The CTXφ repressor RstR binds DNA cooperatively to form tetrameric repressor:operator complexes

Harvey H. Kimsey* and Matthew K. Waldor
Tufts University School of Medicine, Department of Molecular Biology and Microbiology and the Howard Hughes Medical Institute
Boston, MA 02111

*corresponding author, telephone: 617-636-2778, fax: 617-636-2723, email: harveykimsey@yahoo.com

Running title: CTXφ gene regulation
Summary

CTXφ is a filamentous bacteriophage that encodes cholera toxin and integrates into the *Vibrio cholerae* genome to form stable lysogens. In CTXφ lysogens, gene expression originating from the *rstA* phage promoter is repressed by the phage-encoded repressor RstR. The amino-terminal region of RstR contains a helix-turn-helix (HTH) DNA-binding element similar to the HTH of the cI/Cro family of phage repressors, whereas the short C-terminal region is unrelated to the oligomerization domain of cI repressor. Purified his-tagged RstR bound to three extended 50 bp operator sites in the *rstA* promoter region. Each of the RstR footprints exhibited a characteristic staggered pattern of DNase I-accessible regions that suggested RstR binds DNA as a dimer-of-dimers. In gel permeation chromatography and cross-linking experiments, RstR oligomerized to form dimers and tetramers. RstR was shown to be tetrameric when bound to operator DNA by performing mobility shift experiments with mixtures of RstR and a lengthened active variant of RstR. Binding of RstR to the high-affinity *O1* site could be fit to a cooperative model of operator binding in which two RstR dimers associate to form tetrameric RstR:operator complexes. The binding of RstR dimers to left or right halves of *O1* operator DNA was not observed in mobility shift assays. These observations support a model in which protein-protein contacts between neighboring RstR dimers contribute to strong operator binding.
**Introduction**

The molecular mechanisms that regulate bacteriophage lysogeny have been most extensively studied in phage lambda and its close relatives that infect *Escherichia coli* and *Salmonella enterica* serovar Typhimurium (1,2). For these bacteriophages, the bistable switch from lysogenic to lytic development involves the interplay of two antagonistic transcriptional repressors, CI and Cro in the case of λ, that bind to the same set of regulatory sites in the bacteriophage control regions. A similar regulatory network controls lysogeny in Ø80 (3), HK022 (4), and the unrelated bacteriophage P2 (5). However, the regulation of lysogeny in a wide variety of other bacteriophages has not been investigated.

CTXφ encodes the genes for cholera toxin, the virulence factor primarily responsible for the watery diarrhea characteristic of the disease cholera. Non-toxigenic strains of *V. cholerae* can be readily transduced to toxin-producing strains by lysogenization with CTXφ, a process known as lysogenic conversion (6). Like other filamentous bacteriophage, CTXφ does not have a truly lytic growth phase. However, CTXφ lysogens exhibit several characteristics found in other temperate bacteriophages: 1) CTXφ lysogens contain the CTX prophage stably integrated into the *V. cholerae* genome. 2) The CTX prophage expresses a transcriptional repressor, RstR, that represses the expression of CTXφ replication genes and provides immunity to secondary infection by CTXφ (7,8).

CTXφ shares genetic and morphological similarities with the *E. coli* filamentous bacteriophage fd. Similarities in gene sequence and gene order indicate that many elements of the pathway for phage assembly and secretion are conserved (6,9). A novel region of the CTXφ genome known as the ‘RS’ region encodes *rstA* and *rstB*, two genes whose products are required for the replication and integration of the CTXφ chromosome, respectively (7). The RS region
also encodes the transcriptional repressor RstR, which is divergently transcribed from \( rstA \). \( rstA \) is expressed from a strong promoter in intergenic region 2 (\( ig-2 \)), the 138 base-pair (bp) region separating \( rstA \) from \( rstR \) (8) (see Fig. 1A). Expression of \( rstA \) is strongly repressed in CTXφ lysogens and RstR is the only \( V. \text{cholerae} \) factor required for \( rstA \) repression in \( E. \text{coli} \) (7). RstR encodes a 13 kilodalton (Kd) protein with an N-terminal HTH element similar to the HTH present in the Cro/cI superfamily of repressors (7). This sequence similarity extends to regions surrounding the bi-helical HTH motif, suggesting that the entire N-terminal region of RstR folds into an \( \alpha \)-helical domain similar to that of CI repressor (10). The short C-terminal region of RstR is unrelated to the oligomerization domain of CI repressor. Three distinct CTXφ variants have been described. CTX\(^{ET}\)φ (derived from O1 El Tor \( V. \text{cholerae} \)), CTX\(^{CL}\)φ (derived from Classical \( V. \text{cholerae} \)) and CTX\(^{Cal}\)φ (derived from O139 Calcutta \( V. \text{cholerae} \)) are largely identical, but each encodes a unique RstR repressor and adjacent \( ig-2 \) region (8,11). Each RstR allele specifically represses expression of its neighboring \( rstA \) gene, but is unable to repress \( rstA \) expression from a heterologous CTXφ. These observations led to the proposal that each RstR repressor binds exclusively and specifically to cognate operators in its neighboring \( ig-2 \) sequences (8).

Here we show that purified RstR\(^{ET}\) binds to three operator sites in the \( rstA \) promoter region. Unlike many bacterial HTH repressors, RstR binds to extended 50 bp operators, forming tetrameric repressor:operator complexes. The pathway for formation of these complexes appears to involve the cooperative interaction of two RstR dimers bound to neighboring sites in operator DNA. Our model for RstR binding bears similarities, as well as interesting differences, to the ‘pairwise’ cooperative binding of lambda CI dimers to adjacent operator sites in the lambda control region.
Experimental procedures

Plasmids and strains: *E. coli* XL1-Blue (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lacZΔM15/F′::Tn10 lacIq*) (Stratagene, La Jolla CA) was used for routine cloning and for the over-expression of RstR6H. The *ig-2* fragments 312, 309, and 301 were generated by the polymerase chain reaction with the appropriate primers and pCTX-Kn as template DNA. The resulting PCR products were cloned into plasmid PCRII (Invitrogen, Carlsbad CA) and the DNA sequence of the resulting plasmids was determined.

RstR6H expression vector: The *rstR* coding sequence was amplified by PCR from pCTX-Kn using the primers RSTR1 (5′CCCATGGCGAAGATAAAAGAA) and RSTR4 (5′GCGGATCCAGCACCATGATTT) and ligated to PCR2.1 (Invitrogen). The resulting plasmid, pHK297, was digested with *NcoI* and *BamHI* and the *rstR*-containing fragment was isolated and ligated to the expression vector pQE60 (Qiagen, Valencia CA), also digested with *NcoI* and *BamHI*. The resulting plasmid, pHK300, encodes RstR6H with an additional alanine residue after the initiator methionine (due to the introduction of the *NcoI* restriction site) and the sequence GSRSHHHHHH(COOH) after the terminal alanine residue of RstR. RstR-CBD-6H was constructed by cloning a PCR product encoding the 52 amino acid chitin-binding domain of *B. circulans* from plasmid pTXB1 (New England Biolabs, Beverly MA) into the unique *BamHI* site in pHK300.

RstR6H purification: A 10L fermentor culture of *E. coli* strain XL1-blue (Stratagene) containing pHK300 was grown with aeration at 37°C in Luria broth supplemented with 50µg/mL Carbenicillin. RstR6H expression was induced at an OD₆₀₀ of 0.8 by the addition of 1 mM
isopropyl-β-D-thiogalactopyranoside (IPTG). After 16 hours, cells were collected by centrifugation, resuspended in 100 mL chilled buffer A (0.05M NaH₂PO₄, 0.5M NaCl, 0.05% Tween-20, pH 7.0) and lysed by two passes in a French pressure cell at 24,000 psi. Cell debris was sedimented at 100,000xg for one hour (S100 fraction) and RstR6H was purified by Ni⁺²-chelation affinity chromatography as follows: Two mL of a 50:50 slurry of Ni²⁺ NTA agarose (Qiagen) in buffer A was added to thirty five mL of S100 extract and the mixture was gently mixed on ice for one hour. The slurry was transferred to a small chromatography column and washed with approximately 150 mL buffer A, followed by 35 mL buffer A containing 0.1M imidazole. RstR6H was eluted in 5 mL buffer A containing 0.3M imidazole (pH unadjusted). The RstR6H-containing fraction (2.5 mL) was dialyzed overnight at 4°C versus 200 mL buffer A. RstR6H was further purified by a second round of affinity chromatography using the above procedure. The final eluate (1 mL) was dialyzed overnight versus buffer B (20 mM Tris-HCl, 0.25M NaCl, 0.05% Tween-20, pH 8.0) and aliquots were stored frozen at -70°C. The lengthened RstR-CBD-6H protein was purified similarly. RstR6H concentrations were determined by UV absorption at 280 nm, using a molar extinction coefficient of 8,370 M⁻¹cm⁻¹ (12). After SDS-PAGE, proteins bands were transferred to a PVDF membrane and subjected to amino-terminal sequencing. Circular dichroism (CD) spectra were obtained at 25°C on a Jasco 810 spectropolarimeter, using a 0.1cm quartz sample cell. For CD analysis, RstR6H was first dialyzed against 10mM sodium phosphate buffer, pH7.5, 0.25M NaF.

**Gel mobility shift assay:** The ig-2 DNA fragments were liberated from the PCRII cloning vector by EcoRI digestion, fractionated by agarose gel electrophoresis, and purified using a Qiaquick gel extraction kit (Qiagen). DNA was radiolabeled using T4 DNA polymerase and [α-
$^{32}\text{P}$-dATP. Synthetic OL DNA probes were made by radiolabeling one oligonucleotide with T4 polynucleotide kinase and [$\gamma^{32}\text{P}$]ATP and annealing to a complementary unlabeled oligonucleotide. The resulting DNA fragments were purified on non-denaturing polyacrylamide gels and extracted using the “soak and crush” method (13). DNA-binding reactions were as follows: 20 µL binding reactions contained: 20mM Tris-HCl, pH 7.5, 20 mM KCl, 1mM MgCl$_2$, 5mM dithiothreitol (DTT), 5% glycerol, 50 µg/mL bovine serum albumin, 50µg/mL sonicated salmon sperm DNA, and 5,000 cpm probe DNA (1-2 fmol). RstR6H dilutions were prepared in cold reaction buffer immediately prior to use. Binding reactions were incubated at room temperature for 30 minutes. In experiments with 309 probe, an incubation on ice for 20 minutes prior to gel electrophoresis was found to significantly improve the stability of RstR:DNA complexes. Where probes were digested with AluI, the digested DNA was added directly to binding reactions without further purification. For mixing experiments with RstR6H and RstR-CBD-6H, protein mixtures were prepared in binding buffer plus 50mM DTT to prevent covalent dimer formation through the single cysteine residue present in RstR6H and RstR-CBD-6H. RstR mixtures were incubated at 30°C for 30 minutes prior to addition to DNA binding reactions, as described above. One microliter of a buffered dye solution (0.1% bromophenol blue) was added to each sample and aliquots were loaded onto 6% acrylamide DNA retardation gels (0.5X TBE) and electrophoresed at 10V/cm. Gels were dried directly to blotting paper and exposed to autoradiographic film. For quantitation of bound and free DNA, dried gels were scanned with a Molecular Dynamics STORM phosphoimager and quantitated using ImageQuant software. For DNA binding experiments, the fraction of total DNA bound was calculated as counts bound/(counts bound+counts free) without any corrections. Phosphorimager data from a preliminary DNA binding experiment were collected after 16 hours and 62 hours exposure.
These data yielded essentially identical binding curves, indicating that the response of the phosphor-imaging device was linear over a wide range of exposure levels.

**DNase I footprint assay:** DNase I footprint assays were performed as previously described (14). End-labeled DNA probes were generated by PCR with one 5’-radiolabeled primer and a second non-radiolabeled primer. DNA probes were purified by non-denaturing PAGE and eluted using the “crush and soak” method. Reactions were carried out in 30µL containing 50,000 cpm labeled DNA fragment and purified RstR6H, as described above for gel-shift experiments. Binding proceeded for 30 minutes at room temperature. 0.2U DNase I (Ambion, Austin TX) was added and the incubation was continued for one minute. The G+A sequencing ladders were generated as previously described (15). Dried DNA pellets were resuspended in loading buffer (95% formamide, 18 mM Na2EDTA, .025% SDS, .01% xylene cyanol and bromophenol blue) and heated to 90°C for two minutes prior to loading onto pre-run 8% sequencing gels. The gels were dried to blotting paper and exposed to autoradiographic film.

**Gel permeation chromatography:** Gel permeation was performed on a Superdex 75 HR 10/30 column (24 mL total volume), using an Amersham-Pharmacia FPLC system. The column was equilibrated in buffer B plus 5mM DTT and calibrated with a set of molecular weight markers (Fluka), as shown in Fig. 4. RstR6H samples (0.2 mL) were pre-incubated at 30°C for 30 minutes and chromatographed at a flow rate 0.5 mL/min. 0.5 mL fractions were collected and RstR6H was detected in an immobilized immunoassay using mouse anti-RstR antisera. To confirm the presence of RstR, column fractions were analyzed by SDS-PAGE followed by immuno-blotting with anti-RstR antisera.
**Formaldehyde crosslinking:** Formaldehyde was added to RstR6H (0.1-0.3 mg/ml) in buffer B to a final concentration of 1% and incubated for 10 min at 22°C. Cross-linking was stopped by the addition of an equal volume of 2X SDS sample buffer and heating to 95°C for 5 min. Samples were fractionated by SDS-PAGE under reducing conditions (0.1M DTT) on 10% polyacrylamide gels. Gels were stained with a colloidal Coomassie G-250 staining kit (Invitrogen).

**Calculations:** A reaction describing the equilibrium binding of two RstR dimers to DNA to form tetrameric complexes with no intermediates is

\[
K_A = \frac{[R_4O]}{[R_2]^2[O]} \quad (1)
\]

where \(R_2\) are free RstR dimers, \(O\) is free operator DNA, and \(R_4O\) is the bound complex. \(K_A\), the apparent equilibrium association constant, is

\[
K_A = [R_4O]/[R_2]^2[O] \quad (2)
\]

For a highly cooperative binding reaction, it is useful to define

\[
\Theta/(1-\Theta) = K_A[R_2]^2 \quad (3)
\]

where \(\Theta\) is the fraction of total binding sites occupied in DNA binding experiments (16). Preferring to describe binding in terms of the apparent dissociation constant, \(K_D\)

\[
\Theta/(1-\Theta) = [R_2]^2/K_D \quad (4)
\]

Rearranging (4) gives
\[ \Theta = \frac{[R_2]^2}{(K_D + [R_2]^2)} \]  

(5)

We approximated \([R_2] \approx \frac{[R_{Total}]}{2}\), since the dimer:tetramer equilibrium constant is in the micromolar range, whereas DNA binding occurs at nanomolar RstR concentrations. A monomer:dimer equilibrium was not considered here, but could play a role in operator binding. Trial-and-error fitting of theoretical binding curves and linear regression analyses of the Hill plot was carried out using GraphPad Prism 4.0 (GraphPad Software, San Diego CA).
**Results**

**RstR binds to multiple sites in the rstA promoter region:** RstR containing a C-terminal polyhistidinyl tag (RstR6H) was over-expressed and purified from *E. coli* by two rounds of Ni²⁺-NTA affinity chromatography (see Materials and Methods). This material was about 95% pure, as estimated from sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue-staining (see Fig. 4B). As predicted from the gene sequence, the final material yielded a single band of approximately 14 kilodaltons (Kd) after SDS-PAGE. The amino-terminal sequence of the 14 Kd band was found to be AKIKER. Residues 2 through 6 were identical to positions 2 through 6 of the predicted RstR sequence, while the alanine at position 1 is the result of cloning the *rstR* gene into the expression plasmid vector. The apparent lack of an amino terminal methionine residue indicates that the initiator formyl-methionine is proteolytically removed in *E. coli*. Far UV circular dichroism spectra of RstR6H showed clear minima at 208 and 222 nM, indicating that RstR contains α–helical secondary structure.

By analyzing β-galactosidase expression from an *rstA::lacZ* reporter plasmid, we previously showed that a 290 bp DNA fragment containing the *ig-2* region from CTXφ<sub>ET</sub> contained both the P<sub>rstA</sub> promoter and sequences sufficient for *rstR*-mediated transcription repression of P<sub>rstA</sub> (8). A similar DNA fragment, extending from -168 to +88 relative to the *rstA* transcription start site, was used to investigate DNA binding by purified RstR6H in mobility shift assays. As shown in Figure 1, RstR6H retarded the migration of the 256 bp 312 probe, giving rise to three slower-migrating bands. The faster-migrating shifted band probably represents repressor binding to a high-affinity binding site, while the second and third shifted bands indicate successive binding of RstR6H to two additional sites. In further mobility shift experiments, RstR6H did not bind to *ig-2* DNA probes derived from CTXφ<sub>Cl</sub> or CTXφ<sub>Calc</sub> (data not shown).
The rstR-ig-2 regions of all three CTXφ genomes are widely divergent and individual RstR repressors exclusively repress expression of their neighboring rstA promoter (8,17). We conclude from these observations that RstR6H binds specifically to regulatory sequences within the ig-2 region of CTXφ\textsuperscript{ET}.

To map RstR6H binding sites in ig-2, mobility shift assays were performed with two smaller, overlapping probes. RstR6H retarded the mobility of labeled 301 DNA (-77 to +43) to a single slower-migrating species (Fig. 1B). Binding occurred over a relatively small range of repressor concentrations (<10 fold), with half-maximal binding occurring at approximately 3.5 nM RstR6H. RstR6H shifted the 309 probe (-168 to -7) to two slower-migrating bands. Upon titration of RstR6H into the 309 binding reactions, the faster-migrating shifted form (band 1) diminished in intensity and the slower-migrating species (band 2) accumulated, suggesting that RstR6H binds sequentially to two sites in 309. Half-maximal binding to 309 occurred at approximately 20 nM RstR6H. To refine the mapping of RstR binding sites, end-labeled 301 and 309 probes were digested with the restriction endonuclease \textit{Alu}\textsubscript{I} and the resulting sub-fragments were used in mobility shift assays. \textit{Alu}\textsubscript{I} cuts 301 once at position -4, resulting in two labeled sub-fragments. Neither 301 sub-fragment was shifted by RstR6H, indicating that the RstR binding site in 301 overlaps, or is very close to, this \textit{Alu}\textsubscript{I} site (data not shown). Similarly, \textit{Alu}\textsubscript{I} digestion of 309 at position -80 resulted in two sub-fragments, neither of which was shifted by RstR6H (data not shown). One interpretation of this experiment is that the two potential RstR binding sites in 309 both overlap the \textit{Alu}\textsubscript{I} site at -80. Alternatively, only one binding site overlaps the (-80)\textit{Alu}\textsubscript{I} site but RstR binding to a neighboring site requires that RstR first bind to the \textit{Alu}\textsubscript{I}-containing site. These experiments physically map three RstR binding sites in ig-2: one high-affinity binding site overlaps the \textit{Alu}\textsubscript{I} restriction site at position -4, while two lower-affinity
binding sites are located near the AluI site at position -80. The high-affinity binding site in 301 was named $O_1$ and the two promoter-distal binding sites were named $O_2$ and $O_3$.

**DNase I protection analysis of RstR6H binding:** The precise positions of RstR binding sites in $ig-2$ was determined by DNase I protection assay (18). Since full-length $ig-2$ DNA was too large for high-resolution footprint analysis, the above-mentioned 301 and 309 DNA fragments were radiolabeled and used as probes in these experiments. Figures 2A shows DNase I protection data for each strand of the 301 probe. Figure 3 summarizes all of the DNase I protection experiments. RstR6H protected an unusually large 50 bp region (+17 to -32) in 301 corresponding to the high-affinity $O_1$ binding site. Complete occupancy of $O_1$ occurred abruptly with increasing RstR6H concentrations, similar to the binding pattern observed in mobility shift assays (Fig. 1). Two footprinted regions, each approximately 50 bp in length, were identified in 309 as corresponding to the promoter-distal sites $O_2$ and $O_3$. These footprints extended from positions -40 to -150. The region between $O_2$ and $O_3$ contained an unusual number of DNase I hyper-sensitive sites on both strands, suggesting that this region of the DNA is bent or distorted upon RstR binding. Interestingly, the footprints show that only the $O_2$ site spans the single AluI restriction site in 309 (see Fig. 3). Therefore, our previous mobility shift results indicate that RstR cannot stably bind to the isolated $O_3$ site present on one of the AluI sub-fragments. Taken together, these data suggest that RstR can only bind to $O_3$ if RstR is first bound to the neighboring $O_2$ site.

The $O_1$, $O_2$ and $O_3$ footprints included sites that remained DNase I-accessible and positions that became hyper-sensitive to DNase I attack after RstR6H binding. These DNase I-exposed sites clustered in short 2-3 bp regions and were interspersed with strongly protected
regions 6-11 bp in length (Fig. 3). A similar pattern was observed on the opposite DNA strand, although shifted 3 bp in the 3’ direction. This offset pattern is expected if the minor groove of relatively short regions remain exposed after RstR binding, since DNase I cuts in the minor groove and the nearest backbone ribose-phosphate groups are about 3 bp apart in the 3’ direction (19). This protection pattern is seen more clearly when the footprint data are mapped onto a planar projection of DNA, as shown in Figure 2B. Displaying the data in this way revealed another feature of the RstR binding sites: DNase I-exposed regions occurred on one face of the DNA helix in one half of the footprint, as shown by the exposed patches forming a line parallel to the primary DNA helical axis. This pattern shifted approximately 3 bp, representing a 100° turn about the DNA helical axis, to a different face of the DNA helix in the remaining half of the footprint. A similar DNase I protection pattern was observed in the O2 and O3 footprints (data not shown). This pattern is most easily explained if one RstR moiety, probably a dimer, bound to the face of the DNA helix opposite each set of exposed sites. Each half-region is sufficiently large (20-25 bp) to present two consecutive major groove elements required for binding the two HTH elements of an RstR dimer.

**RstR is tetrameric when bound to operator DNA:** After affinity purification, the oligomeric state of RstR6H was probed by gel-permeation chromatography and formaldehyde cross-linking. As shown in Figure 4A, RstR6H eluted as two peaks of \( M_r \approx 24,000 \) and \( M_r \approx 46,000 \). We suggest that these peaks represent a mixture of dimers and tetramers (predicted \( M_r \approx 28,000 \) and 56,000 for dimers and tetramers, respectively). The apparent molecular masses deviated from the predicted values for dimers and tetramers, a discrepancy that may be due to RstR6H folding into a more compact shape than a typical globular protein. When increasingly dilute RstR6H samples
were chromatographed, the peak corresponding to dimers increased in size relative to the tetramer peak, indicating that tetramers dissociate into dimers. We estimated the apparent dimer/tetramer dissociation constant, $K_{tet}$, to be approximately 0.5 µM monomer RstR6H. Therefore, RstR6H would be primarily dimeric - or possibly monomeric - at the nanomolar concentrations where DNA binding is observed. As shown in Figure 4B, RstR6H could be cross-linked to dimers using formaldehyde (apparent $M_r \approx 28$ kD).

The size and pattern of RstR footprints suggested to us that RstR binds operator DNA as a dimer-of-dimers. To determine precisely RstR’s oligomeric state when bound to operator DNA, we employed the mixed-oligomer technique devised by Hope and Struhl (20). A lengthened variant of RstR repressor, RstR-CBD-6H, was mixed with RstR6H and used in mobility shift assays of $O1$ operator binding. RstR-CBD-6H contains 52 additional C-terminal residues and retards $O1$ DNA to a greater extent than RstR6H in mobility shift assays (Fig. 5). RstR activity was unaffected by the C-terminal addition in RstR-CBD-6H, since the purified fusion protein bound $O1$ DNA with an affinity comparable to that of RstR6H.

As shown in Figure 5, a total of five shifted bands were observed when mixtures of RstR6H(short) and RstR-CBD-6H(long) were incubated with labeled $O1$ DNA. The slowest and fastest migrating complexes co-migrated with RstR-CBD-6H and RstR6H complexes, respectively. The bands of intermediate mobility represent complexes that contain distinct mixtures of ‘short’ and ‘long’ forms of RstR. If RstR bound to it’s operator as a tetramer, five complexes corresponding to 4:0, 3:1, 2:2, 1:3, and 0:4 mixtures of short and long RstR would be expected to form. Assuming random assortment of repressor subunits, an equimolar mixture of short and long RstR would be predicted to give rise to these five complexes with relative molar ratios of 1:3:6:3:1. The band intensities that resulted from mixing approximately equimolar
amounts of short and long RstR generally fit the predicted pattern of band intensities for
tetrameric binding (Fig. 5). The simplest conclusion from these data is that RstR is tetrameric
when bound to operator DNA. A tetrameric binding model was also supported by direct
measurements of RstR:DNA stoichiometry. In mobility shift experiments with O1 DNA present
at high concentrations (one tenth micromolar) and using our most active RstR6H preparations,
we observed that approximately 21 picomoles of RstR6H monomer was required to completely
shift 5 picomoles of O1 DNA (data not shown).

Cooperative binding of RstR to operator DNA: Although RstR is dimeric at concentrations
where operator binding is observed, RstR is tetrameric when bound to operator DNA. Operator
binding could occur by the sequential, independent binding of RstR dimers or by the cooperative
binding of two RstR dimers. We favor a cooperative binding model, since mobility shift assays
and DNase I footprinting studies show that RstR binds operator DNA over a narrow range (<10
fold) of repressor concentrations (Figs. 1 and 2). Also, we have never observed an intermediate
shifted band in mobility shift experiments that would indicate dimer-bound complexes with O1
DNA. To examine the possibility of cooperative binding further, detailed mobility shift analyses
were carried out and the resulting data were fit to quantitative models of repressor binding. The
results of these experiments are shown in Figure 6. The fraction of O1 DNA bound by repressor
increased from 0.1 to near unity over an approximately 10-fold range of RstR6H concentrations,
exhibiting half-maximal binding at approximately 3.9 nM RstR6H. These RstR binding data
were successfully modeled by a theoretical binding curve that describes two free repressor
dimers binding DNA cooperatively to form a tetrameric repressor:operator complexes, with a
dissociation constant of $1.5 \times 10^{-17}$ M$^2$ (see “Calculations” in Experimental procedures). A Hill
plot of the data in Figure 6 yielded a best-fit line with a slope of 2.0 ± 0.1, indicating that operator binding occurs cooperatively and that a significant fraction of RstR is dimeric at concentrations where DNA binding occurs.

Finally, we investigated whether RstR6H dimers alone could form stable complexes with left-half or right-half fragments of *O1*. Measuring the equilibrium binding constant for dimer binding would allow us to determine the coupling factor, or the energetic contribution made by dimer-dimer protein interactions to the overall DNA binding reaction (21) (assuming that dimer binding affinity for the left and right halves of *O1* are identical). However, binding of RstR6H to a 33 bp left-half or a 28 bp right-half operator fragment was not detected in mobility shift experiments, even at micromolar RstR concentrations (data not shown). The apparent low affinity of RstR dimers for left and right halves of *O1* DNA suggests that protein-protein interactions between DNA-bound dimers make a significant energetic contribution to operator binding.
Discussion

We have found that the RstR repressor of CTXφ\textsuperscript{ET} binds specifically to three DNA sites surrounding the \textit{rstA} promoter. The high-affinity \textit{O1} site encompasses almost the entire promoter region (Fig. 3), while the two lower-affinity sites, \textit{O2} and \textit{O3}, are located -40 to -150 upstream of the \textit{rstA} promoter. Preliminary in vivo studies with \textit{rstA::lacZ} fusions indicate that \textit{O1} and \textit{O2-O3} function independently to repress expression of \textit{P_{rstA}}. A second function for \textit{O2-O3} is suggested by the observation that \textit{O2} overlaps the divergent \textit{rstR} promoter, indicating that \textit{O2-O3} may also play a role in the auto-regulation of \textit{rstR} expression.

In mobility shift assays, RstR binding to \textit{O3} exhibited an unusual dependence upon a neighboring intact \textit{O2} site, indicating that RstR is recruited to bind \textit{O3} by RstR bound to \textit{O2}. It will be interesting to investigate whether this recruitment is mediated by direct protein-protein contacts, or through an altered DNA conformation resulting from RstR binding to \textit{O2}. It is unlikely that the pattern of RstR binding to \textit{O2-O3} is due to non-specific repressor binding to neighboring DNA, a process also known as “phasing” (22). The DNase I footprint at \textit{O3} was similar to the footprint at the high affinity site \textit{O1}, indicating that \textit{O3} binding involves similar protein:DNA contacts. Also, “phased” binding of RstR was not observed adjacent to \textit{O1}, even at high repressor concentrations.

The primary conclusion of this work is that RstR tetramerizes on operator DNA, using protein-protein interactions between neighboring dimers to stabilize repressor:operator complexes. The approximately 50 bp size of each RstR footprint, together with the staggered pattern of DNase I-accessible regions within each footprint (Fig. 4), suggested that RstR binds DNA as a dimer-of-dimers, with the binding face of one dimer rotated approximately 100° about
the DNA helical axis from the neighboring dimer. RstR was indeed tetrameric when bound to operator DNA, as shown by the ability of ‘short’ and ‘long’ forms of RstR6H to form five distinct complexes with O1 DNA in mobility shift experiments (Fig. 4). This method was previously used to show that Arc repressor of bacteriophage P22 and the mammalian transcription factor LSF also bind DNA as tetramers (23,24). Better studied HTH repressors, such as LacI, GalR, or CI and Cro of phage λ, bind to their operator sites as homodimers and footprint 20-25 bp regions, or about two full turns of the DNA helix (18,25) (26). The two HTH elements (one per repressor monomer) of a CI dimer, for example, are spaced such that each recognition helix can contact adjacent major groove elements on one face of the DNA helix (27). Our model for the staggered arrangement of RstR dimers on operator DNA is strikingly similar to the arrangement of two CI dimers bound to O_R1 and O_R2 of bacteriophage λ (1,28). In spite of this similarity, the protein-protein contacts that mediate the cooperative binding of RstR dimers are probably different from the contacts that mediate the “pairwise” cooperative binding of CI dimers to O_R1 and O_R2 (29). The C-terminal region of RstR is only 52 residues long and does not share sequence similarity to the C-terminal domain of CI repressor, which mediates CI oligomerization. Also, secondary structure predictions indicate that the RstR C-terminal region is largely α–helical, whereas the C-terminal domain of CI consists largely of β–sheet elements (29).

Several lines of evidence suggest it is unlikely that each RstR operator is merely two classical-sized operators situated close together. First, RstR6H bound strongly to a synthetic 50 bp DNA corresponding to the footprinted region at O1, but failed to bind to any sub-fragments of O1 in mobility shift assays, including a 33 bp and 28 bp DNA corresponding to the left and right halves of O1, respectively. Also, RstR failed to stably bind to any of the AluI sub-fragments of
301 and 309 DNA that contain large portions of O1, O2, or parts of both sites. Finally, β-galactosidase expression from an rstA::lacZ fusion that contains the leftmost 26 bp of O1 and the rightmost 39 bp of O2, was not repressed by RstR in vivo (data not shown). The entire 50 bp footprinted region is required for strong repressor binding.

To identify potentially important RstR:DNA contacts, we searched the O1, O2, and O3 sequences for an RstR consensus binding site. At O1, two sets of inverted repeats were identified that contained the potential ‘half-site’ sequence CTNN^/C AAG (see Fig. 3). However, this ‘half-site’ sequence was not readily identified at the expected positions in O2 or O3. A survey of other HTH transcription regulators that have extended binding sites, such as OxyR, OccR, and Rns, indicates that consensus binding sites for these regulatory factors were also difficult to identify, in some cases requiring the identification of numerous natural binding sites or the in vitro generation of many synthetic binding sites (30-32). As others have pointed out, such extended binding sites could be biologically advantageous, allowing for sites with large variations in transcription factor binding affinities (30).

The steep nature of the DNA binding curve (see Fig. 6), plus the absence of any detectable dimer-bound complexes in mobility shift experiments, indicate that RstR binds operator DNA cooperatively. These observations account for the overall reaction but ignores the possible role of dimeric RstR:operator complexes as intermediates in the binding reaction. We imagine two possible pathways for the assembly of tetrameric complexes. In one, dimers first combine to form free tetramers, which then bind to operator DNA in a coupled reaction:

\[ 2 \text{R}_2 + \text{O} \leftrightarrow \text{R}_4 + \text{O} \leftrightarrow \text{R}_4\text{O} \]

Although our gel filtration experiments indicate that tetramers only form at high repressor concentrations, the small amount of tetramers formed at low repressor concentrations could be
rapidly trapped by operator binding. Alternatively, two RstR dimers could bind sequentially to 
operator DNA to form the final tetrameric complex:

\[ 2 \text{R}_2 + \text{O} \leftrightarrow \text{R}_2\text{O} + \text{R}_2 \leftrightarrow \text{R}_4\text{O} \]

Our mobility shift experiments did not show evidence of dimer complexes. However, a dimer 
complex might be a very short-lived intermediate or sufficiently unstable to be seen in our 
 mobility shift assays. Kinetic binding experiments carried out under conditions where RstR is 
primarily dimeric or tetrameric would aid in determining the active form of RstR.

Gene transcription is frequently regulated by factors that bind cooperatively to DNA. 
The energetic contribution of cooperativity is often-though not always-mediated by protein-
protein interactions. These interactions fall into two broad classes. In one class are factors that 
utilize protein-protein interactions to loop DNA between distant sites, as in the case of the LacI 
and GalR repressors (33,34), and the AraC regulator (35). In these systems the spacing of distant 
operator sites can often be altered without disrupting normal regulation. The second class 
consists of regulatory proteins that bind and oligomerize to closely spaced sites on the DNA. 
Examples include the Arc and Mnt repressors of phage P22 (21,23,36), LexA (37), and the 
MCM1/alpha 2 complex of yeast (38). RstR may belong to this later class of regulators. On-
going studies of mutant binding sites indicate that as little as a 1 bp deletion near the center of
\textit{O1} drastically reduces RstR binding, suggesting that RstR dimers must bind to correctly spaced 
sites in order for dimer-dimer interactions to stabilize the tetrameric complexes.
Acknowledgments

We thank R.T. Sauer, G. Kudelka, S. McLeod, and D. Lazinski for advice and valuable discussions. We also thank Anne Kane of the New England Medical Center GRASP Center for protein expression and member of the Baleja lab for help in collecting CD spectra.
References

1. Ptashne, M. (1992) A Genetic Switch, 2nd Ed., Blackwell Scientific, Cambridge
2. Susskind, M., and Youderian, P. (1983) in Lambda II (Hendrix, R. W., Roberts, J. W.,
   Stahl, F. W., and Weisberg, R. A., eds), pp. 347-363, Cold Spring Harbor Press, Cold
   Spring Harbor, NY
3. Ogawa, T., and Ogawa, H. (1988) J. Mol. Biol. 202, 537-550
4. Oberto, J., Weisberg, R. A., and Gottesman, M. E. (1989) J. Mol. Biol. 207, 675-693
5. Saha, S., Haggard-Ljungquist, E., and Nordstrom, K. (1987) EMBO J. 6, 809-814
6. Waldor, M. K., and Mekalanos, J. J. (1996) Science 272, 1910-1914
7. Waldor, M. K., Rubin, E. J., Pearson, G. D., Kimsey, H., and Mekalanos, J. J. (1997)
   Mol. Microbiol. 24, 917-926
8. Kimsey, H. H., and Waldor, M. K. (1998) Proc. Natl. Acad. Sci. USA 95, 7035-7039
9. Davis, B. M., Lawson, E. H., Sandkvist, M., Ali, A., Sozhamannan, S., and Waldor, M.
   K. (2000) Science 288, 333-335
10. Wintjens, R., and Rooman, M. (1996) J. Mol. Biol. 262, 294-313
11. Davis, B. M., Kimsey, H. H., Chang, W., and Waldor, M. K. (1999) J. Bacteriol. 181,
    6779-6787
12. Gill, S. C., and von Hippel, P. H. (1989) Anal. Biochem. 182, 319-326
13. Sambrook, J., and Russel, D. W. (2001) Molecular cloning: a laboratory manual, 3rd Ed.
    3 vols., Cold Spring Harbor Press, Cold Spring Harbor, NY
14. Bruist, M. F., Glasgow, A. C., Johnson, R. C., and Simon, M. I. (1987) Genes Dev. 1,
    762-772
15. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-559
16. Tinoco, I., Sauer, K., and Wang, J. C. (1985) Physical Chemistry: Principles and
    Applications in Biological Sciences, 2 Ed., Prentice-Hall, Englewood Cliffs, N.J.
17. Kimsey, H. H., Nair, G. B., Ghosh, A., and Waldor, M. K. (1998) Lancet 352, 457-458
18. Galas, D. J., and Schmitz, A. (1978) Nucl. Acids Res. 5, 3157-3170
19. Drew, H. R. (1984) J. Mol. Biol. 176, 535-557
20. Hope, I. A., and Struhl, K. (1987) EMBO J. 6, 2781-2784
21. Smith, T. L., and Sauer, R. T. (1995) J. Mol. Biol. 249, 729-742
22. Carlson, N. G., and Little, J. W. (1993) J. Mol. Biol. 230, 1108-1130
23. Brown, B. M., Bowie, J. U., and Sauer, R. T. (1990) Biochemistry 29, 11189-11195
24. Shirra, M. K., and Hansen, U. (1998) J. Biol. Chem. 273, 19260-19268
25. Majumdar, A., and Adhya, S. (1987) J. Biol. Chem. 262, 13254-13262
26. Johnson, A. D., Meyer, B. J., and Ptashne, M. (1979) Proc. Natl. Acad. Sci. USA 76,
    5061-5065
27. Pabo, C. O., and Lewis, M. (1982) Nature 298, 443-447
28. Johnson, A. D., Poteete, A. R., Lauer, G., Sauer, R. T., Ackers, G. K., and Ptashne, M.
    (1981) Nature 294, 217-223
29. Bell, C. E., Frescura, P., Hochschild, A., and Lewis, M. (2000) Cell 101, 801-811
30. Toledano, M. B., Kullik, I., Trinh, F., Baird, P. T., Schneider, T. D., and Storz, G. (1994)
    Cell 78, 897-909
31. Wang, L., Helmann, J. D., and Winans, S. C. (1992) Cell 69, 659-667
32. Munson, G. P., and Scott, J. R. (1999) J. Bacteriol. 181, 2110-2117
33. Oehler, S., Eismann, E. R., Kramer, H., and Muller-Hill, B. (1990) EMBO J. 9, 973-979
34. Adhya, S. (1989) *Annu. Rev. Genet.* **23**, 227-250
35. Schleif, R. (2003) *Bioessays* **25**, 274-282
36. Waldburger, C. D., and Sauer, R. T. (1995) *Biochemistry* **34**, 13109-13116
37. Kim, B., and Little, J. W. (1992) *Science* **255**, 203-206
38. Vershon, A. K., and Johnson, A. D. (1993) *Cell* **72**, 105-112
**Figure Legends**

**Figure 1:** Mobility shift assays of RstR6H binding to *ig-2* DNA. **A)** Diagram of the *ig-2* region of CTXφ and the DNA fragments used to map RstR binding sites. Angled arrows show the approximate positions of the *rstA* and *rstR* transcription start sites. Block arrows represent the *rstA* and *rstR* coding sequences and are not drawn to scale. Numbering is relative to the start of *rstA* transcription at +1. **B)** Autoradiograms of mobility shift assays. The left-most lane in each panel shows probe DNA with no added protein. ‘f’ indicates the position of the free DNA. In the 312 and 301 panels, RstR6H monomer concentrations were increased from 4.8 to 38.4 nM in 2-fold increments. In the 309 panel, RstR6H concentrations increased from 2 to 82 nM in 1.7 fold increments.

**Figure 2:** **A)** DNase I footprints of RstR6H bound to 301 DNA. DNA fragments were $^{32}$P-labeled at the 5’ end of one strand, incubated with freshly diluted RstR6H, and digested with DNase I. Cleavage products were separated on 8% sequencing gels. Top and bottom strands correspond to the DNA sequences as depicted in Figure 3. “G+A” is a marker lane containing the same probe chemically cleaved at guanosine and adenosine residues. Lanes 1 in each panel is probe DNA with no added RstR6H. Beginning with lane 2, RstR6H concentrations start at 2 nM and increase in 2-fold increments. Numbering is relative to the *rstA* transcription start site. **B)** Planar projection of the RstR6H footprint in 301 DNA. The parallel slanted lanes represent the sugar-phosphate DNA backbone and the horizontal lines depict base pairing across the minor groove. Promoter-distal sequences begin at the bottom of the panel. Thick shaded areas depict regions RstR protected from DNase I digestion and are shown in the minor groove since DNase I
binds and cuts in the minor groove. Filled stars represent DNase I hyper-sensitive sites and open circles represent weakly protected sites. The staggered dashed line highlights the approximate face of the DNA that remains susceptible to DNase I attack after RstR binding.

**Figure 3:** Summary diagram of RstR6H footprints in *ig-2*: Shaded areas indicate DNase I protected regions and unmarked bases indicate positions that remain DNase I-susceptible after RstR binding. Stars denote DNase I hyper-sensitive sites; open circles denote partially protected sites. The *rstA* and *rstR* transcription start sites are depicted with angled arrows and their -10 and -35 consensus sites are shown in bold type. The *Alu*I restriction sites at positions -80 and -6 are shown boxed. The four potential ‘half-site’ operator sequences in *OI* are identified by arrows.

**Figure 4:** A) Gel permeation chromatography of RstR6H. RstR6H samples were chromatographed on a Superdex 75HR column (24 mL bed volume) using an FPLC system. Fractions were analyzed for the presence of RstR6H using anti-RstR antibody in a microtiter plate-based immunoassay. Solid squares: 5 µM RstR6H loaded. Open circles: 0.5 µM RstR6H loaded. The elution peaks of molecular weight standards are depicted with arrows. A = blue dextran (>200 kD); B = bovine serum albumin (66 kD), C = bovine carbonic anhydrase (29 kD), D = cytochrome C (12.4 kD), E = aprotinin (6.5 kD). B) Formaldehyde cross-linking of RstR6H. Lane 1 is the molecular weight standards. Lane 2 shows an untreated control. Lane 3 contains RstR6H cross-linked with 1% formaldehyde for 10 min at 22°C. Samples were analyzed by SDS-PAGE and gels were stained with Coomassie.
Figure 5: Mobility shift assays of operator binding by RstR6H and the lengthened variant RstR-CBD-6H. A synthetic 50 bp DNA fragment containing $O1$ was $^{32}$P-labeled and used as probe. RstR6H and RstR-CBD-6H were mixed and equilibrated in binding buffer prior to the addition of probe DNA. Lane 1 is operator DNA with no added protein. Lane 2 shows shifted complexes with RstR6H alone. Lane 6 shows complexes formed with RstR-CBD-6H alone. Lanes 3-5 show complexes formed with mixtures of RstR6H and RstR-CBD-6H at approximately 4:1, 1:1, and 1:4 molar ratios, respectively. Total protein concentration was maintained at approximately 24 nM in each sample.

Figure 6: Binding isotherm for RstR6H binding to $O1$ DNA. The probe is the $^{32}$P-labeled 301 fragment depicted in Fig. 1. A) shows representative data obtained from one mobility shift experiment. The faint band migrating just above the free probe is a minor contaminant in the probe preparation. B) The fraction of DNA bound, $\theta$, was measured at numerous repressor concentrations in two independent experiments. The solid line is a theoretical binding curve describing the dissociation of tetrameric RstR:DNA complexes to two free dimers and free operator DNA with a dissociation constant of $1.5 \times 10^{-17} \text{M}^2$ (see equation (5) in Experimental Procedures). The inset graph depicts a Hill plot of RstR6H binding to $O1$ DNA. Data from duplicate experiments were averaged and plotted as a single point. The line is a linear regression best-fit to the data.
Figure 2
Figure 4
Figure 5
The CTX repressor RstR binds DNA cooperatively to form tetrameric repressor:operator complexes
Harvey H. Kimsey and Matthew K. Waldor

J. Biol. Chem. published online November 10, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M311109200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts