Characterization of DNA-binding specificity and analysis of binding sites of the *Pseudomonas aeruginosa* global regulator, Vfr, a homologue of the *Escherichia coli* cAMP receptor protein

Kristen J. Kanack,† Laura J. Runyen-Janecky,‡ Evan P. Ferrell,§ Sang-Jun Suh|| and Susan E. H. West

Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, 2015 Linden Drive, Madison, WI 53706, USA.

Vfr, a global regulator of *Pseudomonas aeruginosa* virulence factors, is a homologue of the *Escherichia coli* cAMP receptor protein, CRP. Vfr is 91 % similar to CRP and maintains many residues important for CRP to bind cAMP, bind DNA, and interact with RNA polymerase at target promoters. While vfr can complement an *E. coli crp* mutant in β-galactosidase production, tryptophanase production and catabolite repression, *crp* can only complement a subset of Vfr-dependent phenotypes in *P. aeruginosa*. Using specific CRP binding site mutations, it is shown that Vfr requires the same nucleotides as CRP for optimal transcriptional activity from the *E. coli lac* promoter. In contrast, CRP did not bind Vfr target sequences in the promoters of the *toxA* and *regA* genes. Footprinting analysis revealed Vfr protected sequences upstream of *toxA*, *regA*, and the quorum sensing regulator *lasR*, that are similar to but significantly divergent from the CRP consensus binding sequence, and Vfr causes similar DNA bending to CRP in bound target sequences. Using a preliminary Vfr consensus binding sequence deduced from the Vfr-protected sites, Vfr target sequences were identified upstream of the virulence-associated genes *plcN*, *plcHR*, *pbpG*, *prpL* and *algD*, and in the *vfr*orf*X*, *argH*/*limS*, *pilM*/*ponA* intergenic regions. From these sequences the Vfr consensus binding sequence, 5'-ANWWTGNGAWNY : AGWTCACAT-3', was formulated. This study suggests that Vfr shares many of the same functions as CRP, but has specialized functions, at least in terms of DNA target sequence binding, required for regulation of a subset of genes in its regulon.

**INTRODUCTION**

Regulation of virulence factors in *Pseudomonas aeruginosa* is necessarily complex and multifactorial, allowing adaptation to different environmental conditions and a wide range of infection sites. Vfr, a global regulator of gene expression in *P. aeruginosa*, affects the production of several virulence factors and/or their regulators. Vfr activates expression of both *toxA* (the structural gene encoding exotoxin A) and *regA* (a gene encoding a regulator of *toxA* expression, also called *toxR*), and production of protease IV (*prpL*) (West *et al.*, 1994). Vfr has also been found to regulate quorum sensing by directly enhancing production of the quorum sensing regulator *LasR*, and activating expression of the quorum sensing regulator *RhlR*, in a manner independent of the *LasR* effect on *rhlR* transcription, thus affecting the expression of multiple virulence factors positioned downstream in the quorum sensing cascade (Albus *et al.*, 1997; Medina *et al.*, 2003). A recent study has also suggested that Vfr negatively regulates the PQS quorum sensing system (Whitchurch *et al.*, 2005). Additional studies have defined the role of Vfr in the regulation of twitching motility mediated by type IV pili, the type II secretion system, the type III secretion system, flagellar biogenesis, and production of the stationary-phase sigma factor RpoS (Beatson *et al.*, 2002; Bertani *et al.*, 2003; Dasgupta *et al.*, 2002; Wolfgang *et al.*, 2003). Genomic and proteomic analyses of *P. aeruginosa* wild-type and Vfr mutant strains suggest that...
Vfr is ultimately responsible for positively or negatively affecting the expression of more than 100 genes and the production of more than 60 proteins in *P. aeruginosa* (Suh et al., 2002; Wolfgang et al., 2003). A recent study by Smith et al. (2004) determined that Vfr is necessary for lung colonization and dissemination in a mouse model of acute pneumonia.

Vfr is 67% identical and 91% similar to the cAMP receptor protein (CRP), a cAMP-responsive global regulator of gene expression in *Escherichia coli* (West et al., 1994). Genes regulated by CRP in *E. coli* include, but are not limited to, those involved in catabolite repression, flagellum biogenesis, iron acquisition, enterotoxin production and the heat shock response (Botsford, 1981). CRP functions as a dimer that, when complexed with its allosteric effector cAMP, binds to specific sequences in or near the promoters of the genes it regulates (Berg & von Hippel, 1988). Analysis of 25 CRP-binding sequences yielded the twofold symmetrical consensus CRP-binding sequence (CCS), AAATGTTGATCT:AGATCACATT, with positions 4 to 8 (underlined) from the centre being the most conserved residues (Berg & von Hippel, 1988). CRP was found to bind with 450 times greater affinity to this consensus sequence than to its natural binding sequence present upstream of the *lac* operon (Ebright et al., 1989). All identified naturally occurring CRP-binding sequences diverge from the consensus to varying degrees. This may create a hierarchy of affinity for CRP, allowing a mechanism for fine control over the regulon (Busby & Kolb, 1996).

Analysis of the crystal structure of a CRP–DNA complex revealed that CRP induces an approximately 90° bend in the DNA resulting primarily from two 40° kinks occurring between positions 5 and 6 on each side of the dyad axis of the sequence (Schultz et al. 1991). The sequence in each half-site along with the flanking sequence determines the geometry of the bent DNA. Pyles & Lee (1998) found that while the DNA appears to be symmetrically bent in the *lac* promoter, it appears to be asymetricaly bent in the *gal* promoter. Changes in DNA bend geometry may reflect differences in the mechanism of CRP regulation at different promoters.

CRP-dependent promoters have been grouped into three classes based upon the mechanism by which CRP facilitates transcriptional activation (reviewed by Busby & Ebright, 1999). At class I promoters, CRP binds upstream of RNA polymerase and interacts with the RNA polymerase α-C-terminal domain (α-CTD). This interaction involves ‘activating region 1’ (AR1) of CRP (residues 156–164). At class II promoters, CRP binds to a site overlapping that of RNA polymerase. CRP interactions with the RNA polymerase α-CTD at class II promoters require not only AR1 but an additional activating region, AR2 (residues His-19, His-21, Glu-96 and Lys-101). At class III promoters CRP synergistically activates transcription with additional CRP dimers and/or other activators. CRP-binding sequences may occur at far distances upstream from the transcriptional start site at class III promoters. Finally, CRP acts as a negative regulator at numerous other promoters by preventing RNA polymerase–promoter interactions, by mediating co-repression with another regulator, or by more complex mechanisms.

The extensive amino acid similarity of Vfr and CRP suggests that the two proteins may be exceedingly similar in structure and function. Closer analysis revealed that residues found to be important for binding cAMP, binding DNA and interacting with RNA polymerase in CRP are identical or conserved in Vfr (West et al., 1994). In fact, the cloned *vfr* gene can complement an *E. coli crp* deletion mutant in β-galactosidase production, tryptophanase production and cAMP-modulated catabolite repression, although not to the same levels as *crp* (West et al., 1994). Therefore, Vfr can respond to cAMP, bind to CRP-dependent promoters and productively interact with RNA polymerase to activate transcription of several *E. coli* genes. However, the cloned crp gene was unable to complement the exotoxin A and protease IV deficient phenotypes of *vfr* mutants (West et al., 1994). These data suggest that CRP is unable to perform one or more functions performed by Vfr at several Vfr-dependent promoters. These functions may include binding to the Vfr-binding sequence, properly affecting DNA geometry, interacting productively with *P. aeruginosa* RNA polymerase or other regulatory proteins, or binding to a *P. aeruginosa*-specific alternative effector molecule. In this study, we determine that Vfr interacts with the same conserved nucleotides as CRP in the *lacZ* promoter. While CRP is able to activate transcription from the *lasR* promoter containing a CCS (Albus et al., 1997), we demonstrate that it is unable to bind to divergent sequences recognized by Vfr in the *toxA* and *regA* promoters. Pattern searches of the *P. aeruginosa* genome with a preliminary Vfr consensus binding sequence (VCS) revealed Vfr-binding sequences in the promoters of additional genes and allowed the formulation of a more informed Vfr VCS.

**METHODS**

**Bacterial strains, plasmids and culture conditions.** Strains and plasmids used in this study are listed in Table 1. All strains were routinely maintained using Luria–Bertani (LB) medium (Miller, 1992) at 37°C. Antibiotics were used as necessary at the following concentrations: 100 μg ampicillin ml⁻¹ for *E. coli*; 200 μg carbenicillin ml⁻¹; 30–50 μg kanamycin ml⁻¹ and 30–100 μg gentamicin ml⁻¹ for *P. aeruginosa*.

**Recombinant DNA techniques.** Standard recombinant DNA techniques were performed as described by Sambrook et al. (1989) and Ausubel et al. (1990). *E. coli* SURE (Stratagene), XL-1 Blue (Stratagene) and HB101 (Sambrook et al., 1989) were used as host strains in the cloning experiments. Plasmid DNA was electroporated into *P. aeruginosa* as described by Smith & Igleswki (1989). Triparental matings using pRK2013 as the helper plasmid were performed as described by Figurski & Helinski (1979).

**β-Galactosidase assays.** β-Galactosidase production was measured in *E. coli* tester strains (Ebright et al., 1984b) containing pWNP28 or pRZ1306, which carry the *vfr* and *crp* genes, respectively, under control of the inducible tac promoter. The tester strains
Table 1. Bacterial strains and plasmids

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Relevant characteristics</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P. aeruginosa</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAO1</td>
<td>Prototrophic laboratory strain</td>
<td>Holloway et al. (1979)</td>
</tr>
<tr>
<td>PA103</td>
<td>Hypertoxicogenic laboratory strain</td>
<td>Holloway et al. (1979)</td>
</tr>
<tr>
<td>XAE102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RZ1330</td>
<td>MG1655 rpsL Δcrp lacP1 + ivr::Tn10 Δcya</td>
<td>W. S. Reznikoff, Univ. Wisconsin</td>
</tr>
<tr>
<td>XAE400</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE451</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-68A); -55T)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE522</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-68C); -55G)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE535</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-68T); -55A)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE461</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-67A); -56T)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE462</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-67G); -56G)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE463</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-67G); -56CG)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE471</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-66A); -57T)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE472</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-66C); -57G)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>XAE473</td>
<td>XAE102 Δcrp45 strA frn1 zci::Tn10 [i43plac5-P1(-67G); -57C)]</td>
<td>Zhang &amp; Ebright (1990b)</td>
</tr>
<tr>
<td>SURE</td>
<td>Used for cloning, wtC, umuC</td>
<td>Stratagene</td>
</tr>
<tr>
<td>XL-1 Blue</td>
<td>Used for cloning; K12 supE44 hsdR17 recA1 endA1 gyrA96 thi relA lac– F’ [proAB’ lacPlac2AM15 Tn10(TetR)]</td>
<td>Stratagene</td>
</tr>
<tr>
<td>HB101</td>
<td>Used for cloning; rII mII leuB6 thi-1 galK2 lacZ4 proA2 recA13 supE44 SmR</td>
<td>Sambrook et al. (1989)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRRK213</td>
<td>pMB1 derivative carrying the RK2 transfer functions; KmR</td>
<td>Figurski &amp; Helinski (1979)</td>
</tr>
<tr>
<td>pWNP28</td>
<td>pMMB66HE carrying the promoterless vfr gene; ApR CbR</td>
<td>Suh et al. (2002)</td>
</tr>
<tr>
<td>pRZ1306</td>
<td>pMMB66HE carrying the promoterless E. coli crp gene; ApR CbR</td>
<td>W. S. Reznikoff</td>
</tr>
<tr>
<td>pMI21</td>
<td>pUC18 carrying the toxA gene from P. aeruginosa</td>
<td>Hamood et al. (1990)</td>
</tr>
<tr>
<td>pDFA2</td>
<td>Derivative of pDF1918.202 (Frank et al., 1989) containing PstI and Smal–EcoRV deletions</td>
<td>This study</td>
</tr>
<tr>
<td>pRL88AP</td>
<td>Derivative of pRL88 (Storey et al., 1990) containing the regA promoter region</td>
<td>This study</td>
</tr>
<tr>
<td>pMMMB66HE</td>
<td>Broad-host-range tacP expression vector derived from RSF1010; lacP, ApR</td>
<td>Furste et al. (1986)</td>
</tr>
</tbody>
</table>

(listed in Table 1) contain deletions of the genes encoding wild-type CRP and LacZ. In XAE400, lacZ is placed under the control of the wild-type lac promoter, while in the other tester strains, lacZ is placed under the control of derivatives of the lac promoter containing specific substitutions in the CRP-binding sequence. Bacterial cultures were grown to an OD600 of ~0.5 in LB medium containing 1 mM IPTG and the appropriate antibiotics. β-Galactosidase activity in cell lysates was quantified as described by Miller (1992).

**CRP and Vfr purification.** CRP was purified from RZ1330/prZ1306 by cAMP-affinity chromatography as described by Ghosaini et al. (1998). Vfr purification from PA103/pWNP28 was based on cAMP-affinity chromatography and additional procedures as described by Albus et al. (1997), with modifications. Briefly, cultures were grown to an OD600 of ~0.5, at which time IPTG was added to a final concentration of 1 mM to induce overexpression of Vfr. After 6 h further incubation, the cells were harvested by centrifugation at 5000 g for 15 min. The cell pellet was resuspended in 100 mM Tris/HCl (pH 8.0), 50 mM KCl, 2 mM EDTA, 5 mM β-mercaptoethanol, 1 mM sodium azide. A crude lysate was prepared by incubation of the solution with 200 μg lysozyme ml⁻¹ for 20 min at room temperature. The lysate was adjusted to 125 mM Tris/HCl (pH 8.0), 25 mM MgSO4, 1% Brij 35, incubated at 4°C for 10 min, then passed through a 40 kPa French pressure cell at approximately 15,000 p.s.i. (~100 MPa) Cellular debris was removed by centrifugation at 16,000 g for 30 min at 4°C. An ammonium sulfate fractionation step was next performed; most of the Vfr protein precipitates between 50 and ~55% ammonium sulfate saturation. This fraction was centrifuged at 27,000 g for 30 min at 4°C. The pellet was resuspended in 50 mM Tris/HCl (pH 8.0), 2 mM EDTA, 5% (v/v) glycerol, 5 mM β-mercaptoethanol and dialysed against the same buffer to remove the ammonium sulfate. The resulting solution was then applied to a cAMP-agarose column (Sigma) as described by Ghosaini et al. (1998). Eluted fractions containing Vfr were pooled and dialysed into 10 mM sodium phosphate (pH 7.2), 2 mM EDTA, 5 mM β-mercaptoethanol, 5% glycerol. The resulting solution was applied to an SP-Sephadex cation-exchange column. Vfr-containing fractions were pooled and dialysed into 50 mM Tris/HCl (pH 7.2), 500 mM KCl, 0.2 mM EDTA, 0.2 mM DTT and 50% glycerol for storage.

**Preparation of fragments for electrophoretic mobility shift and DNase I footprinting assays.** For footprinting, DNA fragments containing the toxa and regA P1 promoters were generated by endonuclease digestion of appropriate plasmids. A 202 bp NotI–BflI fragment of the toxa promoter was isolated from pMI21; and a ~150 bp Ksl–EagI fragment containing the regA P1 promoter was isolated from pRL88AP. The fragments were end-labelled using the Klenow fill-in reaction with α-32P-labelled deoxyguanosine triphosphate.

For electrophoretic mobility shift assays, a NotI–BflI DNA fragment from pMI21 containing the toxa promoter and an EcoRI–EagI fragment from pDFA2 containing the regA P1 promoter were end-labelled using the Klenow fill-in reaction with α-32P-dNTPs. Other DNA fragments containing the algD, argY, fimS, pbpG, plcN, plcHR, prpl, pilM, ponaA, pilQ, pslA, lecA, pvdS and vfr/orfX promoters were generated by PCR of genomic DNA from P. aeruginosa PA01. Sets of
primers (Table 2) were designed to amplify regions of approximately 300 bp that contained putative Vfr-binding sites upstream of the genes of interest. The high-fidelity polymerases Pfu (Stratagene) or Pfx (Invitrogen) were used in the amplification reactions according to the manufacturers’ specifications. Radiolabelling of the DNA fragments was accomplished by incorporation of [a-32P]dCTP during PCR.

Electrophoretic mobility shift assays (EMSA). Binding reactions were performed as described by DeVault et al. (1991) with minor modifications. Briefly, the radiolabelled DNA fragments, standardized to approximately 1000 c.p.m. per reaction (15–40 pM), and various concentrations of purified Vfr or CRP were incubated in 10 mM Tris/HCl (pH 7.4), 100 mM KCl, 1 mM EDTA, 5% glycerol, 1 mM DTT, 50 µg ml⁻¹ BSA, 20 µM cAMP and 10 µg ml⁻¹ poly(dI-dC)-poly(dI-dC) for 20 min at room temperature. Competition assays with unlabelled target PCR fragments were performed to confirm specificity of the binding reactions (data not shown). The samples were electrophoresed on native 6% polyacrylamide gels at 10 V cm⁻¹ at 4 °C. The gels were dried and exposed to Kodak X-OMat AR5 film overnight at −70 °C with an intensifying screen.

DNase I footprinting. Binding reactions were performed by incubating purified Vfr with the radiolabelled DNA fragments in 10 mM Tris/HCl (pH 7.4), 100 mM KCl, 10 mM MgCl₂, 5% glycerol, 1 mM dithiothreitol, 50 µg ml⁻¹ BSA and 20 µg ml⁻¹ cAMP for 20 min at room temperature. DNase I (0.75–1.5 µg ml⁻¹) was then added to the samples. The DNase I digestion reactions were incubated for 30 s at room temperature and then terminated by the addition of EDTA to 12.5 mM, sodium acetate to 200 mM and phenol to 30% (v/v) final concentrations. The DNase I digestion products were ethanol-precipitated after the addition of 10 mg glycogen and then resuspended in 0.5× TBE buffer (45 mM Tris/ borate, 1 mM EDTA, pH 8.3) containing 8 M urea, 0.5% bromophenol blue and 0.5% xylene cyanol. The samples were analysed on a 7 M urea/8% polyacrylamide gel to determine the regions protected by Vfr. Maxam–Gilbert chemical sequencing reactions (Maxam & Gilbert, 1980) were also performed with the radiolabelled DNA fragments to provide sequence ladders for comparison with the footprinting reactions.

Sequence analysis. Sequence searching for putative Vfr-binding sites was performed using the program FindPatterns of the Wisconsin Package Genetics Computer Group (GCC) Version 9.1 and the complete P. aeruginosa genome sequence available at the Pseudomonas Genome Project, http://www.pseudomonas.com (Stover et al., 2000). A Vfr-binding site sequence logo was created from aligned Vfr-binding sequences and flanking regions using the WebLogo program available from the University of California, Berkeley (http://weblogo.berkeley.edu) (Crooks et al., 2004; Schneider & Stephens, 1990).

### RESULTS

Vfr binds to the lacZ promoter CRP-binding site with a similar specificity to that of CRP

It was previously shown that P. aeruginosa Vfr can replace CRP to activate β-galactosidase production in E. coli (West et al., 1994). Based on the considerable identity between the two molecules, especially in the helix–turn–helix DNA-binding motif, we hypothesized that Vfr may interact with the consensus CRP-binding sequence (CCS) in a manner similar to CRP. To test this hypothesis, we examined the effect of basepair substitutions in the CRP-binding sequence on the ability of Vfr to activate lacZ transcription. The E. coli ‘tester’ strains used in this analysis (Table 1) contain a deletion of the crp gene, making β-galactosidase production dependent on provision of either crp or vfr in trans (Ebright et al., 1984b). Additionally, each strain contains symmetric basepair substitutions at position 5, 6 or 7 in each half of the palindromic CRP-binding sequence present in the lacZ promoter. We measured β-galactosidase production in each ‘tester’ strain after introduction of either pWNP28 or pH31306, carrying the cloned vfr or crp gene, respectively (Fig. 1). Strains containing the vector control, pMMB66HE, produced negligible amounts of β-galactosidase. We found that, with CRP, Vfr-dependent activation of the lacZ promoter decreased when mutations were present at positions 5, 6 or 7 of the CRP-binding sequence half-site. Vfr preferred the G:C basepair (G>T or A>C) at position 5, the T:A basepair (T>G or A>C) at position 6, and the G:C basepair (G≡A=C≡T) at position 7. Based on these data, we conclude that in E. coli, P. aeruginosa Vfr interacts with the same basepairs as CRP in the lacZ promoter CRP-binding site.
The toxA or regA promoters are bound by Vfr but not by CRP

Since Vfr has been implicated in regulating the expression of exotoxin A and RegA (West et al., 1994), EMSA using DNA fragments from the promoters of these genes were performed to confirm binding of Vfr. Incubation with purified Vfr altered the mobility of the DNA fragments carrying the toxA and regA P1 promoter regions (Fig. 2a), respectively. We have previously shown that Vfr is required for the activation of the lasR promoter and binds the CCS present in the lasR promoter (Albus et al., 1997). CRP is also able to bind to this CCS in vitro and can also complement LasR production in a vfr mutant (data not shown). However, CRP does not complement the exotoxin A deficient phenotype of a Vfr mutant (West et al., 1994) and does not appear to activate expression from the regA P1 promoter (data not shown). To determine if these effects are due to lack of binding of CRP to these promoters, EMSA was performed using purified CRP and the toxA and regA P1 promoter fragments. E. coli CRP did not alter the mobility of either toxA and regA P1 promoter DNA fragments (Fig. 2b). Since both Vfr and CRP can bind the lasR and lacZ promoters in EMSA, but only Vfr can bind the regA P1 and toxA promoters (Fig. 2), it is possible that in P. aeruginosa Vfr recognizes a binding site that is different from the CCS in the promoters of some of the genes that it regulates.

Vfr recognizes a degenerate CCS in the toxA and regA promoters

To determine the sequence to which Vfr binds in the promoters it regulates, DNase I footprinting analyses were performed using single-stranded DNA fragments representing the toxA, regA P1 and lasR promoters. Vfr protected a 28 bp sequence (5'-CCACTCTGCAATCCAGTTCATAAATCC-3') located from 53 to 78 bp 5' of the S1a transcriptional start site of toxA (Fig. 3a). This sequence corresponds to a region in which Tsaur & Clowes (1989) found that linker-substituted mutations decreased exotoxin A production. Based on the localization of the Vfr-protected sequence to this region, we propose that the decreased synthesis of exotoxin A in these toxA promoter linker substituted mutants was due to the alteration of the Vfr-binding sequence. Vfr also protected a 27 bp sequence (5'-GTAACAGCGGAACCACTGCACAGACCA-3') located 51–77 bp 5' of the regA P1 transcriptional start site (Fig. 3b). Footprinting of the opposite DNA strands for each promoter showed similar corresponding areas of protection (data not shown).
Within the toxA and regA Vfr-binding sequences, we observed sites with increased sensitivity to DNase I when Vfr was bound to the DNA fragments. The presence of these Vfr-dependent DNase I hypersensitivity sites is indicative of a structural perturbation in the helical structure of the DNA such as a kink or bend (Baichoo & Heyduk, 1999; Brierley & Hoggett, 1992; Giraud-Panis et al., 1992). Vfr-dependent DNase I hypersensitivity sites were seen at phosphodiester bonds located one helical turn apart within Vfr-protected sequences (Fig. 3a, b, and data not shown). Vfr protected the CCS present in the lasR promoter (data not shown), whose alteration was previously seen to decrease lasR expression by Vfr (Albus et al., 1997). Vfr-dependent DNase I hypersensitivity sites were also apparent at phosphodiester bonds located one helical turn apart in the Vfr-protected lasR promoter sequence (data summarized in Fig. 3c).

To identify a preliminary consensus Vfr-binding site, we aligned the sequences protected by Vfr in each promoter based on the position of the conserved Vfr-dependent DNase I hypersensitivity sites observed between the T and G nucleotides in the bottom strand (Fig. 3c). This 5’-XTG-3’ sequence is highly conserved in different CRP-binding sites in E. coli promoters and is the site of an observed 40° kink in DNA bound by CRP (Schultz et al., 1991). Based on this alignment, 6 of 22 basepairs were identical between the overlapping regions of all three sequences (Fig. 3c). A preliminary Vfr consensus sequence (VCS) (5’-ANA-TGNGATCCACTTCACATTT-3’) was deduced from this comparison. This preliminary consensus binding sequence is quite similar to the CCS, having only 7 differences in the 22 overlapping basepairs. Thus, the regA P1 and toxA promoters appear to be bound by Vfr at a
degenerate CRP-binding sequence not able to be bound by CRP itself.

**Vfr-binding sites are present in the promoters of other *P. aeruginosa* genes**

To identify additional Vfr-binding sites upstream of *P. aeruginosa* genes, the full *P. aeruginosa* sequenced genome (Stover et al., 2000) was searched with the deduced preliminary VCS. Several putative matches upstream of virulence-associated genes, including *algD*, *argH/fimS*, *pbpG*, *plcN*, *plcHR*, *prpL*, *plIM/ponA*, *plIQ*, *psrA/lexA*, *pvdS* and *vfr/orfX* (Table 3), were assessed for in vitro Vfr binding using EMSA (Fig. 4). The DNA fragments used in these assays were amplified and radiolabelled using PCR. The primers are given in Table 2. In a previous study, CRP was shown to bind to only one of the two regions containing a putative CRP-binding sequence upstream of the *algD* operon (De Vault et al., 1991), which comprises 12 genes whose products are involved in the biosynthesis of the exopolysaccharide alginate (Chitnis & Ohman, 1993; Gacesa, 1998). In this study, the same result was obtained using Vfr. Vfr bound to the far upstream (FUS) promoter region containing a putative Vfr-binding sequence centred at −362:5 bp relative to the transcriptional start site (Fig. 4) and not to a putative binding sequence in the proximal promoter region. Vfr was also able to retard the mobility of DNA fragments containing putative Vfr-binding sequences identified upstream of the following genes: *plcN*, which encodes a non-haemolytic phospholipase C (Ostroff et al., 1990); *plcHR*, encoding a haemolytic phospholipase C (*PlcH*) in an operon with two overlapping genes encoding the PlcR1 and PlcR2 regulators (Ostroff et al., 1990; Shen et al., 1987); *pbpG*, which encodes penicillin-binding protein 7 (Song et al., 1998); and *prpL*, encoding the PvdS-regulated endoprotease lysyl class (protease IV) (Wilderman et al., 2001) (Fig. 4). Vfr was also found to bind to fragments containing the intergenic region between *argH* and *fimS*, encoding argininosuccinate and a regulator of twitching motility, respectively (Haas et al., 1977; Whitchurch et al., 1996), and the intergenic region between *plIM* and *ponA*, encoding the first gene of an operon (*plIM–P*) involved in biogenesis of type 4 fimbriae, and the high molecular mass penicillin-binding protein 1A (PBPA) respectively (Handfield et al., 1997; Martin et al., 1995) (Fig. 4). Finally, Vfr was observed to bind to a DNA fragment representing the intergenic region between its own transcriptional start site and that of the divergently transcribed *orfX* gene (Runyen-Janecky et al., 1997), whose role is unknown. Vfr-binding was not demonstrated for DNA fragments containing putative sites in the intergenic region between *prsA* and *lexA* (Garriga et al., 1998), or the *pvdS* (Cunliffe et al., 1995) or *plIQ* (Martin et al., 1993) promoter regions.

**Identification of a VCS**

Ten of the Vfr-binding regions so far identified for Vfr in *P. aeruginosa* were aligned based upon sequence similarity of the identified putative Vfr-binding sites. The multiple putative Vfr-binding sites identified in the intergenic regions of the *plIM/ponA* and the *vfr/orfX* divergent gene pairs were not included in the analysis. A sequence logo (Schneider & Stephens, 1990) was created from the ten aligned Vfr-binding sequences and small flanking sequences

---

**Table 3. Sequences and locations of putative Vfr-binding sites in promoters of virulence-associated genes**

Putative Vfr-binding sites were identified by searching the *P. aeruginosa* genome (Stover et al., 2000) for sequences similar to the preliminary VCS: 5′-ANATGNGATCC:ACTTCACANAT-3′. Sequences corresponding to the putative conserved half-sites are indicated in bold. The location of the pseudodyad axis (−) of each putative binding site within the sequence and in relation to known transcriptional start sites (TX) or translational start sites (TL) of the corresponding promoters is indicated.

<table>
<thead>
<tr>
<th>Promoter</th>
<th>Putative Vfr-binding sequences</th>
<th>Location (bp) upstream of transcriptional (TX) or translational (TL) start</th>
<th>Shifted by Vfr</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>algD</em> FUS</td>
<td>CATGCCGAACCT:GCACTACATTT</td>
<td>362:5 (TX)</td>
<td>+</td>
</tr>
<tr>
<td><em>algD</em> proximal</td>
<td>AAGTGATATCA:AAAGGAATATT</td>
<td>77:5 (TX)</td>
<td>-</td>
</tr>
<tr>
<td><em>argH/fimS</em></td>
<td>TTCTTTCGGC:AGTGCCATTT</td>
<td>143:5/81:5 (TL)</td>
<td>+</td>
</tr>
<tr>
<td><em>pbpG</em></td>
<td>TATGGGACC:AGGCCACTG</td>
<td>54:5 (TL)</td>
<td>+</td>
</tr>
<tr>
<td><em>plIM/ponA</em></td>
<td>GCGAATGCT:AGGTCACAGAA</td>
<td>87:5/97:5 (TL)</td>
<td>+</td>
</tr>
<tr>
<td><em>plIQ</em></td>
<td>TAAATGGAACA:GGCTGGCCTT</td>
<td>119:5/65:5 (TL)</td>
<td>+</td>
</tr>
<tr>
<td><em>plcHR</em></td>
<td>GAGCCTGATCG:AAATCGTTCT</td>
<td>85:5 (TL)</td>
<td>-</td>
</tr>
<tr>
<td><em>plcN</em></td>
<td>TATGGGATG:AGGCCAGTA</td>
<td>240:5 (TX)</td>
<td>+</td>
</tr>
<tr>
<td><em>prpL</em></td>
<td>TACTGGATACC:GATCCACATG</td>
<td>160:5 (TL)</td>
<td>+</td>
</tr>
<tr>
<td><em>psrA/lexA</em></td>
<td>GATGATTTCC:GCAGCACTTT</td>
<td>70:5/128:5 (TL)</td>
<td>-</td>
</tr>
<tr>
<td><em>pvdS</em></td>
<td>CAGCTGATGG:TTGGATATTT</td>
<td>54:5 (TX 1), 80:5 (TX 2)</td>
<td>-</td>
</tr>
<tr>
<td><em>orfX/vfr</em></td>
<td>GCACGCTAGG:GCATCACAGT</td>
<td>38:5/67:5 (TX)</td>
<td>+</td>
</tr>
</tbody>
</table>
using the WebLogo program (Crooks et al., 2004) available from the University of California, Berkeley (http://weblogo.berkeley.edu/) (Fig. 5). The region of conservation corresponding to the aligned Vfr-binding sites extends farther upstream of the pseudodyad axis than it does downstream; however, the region of highest conservation is located in the downstream half-site (TCACA). The frequencies of individual bases were calculated for each position in the highly conserved Vfr-binding region identified in Fig. 5. From this analysis, we deduced the following 21 bp VCS: AnWWTGnGAWnY : AGWTCACAT (Fig. 5). The efficacy of pattern searching of the P. aeruginosa genome with the new VCS was validated by identifying six additional putative Vfr-binding sequences upstream of genes shown by Wolfgang et al. (2003) to be regulated by Vfr and/or cAMP using microarrays (Table 4). These six putative binding sequences range from four to seven mismatches from the newly derived VCS. The majority of sequences used to formulate the new VCS differ from the VCS by four mismatches (algD, plcN, plcH, lasR, fleQ, toxA, ppcG and prpL).

**DISCUSSION**

Vfr shows striking sequence similarity to its homologue, CRP, having 67% identical residues and 91% overall similarity (West et al., 1994). This suggests that Vfr also maintains the structural and functional characteristics of CRP. Previous data have shown that Vfr can restore β-galactosidase activity in an E. coli crp mutant (West et al., 1994). In this study, we determined that Vfr interacts with the same residues in the lacZ promoter that are used by CRP. Basepair substitutions at positions 5, 6 and 7 in the near-palindromic half-sites of the binding sequence decreased Vfr-dependent activation of the lacZ promoter. The residues Arg-180, Glu-181 and Arg-185, found in the helix–turn–helix DNA binding motif of CRP, make direct contact with some of the conserved bases in the CRP-binding sequence (Ebright et al., 1984a, 1987; Zhang & Ebright, 1990a). All three of these residues are identical in Vfr. In fact, the helix–turn–helix motif of Vfr only differs by four amino acids from the same motif in CRP. Therefore, Vfr appears to form the same hydrogen bonds as CRP with bases at positions 5 and 7 in the binding sequence. These data agree with previous findings showing a significant reduction in expression from the lasR promoter when positions 5 and 7 in the CCS present in this promoter were mutated (Albus et al., 1997).

While both Vfr and CRP can activate transcription from the lacZ promoter, we found that only Vfr can activate transcription from the regA and toxA promoters. There are several possible explanations for why the homologous CRP and Vfr proteins may not activate transcription from the same promoters. First, CRP and Vfr may not be able to bind to all of the same sequences. Although strikingly similar, the helix–turn–helix DNA binding motifs of Vfr and CRP are not identical. Differences at other positions within the proteins may also affect binding by restricting or allowing conformational changes required for efficient binding at specific sites. Secondly, CRP and Vfr may require different characteristics in the sequences flanking the binding sites to mediate proper DNA geometry or proper spacing for activation at individual promoters. Thirdly, it is
also possible that Vfr and CRP use different mechanisms to interact with RNA polymerase. Vfr contains identical AR1 and AR2 activating regions as are present in CRP. However, unlike wild-type CRP, Vfr contains an intact AR3 activating region (West et al., 1994). Fourthly, Vfr and CRP may not be able to interact with the same coregulatory molecules at specific promoters. Fifthly, Vfr and CRP may respond to different levels of cAMP or may respond to different effector molecules in some instances. Finally, a combination of the above differences may affect the ability of Vfr and CRP to activate transcription at the same promoters.

In the cases of the regA and toxA promoters, we determined that CRP was not able to activate transcription, most likely because it could not bind to the Vfr-binding sequences. Compared to the CCS, the regA and toxA Vfr-binding

Table 4. VCS pattern searching reveals putative Vfr-binding sites within promoters of Vfr regulon-associated genes

The P. aeruginosa genome (Stover et al., 2000) was searched for sequences similar to the newly derived VCS: 5’-ANW-WTGNGAWNY:AGWTCACAT-3’. Putative Vfr-binding sequences were found within the promoter regions of several genes previously identified by Wolfgang et al. (2003) as being transcriptionally modified in P. aeruginosa strains containing mutations of the vfr and/or cyaAB genes. The number of mismatches from VCS and corresponding gene promoter is listed for each putative binding site. Sequences corresponding to the putative conserved half-sites are indicated in bold. The location of the pseudodyad axis (:) of each putative binding site is also indicated.

<table>
<thead>
<tr>
<th>No. of mismatches</th>
<th>Putative Vfr-binding sequence</th>
<th>Promoter</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>ACAAGGAGAGCT:AGCTTACAT</td>
<td>Overlaps start codon of pa5291 (probable choline transporter betT2)</td>
</tr>
<tr>
<td>4</td>
<td>AAATTCGGATTG:AGAGCAAC</td>
<td>pa4697 (unknown)</td>
</tr>
<tr>
<td>4</td>
<td>GCACGTCGCGCC:AGATCAGAA</td>
<td>pa5402 (unknown)</td>
</tr>
<tr>
<td>6</td>
<td>CGCATGTAAGC:CACTCAGAT</td>
<td>pa3452 (probable malate:quinone oxidoreductase mqoA)</td>
</tr>
<tr>
<td>7</td>
<td>TCTCTGTGACGA:ATGTTCAT</td>
<td>pa2560 (unknown)</td>
</tr>
<tr>
<td>7</td>
<td>AACATGTGACAG:GCACTCACCC</td>
<td>pa2883 (unknown)</td>
</tr>
</tbody>
</table>
sequences contained 10 of 22 and 11 of 22 of the same basepairs, respectively. Among the 129 enterobacterial CRP-binding sequences that have been identified, the average number of conserved basepairs is 13 (Busby & Kolb, 1996). However, the binding of CRP to promoters containing CRP-binding sequences with 11 or fewer conserved basepairs has been documented (Busby & Kolb, 1996). Therefore, CRP’s inability to bind to the _toxA_ and _regA_ Vfr-binding sites may be due to the specific combination of basepairs or the flanking sequences of the binding sites.

We aligned the nucleotides in the Vfr-protected sequences and flanking regions in the _toxA_, _regA_ and _lasR_ promoters based on the DNase I hypersensitivity sites and the pseudo-dyad axis. We found that the patterns of A-T or G-C richness in the regions flanking the Vfr-binding sequences in the _toxA_ and _regA_ promoters were the opposite of those considered optimal for CRP binding to the CRP-binding sequence (Liu-Johnson _et al._, 1986). The regions flanking the Vfr-binding sequences may either prevent binding of CRP or promote binding of Vfr via an interaction with a unique structural element in Vfr. In contrast to the _E. coli_ genome (Blattner _et al._, 1997), which has a G+C content of approximately 50 mol%, the _P. aeruginosa_ genome has a G+C content of approximately 67 mol% (Stover _et al._, 2000). The Vfr protein and regulatory mechanisms may have evolved to successfully utilize sequences with the higher G+C content within the _P. aeruginosa_ genome.

When bound to a binding sequence, Vfr also appears to induce some of the same perturbations in DNA geometry as does CRP. DNase I hypersensitivity sites were apparent in each of the _toxA_, _regA_ and _lasR_ Vfr-protected footprints. These hypersensitivity sites are positioned approximately one helical turn apart, and align perfectly with those produced when CRP is bound to its target DNA sequences. The previously published Vfr-protected footprint in the _fleQ_ promoter region also contains these hypersensitivity sites (Dasgupta _et al._, 2002). Based upon our data, we hypothesize that Vfr wraps the DNA around itself, producing approximately 40° kinks between positions 5 and 6, and stabilizing the Vfr-DNA complex by allowing additional contacts between the DNA and the ‘sides’ of the bound Vfr dimer.

Alignment of the three Vfr-binding sequences that we confirmed with DNase I footprinting allowed us to determine a preliminary Vfr consensus sequence (VCS). We used this preliminary VCS to successfully search the complete genomic sequence of _P. aeruginosa_ for previously unrecognized Vfr-binding sequences upstream of several virulence-related genes. Evidence for Vfr regulation of several of these genes, specifically _plic_ and _pbgG_, has recently been demonstrated by Wolfgang _et al._ (2003) by microarray comparison of the transcriptomes of wild-type versus _vfr_-mutant _P. aeruginosa_ strains. Additionally, Vfr has been implicated in regulating type IV _pil_ formation and twitching motility, which involve the _fimS_ and _pilM_ genes/operons (Beatson _et al._, 2002; Wolfgang _et al._, 2003). Interestingly, we did not observe binding of Vfr to the promoter of _pilQ_, which was previously identified as a Vfr-regulated gene by Wolfgang _et al._ (2003). The observed regulation may be indirect, or binding of Vfr to the _pilQ_ promoter may require conditions different from those used in this study. Although there is evidence for Vfr-dependent regulation at several of the _in vitro_-bound target sequences, binding of a transcriptional regulator to a DNA sequence does not always correlate with _in vivo_ regulation, and future studies will aim to address the role of Vfr in the regulation of specific genes. The identified Vfr-binding sites are located at varying distances from the transcriptional start sites or coding regions of the corresponding genes, which indicates that Vfr-dependent promoters may be divided into different classes based upon different mechanisms of activation/repression.

By aligning 10 known Vfr-binding sequences (eight determined in this study and previously identified sites present in the _lasR_ and _fleQ_ promoters: Albus _et al._, 1997; Dasgupta _et al._, 2002), we developed a more informed VCS (5’-ANWWTGNGAWNY_AGWTCACAT-3’). As in the CCS, two conserved half-sites (underlined) were evident in the alignment used to develop the VCS. However, in CRP-dependent promoters, the 5’-TGTGA half-site is more conserved (Pyles _et al._, 1998): this is the opposite of our deduced VCS, with the TCACA half-site closest to the promoter being more conserved. Interestingly, this VCS is not completely symmetrical as a whole. Although there is a pseudo-dyad axis between the two highly conserved half-sites, there appears to be a preference for an A/T-rich sequence (ANWW) immediately upstream of the 5’-half-site of the VCS. Only a T residue is preferred immediately downstream of the 3’ half-site. Pattern searching of the _P. aeruginosa_ genome with the newly derived VCS was also successful in identifying putative Vfr-binding sequences upstream of several genes identified by Wolfgang _et al._ (2003) to be transcriptionally affected by _vfr_ and/or _cyaAB_ mutation. Confirming Vfr binding to these and other sequences will allow the VCS to be further refined. These data serve to emphasize the utility of pattern searching to identify genes within the regulon of DNA-binding transcriptional regulators. With the recent and continuing completion of numerous bacterial genome sequences, the utility of this straightforward but robust method of elucidating transcriptional regulons should not be overlooked.

The identified Vfr-binding sequences diverge from the consensus to varying degrees, which may imply that there is a hierarchy of regulation, where changes in the concentration of Vfr or levels of its allosteric effector _cAMP_ may cause different sets of genes to be activated/repressed at different times, in accordance with different external and/or internal signals. Like CRP, it appears that Vfr may use multiple mechanisms of transcriptional control at a wide range of promoters. However, it also appears that although Vfr shares many of the functions of CRP, it has also developed functions not shared by CRP that are specific to its role as a
global virulence factor regulator in *P. aeruginosa*. Derouaux *et al.* (2004) similarly demonstrated that the CRP homologue in *Streptomyces coelicolor* (CRP<sup>SCO</sup>) could not complement the lactose- and maltose-negative phenotypes of an *E. coli* CRP mutant. Comparison of shared and divergent functions may add to our knowledge about the individual regulators in this family and the evolution of prokaryotic regulatory pathways.

**ACKNOWLEDGEMENTS**

This study was supported by Public Health Service grant AI31477 from the National Institute of Allergy and Infectious Diseases awarded to S. E. H. W.; J. J. R.-J. was a trainee under Public Health Service grant T32 GM 07215. We would like to thank W. Hendrickson, N. A. Zielinski and Wilma E. Ross for their helpful discussions, W. S. Reznikoff for *E. coli* RZ1330 and pRKZ1306, and R. H. Ebright for the *E. coli* ‘tester’ strains.

**REFERENCES**


Derouaux, A., Deharenge, D., Lecocq, E. & 9 other authors (2004). Crp of *Streptomyces coelicolor* is the third transcription factor of the large CRP-FNR superfamily able to bind cAMP. *Biochem Biophys Res Commun* 325, 983–990.


