target.

**METHODS**

**Bacteria and plasmids.** All the bacteria were used as derivatives of E. coli strain AB1157 (Bachmann, 1972). DS941 is AB1157 recF lacI lacZAM15. JC8679 (AB1157 recBC sbcA) was obtained from Professor D. J. Sherratt, Department of Biochemistry, University of Oxford. Plasmid pKS490 is a pUC8 derivative which has a cer site inserted into the multiple cloning site (Summers & Sherratt, 1988). pKS494 and pKS496 are derivatives of pKS490. In pKS494 the Pcer promoter has been inactivated by mutation of the invariant T and in pKS496 a mutation within Pcer has increased the level of Rcd expression (Patient & Summers, 1993). Bacteriophages M13mp8cer and M13mp19cer are derivatives of M13mp8 and M13mp19 in which the cer site from pKS490 has been inserted as an EcoRI–HindIII fragment into the polylinker of the phage. Plasmids pHW115 and pHW118 are derivatives of the transcription fusion vector pKO1 (McKenny et al., 1981); their construction is described in Results.

**Media and bacterial transformation.** For routine growth of bacteria we used L-broth (Kennedy, 1971) and Oxoid iso-sensitest agar. Where appropriate, ampicillin (50 μg ml−1) was added to the medium. Plasmid transformation of CaCl2-treated cells was by the method of Cohen & Hsu (1972). When transformants were to be grown at 42 °C, all stages of the procedure normally performed at 37 °C were performed at 42 °C.

**Hydroxyamine mutagenesis.** Plasmid DNA (40 μl in sterile distilled water) was mixed with 20 μl 0.1 M potassium phosphate buffer (pH 6.0) containing 5 mM EDTA and prewarmed to 75 °C. Forty microlitres of a freshly prepared solution of 1 M hydroxyamine (0.56 ml 4 M NaOH, 0.35 g hydroxyamine hydrochloride, made up to 5 ml in water) was pre-heated for 5 min at 75 °C and the two solutions were mixed and incubated at 75 °C for 60 min before induction with 100 μl cold 2-propanol (500 μl) were added and the mixture was placed on ice for 10 min. After centrifugation (13000 r.p.m. for 10 min) the supernatant was discarded and the DNA pellet washed twice with 70% ethanol. After drying the pellet was dissolved in 20 μl TE.

**Isolation of RNA and Northern blotting.** Methods described in Patient & Summers (1993) were used to isolate RNA, to prepare the oligonucleotide probe and to perform Northern blotting.

**3’ and 5’ end-mapping of RNA.** 3’ end-mapping of RNA was performed according to the method of Maniatis et al. (1982). The DNA probe was made by primer extension on M13mp8cer from the −40 forward primer. This produced a labelled strand complementary to Rcd. The double-stranded primer extension products were cleaved with PvuII (CGATCG) which cuts 9 bp downstream of the Rcd transcription start (see Fig. 2). The probe was eluted from a polyacrylamide gel by the ‘crush and soak’ method (Maniatis et al., 1982). For digestion of the probe–RNA hybrid Maniatis et al. (1982) recommend 100–1000 units S1 nuclease ml−1 but to ensure complete digestion of single-stranded regions we explored a range of 1000–3500 units S1 nuclease ml−1 at an incubation temperature of 40 °C. Our results indicated that under our conditions digestion was complete at above 2000 units S1 nuclease ml−1. Digestion products were subjected to electrophoresis on an 8% polyacrylamide/7 M urea gel beside a standard dideoxy sequencing reaction using M13mp19cer as template and a sequencing primer whose 5’ end was exactly complementary to the 3’ end of our S1 mapping probe (5’-ATCGCGCCAGTTTTTCG-3’).

To map the 5’ end of Rcd, reverse transcriptase was used to synthesize DNA from a primer complementary to an internal region of the transcript. The DNA strand terminates at the 5’ end of the transcript. Labelled oligonucleotide primer (5–10 ng) was annealed to 5 ng total cellular RNA in the presence of 1 M KCl (1 μl). The mixture was heated to 90 °C for 5 min and cooled slowly to 42 °C over a 2 h period (all subsequent procedures were performed at 42 °C to minimize the problems of secondary structure). Four microlitres of 5 × reverse transcriptase buffer (250 mM Tris/HCl, pH 8.3, 375 mM KCl, 50 mM DTT, 15 mM MgCl2) and 1 μl distilled water were added to the annealed mixture and incubated at 42 °C for 30 min. Primer extension was performed by adding 4 μl dNTP mixture (2.5 mM each dNTP adjusted to pH 7.5) and 200 units M-MLV reverse transcriptase and continuing the incubation
for 2 h at 42 °C. After 1 h a further 200 units reverse transcriptase was added. Twenty microlitres of STOP mix (95 % formamide, 20 mM EDTA, 0.05 % bromophenol blue, 0.05 % xylene cyanol) was added to terminate the reaction. An ATCG ladder was obtained with the same primer oligonucleotide complementary to the 15 base conserved region of Rcd was transcribed with the same primer oligonucleotide using plasmid DNA as the template for sequencing. The products of both the primer extension analysis and sequencing reactions were denatured prior to loading on a sequencing gel. Autoradiography was at -70 °C.

**Probing the E. coli Gene Mapping Membrane.** In a direct approach to identify the target of Rcd, an oligonucleotide complementary to the 15 base conserved region of Rcd was end-labelled with [γ-32P]ATP and hybridized to the Gene Mapping Membrane (Noda et al., 1991). The hybridization was carried out under conditions of very low stringency (i.e. at 40 °C with the melting temperature of probe estimated to be 48 °C) in a buffer consisting of 6 × SSC (0.9 M sodium chloride, 0.09 M sodium citrate) and 0.5 % SDS without blocking agents. Following hybridization, the membrane was given a 5 min rinse in 6 × SSC, 0.1 % SDS before exposing to X-ray film at -70 °C.

**RESULTS**

5' end-mapping of Rcd

Rcd appears to be produced from the P_{rer} promoter since mutation of the invariant T in the −10 sequence abolishes Rcd expression (Patient & Summers, 1993). To establish the transcription start site we used primer extension by reverse transcriptase to map the 5' ends of Rcd expressed from plasmids pKS490 and pKS496. The cer site in pKS490 contains the wild-type P_{rer} promoter and rcd sequence; pKS496 is identical except for an up-promoter mutation in P_{rer} (Fig. 1a; Patient & Summers, 1993). Since P_{rer} is activated by plasmid multimerization, total RNA was prepared from pKS490 or pKS496 in JC8679; a recBC s6cA strain in which both plasmids are extensively multimeric. The primer for reverse transcriptase was an oligonucleotide complementary to a sequence approximately 30 nt downstream of P_{rer} (Fig. 2). Primer extensions on RNA from strains carrying either plasmid yielded major products of 25, 26 and 27 nt. The result for JC8679(pKS496) is shown in Fig. 1(b). The intensities of the bands suggest that three start points are used at similar frequencies. The corresponding 5’ ends of Rcd were identified by comparing the primer extension products with the products of a DNA sequencing reaction using the same primer and are marked with asterisks in Fig. 1(a).

3' end-mapping of Rcd

In an earlier study (Patient & Summers, 1993) it was proposed that rcd transcription terminates at a sequence resembling a p-independent terminator 70–80 bases downstream of the transcription start. To establish the termination site more directly, S1 mapping of the Rcd 3' end was performed on total RNA from JC8679(pKS490) (wild-type P_{rer}). We used a radiolabelled ssDNA probe complementary to Rcd, whose 3' end was 9 nt downstream of the transcription start. We chose a probe which does not cover the 5' end of Rcd so that variability of the transcription start site did not influence the result of the 3' end-mapping. S1 nuclease digestion of the Rcd-probe hybrid should yield a labelled DNA fragment whose 5' end is complementary to the 3' end of the Rcd transcript. The products of digestion were run on a sequencing gel alongside a cer sequence primed from an oligonucleotide whose 5' end was exactly complementary to the 3' end of the DNA probe. The 3' ends of Rcd could thus be inferred directly from the products of the sequencing reaction. The result of the analysis was complex but highly reproducible over eight independent

![Fig. 1. (a) Sequences of the P_{rer} promoter and its derivatives. The −10 and −35 sequences are underlined and the binding sites of ArgR (Arg boxes) are in bold type. Transcription start sites are indicated by asterisks and the altered bases in the up- and down-promoter derivatives are enlarged. (b) 5' end-mapping of Rcd transcribed from the P_{rer} up-promoter derivative in pKS496. Equivalent results were obtained with the wild-type promoter (data not shown).](image-url)
Predicting the secondary structure of Rcd

The transcript mapping results show heterogeneity at both the 3' and 5' ends and suggest that cells expressing Rcd contain a family of transcripts between 68 and 95 nt. This is consistent with the previous observation that Northern analysis of Rcd on low resolution gels gives a smear of approx. 70-90 nt (Patient & Summers, 1993; but note that in Fig. 3 of this publication the size markers were misaligned). We used MFOLD from the GCG package to predict the secondary structures of three representative Rcd transcripts (Fig. 4). The input sequences started at the first base of the AGG initiation triplet and terminated at bases 95 (Rcd95), 79 (Rcd79) and 70 (Rcd70), which correspond to the slowest-running band in each of the three clusters of 3' ends identified by S1 mapping (Fig. 3). The predicted structures for Rcd70 and Rcd79 contain a single stem-loop while an additional minor stem-loop appears in Rcd95. The large GC-rich stem-loop is identical in all three structures and is interrupted by an internal loop and an asymmetrical bulge. MFOLD offered alternative structures of very similar energy but with a loop of 13 nt (see, for example, Fig. 4 of Patient & Summers, 1993). Such a large loop seemed less plausible than the 5 nt loop structure which is close to the optimum for regulator-target interactions (Hjal & Wagner, 1992). However, final confirmation of Rcd secondary structure must await physical mapping of the transcript.

Characterization of mutations which reduce Rcd effectiveness

In an attempt to learn more about the relationship between Rcd structure and function, we screened for mutations in the transcript which reduce its effectiveness

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Rcd95</th>
<th>Rcd79</th>
<th>Rcd70</th>
</tr>
</thead>
<tbody>
<tr>
<td>T (mut-38)</td>
<td>GTGGATTAAGTGGGGGGTAATGTGGCCTGTACCCTCTGGTTATAGGTATTCATACGGTTTTTATC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T (up-promoter mutation)</td>
<td>*** Thal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (mut26)</td>
<td>+1 +10 +20 +30 +40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (mut49)</td>
<td>+70</td>
<td>###*</td>
<td></td>
</tr>
<tr>
<td>A (mut46)</td>
<td>+80 +90 +100 +110 +120 +130 +140</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Part of ColE1 cer showing Pcer (−10 and −35 sequences in bold type), the Rcd coding region and the XerC and XerD binding sites. Asterisks indicate transcription start sites (**). The results of 3' end-mapping of Rcd (see Fig. 3) are summarized as follows: *, a strong signal; ∗, weak signal. Mutations which abolish Rcd-mediated growth inhibition are shown above or below the main sequence, followed by the mutant number in parentheses. The Thal site immediately downstream of Pcer, used for cloning promoter fragments into pKO1 is shown and the sequence complementary to the oligonucleotide used for 5' end-mapping of Rcd is underlined.

Fig. 3. 3' end-mapping of Rcd. Lanes 1-4 are the results of four independent mapping experiments. Rcd95, Rcd79 and Rcd70 are the largest members of each of the three clusters of 3' ends.

experiments; typical results are shown in Fig. 3. We observed one major cluster of bands (Rcd95) and two minor clusters (Rcd79 and Rcd70). At least half of the 3' termini appear to be within the Rcd95 cluster. The distribution of Rcd 3' termini is summarized in Fig. 2.
Rcd transcript of ColEl

Fig. 4. Computer-predicted secondary structures for Rcd. The structures have a common 5' end and 3' termini at nt 95 (Rcd95), 79 (Rcd79) or 70 (Rcd70). The 15 nt sequence conserved among six hypothetical Rcd-like transcripts is indicated by a thick line.

as a growth inhibitor. Plasmid pKS496 overexpresses Rcd because of an up-promoter mutation in Pcer and consequently DS941(pKS496) can only form minute colonies on L-agar after overnight incubation at 42 °C (Patient & Summers, 1993). Hydroxylamine-muta-
genized pKS496 was transformed into DS941 and the selection plates were incubated overnight at 42 °C when a small number of normal-sized colonies appeared. Plasmid DNA from these colonies was transformed into DS941 and the normal growth of transformed colonies at 42 °C confirmed that a plasmid mutation was responsible for the loss of Rcd-mediated growth inhibition. Four mutant plasmids were sequenced and a single base change (in addition to the Pcer up-promoter mutation) was identified in each (Fig. 2). They were all within cer; three (mut26, mut46 and mut49) within rcd and one (mut-38) upstream of Pcer. Each mutation is identified by a number which refers to its position relative to the first base of the AGG trinucleotide where transcription initiates. The cer sites in all of the mutant plasmids remained proficient in dimer resolution and measurement of plasmid copy number confirmed that the loss of growth inhibition was not due to a reduction in the number of Rcd-producing plasmids (data not shown).

All three mutations in the rcd coding region alter bases in the upper part of the major stem-loop of the predicted structure. mut26 and mut46 change bases which are paired in the wild-type structure. Hypothetical structures for the three mutant Rcd transcripts were generated using MFOLD (Fig. 5a). Despite the differences in their primary sequences, the predicted structures for the mutant transcripts are identical. Compared to the wild-type structure prediction, the upper part of the major stem is shorter and the bulge and internal loop are replaced by a single large internal loop. The loss of Rcd-mediated growth inhibition in these three mutants could result from instability of the transcript or a reduction in the effectiveness of its target interaction. Instability would result in a reduction of the steady-state level of Rcd transcript, so we compared levels of Rcd in total RNA from JC8679(pKS496) and each of the mutant plasmids. Equal amounts of total RNA were run on a low resolution gel and a Northern blot was probed with the
Table 1. Galactokinase activity expressed from Pcer, fused to galK

<table>
<thead>
<tr>
<th>Strain</th>
<th>Galactokinase activity a (mean ± SD)</th>
<th>[no. of replicas]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS941(pKS463)</td>
<td>61 ± 24 [31]</td>
<td></td>
</tr>
<tr>
<td>DS941argR(pKS463)</td>
<td>46 ± 14 [14]</td>
<td></td>
</tr>
<tr>
<td>DS941(pHW118)</td>
<td>158 ± 63 [12]</td>
<td></td>
</tr>
<tr>
<td>DS941argR(pHW115)</td>
<td>411 ± 91 [12]</td>
<td></td>
</tr>
<tr>
<td>DS941argR(pHW118)</td>
<td>208 ± 122 [8]</td>
<td></td>
</tr>
</tbody>
</table>

a Galactokinase units as defined by McKenny et al. (1981).

oligonucleotide used for 5' end-mapping of Rcd. A strong Rcd band was present in the pKS496 track but was barely detectable in RNA from cultures containing the mutant plasmids (Fig. 5b).

Mut-38 reduces the activity of Pcer

The fourth mutation (mut—38) was a C to T change upstream of Pcer at ColE1 co-ordinate 3790 (Fig. 2). Northern analysis of total RNA showed a dramatic reduction in the steady-state level of Rcd (Fig. 5b). It has been known for some time that sequences upstream of Pcer influence transcription (Summers & Sherratt, 1988) so it seemed plausible that mut — 38 down-regulates Rcd expression. To test this, transcription fusion plasmids were constructed containing Pcer from pKS496 and its mut—38 derivative upstream of a promoterless galK gene in pKO1 (pHW115 and pHW118, respectively). The inserts in these fusion plasmids contain the whole of the cer site upstream of Pcer, and stop at the Thal site immediately downstream of the transcription start (Fig. 2). In DS941, pHW118 showed a lower level of transcription than pHW115 (158 ± 63 and 242 ± 65 galactokinase units, respectively). Although the reduction in transcription detected in the gene fusion experiment is small compared with the much greater decline in Rcd concentration seen by Northern blotting, it seems likely that reduced transcription from the mut—38 promoter is responsible for the lifting of growth inhibition.

ArgR is a transcriptional repressor of genes for arginine biosynthesis and an essential accessory factor in ColE1 dimer resolution. The ArgR binding site (the Arg box) overlaps the —35 consensus of Pcer (Fig. 1a) and, noticing the proximity of the mut—38 mutation to the Arg box, we wondered whether it might exert its effect by altering the binding of ArgR to Pcer. Experiments with pKS463 (Summers & Sherratt, 1988) in which wild-type Pcer is fused to galK (Table 1) showed that the presence or absence of ArgR had little effect. To investigate the effect of ArgR on the mutant promoters, transcription from Pcer in pHW118 and pHW115 was compared in DS941 and DS941argR. The cer sites in both of these plasmids contain the Pcer up-promoter mutation and in addition pHW118 contains mut—38. The absence of ArgR increased promoter activity for both plasmids (Table 1) but transcription from pHW118 was still lower than from pHW115. Thus it seems unlikely that the effect of mut—38 is due simply to altered binding of ArgR to the promoter.

rcd-like genes in other multimer resolution sites

In an attempt to identify the active site of Rcd (i.e. that part of the molecule which interacts with the target) we have investigated Rcd analogues among 24 cer-like multimer resolution sites available on DNA sequence databases. Although functional Rcd analogues might diverge at the primary sequence level, we would expect to see conservation of secondary structure. Furthermore, there should be conservation of primary sequence within the active site. As a first step we searched the 24 multimer resolution sites for Pcer-like promoters. In the region corresponding to the —10 sequence of Pcer only six of the sites (those from plasmids ColE1, pWQ799, pRI13, ColK, CoLA and CloDF13) retained the invariant T (Fig. 6). Mutation of this base abolishes transcription from Pcer (Patient & Summers, 1993) so the remaining 18 sites seem unlikely to produce a transcript (ColN and pNTP16 are included in Fig. 6 as representatives of these sites). The six T-containing sites show complete homology to Pcer within their —10 sequences (TAAAAT) and strong conservation in the —35 sequence (NTGcAT). There is more sequence variation among the spacer regions but all are 15 nt long; a characteristic feature of the Pcer promoter. Assuming the same 5' and 3' ends as Rcd, the putative transcribed regions in the six sites show considerable sequence variation but it is notable that one 15 nt block (5'-CGGGTGGTTTG-TTGC-3') is totally conserved, except in CloDF13 where it differs at one position. There is far less conservation of this sequence among the remaining 18 sites (exemplified by ColN and pNTP16 in Fig. 6).

In the Rcd95 secondary structure prediction the conserved 15 nt block lies within the major stem–loop, comprising the loop and the 5' side of the stem (Fig. 4). For comparison, MFOLD was used to predict secondary structures for the hypothetical transcripts from CoLA, ColK and CloDF13. Despite considerable primary sequence divergence (ColE1 and CloDF13 show only 63% identity in Rcd95) the secondary structures are similar, each displaying a major and a minor stem–loop (Fig. 7). The lower parts of the major stems vary in length from 14 to 17 bp but are invariably GC-rich. Each major stem contains two regions where base pairing is disrupted; a bulge with unpaired bases on one strand and an internal loop involving unpaired bases on both strands. In all four structures the conserved 15 nt block begins at the internal loop on the 5' side of the stem, extending through the terminal loop and ending at the bulge on the 3' side of the stem.
Red transcript of CoE1

Identification of possible Rcd targets

By virtue of its sequence conservation and position in the predicted secondary structures, the 15 nt block has the features we would expect for the active site of Rcd. If Rcd regulates translation in a manner similar to MicF and DicF, we would expect the 15 nt sequence to be complementary to the 5' end of the target transcript in the vicinity of the ribosome-binding site. The regions of MicF and DicF which bind their targets contain blocks of C+U-rich sequence which bind to the G+A-rich ribosome-binding site, but no such sequence is apparent in the 15 nt conserved block in Rcd. However this does not exclude the possibility of translational regulation since it is possible that Rcd might bind adjacent to (rather than overlapping) the ribosome site. We therefore probed an E. coli chromosome library and searched the E. coli genome sequence in an attempt to identify regions complementary to the conserved sequence. A radiolabelled DNA probe matching the 15 nt conserved sequence was hybridized to the E. coli Gene Mapping Membrane (Noda et al., 1991). The membrane consists of an ordered array of overlapping phage λ clones containing E. coli DNA from the Kohara library, allowing the map position of any clone to be determined rapidly. The membrane was probed on three separate occasions under conditions of very low stringency but no specific hybridization was detected.

We searched the E. coli genome sequence for matches to the conserved 15 nt block but no chromosomal locus showed full homology. A contiguous 14 base match was found in the sense strand of gltS at 82.4 min, although a sense strand match is inconsistent with antisense regulation of translation. Thirteen base matches were found in the antisense strand of hns which maps at 34.2 min. Preliminary resolution sites in various multicopy plasmids. Asterisks (*** indications) indicate the Rcd transcription start site.

Fig. 6. Alignment of regions corresponding to Pcer and rcd from CoE1 and multimer resolution sites in various multicopy plasmids. Asterisks (*** indications) indicate the Rcd transcription start site.

At 35.4 min the 13 base match to the Rcd conserved region lies between two converging ORFs, an unlikely situation for the action of an anti-sense RNA. Closer inspection of the chromosome sequence revealed that the match extends beyond rcd and includes the Arg box and XerC-XerD binding site of cer, suggesting that we had identified a cer-like site in the chromosome terminus region (Fig. 8b). This site is distinct from dif (Kuempel et al., 1991) which maps at 34.2 min. Preliminary experiments in this laboratory (M. Martin, H. Griffiths & D. K. Summers, unpublished) demonstrate that this site supports XerCD-mediated recombination when located on a plasmid.
DISCUSSION

The structure of Rcd

Patient & Summers (1993) proposed a structure for Rcd based upon the assumptions that transcription starts from the P\textsubscript{rer} promoter and stops at a putative \(\phi\)-independent terminator (Fig. 6) approximately 75 bases downstream. 5' end-mapping of Rcd has confirmed that transcription starts at P\textsubscript{rer}, but S1 mapping revealed a population of Rcd molecules with heterogeneous 3' termini. The simplest interpretation of the data is that Rcd95 is the primary transcription product and the Rcd79 and Rcd70 clusters represent relatively stable intermediates during Rcd turnover. Heterogeneity within the three clusters may reflect exonuclease action at their 3' ends or even polyadenylation of the transcripts. Termination of Rcd95 must occur beyond the putative terminator which, in any case, is not conserved among the six related multimer resolution sites which contain P\textsubscript{rer}-like promoters. Interestingly, there is no terminator-like sequence corresponding to the 3' end of Rcd95 but termination occurs within the XerCD binding site (Fig. 2) and might involve the interaction of RNA polymerase with the recombinase heterodimer bound to the DNA. Alternatively, termination may occur at a downstream terminator followed by very rapid processing of the 3' end.

A bioinformatic analysis identified five possible Rcd analogues in cer-like multimer resolution sites. We are optimistic about the validity of this exercise for two reasons. First, all of the sites containing a P\textsubscript{rer}-like promoter sequence also contained a conserved 15 base sequence within the rcd coding region. This sequence was not conserved among promoterless sites. Assuming that the transcripts have a common function, the 15 nt sequence is a strong candidate for the transcript domain which makes the primary interaction with the target. Second, it is notable that secondary structure is largely conserved among the putative transcripts, despite variation in primary sequence. For example, nt 5-18 of Rcd and the corresponding region of the hypothetical CloDF13 transcript both form the lower part of the major stem. The ColE1 and CloDF13 sequences differ in four positions but in each case complementary changes in the other half of the stem ensure that base pairing is not lost. This evidence for concerted evolution of the stem sequence supports the idea that the secondary structure of Rcd and its analogues is of functional significance.

In view of the conservation of secondary structure among Rcd and its putative analogues it is interesting that the three mutations which abolish growth inhibition map to the upper stem of Rcd and computer predictions suggest that they change its structure in identical ways. Northern blot analysis of total cellular RNA from cultures bearing the mutant plasmids showed that the steady-state concentrations of these three mutant transcripts were extremely low. The steady-state concentration is determined by the rate of transcription and the half-life of the transcript but since the mutant transcripts have unaltered promoters (and therefore, presumably, unaltered rates of transcription initiation) their lower concentration must be due to more rapid turnover. Turnover rates are determined by the thermodynamic stability of the RNA, particularly the stem domain (Case \textit{et al.}, 1990), and susceptibility to exoribonuclease and endoribonuclease attack (Belasco \textit{et al.}, 1986). Since the predicted structures for the mutant transcripts are identical, this structure may be particularly susceptible to attack by cellular RNase. It is possible that wild-type Rcd may adopt this structure during turnover (e.g. as a result of interaction with an RNA-binding protein).

What is the target of Rcd?

The structure of Rcd is consistent with a role as a trans-acting anti-sense RNA. In a previous study the observation of 64% sequence complementarity between Rcd and the leucine attenuator led to the tentative suggestion that Rcd might exert its effect through modulation of intracellular leucine concentrations (Patient & Summers, 1993). This now seems implausible because strain DS941 used in this study is incapable of leucine biosynthesis (it carries the \textit{leuB6} mutation) and
yet remains Rcd-sensitive. The search for alternative potential targets has been hampered by the absence of knowledge of the extent of complementarity between Rcd and its target, or which part of the transcript is involved in the interaction. However, the discovery that a 15 nt sequence within Rcd is conserved among its putative analogues suggested that this might be the active site of the transcript. Further support for this hypothesis came from the location of the sequence in the loop and upper stem of Rcd, which is the region of antisense RNAs which most often makes the primary interaction with the target.

Database searching identified three loci in the *E. coli* chromosome which matched at least 13 of the 15 conserved bases. Among these, *hns* mRNA looked the most plausible target. This was particularly interesting because *hns* mutants have been shown to have increased sensitivity to Rcd (Summers & Rowe, 1997; Rowe & Summers, 1999). However, the hypothesis that *hns* mRNA might be a target of Rcd was ruled out by experiments which showed the transcript had no effect on *hns-lacZ* transcriptional and translational fusions. Neither of the two remaining chromosomal loci provided a credible mRNA target for Rcd so its mechanism of action remains ill-defined. Our failure to find a functional target could be due to differences between the AB1157 derivatives used in this work and the *E. coli* genome project strain MG1655 (Blattner et al., 1997). However, MG1655 is Rcd-sensitive (our unpublished data) so it seems unlikely that the strain lacks the Rcd target.

If the 15 nt conserved region is indeed the target interaction domain of Rcd our results imply that Rcd is not an antisense RNA. An alternative mechanism of action is a direct interaction between Rcd and a protein involved in cell division or its regulation.

### A chromosomal analogue of Rcd

The discovery of a good match to the Rcd conserved sequence within a cer-like site near the replication terminus (35.4 min) raises the intriguing possibility that...
the chromosome encodes an Rcd analogue with a role in cell cycle control. There is poor conservation of the \( P_{cer}\) promoter between \( cer \) and the 35.4 min site but one would not expect plasmid and chromosomal versions of Rcd to be under identical transcriptional control. Certainly the region of the 35.4 min site corresponding to Rcd can be folded into an Rcd-like structure (data not shown) but we do not yet have any evidence whether a transcript is expressed or whether, if expressed, it would inhibit growth or cell division. Indeed, we cannot exclude the possibility that the 35.4 min site is the result of a chance plasmid integration event and is of no functional significance. However, if the chromosome does indeed encode an Rcd-like transcript with a role in cell cycle regulation it would explain the ease with which ColEl seems able to control the cell cycle of its host. In retrospect, our failure to detect this locus when we used a conserved region probe against the gene mapping membrane is surprising. However, it is possible that the Kohara library (used to produce the membrane) is incomplete or that the chromosomal cer-like site is not universal and is absent from the Kohara source strain.

ACKNOWLEDGEMENTS

This work was supported by research studentships to M. E. S. and C. M. from the Medical Research Council and to H. L. W. from BBSRC. H. C. was supported by an MRC project grant to D. K. S. We thank Ian Viney, Charlie Hodgman and Marion Martin for their assistance with various aspects of the work.

REFERENCES


DiC antisense RNA. Division does not require a fixed number of FtsZ molecules. Mol Microbiol 6, 615–620.