Relevance of nucleotides of the PcaU binding site from *Acinetobacter baylyi*

Bettina Jerg† and Ulrike Gerischer

Institute for Microbiology and Biotechnology, University of Ulm, D-89069 Ulm, Germany

Results from a random mutagenesis procedure on the PcaU binding site from *Acinetobacter baylyi* followed by *in vivo* and *in vitro* screening are presented. PcaU is an IclR-type transcriptional regulator from the soil bacterium *A. baylyi* and is required for the regulated expression of enzymes for protocatechuate and quinate degradation encoded by the *pca-qui* operon. It binds to a 45 bp area located in the *pcaU–pcaI* intergenic region which consists of three perfect 10 bp sequence repeats forming one palindrome (R1, R2) and an additional direct sequence repeat (R3). *In vivo* selection for *pca-qui* gene expression revealed that mutations within the three sequence motifs are tolerated to different extents. The functional requirement for conserved nucleotides was greatest in the external half of the palindrome (R1). Four positions within and directly adjacent to this 10 bp sequence never acquired a mutation, and are therefore considered to be the most important for transcriptional regulation by PcaU. Transcriptional output is affected in different ways; for some of these changes there is a correlation with a reduction in the affinity of PcaU for these sites. Two of these positions were also preserved when *in vitro* screening was performed for PcaU binding alone. Additional conserved residues are detected by the *in vitro* approach, indicating that the regions of the PcaU binding site involved in binding differ, at least in part, from those required for functional gene expression.

INTRODUCTION

The interaction between a DNA-binding protein and its specific binding site on the DNA molecule is crucial in gene regulation, and this is closely related to the question of how these proteins find their binding sites (Freemont et al., 1991; Halford & Marko, 2004). Some regulatory proteins recognize multiple specific binding sites, which in most cases have a similar but not identical sequence. Aligning such known binding sites makes it possible to deduce which positions of the binding sites are crucial, as has been done, for example, for the cAMP receptor protein (CRP) of *Escherichia coli* or more recently for the Vfr protein from *Pseudomonas aeruginosa* (Kanack et al., 2006; Kolb et al., 1993). Other regulators have only one or a few known binding sites, which is not sufficient to determine a consensus sequence. In such cases the relevance of certain nucleotides within a regulator binding site can be evaluated in a powerful experimental approach whereby randomly modified binding sites are offered to the protein in different screening setups (SELEX: systematic evolution of ligands by exponential enrichment). This approach has been used, for example, to determine the binding site of the CopR repressor from *Bacillus subtilis* (Freede & Brantl, 2004) and to define promoter recognition by σ^F^ in the same bacterium (Amaya et al., 2001).

The PcaU protein from *Acinetobacter baylyi* is an IclR-type transcriptional regulator involved in the control of gene expression in aromatic compound degradation (Gerischer et al., 1998). In this bacterium, aromatic compounds are converted to protocatechuate or catechol by a number of funnelling pathways. These compounds are the starting molecules for the breakdown of the aromatic core by the ortho cleavage (also referred to as the β-ketoadipate) pathway (Harwood & Parales, 1996). All enzymes necessary for protocatechuate degradation are encoded by the *pca* genes, which form an operon. This operon includes the cluster of four downstream *qui* genes involved in the conversion of the hydroaromatic compounds quinate and shikimate (*pca-qui* operon, Fig. 1; Dal et al., 2005). The *pcaU* gene is located upstream of this large (14 kb) operon and transcribed divergently from it. The 309 bp intergenic region between *pcaU* and the *pca-qui* operon (PcaU is shorter than originally determined; B. Jerg and U. Gerischer, unpublished observation) contains the large (45 bp) PcaU binding site (Fig. 1) (Popp et al., 2002). It contains three perfectly conserved 10 bp repeats, here referred to as R1, R2 and R3. R1 and R2 form a palindrome; R3 is a direct sequence repeat of R2 separated by 10 bp. This architecture makes the PcaU binding site exceptional in comparison with binding sites of other regulator proteins, which are most frequently formed by a sequence of less perfectly conserved sequence motifs with dyad symmetry (Tropel & van der Meer, 2004).
Binding of PcaU to this site is an obligatory step in the protocatechuate-dependent induction of the pca-qui operon. In the absence of protocatechuate, binding of PcaU to its intergenic binding site is necessary to reduce the basal expression to the wild-type level. In the absence of PcaU this expression level increases significantly. Thus, PcaU has been shown to be an activator as well as a repressor at the same promoter (pcaIp), depending on the concentration of the effector protocatechuate (Siehler et al., 2007). The same pattern of expression has been observed for the pcaU gene itself. Besides the PcaU-dependent expression control of the pca-qui operon, there are at least two more systems controlling pca-qui operon and pcaU expression which act at higher levels and in both cases reduce expression (Gerischer et al., 2008). Cross-regulation applies when in addition to substrates degraded via protocatechuate, substrates of a parallel branch (with catechol as the central compound) of the pathway are offered simultaneously. This situation leads to repression of pcaU gene and pca-qui operon expression (Brzostowicz et al., 2003; Siehler et al., 2007). The two regulators BenM and CatM are involved in this mechanism. They are primarily responsible for the activation of expression of the ben and cat genes, which encode enzymes of the second branch of the pathway for the breakdown of benzoate and catechol. The third regulatory level recognized is carbon catabolite repression of pcaU and the pca-qui operon upon the simultaneous presence of an aromatic substrate and the organic acids acetate and succinate (Dal et al., 2002; Siehler et al., 2007).

Given the central relevance of PcaU binding to the expression of two transcripts, as well as the fact that understanding of the molecular mechanism of the function of IclR-type regulators is still limited (Molina-Henares et al., 2006), it seemed interesting to analyse the relevance of the PcaU binding site for gene expression and for PcaU binding alone. Here we describe experiments in which the PcaU binding site has been randomly modified and incorporated into the organism. Derivatives of the PcaU binding site that enabled the cell to express the pca-qui operon at enhanced levels were selected and characterized. In a further approach, in vitro screening of randomly modified PcaU binding sites was performed, searching for sequences that were still bound by PcaU.

**METHODS**

**Bacterial strains and growth conditions.** The bacterial strains used in this study are listed in Table 1. A. baylyi strains were cultivated at 30 °C in mineral medium, as described previously (Trautwein & Gerischer, 2001), supplemented with the carbon sources lactose (10 mM), succinate (10 mM) or p-hydroxybenzoate (5 mM as sole carbon source or 1 mM in addition to lactose for induction). Antibiotics for A. baylyi strains were used at the following
concentrations: kanamycin, 12 µg ml⁻¹; spectinomycin, 100 µg ml⁻¹; streptomycin, 10–50 µg ml⁻¹. *E. coli* strains were grown in Luria broth supplemented with the appropriate antibiotic (ampicillin, 100 µg ml⁻¹) at 37 °C.

**Plasmid and strain construction.** Inactivation of the *pcaU* gene in strains of *A. baylyi* was done by deleting a *pcaU* internal fragment and replacing it with a cassette designed to terminate transcription as well as translational activities (the Ω fragment from plasmid pHP45Ω). Plasmid pAC64 was cleaved with MscI and PmlI, releasing a 699 bp *pcaU* internal fragment. The Ω fragment (Smal-cut) was inserted and insertion was confirmed by the streptomycin resistance of the respective clones. The resulting plasmid (pAC94) was verified by restriction analysis. A fragment containing the modified *pcaU* gene was released with the restriction enzymes SnaBI/AflII and used to transform A. baylyi strains, followed by selection on mineral medium with succinate supplied with streptomycin (10 µg ml⁻¹). The specific integration of the DNA fragment into the *Acinetobacter* genomes was confirmed by whole-cell PCR with primers ON3000 and UI-2604 (Table 2).

Plasmid pAC53 (used as PCR template, Fig. 2) contained a randomized PcaU binding site (41 nucleotides containing R1, R2 and R3, plus one additional nucleotide on either side replaced by GGATCATTTGTAATTGGTACGATATTTAGTATCTATCATGACG). It was constructed by overlap-extension PCR using plasmid pZBr as a template; the resulting fragment was cleaved with EcoRI and NotI (at sites native to *A. baylyi*), resulting in a 335 bp fragment, which was cloned into the insert in pAC40, resulting in plasmid pAC53.

For the creation of transcriptional *pcaU-*luc fusions, strains of *A. baylyi* were transformed with an *EcoRI/Sall* fragment cleaved out of plasmid pAC64 containing a *lac-aad9* cassette in *pcaU*. Transformants were screened for their ability to grow on mineral medium with succinate and spectinomycin. Colony PCR using primers *pca-I/+* and *Sall*-cut was used to confirm the identity of the constructed strains.

**Genetic modification of *Acinetobacter* by natural transformation.** Recipient strains of *A. baylyi* were transformed with restriction fragments or DNA fragments produced by PCR, as described previously (Trautwein & Gerischer, 2001).

**Determination of luciferase activity.** All data reported are representative of at least three independent experiments. Aliquots of cell suspensions were withdrawn every hour, covering the complete growth curve, and the OD₆₀₀ was measured. The reservoirs of the luminometer (Flash’n Glow, Berthold Detection Systems) were filled with 2.5 x assay buffer (62.5 mM glycglycine, pH 7.8, 25 mM MgCl₂) and 33 µM t-luciferin (freshly dissolved in water). After placing the sample tube with 150 µl cell suspension into the luminometer, 150 µl of each solution was added and the luminescence was immediately measured for 15 s at 22 °C. Light emission was given in relative light units (RLU) and normalized to the OD₆₀₀.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. baylyi strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADP1</td>
<td>Wild-type (strain BD413, ATCC 33305)</td>
<td>Juni &amp; Janik (1969); Vaneechoutte et al. (2006)</td>
</tr>
<tr>
<td>ADPU1</td>
<td>ΔpcaP</td>
<td>Dal et al. (2005)</td>
</tr>
<tr>
<td>ADPU47</td>
<td>*pcaI-*luc transcriptional fusion inserted into <em>pcaU</em></td>
<td>Siehler et al. (2007)</td>
</tr>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>F⁻, dhdZAM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, hsdR17(λ⁻, m₉g, m₉u), phoA, supE44, λ⁻, thi-1, gyrA96, relA</td>
<td>Hanahan (1983)</td>
</tr>
<tr>
<td>BL21 AI</td>
<td>F⁻, ompT, hsdR30 (λ⁻, m₉g, m₉u), gal, dcm, araB::Tn5::R6K-2a tetA</td>
<td>Grodberg &amp; Dunn (1988); Invitrogen</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pHp45Ω</td>
<td>Ω fragment for <em>in vitro</em> insertional mutagenesis, <em>aadA</em>+ (Sm², Spc²)</td>
<td>Prentki &amp; Krisch (1984)</td>
</tr>
<tr>
<td>pDrive</td>
<td>TA cloning vector</td>
<td>Qiagen</td>
</tr>
<tr>
<td>pBluescript II SK(+)</td>
<td>Ap⁴, lacZ</td>
<td>Alting-Mees &amp; Short (1989); Stratagene Europe</td>
</tr>
<tr>
<td>pET29(+)</td>
<td>T7 promoter expression plasmid</td>
<td>Studier et al. (1990); Novagen (Merck)</td>
</tr>
<tr>
<td>pAC7</td>
<td>2792 bp <em>EcoRI</em> fragment from <em>A. baylyi</em> containing <em>pcaU</em> in <em>pET21(+)</em> for expression of native PcaU</td>
<td>Popp et al. (2002)</td>
</tr>
<tr>
<td>pAC40</td>
<td>2618 bp *HindIII-*Hpall fragment from <em>A. baylyi</em> containing <em>pcaU</em> in vector *pET21(+)</td>
<td>Siehler et al. (2007)</td>
</tr>
<tr>
<td>pAC46</td>
<td>2618 bp *HindIII-*Hpall fragment from <em>A. baylyi</em> containing <em>pcaU</em>, *pcaI-*lacz transcriptional fusion, in vector *pET21(+)</td>
<td>This study</td>
</tr>
<tr>
<td>pAC53</td>
<td>Randomized PcaU binding site in pAC40</td>
<td>This study</td>
</tr>
<tr>
<td>pAC64</td>
<td>*pcaI-*luc transcriptional fusion</td>
<td>This study</td>
</tr>
<tr>
<td>pAC94</td>
<td>ΔpcaU (699 bp MscI/PmlI) in pAC46, integration of Ω fragment from pHp45Ω (Smal)</td>
<td>Siehler et al. (2007)</td>
</tr>
<tr>
<td>pAC115</td>
<td>844 bp <em>NdeI</em>/<em>NotI</em> fragment containing <em>pcaU</em> in pET28(+) producing a C-terminal 6His tag</td>
<td>This study</td>
</tr>
<tr>
<td>pZBr</td>
<td>5407 bp Sau3A fragment <em>A. baylyi</em> (<em>pcaU</em>-pcaIJFBDΔ) in pUC19</td>
<td>Gerischer et al. (1998)</td>
</tr>
</tbody>
</table>

*Details exactly as for pET29(+)*. 

Table 1. Bacterial strains and plasmids
of the bacterial culture at the respective time. The SD for multiple measurements from the same cell suspension was no more than 2%. To characterize expression levels, we used the reporter gene activity in the mid-exponential growth phase. The SD for this value was up to 30% between different cultures grown under identical conditions.

Expression and purification of His-tagged and native PcaU from *E. coli*. The *pcaU* gene, including its ribosome-binding site, was amplified from *A. baylyi* chromosomal DNA using primers Chitin1 and His1. The PCR product was ligated into the TA cloning vector pDrive. The *pcaU*-containing fragment was subsequently isolated by *Not*I/*Nde*I cleavage and ligated into the *Not*I/*Nde*I-digested expression vector pET29(+) to produce plasmid pAC115. Sequencing of the insertion confirmed the absence of accidental mutations. The PcaU fusion protein possesses six histidyl residues at the C terminus and is therefore referred to as PcaU–6His. Synthesis of the fusion protein was induced in recombinant *E. coli* BL21-AI by the addition of 1 mM IPTG, L(+)-arabinose (0.2 g l⁻¹) and glucose (0.1 g l⁻¹) after the culture had reached OD₆₀₀ 0.6, and cells were then incubated for 4 h at 30 °C. The cells were disrupted with a French pressure cell, and purification of the fusion protein was performed by Ni²⁺ affinity chromatography. Native PcaU was produced and purified using plasmid pAC7, as described previously (Popp et al., 2002).

**Gel retardation assay.** The DNA probes containing the PcaU binding site were generated by whole-cell PCR of *A. baylyi* strains ADPU100–ADPU124 with the primer pair UI-2604 and pZR-Seq1. The resulting DNA fragments (193 bp) were radiolabelled using T4 polynucleotide kinase. The labelled fragments were purified via size exclusion by Sephadex G-25 (MicroSpin G-25 columns, Amersham Biosciences). The assay was performed as described elsewhere (Popp et al., 2002). For the determination of the dissociation constant (*Kₐ*) the quantity of the unretarded probe as a percentage of the total amount of probe employed in one assay was plotted against the concentration of PcaU protein in the assay. The *Kₐ* was the PcaU concentration at which 50% of the probe had been retarded. In all cases, the concentration of the DNA probe was significantly lower than the PcaU concentration (Carey, 1991). Each *Kₐ* determination was repeated at least three times, and the maximum SD was 30%.

**Mutagenesis of the PcaU binding site in vivo.** Random mutations were introduced into the PcaU binding site using a mutagenic oligonucleotide (ON1; Table 2, Fig. 1a). At either end, the oligonucleotide carried wild-type sequence (Table 2). A 1540 bp DNA fragment was created in a three-step PCR procedure, starting with the mutagenic oligonucleotide ON1 (Fig. 2) to enable subsequent incorporation into the chromosome of *A. baylyi* by homologous recombination. The PCRs were designed to avoid

---

**Table 2. Oligonucleotides used in this study**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UI-2604</td>
<td>CGTCAATATTGTACCACATTCC</td>
</tr>
<tr>
<td>Lisa1</td>
<td>GTATACGAGAGCCAGA</td>
</tr>
<tr>
<td>ON3000</td>
<td>ATAGATCTGAGGCGGCGCTTGGCGGCTGTTCC</td>
</tr>
<tr>
<td>pcaJ-4293</td>
<td>TCGACGATGTATGATG</td>
</tr>
<tr>
<td>ON1</td>
<td>CGGAATGACTATAACTAATCATAATGTACAGGGTTGTGGATATTATCGAACAAGAAATTCACAA</td>
</tr>
<tr>
<td></td>
<td>AAATATCGAACAACACGACTAATAACTCCTGATTAATCTGGAACCTG*</td>
</tr>
<tr>
<td>M13/pUC primer</td>
<td>GTTAGAACGAGGCGCAG</td>
</tr>
<tr>
<td>ON2740</td>
<td>CTAAACTCTTGGAACTG</td>
</tr>
<tr>
<td>adpU6-5</td>
<td>AAACCACAACTCAGAGT</td>
</tr>
<tr>
<td>adpU6-4</td>
<td>TGAAATGACTGATATG</td>
</tr>
<tr>
<td>Chitin1</td>
<td>GGTATGCAATATGAAAGAAAAATTCTTCACAA</td>
</tr>
<tr>
<td>His1</td>
<td>CGGATGCGGCGGCGAAGAAATGTTGCGCAATTTC</td>
</tr>
<tr>
<td>pZ8-Seq1</td>
<td>AATTCCTTTCTGATC</td>
</tr>
<tr>
<td>ON2669/Xho</td>
<td>GCGTACTCTGAGGCGGCAATGACTATAACTAACTAATC</td>
</tr>
<tr>
<td>ON2758</td>
<td>CAGTTCCAGAGTATGTTAG</td>
</tr>
<tr>
<td>ON2758/HindIII</td>
<td>ATTCGAAGCTCTTCAGGCTAAGTAGTTTAG</td>
</tr>
</tbody>
</table>

*In the area marked by bold type the respective alternative bases were introduced at a frequency of 0.05.*
incorporation of the wild-type PcaU binding site sequence into the fragment. Therefore, in PCR 1, a fragment was produced from chromosomal DNA of A. baylyi that was adjacent to the PcaU binding site but did not contain it. The resulting 245 bp fragment was used as a template for PCR 2, which ON1 was used as one of the primers, resulting in the incorporation of the mutations. PCR 3 was used to extend the fragment. Plasmid pAC33 (containing a randomized, dysfunctional PcaU binding site) was used as a template; the 316 bp fragment that had been produced in PCR 2 served as one of the primers. The resulting 1540 bp fragment consisted of the randomly mutated PcaU binding site and unmodified DNA corresponding to the wild-type sequence at either side (1245 bp in the direction of pcaU, 245 bp in the direction of pcaI).

To introduce these fragments carrying the randomly mutated PcaU binding site into the genome of A. baylyi, we took advantage of a strain with a 220 bp deletion in the pcaU-pcaI intergenic area which removes pcaIp (A. baylyi strain ADPU1) (Fig. 1b) (Dal et al., 2005). This strain is not able to grow on substrates that are degraded via protocatechuate because of the lack of the main promoter pcaIp, which drives the expression of the pca-qui operon. Furthermore, it lacks part of the PcaU binding site. Using the mutagenic PCR product described (Fig. 2), strain ADPU1 was transformed successfully, even though the overlapping DNA available for base pairing on one side of the deletion (within the pcaI gene) was only 47 bp. Growth on an aromatic substrate degraded via protocatechuate (p-hydroxybenzoate) indicated that the transformants had incorporated a DNA sequence that allowed pcaI gene expression. Whole-cell PCR revealed the incorporation of the 220 bp that were missing in the recipient strain ADPU1.

**Isolation of modified PcaU binding sites by in vitro selection.**

The 90 bp oligonucleotide ON1 containing a randomly mutagenized PcaU binding site was converted into double-stranded DNA and simultaneously radiolabelled by incorporation of [γ-32P]dATP using Klenow enzyme and ON2758 as a primer. After performing a gel retardation assay with purified native PcaU protein, the mixture was separated on a native gel (12% polyacrylamide; Fig. 3). The native PcaU binding site on a 214 bp EcoRI/XmnI restriction fragment labelled with Klenow fragment was included as a control. DNA was isolated from the area containing the retarded ON1 material (Sambrook & Russell, 2001) and amplified by PCR using primers ON2669/Xhol and ON2758/HindIII. The resulting fragment was cleaved with Xhol and HindIII and cloned into pBluescript II SK(+) prepared with the same enzymes. The sequences of the insertions were determined using IRD-labelled M13/pUC primer on a LI-COR 4000L Sequencer (MWG Biotech) and a SequiTherm EXCEL II DNA Sequencing Kit-LC (Epicentre).

**DNA sequence determination of the PcaU binding site in strains of A. baylyi.**

The mutated binding site was amplified by whole-cell PCR using oligonucleotides UI-2604 and Lisa1, producing a 632 bp fragment. A second nested PCR with UI-2604 and Lisa1, producing a 397 bp fragment, which was sequenced by MWG Biotech.

**RESULTS**

**Modified PcaU binding sites allowing PcaU-dependent transcriptional regulation**

The range of DNA sequences within the PcaU binding site allowing significant expression from pcaIp was determined by oligonucleotide selection in vivo in A. baylyi. Strains containing modified PcaU binding sites were produced and selected as described in Methods. In theory, as a result of the synthesis of the mutagenic oligonucleotide ON1 was combined with the complementary strand and labelled. Ten picomoles of this material was incubated with different amounts of native PcaU protein, as indicated. In the two control lanes on the right, a restriction fragment containing the wild-type PcaU binding site was incubated with native PcaU.

In all cases, mutations were observed, ranging from 1 to 15 nucleotide exchanges per molecule. Since the wild-type sequence was not observed in the isolated strains, the PCR procedure worked successfully to permit the selection and analysis of variant regulatory sequences. In a few cases, 1 bp insertions or deletions had occurred. This had not been intended, but is bound to happen as a result of the nature of the chemical DNA synthesis. On average, 6.4 base pair changes were observed per molecule; thus, the number of mutations detected after in vivo screening was slightly lower than the expected number of 7.5 mutations per molecule. Comparing the frequency of mutations within the three 10 bp sequence repetitions and the 10 bp spacer region between R2 and R3, there were significant differences (R1, 15 mutations; R2, 54 mutations; spacer, 30 mutations; R3, 23 mutations). The R2 area seems to be most tolerant of changes without losing the capacity of causing expression from pcaIp. In contrast, the lowest number of changes was observed in the R1 area, which is therefore considered to be the most important for the function of PcaU, closely followed by R3. The spacer region...
displayed an intermediate level of mutation frequency. Seven positions never carried a mutation and are therefore likely candidates for being indispensable for transcriptional regulation by PcaU (Fig. 4). All of these had significant numbers of mutations in the control group of sequences (between 11 and 21% with respect to all analysed sequences determined in the ON1 pool of molecules prior to selection, as described below). Only the cytosine within R3 and the thymine directly in front of R3 cannot be considered; this is because in the control group there were only one and two mutations, respectively, at these positions (2 and 4%, respectively, open arrows in Fig. 4). Four of these conserved positions are located within R1 (or directly adjacent to R1 in one case) and one is located within R3, which fits with the overall frequency of mutations described above.

Another interesting observation is that the insertions and deletions (in strains ADPU114, ADPU116, ADPU118 and ADPU121) did not eliminate PcaU-dependent expression, even though this kind of mutation obviously modifies the spacing of sequence elements.

**Effect of the mutated PcaU binding sites on pca-qui gene expression**

To analyse how the expression from the pcaIp promoter is affected by the characterized PcaU binding-site derivatives under inducing and non-inducing conditions, a luc reporter gene (encoding firefly luciferase) was introduced into the pcaI gene of each of the 24 strains creating a transcriptional fusion between the first gene of the pca-qui operon and the luc gene. Furthermore, we inactivated the pcaU gene in all 24 mutant strains containing the pcaIp–luc fusion to analyse the role of PcaU in activation and repression of the pca-qui operon with the modified PcaU
due to the insertion of the luciferase cassette into pcaI, blocking the pathway at the level of protocatechuate. Thus, the inducer protocatechuate is present throughout the course of the growth experiment. Lactate was chosen because it has no repressing effect (Dal et al., 2002).

Despite the presence of the mutations in the PcaU binding site, induction of the pca operon was still possible, but in almost all cases the level of induction was lower than in the wild-type (40,000 relative light units at OD600 1). The lowest induction observed among the mutant strains was fivefold for strain ADPU108.

Under non-inducing conditions, the basal expression level is negatively controlled by PcaU (columns B and R in Fig. 4). Most of the mutants displayed only minor changes in basal expression level, except for five strains, in which it increased between fivefold and 46-fold. Comparison of the basal expression level of strains containing PcaU with the strains without PcaU allowed the calculation of the repressive effect of PcaU. Most of the mutant strains had lost part or all of this repression. In a few cases repression was as strong as or even stronger than that of the wild-type, which coincided with a decreased basal expression level in the strains with PcaU. We also measured the strains without PcaU under inducing conditions to see if induction had ceased or was being mediated by an unknown component. Almost all PcaU-negative strains displayed no significant change in the presence of inducer. Two strains (ADPU107, ADPU108) showed a 3.5-fold increase, which might be an indication of an unknown alternative regulator causing this effect (data not shown).

Interaction of PcaU with modified binding sites

Gel retardation assays were used to analyse the effect of different mutations in the PcaU binding site on the PcaU–DNA interaction. Therefore, the affinity of purified PcaU–6His protein for the modified PcaU binding sites was determined (Fig. 4). PcaU–6His has been shown to bind to the unmodified PcaU binding site with a somewhat lower affinity than the native PcaU [Kd 1.9 versus 0.16 nM (Popp et al., 2002)]; therefore, all comparisons of affinities were carried out for measurements made with the same PcaU preparation. All mutated PcaU binding sites still bound PcaU–6His specifically. The presence of modified binding sites resulted in a decrease in the binding affinity in most cases (down to ninefold weaker binding in strain ADPU101), but in four cases the binding affinity was close to that observed in the wild-type. A correlation between reduced binding affinity and lower induced expression level could be observed in a number of strains (for example ADPU112), indicating the relevance of the binding affinity of PcaU for its binding site to transcriptional response. In other cases this correlation was not observed (for example strain ADPU101); here other factors must compensate for the loss of binding site affinity of PcaU.

Identification of modified PcaU binding sites based on their affinity for PcaU

In the approach described above, modified PcaU binding sites were identified that allowed PcaU binding and transcriptional activity (recruitment of RNA polymerase and enhancement of its activity at pcaAlp). The results led to the question as to whether sequence conservation is less important when considering PcaU binding alone versus regulation of pca-qui operon expression. To answer this question, mutated PcaU binding sites were screened for their ability to bind to PcaU in vitro. To achieve this, the same synthetic oligonucleotide as in the in vivo strategy (ON1) was used. After synthesis of the second strand and radioactive labelling, the pool of randomly modified PcaU binding sites (10 pmol per assay) was applied to an electrophoretic mobility shift assay to select sequences that were capable of binding to native PcaU. PcaU protein was applied in excess (20 and 60 pmol), which led to specific binding of 0.1 and 0.5 pmol of DNA molecules, as revealed by quantification of the retarded band (Fig. 3). Despite a considerable surplus of PcaU, only a small proportion of the DNA molecules were bound (1 and 5 %, respectively) indicating that the majority of modified binding sites were not specifically recognized by PcaU and therefore were not retarded in the gel. In contrast, a fragment containing the wild-type PcaU binding site was retarded efficiently by 20 pmol protein (Fig. 3).

The bound DNA was eluted from the gel and amplified via PCR. The primers contained sites for restriction cleavage, which allowed subsequent cloning and sequencing. None of the 57 plasmids analysed contained the wild-type sequence. The number of mutations at each individual position varied between one and 20 with an average of 11. The average number of mutations per molecule was 9.6.

As a control for the distribution of mutations in the mutagenic oligonucleotide, it was amplified without prior selection, cloned into the sequencing vector and sequenced. A total of 48 sequences were determined. The average number of mutations per molecule was 8.3, which was slightly higher than the expected number of 7.5 mutations per molecule. The distribution of mutations at each position varied (one–17 mutations per position among 48 molecules, average eight). Therefore, the mutation frequency after selection by PcaU was normalized against the mutation frequency prior to selection. Positions with a significantly lower occurrence of mutations after selection by PcaU binding as compared to the control pool without selection are considered to be important for the PcaU–DNA
interaction, and therefore conserved (open arrows in Fig. 5). The normalization was intended to focus the analysis on the changes between the two pools of sequences and not on the variation in mutation between the different positions caused by the different coupling efficiencies of the oligonucleotide synthesis. This normalization was not applied to the in vivo experiment described above because the selection criteria applied there (protocatechuate-dependent induction at pcaIp) were much more stringent than attachment to a DNA fragment.

Some of the conserved positions found in the in vitro approach are identical to positions defined by the in vivo approach. The two guanine nucleotides in R1 are strongly conserved in both approaches. Of the other five positions found to be conserved with the in vitro approach, only one (the first cytosine on R3) is moderately conserved in the in vivo experiment. All other conserved positions in the in vivo experiment are neutral or moderately conserved in vitro. One exception is the second cytosine in R3, but this result should not be considered significant due to the low number of mutations in the non-selected pool at this position. Taken together, two positions in R1 were found to be strongly conserved, in both the in vivo and the in vitro experiments, underlining the relevance of these residues for both PcaU binding alone and transcription activation. Other positions have different significances in the in vivo and in vitro approaches.

**DISCUSSION**

In the current investigation the complete binding site of the PcaU protein from *A. baylyi*, including a few neighbouring nucleotides, was analysed by random mutagenesis. Because of the size of the mutagenized area (50 bases), rather than using saturating random mutagenesis, a synthesis procedure was chosen that in theory introduced mutations with a likelihood of 0.15 per position. The experiments dealt with the question of whether and how the binding site of the PcaU protein from *A. baylyi* can be modified without losing function. Furthermore, the screening was expected to reveal modifications of the PcaU binding site that significantly changed the effect of PcaU on gene expression at pcaIp. In one approach (the in vivo approach), the function screened for was the native function in the organism. In a second approach, the screening parameter was PcaU binding.

![Fig. 5. In vitro selection for PcaU binding by modified PcaU binding sites. Two sets of modified PcaU binding sites were compared, prior to selection (47 sequences) and after selection by PcaU binding (56 sequences). Each column displays the increase or decrease in mutations after selection by PcaU versus without selection. This calculation was necessary because the distribution of mutations for each position was not uniform. The 100% value is indicated by a dashed line, indicating no change as a consequence of selection by PcaU binding. Values significantly lower than 100% are indicated by open vertical arrows. These positions are suggested to be relevant for binding by PcaU because they tolerate changes less frequently. Values significantly higher than 100% indicate an increase in mutations after selection by PcaU; in these cases the nucleotide occurring most frequently is indicated above the column. These changes are proposed to allow better binding of PcaU. On numerous occasions, the number of mutations in the set prior to selection was lower than five and therefore considered to be too low to reveal reliable data (labelled with a white asterisk). The numbers below the columns give the absolute numbers of mutations at each position in the unselected ON1 pool.](http://mic.sgmjournals.org)
quality alone (referred to as the in vitro approach). A potential ambiguity of the approach lies in the creation of multiple mutations per molecule, a situation which could lead to compensation or suppression effects. Therefore, investigation of important positions by creating single mutations is an important follow-up experiment.

The 24 mutant strains identified by the in vivo approach revealed that the three sequence repetitions R1, R2 and R3, and the spacer region between R2 and R3, undergo mutations at significantly different frequencies. R1 is clearly the area with the lowest number of mutations and therefore is indicated to be most relevant to the biological function of PcaU. The R2 sequence accumulated the highest number of mutations and therefore is indicated to be less important for the function of PcaU. It is unlikely that this 10 bp sequence has no relevance because in this case its sequence would not be perfectly identical with the R1 and R3 sequences. R3 seems to have an intermediate level of importance for in vivo PcaU function. The observation that the spacer region between R2 and R3 acquires a significantly lower number of mutations than R2 is unexpected and is an indication that this region fulfils a function in the binding site. Five positions could be identified that never carried a mutation, indicating the relevance of these positions in the process of PcaU-dependent gene regulation.

The effect of the mutated PcaU binding sites on gene expression was investigated. All 24 mutant strains showed an increase in pca-qui expression in the presence of protocatechuate, even though in most cases to a lesser extent than in the wild-type. The diminished induction was due to (i) a reduced maximal expression level or (ii) an increased basal expression level, or both. In these strains the positioning of the regulator is probably not optimal, allowing only a reduced expression level. No binding site modification resulted in a significantly higher induction than in the wild-type. Clearly the native PcaU binding site has a sequence optimized for high induction upon protocatechuate binding.

The uninduced expression levels (basal levels) on the other hand stayed at a similar level (0.5–2.0-fold increase in roughly half of the strains) or increased significantly (up to 46-fold). Here, two things may have happened. Firstly, the mutations may have potentiated the nearby promoter of the pca-qui operon. Secondly, the repressing effect of PcaU at the wild-type DNA binding site may have been diminished. This could be observed for most strains with a significantly increased level of basal expression.

In addition to the expression data, the affinity of PcaU for the modified binding sites was determined. In all cases PcaU bound specifically, an expected observation, since the gene expression data had already indirectly shown that a PcaU-dependent activity could be observed. None of the modified binding sites bound PcaU significantly more tightly than the wild-type sequence, an observation that has also been made for other regulators like CopR (Freede & Brantl, 2004).

Of interest with respect to the spatial organization of the binding site are four binding sites in which insertions and/or deletions (of one nucleotide in each case) had occurred. Due to the nature of the screening, these changes must still allow gene expression, demonstrating that certain rearrangements are possible.

In the in vitro approach, eight positions were identified at which mutations occurred to a significantly lesser extent than in the control group. These positions are indicated to be relevant for PcaU binding alone. Clearly, there are differences between these positions and the ones found to be most conserved in the in vivo approach. This may be a reflection of several additional parameters that influence the performance of the PcaU binding site in vivo: (i) for induced expression, protocatechuate is bound by PcaU, and this complex may have a different DNA binding requirement from that of PcaU without inducer; (ii) the interaction with RNA polymerase is likely to be part of the regulatory process, and this interaction may also change the DNA binding requirement. Despite this, in two cases the same positions were seen to be conserved with both approaches, namely the two guanosines within R1. These positions seem to be uniquely important, in both PcaU binding alone and the complete process of PcaU-dependent gene regulation.

The PcaU protein has a second binding site directly downstream of its own gene. It contains all three sequence repetitions, although they are less well conserved. Binding of PcaU to this site has been demonstrated. In this case, only the palindromic area is bound and the binding is weaker than that to the intergenic binding site (Popp et al., 2002). Despite these differences, this second PcaU binding site displays all the residues that have been identified as conserved in the current investigation by both the in vivo and the in vitro approaches, supporting their relevance in both binding sites.

Of all 540 recognized members of the IclR family (Bairoch et al., 2005; Krell et al., 2006; Mulder et al., 2005) only a small number (20 regulators) have been investigated experimentally, and only for some of them is the binding site known. A conserved IclR-type DNA binding site cannot be found, but for a subgroup within the family (namely the PcaU/PobR/PcaR group) there is a conserved sequence motif (Fig. 6) covering a palindromic 17 bp sequence. The two guanosine nucleotides in R1 identified in the current study as being strongly conserved by both the in vivo and the in vitro approach also turn out to be moderately to completely conserved in comparisons between different binding sites (Fig. 6). The other arm of the palindrome (corresponding to R2) did not display completely conserved residues with the in vivo approach, but the guanosine residue found to be conserved by the in vitro approach is also completely conserved in the alignment shown in Fig. 6. Thus, three positions shown to be important for PcaU binding and/or activity are also conserved in binding sites of related regulators of the IclR subfamily.
The IclR protein itself is one of the best-investigated proteins within the family. Its binding site has been studied using a random mutagenesis approach followed by in vitro screening (Pan et al., 1996). The core sequence is a highly conserved 15 bp palindromic motif, which overlaps the −35 motif of the promoter of the first regulated gene (aceB).

Among the six structures of IclR family members that are conserved 15 bp palindromic motif, which overlaps the site. Also, the constellation that is partially found in the PcaU binding site of A. baylyi. Such binding sites resemble a simple palindrome thus indicating that all members of the PcaU/PobR/PcaR group have binding sites that exceed a simple palindrome thus suggesting that the protein is active as a tetramer. In the case of the IclR member AllR from E. coli IclR interacting with the T. maritima tetramer. This resulted in the suggestion that T. maritima IclR is likely to interact with a binding site consisting of two palindromes, a constellation that is partially found in the PcaU binding site. Also, A. baylyi PobR and Pseudomonas putida PcaR have binding sites that exceed a simple palindrome thus indicating that all members of the PcaU/PobR/PcaR group may act as tetramers. In the case of the IclR member AllR from E. coli, residues Cys135–Cys142 have been predicted to be involved in the tetramerization, and mutations in this region made the respective proteins functionally inactive in vivo (Walker et al., 2006).

Taken together, the mutagenesis of the complex binding site of A. baylyi PcaU revealed that the three sequence motifs R1, R2 and R3 have differential relevance to pca-qui operon regulation in vivo, with R1 being the least susceptible to mutations and therefore most important, and R2 being least important. Four positions within (or adjacent to) the 10 bp sequence motif R1 were conserved and therefore required for in vivo function, whereas other positions could be mutated. All modified binding sites changed the transcriptional response of the controlled promoter, pcaIp, by changing its activating or its repressing quality, or both. All mutant binding sites displayed a lower affinity for PcaU, which was in part, although not systematically, correlated with changes in the transcriptional output. Other effects of the mutated DNA binding site seem to be more relevant to transcriptional regulation; for example, the exact positioning of the regulator with respect to the RNA polymerase–promoter complex. The results of the in vitro screening partially overlapped with those of the in vivo screening, confirming the relevance of the respective positions for both PcaU binding alone and the complete process of PcaU-dependent gene regulation. Considering the unusual qualities of the binding site as well as the dual functionality (positive as well as negative regulation of the same promoter), it will be most interesting to unravel more details of the function of this regulator.

ACKNOWLEDGEMENTS

The brilliant technical assistance of Iris Steiner is gratefully acknowledged. This study was supported by funding from the Konrad-Adenauer-Stiftung to B.J. and by a grant from the Rudolf und Clothilde Eberhardt-Stiftung.

REFERENCES


Freede, P. & Brantl, S. (2004). Transcriptional repressor CopR: use of SELEX to study the copR operator indicates that evolution was edged. This study was supported by funding from the Konrad-Adenauer-Stiftung to B.J. and by a grant from the Rudolf und Clothilde Eberhardt-Stiftung.

REFERENCES


Edited by: M. A. Kartesz