Identification of Operators and Promoters That Control SXT Conjugative Transfer

John W. Beaber and Matthew K. Waldor*

Department of Microbiology, Tufts University School of Medicine, and Howard Hughes Medical Institute, Boston, Massachusetts

Received 1 April 2004/Accepted 24 May 2004

Transfer of SXT, a Vibrio cholerae-derived integrating conjugative element that encodes multiple antibiotic resistance genes, is repressed by SetR, a λ-cr-related repressor. Here we identify divergent promoters between setR and setCD that drive expression of the regulators of SXT transfer. One transcript encodes the activators of transfer, setR and setD. The second transcript codes for SetR and, like the cr transcript of lambda, is leaderless. SetR binds to four operators located between setR and s086; the locations and relative affinities of these sites suggest a model for regulation of SXT transfer.

SXT is a 100-kb integrating conjugative element (ICE) derived from Vibrio cholerae. SXT encodes resistance to multiple antibiotics (27), and in the past decade, this element or closely related elements have become widespread among V. cholerae clinical isolates in Asia (1, 13) and Africa (10). SXT is part of a larger family of ICEs previously referred to as IncI elements that includes R391 (9), R997 (12), and pMERPH (17). These ICEs all contain a common set of genes coding for their conjugative transfer, integration and excision, and regulation (2–6).

We previously identified several SXT genes involved in controlling expression of the SXT integrase, int, and conjugation-associated loci (3, 4). Two of these genes, setC and setD, are orthologues of the flagellar activators flhC and flhD, and their products activate transcription of int and conjugation-associated operons (3). setC and setD expression is repressed by SetR, a protein similar to the bacteriophage lambda repressor, crI (4). Repression of setC and setD is alleviated by induction of the SOS response, the bacterial response to DNA damage. We observed a marked increase in SXT transfer when donor cells were grown in the presence of agents such as mitomycin C and ciprofloxacin that induce SOS, suggesting that the SOS response stimulates SXT transmission (4).

SetR represses expression from a promoter upstream of s086 (Fig. 1A), which we hypothesize is part of an operon that includes setC and setD. SetR repression of this promoter appears to be direct, since SetR binds to this region of DNA at multiple sites (4). setR is divergently transcribed from s086. In this report, we characterize the setR and s086 promoters and define the SetR operators located in the region between these genes. Interestingly, like the mRNA encoding lambda crI, we found that the mRNA encoding SetR is leaderless. We identified four SetR operators between s086 and setR, and their locations suggest a model for SetR control of s086 and setR transcription.

Identification of the s086 and setR promoters. The strains and plasmids used in this study are detailed in Table 1. The sequences of the primers used are given in Table 2.

The arrangement of genes at the 3′ end of the integrated SXT suggested that there were two divergent promoters in the intragenic region between s086 and setR (Fig. 1A). We previously found that P1, which lies upstream of s086, is repressed by SetR (4). Since setCD expression is also repressed by SetR (4), and since the genes from s086 through s079 (which includes setCD) are predicted to form an operon (3), we used reverse transcription-PCR (RT-PCR) to assess whether these gene products are encoded within a single transcript. Using primers in setC and s086, we amplified a 3.1-kb product (Fig. 1B), which is consistent with cotranscription of setCD and s086 initiating at P1. This result suggests that SetR-mediated repression of setCD occurs at P1, but does not rule out the possibility that other promoters for setCD expression lie downstream of P1.

A second promoter, located upstream of setR, was also found to be repressed by SetR. A plasmid containing a transcriptional fusion of the setR promoter to lacZ was introduced into Escherichia coli either lacking SXT or containing SXT or SXT mutant derivatives. In the no-SXT, SXT+, SXT–ΔsetCD, and SXT–ΔsetCD ΔsetR strains, the β-galactosidase activities were 275 ± 3, 193 ± 7, 181 ± 7, and 254 ± 2 Miller units, respectively (mean ± standard deviation from at least three experiments). These strains were all derivatives of BW25113 (11) containing pRRepR. Note that expression from this promoter (designated P2R) was approximately 30% lower in cells containing SXT than in cells lacking SXT. Deletion of setR raised P2R expression to levels comparable to those observed in the strain lacking SXT, implicating SetR as the SXT-encoded repressor of P2R. These results may underestimate the degree of SetR regulation, since the cellular levels of SetR are very low (unpublished observations) and the setR::lacZ fusion is present on a multicopy plasmid. Therefore, there may not be enough SetR present in the cell to fully repress the P2R reporter.

Computer algorithms and 5′ random amplification of cDNA ends (RACE) were used to define the setR and s086 transcription start sites. Software for the identification of bacterial pro-
motors (http://www.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb) identified putative −10 and −35 elements for both P_L and P_R (Fig. 2) (23, 24). A putative Shine-Dalgarno sequence was also identified upstream of s086 (Fig. 2). The results from mapping the 5′ end of these transcripts by 5′RACE experiments exactly matched the bioinformatic predictions of the P_L and P_R promoters (Fig. 2) (data not shown). In this technique, cDNA representing the 5′ end of an mRNA is tailed with terminal deoxynucleotidyl transferase and subsequently amplified by PCR. These products are then sequenced to determine the putative start site of transcription. Unexpectedly, the +1 position for setR was predicted to be 2 bases upstream of the A residue of the setR start codon, suggesting that setR transcripts may not encode a Shine-Dalgarno sequence upstream of the site of translation initiation. Substitution of 3 bases in the predicted −10 region of the

**FIG. 1.** RT-PCR analysis of the s086 transcript. (A) Schematic representation of the open reading frames at the 3′ end of SXT. Thick arrows represent open reading frames. Promoters are designated by bent arrows. The thin arrows indicate the positions of the primers used for RT-PCR. (B) Amplification of a transcript via RT-PCR with primers specific for s086 and setC. cDNA was produced with a primer specific to the 3′ end of setC. Lane 1, sample amplified after RT; lane 2, control PCR on RNA that was not reverse transcribed; lane L, molecular weight markers.

### TABLE 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant information</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW25113</td>
<td>MG1655 lacF' mBB T14 ΔlacZ W316 hsdR514 ΔaraBAD ΔrhaBAD ΔK7878</td>
<td>11</td>
</tr>
<tr>
<td>Jo193</td>
<td>BW25113 SXT'</td>
<td>3</td>
</tr>
<tr>
<td>Jo212</td>
<td>Jo193 ΔlinR</td>
<td>3</td>
</tr>
<tr>
<td>Jo313</td>
<td>Jo212 ΔsetCD</td>
<td>4</td>
</tr>
<tr>
<td>Biot95</td>
<td>Jo313 ΔsetR</td>
<td>4</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pCB182</td>
<td>Transcriptional fusion vector</td>
<td>21</td>
</tr>
<tr>
<td>pPRelR</td>
<td>Intragenic region between s086 and setR cloned into the BamHI and XbaI sites of pCB182</td>
<td>This study</td>
</tr>
<tr>
<td>pPRePL</td>
<td>Intragenic region between setR and s086 cloned into the BamHI and XbaI sites of pCB182</td>
<td>4</td>
</tr>
<tr>
<td>pPReR10</td>
<td>pPRelR containing three substitutions in the setR −10 region</td>
<td>This study</td>
</tr>
<tr>
<td>pRS414</td>
<td>Translational lacZ fusion vector</td>
<td>22</td>
</tr>
<tr>
<td>pRLacAUG</td>
<td>Translational fusion of setR to lacZ</td>
<td>This study</td>
</tr>
<tr>
<td>pRLacCUG</td>
<td>Translational fusion of setR to lacZ with C-to-A substitution in the setR start codon</td>
<td>This study</td>
</tr>
</tbody>
</table>
setR promoter (Fig. 2) reduced β-galactosidase activity of the setR::lacZ transcriptional fusion more than 20-fold (to 10 Miller units), supporting the identification of the setR promoter.

**setR mRNA is leaderless.** The setR promoter fusion vector used to examine setR transcription encodes the lacZ Shine-Dalgarno sequence immediately upstream of the β-galactosidase gene; therefore, it was not useful for analysis of translation of setR mRNA. To confirm that a Shine-Dalgarno sequence was not required for setR translation, a setR-lacZ translational fusion containing the setR promoter and the lacZ promoter and the lacZ coding sequence was used. A segment that expresses from the transcriptional fusion vector: for pRLacAUG the seven codons of SetR fused to the ninth codon of setR sequence was not required for lacZ expression. To examine setR expression, we have used derivatives of pRS414 (22) in BW25113 (11). This demonstrates that the AUG annotated as the start site for SetR translation is required for maximal expression from this construct. None of the other ATG or TTG codons in the coding sequence is preceded by sequences resembling Shine-Dalgarno sequences, further suggesting that the predicted AUG start codon that we mutated is the normal translation start codon.

These findings suggest that SetR translation initiates directly adjacent to the transcription start site and that the setR mRNA lacks a leader sequence. While uncommon, there are several examples of leaderless mRNAs in prokaryotes as well as in eukaryotes and archaea (28). The lambda CI transcript originating from P_RAM is leaderless, as are those of several other λ CI-related repressors (8, 16, 20). The precise mechanism accounting for initiation of translation from leaderless transcripts is not understood, but it is thought that the 70S ribosome, in complex with the fMet-tRNA, is able to bind and translocate to the ORF without dissociating into its 30S and 50S subunits (25). Leaderless transcripts are translated at lower levels than canonical leadered transcripts (26). Presumably this form of posttranscriptional regulation of SetR synthesis helps maintain low levels of SetR within the cell.

**SetR binds to four operators between s086 and setR.** Previous band shift experiments using the region between s086 and setR as a probe revealed that SetR bound to several sites in this region (4). SetR is most similar to CI-related repressors (8, 16, 20). The precise mechanism accounting for initiation of translation from leaderless transcripts is not understood, but it is thought that the 70S ribosome, in complex with the fMet-tRNA, is able to bind and translocate to the ORF without dissociating into its 30S and 50S subunits (25). Leaderless transcripts are translated at lower levels than canonical leadered transcripts (26). Presumably this form of posttranscriptional regulation of SetR synthesis helps maintain low levels of SetR within the cell. The conservation of this mechanism in this family of CI-related repressors suggests that it may be important to ensure maintenance of low repressor levels, thereby facilitating rapid and sensitive responses to changes in cellular conditions. The putative s086 Shine-Dalgarno sequence is shown in gray. Putative −10 and −35 regions are shown in boldface. Bent arrows indicate the putative s086 and setR transcription start sites that were defined by 5′ RACE. For these experiments, an overnight culture of Jo193 cells was diluted 1:100 into fresh Luria-Bertani medium. Cells were then grown to an OD_600 of 0.3, and mitomycin C was added at a concentration of 200 ng/ml for a final hour of growth. mRNA was prepared with the RNeasy Mini kit (Qiagen, Md.). 5′ RACE was performed as described by the manufacturer (Invitrogen). The arrows show the substituted bases in the P_RAM region. Boxed regions show the 14-bp regions with dyad symmetry that likely represent SetR binding sites. The lines above and below the sequence show the regions that were protected from DNase I cleavage and correspond in style to lines depicted in Fig. 3. Lines above the sequence denote regions protected on the top strand, and lines below the sequence denote regions protected on the bottom strand. Although protection was observed, the boundaries of OL could not be resolved on the bottom strand.
Repeats could represent SetR operators. DNase I protection spacer regions (19). We hypothesized that these imperfect tors, each of the four putative SetR sites contains AT-rich containing similar symmetry and found four similar 14-bp se-
s086 and setR. Regions of protection are denoted by vertical lines beside the gel. The vertical line styles correspond to the lines shown in the schematic in Fig. 2. DNase I protection assays were performed as previously described (14). Briefly, end-labeled DNA probes were generated by PCR with one 5’-radiolabeled primer and a second nonradiolabeled primer. Binding reactions were carried out in 40 μl containing 120,000 cpm of the labeled DNA fragment and purified SetR-H₆ under the same conditions as previously described gel shift experiments (4). Binding was carried out for 10 min at 4°C. Samples were then brought to room temperature, 0.4 U of DNase I (Ambion, Austin, Tex.) was added, and the mixture was incubated for 30 s. The G+A sequencing ladders were generated as previously described (15). Dried DNA pellets were resuspended in a formamide loading buffer and loaded on 6% sequencing gels. The gels were dried and exposed to autoradio-
graphic film. Lane 1 in each panel is probe DNA with no added SetR-H₆. SetR-H₆ concentrations were as follows: lane 2, 120 nM; lane 3, 180 nM; lane 4, 210 nM; lane 5, 300 nM; and lane 6, 600 nM. lamdoid phage 434, ct₄³⁴ has been shown to bind to a 14-bp sequence with dyad symmetry (18). We examined the sequence of the region between s086 and setR for possible repeats containing similar symmetry and found four similar 14-bp se-
tions with partial dyad symmetry (Fig. 2). Like 434 opera-
tors, each of the four putative SetR sites contains AT-rich spacer regions (19). We hypothesized that these imperfect repeats could represent SetR operators. DNase I protection assays using a C-terminally His₆-tagged SetR (SetR-H₆) were carried out to experimentally define the SetR operators located between s086 and setR.

At the lowest SetR-H₆ concentrations tested (120 nM), we observed protection of a 24-bp region, designated O1 (Fig. 2 and 3) that overlaps with the −10 region of P₄ (Fig. 2). At higher SetR-H₆ concentrations, two additional protected regions were observed (Fig. 2 and 3). The region denoted by the dotted line is approximately 48 bp in length and, given its size, likely represents two additional SetR operators, designated O2 and O3. The other protected region, shown as the solid line in Fig. 2 and 3 (designated OL), corresponds to sequences down-
stream of P₄. Each of the protected regions includes the 14-bp sequences discussed above, supporting the idea that they represent SetR binding sites. A fifth site that is nearly identical in sequence to the high-affinity site, O1, lies 800 bp downstream of P₄; SetR binding at this site was not measured.

Conclusions. The binding of SetR to its four operators be-
tween the divergently transcribed s086 and setR genes provides a basis for understanding the control of SXT transfer. SetR bound to O1 with the highest apparent affinity. O1 overlaps the −10 element of P₄; therefore, SetR binding to O1 would likely interfere with RNA polymerase binding to P₄ and inhibit trans-
scription of the transcript containing setCD. This could account for the low basal SXT transfer frequency. Presumably, dimin-
ished SetR levels that result from its RecA-stimulated auto-
cleavage during an SOS response would derepress P₄, leading to expression of setDC. While we have not been able to directly observe cleavage of SetR, an SXT element expressing a puta-
tive noncleavable mutant SetR was not induced to transfer in the presence of SOS-inducing stimuli (4). SetR binding to OL may also aid in repression expression from P₃L. Furthermore, lambda family repressors stimulate their own expression, presumably by promoting RNA polymerase binding (7, 19). It is possible that SetR binding to O1, which is approximately 50 bp upstream of P₃L, activates expression from this promoter and thereby promotes SetR production. Binding of SetR to O2 and O3, which overlap the P₃L −35 and −10 elements, likely ac-
counts for the mild SetR autorepression we observed. Thus, SetR binding to O1, O2, and O3 should repress transcription from both P₃L and P₃R as we have observed through lacZ trans-
criptional fusions to both of these promoters. In the absence of induction, we hypothesize that SetR levels are maintained at a low level due to autorepression. SetR levels would be suf-
cient to repress P₄ but low enough to respond quickly to inducing stimuli.

Control of SXT transfer by its repressor binding to regula-
tory sites between divergently transcribed promoters is similar to the control of the lambda lytic/lysogenic switch. CI binds to three operators, O₆₁, O₆₂, and O₆₃, between the divergently transcribed promoters P₆ and P₆M (19). P₆M controls expres-
sion of CI, and P₆ controls expression of cro, a gene whose product blocks CI expression. In lambda lysogens, CI binding to O₆₁ and O₆₂ represses P₆ and activates P₆M. During an SOS response, CI levels are reduced and repression of P₆ is relieved, allowing production of Cro. Cro binds to O₆₃ and thereby blocks CI expression. The shift from CI to Cro production irreversibly “flips a genetic switch,” and λ begins lytic growth (19). We do not know if the switch from P₃R to P₃l expression in SXT is irreversible. Although overexpression of setDC is toxic, it is possible that P₄ activation is transient, allowing for a controlled burst of SetDC production. Resynthesis of SetR could be sufficient to reestablish repression of P₃R; alterna-
tively, other cellular factors may be important in reestablishing repression of P₃R.

We thank B. Davis, H. Kimsey, and S. McLeod for critical reading of the manuscript. We thank the New England Medical Center GRASP Center for preparation of media and SetR-H₆.

This work was supported in part by funds from NIH training grant AI07422 (J.W.B.), NIH grant AI42347, the Howard Hughes Medical
REFERENCES


