Identification of the nucleotide sequence recognized by the cAMP-CRP dependent CytR repressor protein in the *deoP2* promoter in *E.coli*

Peter Birk Rasmussen⁺, Lotte Sogaard-Andersen and Poul Valentin-Hansen*

Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark

Received November 20, 1992; Revised and Accepted January 14, 1993

ABSTRACT

In *E.coli* repression of transcription initiation by the CytR protein relies on CytR–DNA interactions as well as on interactions between CytR and the cAMP-CRP activator complex. To identify the nucleotide sequence recognized by CytR, mutants of the *deoP2* promoter with a reduced regulatory response to CytR have been isolated. Five single bp mutation derivatives of *deoP2* with a 2–5-fold decrease in CytR regulation have been characterized. *In vitro*, the only effect of the mutations was a decrease in the binding affinity of CytR, and a clear correlation was observed between the reduction in CytR regulation *in vivo* and the reduction in CytR binding *in vitro*. The mutations all reside in a sequence element that contains an imperfect direct as well as an imperfect inverted repeat. As the active form of CytR, most likely, is an oligomer with two-fold rotational symmetry, CytR probably interacts with the inverted repeat. Degenerate versions of the inverted repeat are present in all CytR binding sites characterized so far, however, the distance between the half-sites varies.

INTRODUCTION

cAMP-CRP and CytR in combination regulate initiation of transcription from at least eight operons that code for proteins involved in nucleoside and deoxynucleoside catabolism and transport in *E.coli* (1). cAMP-CRP activates transcription and CytR represses transcription, however, cAMP-CRP also plays a crucial role in the repression mechanism (2). CytR repression is confined to cells containing the cAMP-CRP complex (3) and depends on proper DNA binding of cAMP-CRP (2–5). Consistently, *in vitro* analyses of four cAMP-CRP/CytR regulated promoters, *deoP2*, *cddP*, *cytRP* and *tsxP2*, have shown that CytR alone only interacts weakly with its cognate promoters (6–9). In presence of cAMP-CRP, however, the two proteins bind cooperatively to the DNA and the cooperative DNA binding is mediated by direct protein–protein interactions between the cAMP-CRP and CytR (6, 10). Hence, repression by CytR relies on contacts to both the DNA and the cAMP-CRP complex, and cAMP-CRP may be regarded as an adaptor for CytR.

In this report, we have used the *deoP2* promoter to investigate the importance of CytR–DNA interactions for repression and to identify the DNA sequence motif recognized by CytR. This promoter contains two binding sites for cAMP-CRP, a low affinity site, CRP-1, centered at position –40 and a high affinity site, CRP-2, centered at –93 (11) (Fig. 1). CytR binds to the region between the CRP sites (6). Previously isolated *deoP2* mutants with reduced CytR regulation all contain mutations in either of the CRP sites (3–5). *In vitro* analyses have shown that the effect of these mutations on CytR repression is indirect and can be traced back to their inhibitory effect on DNA binding of cAMP-CRP (2). Thus, despite a growing knowledge of the interaction of CytR with its cognate promoters, little is known about the specific CytR–DNA interactions. Here, we have isolated *deoP2* mutants containing mutations that directly interfere with CytR–DNA interactions in order to locate the positions important for sequence specific DNA binding of CytR more precisely. On basis of the results we discuss how CytR may interact with the DNA.

MATERIALS AND METHODS

All enzymes used for DNA manipulations were purchased from Boehringer, Mannheim except exonuclease III, AMV reverse transcriptase and modified T7 polymerase (Sequenase) which were purchased from Amersham, Life Sciences Inc. and USB Corp., respectively. ³²P-labelled nucleotides were obtained from NEN. Isolation of plasmid DNA, cloning, transformation of *E.coli* and gel analyses of recombinant plasmids were performed as described (13). S1 nuclease mapping of start sites for transcription and purification and sequencing of ³²P-labelled fragments were performed as previously described (14).

Bacterial strains

All strains are *E.coli* K12 derivatives: SO928(Δ deo, Δ lac, *cytR*⁺), SO929 (as SO928 but *cytR*) (15), SO2929 (as SO929 but Δ crp-96 zhd-732::Tn10) (10), SO1316/F' *lacI*^q (*araD*139, *lacU*169, *cod*, *thi*, *rpsL*, *cdd*:: Δ Mu:: λ pl(209)/F' *proA*⁺*B*⁺, *lacI*^q, *lacZ* Δ M15::Tn10) (4), TG1 (Δ (*lac-pro*), *supE*, *thi*, *hsdD5*/F' *traD*36, *proA*⁺*proB*⁺, *lacI*^q, *lacZ* Δ M15 (obtained from Amersham).

* To whom correspondence should be addressed

⁺ Present address: INSA-CTBM, Universite Paul Sabatier, 31077 Toulouse Cedex, France

Plasmid constructions

pJEL152 is a derivative of the low copy number R1 *lacZ* transcriptional fusion plasmid pJEL126 (3, 16). pJEL152 contains *deoP2* sequences from position -116 to +18 (Fig. 1) fused to *lacZ* and the *deoP2* sequence is flanked on both sites by an *EcoRI* site. pUC13 derivatives: The *EcoRI* fragments from the pJEL152 derivatives were cloned in the *EcoRI* site in pUC13 (17). The orientation of the *EcoRI* fragments in all the constructs used in the titration analysis is so that the *deoP2* promoter in the cloned fragment and the *lac* promoter in pUC13 transcribe convergently.

NaNO₂ mutagenesis

In order to create a region of single stranded *deoP2* DNA, two pUC13 derivatives in which the *EcoRI* fragment containing *deoP2* had been cloned in either orientation were digested with BamHI that cleaves pUC13 immediately downstream of the *EcoRI* site. Next, the BamHI restricted DNA was subjected to exonuclease III digestion for 15 min under conditions at which exonuclease III removes 10 nucleotides per min. Subsequently, the DNA was exposed to the mutagen NaNO₂ as described (18). After removal of the mutagen, the double stranded DNA was extended with AMV reverse transcriptase. Finally, the mutagenized plasmids were restricted with *EcoRI* and the *EcoRI* fragments with the mutagenized *deoP2* promoters cloned in pJEL126. Plasmid DNA was isolated from clones with reduced CytR regulation and the *EcoRI* fragments containing the putative mutations recloned in pJEL126. Altogether, 7 clones exhibited reduced CytR regulation following the second cloning. After cloning in mp19 (19) of the corresponding *EcoRI* fragments, the fragments were sequenced by the chain terminating method (20) with modified T7 polymerase as recommended by the supplier.

Site directed mutagenesis

The *EcoRI* fragment containing the entire *deoP2* promoter was cloned in mp19 (19). Selected mutations were introduced using an oligonucleotide-directed *in vitro* mutagenesis system purchased from Amersham (21).

Gel retardation analyses

The *MamI-SfaNI*, *MamI-Sau3A* and *XbaI-PvuII* *deoP2* fragments were purified by electrophoresis on polyacrylamide gels after labelling of the 5'-end of the *MamI* site with polynucleotide kinase in the presence of γ -³²P-ATP and the 3'-end of the *XbaI* site with Klenow fragment in the presence of α -³²P-dCTP. 0.8 nM of the labelled fragments and various concentrations (indicated in the Legend in Fig. 3) of CytR or CRP were incubated in binding buffer (10 mM Tris HCl (pH 7.8), 50 mM KCl, 1 mM EDTA, 50 μ g/ml acetylated bovine serum albumine, 1 mM dithiothreitol, 0.05 % Nonidet-P40) containing 20 μ g/ml competitor DNA (pGEM4 obtained from Promega) for 30 min. at 37°C in a total volume of 10 μ l (6). cAMP was included in the experiments with CRP at a final concentration of 10 μ M. Immediately before loading, 2 μ l of loading buffer (binding buffer containing 50% glycerol and 0.1 mg/ml bromophenol blue) were added to the samples. Samples were loaded with current on. Electrophoresis was conducted at 100 V. 5 % acrylamide gels prepared from a 44:0.8 (acrylamide:N,N'-methylenebisacrylamide) stock were used throughout. The electrophoresis buffer was 10 mM Tris HCl (pH 7.8), 1 mM EDTA containing 25 mM KCl in the experiments with CytR and 10 mM Tris HCl

(pH 7.8), 1 mM EDTA containing 10 μ M cAMP in the experiments with CRP. Following electrophoresis, the gels were dried and autoradiographed. CytR and CRP were purified as described (6, 22).

Hydroxyl radical footprinting

³²P-labelled fragments (final concentration 1 nM), proteins at the concentrations stated in the legend to Fig. 4 and competitor DNA were incubated as in the gel retardation experiments. cAMP was added to a final concentration of 50 μ M. Otherwise the footprints were performed as described by Tullius and Dombroski (23) with the modifications developed by O'Halloran *et al.* (24).

RESULTS

NaNO₂ mutagenesis of *deoP2*

An *EcoRI* fragment containing *deoP2* sequences from position -116 to position +18 with respect to the start site for transcription (+1) contains all the information required for cAMP-CRP and CytR regulation of *deoP2* (3). In order to isolate promoter mutants with decreased CytR regulation, this *EcoRI* fragment was subjected to NaNO₂ mutagenesis *in vitro* (See Materials and Methods) and subsequently cloned in the *EcoRI* site of the low copy number transcriptional fusion plasmid pJEL126 (16) to give the plasmid pJEL152. The resulting plasmids were transformed to a Δ *lac*, Δ *deo*, *cytR*⁺ strain followed by a screening on indicator plates for colonies with an elevated *lacZ* expression. Seven independent colonies were isolated and the mutations identified by sequencing (Materials and Methods) (Fig. 1). Two promoters had obtained the same A→G (all mutations are indicated for the upper strand) substitution at position -65, two promoters had obtained an A→G substitution at -72, three promoters had obtained double mutations: A→G at -65 together with C→T at -1, A→G at

Figure 1. A. Schematic representation of *deoP2*. Coordinates are in bp, +1 refers to the start site for transcription as indicated by the arrow. The open boxes indicate the two cAMP-CRP binding sites, CRP-1 and CRP-2, and the hatched bar indicate the CytR binding site. B. Nucleotide sequence of the upper strand of *deoP2*. The boxed sequences indicate the two CRP sites and the arrows below the boxes indicate the sequences sharing homology with the consensus CRP site (12). In the CytR binding site, the substitutions in the five mutant promoters are indicated, the inverted arrows indicate the inverted repeat in the CytR binding site and the tandem arrows indicate the direct repeat.

–65 with C→T at +4 and A→G at –65 together with A→G at –33.

cAMP-CRP activation and CytR regulation of the five mutant promoters were assessed by monitoring activities of β -galactosidase in *cytR*⁺, *cytR* and Δ *crp* strains. As shown in Table 1, CytR regulation is reduced from approximately 10-fold in the wt-promoter to approximately 2- and 3-fold in the mutant promoters whereas cAMP-CRP activation remains unchanged. Activation and CytR regulation in the three promoters containing double mutations were identical with that of the promoter only containing the –65 mutation (data not shown), hence, positions +4, –1 and –33 do not appear to be important for CytR regulation. The mutations at –65 and –72 are positioned in the region of *deoP2* protected by CytR in DNase I footprint experiments. Thus, these two positions may be important for sequence specific DNA binding of CytR.

Site directed mutagenesis of the CytR binding site

As most protein binding sites on DNA consist of either inverted or direct repeats, we examined the CytR binding sequence in *deoP2* for such structures. Positions –65 and –72 are part of an inverted as well as a direct repeat (Fig. 1). Three additional mutations were introduced in this region by site directed mutagenesis (See Materials and Methods) (T→A at –68, A→G at –73 and G→A at –75) in order to test in more details the importance of this DNA segment for CytR regulation. Following cloning of the mutated promoters in pJEL126 the effect of the mutations was assessed by measuring the activity of β -galactosidase. All three mutations result in decreased CytR regulation with the mutation at –72 exhibiting the largest effect and, once again, neither of the mutations had any effect on cAMP-CRP activation of *deoP2* transcription (Table I).

In order to analyse the effect of the mutations on CytR interaction with *deoP2* in an alternative approach, we performed *in vivo* titration experiments. In this analysis, the different *deoP2* promoters are cloned on a high copy number plasmid (pUC13) and then introduced into a CytR⁺ strain that contains a gene fusion between the cAMP-CRP/CytR regulated *cdd* promoter and *lacZ* on the chromosome. This strain, in addition, contains an F plasmid that encodes the *lacI*^q allele. The presence of *lacI*^q ensures that the *lac* promoter on the pUC13 derivatives is repressed, otherwise transcription from *lacP* interferes with titration of CytR (4). The ability of the different mutant promoters to titrate CytR and, thus, alleviate CytR regulation of the *cdd*-

lacZ fusion was found to be reduced (See Table II). Furthermore, the reduced titration of CytR parallels the decreased CytR regulation in the different promoter mutants (Compare Table I and II).

Finally, the start site for transcription in the mutant promoters was mapped in an S1-nuclease mapping analysis. In all cases, transcription was found to initiate at position +1 (Fig. 2), ruling out the possibility that the decrease in CytR regulation is caused by the introduction of a new promoter.

The mutations between the CRP sites reduce the affinity of CytR for *deoP2* *in vitro*

The only way to directly assess the effect of the mutations on protein–DNA interactions is to analyse both CytR–DNA as well as cAMP-CRP–DNA interactions separately *in vitro*. The *in vitro* binding of CytR to the mutated promoters was characterized in a gel retardation assay using purified CytR protein and purified ³²P-endlabelled DNA fragments carrying either the wt- or the mutated promoters. It was previously observed that CytR has a strong tendency to bind non-specifically to the DNA in gel retardation assays that are carried out in low ionic strength electrophoresis buffer (6). To increase the specificity of CytR binding in the gel retardation assay, 25 mM KCl was added to the starting low ionic strength electrophoresis buffer (6) (See Materials and Methods). At this ionic strength, sequence specific binding of CytR is only weakly reduced whereas non-specific binding is strongly reduced (Søgaard-Andersen and Valentin-Hansen, unpubl.). The region in *deoP2* protected by CytR in DNase I footprinting experiments is contained on a 47 bp *MamI*-*Sfa*NI fragment extending from position –41 to position –87 and on a 52 bp *MamI*-*Sau*3A fragment extending from –41 to –92. By using such relatively small DNA fragments in the gel retardation assay in combination with the changed electrophoresis buffer, non-specific binding of CytR to the labelled fragment was practically abolished. The 47 bp fragments were purified from the wt- and mutated promoters except for the promoter with the mutation at –75 which removes the *Sfa*NI site; for the latter promoter the 52 bp fragment was employed.

Binding of CytR to either of the two fragments result in the formation of only one retarded complex (Fig. 3 upper panels, lane 2–5 for the wt-promoter and lane 2–5 for the promoter with the mutation at –75)). DNase I footprinting experiments

Table I. Effect of mutations in the CytR binding site in *deoP2* on cAMP-CRP activation and CytR regulation.

	Activity of β -galactosidase ^a			Fold of CytR regulation
	CytR ⁺	CytR	Crp	
pJEL152-wt	0.19	1.84	0.07	9.7
–65 ^b	0.52	1.79	0.07	3.4
–68 ^b	0.37	1.81	0.07	4.9
–72 ^b	0.80	1.85	0.07	2.3
–73 ^b	0.75	1.88	0.07	2.5
–75 ^b	0.36	1.77	0.07	4.9

a) Enzyme levels are measured during exponential growth in AB minimal medium supplemented with glycerol as described (25). Activities of β -galactosidase are expressed as OD₄₂₀/OD₄₅₀/ml/min (26). The values are the average of three independent experiments (the observed variations do not exceed 10%).

b) Numbers indicate position of mutations.

Table II. Effect of mutations in the CytR binding site in *deoP2* on titration of CytR.

	Activity of β -galactosidase ^a in SO1316/F' <i>lacI</i> ^q	Fold of titration ^b
pUC16 ^c	0.05	
p13-152-wt	1.16	23.2
–65 ^d	0.64	12.8
–68 ^d	0.87	17.4
–72 ^d	0.46	9.2
–73 ^d	0.56	11.2
–75 ^d	0.80	16.0

a) Activities were determined as described in Table I except that the growth medium was supplemented with 0.1% casamino acids.

b) Fold of titration is calculated as the activity in the presence of the titrator plasmid divided by the activity in the presence of pUC13.

c) pUC16 is a pUC13 derivative in which the α -*lacZ* gene has been inactivated to avoid α -complementation of the α -acceptor encoded by the F plasmid (4).

d) Numbers indicate position of mutations.

performed in presence of saturating concentrations of CytR, show that the retarded complex consists of CytR bound to the sequence that has previously been identified as the CytR binding site on larger *deoP2* fragments (6; data not shown). Gel retardation experiments performed in presence of decreasing concentrations

Figure 2. S1-nuclease mapping analysis of the start site of transcription in the *deoP2* mutants. The probe is a 350 bp *RsaI*-*Clal* fragment comprising the entire *deoP2*, ³²P-end-labelled at the 5'-end of the *Clal* site with polynucleotide kinase and isolated from pJEL152. Lane 1 is the A+G sequence of the hybridization probe. Lane 2: RNA isolated from SO929 harbouring pJEL152-wt, lane 3: pJEL152-65, lane 4: pJEL152-68, lane 5: pJEL152-72, lane 6: pJEL152-73, and lane 7: pJEL152-75.

of CytR clearly show that all the mutations result in a decreased affinity of CytR for *deoP2* (Fig. 3, upper panels). The effect of the mutations on binding of CytR parallels the effect of the mutations *in vivo*, i.e. the -72 and -73 mutations which have the largest effect on regulation and titration *in vivo* also exhibit the largest effect on binding of CytR *in vitro*. Moreover, a prominent feature in the gel retardation experiments is the presence of a smear between the two well-defined bands. The smear is more pronounced in the experiments with the mutant promoters and it is, most likely, the result of dissociation of CytR-DNA complexes during electrophoresis.

In order to analyse the possibility that the effect of the mutations on CytR regulation was an indirect effect mediated by changed binding of cAMP-CRP, gel retardation experiments with ³²P-endlabelled DNA fragments containing both CRP sites in the context of the different mutations were also performed (Fig. 3, lower panels). As previously shown, binding of cAMP-CRP to the wt-promoter resulted in the formation of two retarded complexes: The complex with the highest mobility contains cAMP-CRP bound at CRP-2 and the complex with the lowest mobility contains cAMP-CRP bound at both CRP-2 and CRP-1 (6). The mutated promoters all bound cAMP-CRP with the same affinity as the wt-promoter and the mobilities of the two retarded complexes was identical to that observed for the wt-promoter. Taken together, these data show that the effect of the mutations is a direct consequence of changed CytR-DNA interactions.

Mapping of the contact points between the deoxyriboses in the DNA backbone and cAMP-CRP and CytR

In order to characterize further the binding of CytR, the DNA backbone contacts of the proteins within the repression complex were determined in hydroxyl radical footprinting experiments. A 250 bp ³²P-end-labelled DNA fragment containing *deoP2* sequences from -116 to +18 was mixed with CRP and/or CytR, protein-DNA complexes allowed to form and then reacted with hydroxyl radicals. In the absence of any added protein, strand

Figure 3. Gel retardation analysis of the binding of CytR and cAMP-CRP to wt- and mutant *deoP2* promoters. In the upper panels, the ³²P-end-labelled fragments used were the 47 bp *MamI*-*Sfa*NI fragment for the wt-, -65, -68, -72 and -73 mutant promoters and the 52 bp *MamI*-*Sau*3A fragment for the -75 mutant promoter. CytR binds with the same affinity to the two probes (data not shown). The promoter from which the fragment was isolated is indicated above each panel. The final concentration of the fragments was 0.8 nM. In all the upper panels, CytR was added to the following final concentrations: Lane 1: 0 ng/μl, lane 2: 33.3 ng/μl, lane 3: 11.1 ng/μl, lane 4: 3.7 ng/μl, lane 5: 1.2 ng/μl. The composition of the different bands is indicated. In the lower panels, the ³²P-end-labelled fragments used were 250 bp *XbaI*-*Pvu*II fragments isolated from the p13-152 derivatives as indicated above each panel. The final concentration of the fragments was 0.8 nM. In all the lower panels, CRP was added to the following final concentrations: Lane 1: 0 ng/μl, lane 2: 2 ng/μl, lane 3: 0.2 ng/μl, lane 4: 0.02 ng/μl. cAMP was added to a final concentration of 10 μM in both the binding reactions and in the electrophoresis buffer. KCl was not included in the electrophoresis buffer in the gel retardation experiments with CRP. The composition of each band is indicated.

breakage occurs at each position with approximately the same efficiency (Fig. 4, lanes 2 and 7). When CRP is added, three protected patches each 3–4 nucleotides long and separated by 5–6 efficiently cleaved nucleotides are observed in both CRP sites (Fig. 4, lanes 3 and 8). The protection patterns on the two

strands are identical but offset from each other in the 3'-direction by 3 nucleotides. CytR protects two regions on both strands (Fig. 4, lanes 5 and 10), and the protection pattern is also in this case offset from each other in the 3'-direction by 3 nucleotides. In the presence of both proteins, a clearly defined set of regularly spaced, protected regions appear on both strands and the protection pattern is the sum of those obtained with the two regulators independently (Fig. 4, lanes 4 and 9).

Figure 4. Hydroxyl radical footprints of cAMP-CRP and CytR bound to *deoP2*. The fragment used is a 250 bp *XbaI*-*PvuII* fragment encoding *deoP2* sequences from -118 to +16. Both strands were labelled in the *XbaI* site, the lower strand (lanes 1–5) was 3' labelled with Klenow in the presence of α - 32 P dCTP and the upper strand (lanes 6–10) was labelled with polynucleotide kinase in the presence of γ - 32 P ATP. CRP was added to a final concentration of 0.6 ng/ μ l, cAMP to a final concentration of 50 μ M, CytR was added to a final concentration of 3.3 ng/ μ l in the reactions in lanes 4 and 9, and a final concentration of 44 ng/ μ l in the reactions in lanes 5 and 10. The regions protected by cAMP-CRP and CytR in DNase I footprints are indicated to the left. The large arrows indicate the positions protected by cAMP-CRP from attack by the hydroxyl radicals and the small arrows the positions protected by CytR. Lanes 1 and 6 are the A+G sequences of the probes. The sequence is numbered with respect to the start site of transcription (+1).

DISCUSSION

Sequence specific recognition of DNA by gene regulatory proteins is fundamental to the control of initiation of transcription. The identification of the nucleotide sequence recognized by a protein generally relies on the characterization of mutations that interfere with binding of the protein, analyses of protein–DNA complexes using various footprinting techniques and on a comparison of the sequences of different binding sites. In this report, we have used these approaches to identify the nucleotide sequence recognized by CytR in the *deoP2* promoter.

Previous attempts to determine the binding specificity of CytR to DNA sites have been complicated by the fact that repression by CytR relies on both CytR–DNA interactions as well as on direct interactions with cAMP-CRP (2, 10). Hence, CytR regulation of *deoP2* can be drastically reduced by mutations that interfere with binding of cAMP-CRP to the two CRP sites (3–5). Here, we have approached this problem by using an efficient *in vitro* method for generating mutations at defined regions. Five mutant *deoP2* promoters were isolated that all specifically exhibit reduced CytR regulation and titration *in vivo*. All mutant promoters contain single bp substitutions in the region between the two CRP sites that has previously been identified as the CytR binding site (6). Furthermore, *in vitro* binding analyses demonstrated that the mutations interfered directly with CytR–DNA interactions whereas binding of cAMP-CRP remained unchanged.

The binding site for CytR was also narrowed down in DNA footprinting experiments with hydroxyl radicals. This chemical attacks deoxyriboses in the DNA backbone and produces a high resolution map of protein–backbone interactions due to its small size (22). Using protection data for cAMP-CRP and CytR alone and in combination, the contributions from each of the regulators in the repression complex was delineated (Fig. 4). In full agreement with the crystal structure of the cAMP-CRP/DNA complex (27), cAMP-CRP yielded a tripartite footprint at both binding sites in which the centers of the protected regions were separated by approximately 10 bp. CytR produced a bipartite

Figure 5. A helical representation of the positions protected by cAMP-CRP (filled circles) and CytR (filled triangles) from attack by hydroxyl radicals. The helix is drawn with 10.5 bp/turn and each vertical line represents a bp. The extent of the regions protected by cAMP-CRP and CytR in DNase I footprints is indicated above the helix. The nucleotide sequence is that of the upper strand numbered with respect to the start site of transcription (+1), each nucleotide is written above its corresponding bp. The arrows below the sequence indicate the parts of the CRP sites that share homology with the consensus CRP site and the boxed sequences indicate the CytR recognition sequence.

Figure 6. Comparison of the nucleotide sequences of four CytR binding sites in *deoP2*, *cddP*, *cytRP* and *tsxP2*. The arrows indicate the sequences in the CRP sites that share homology with the consensus CRP site, the boxes indicate the sequences sharing homology with the inverted repeat in the CytR recognition sequence in *deoP2*, the numbers indicate coordinates with respect to the start site for transcription (+1). For *cddP* the sequence of the lower strand is displayed in order to more clearly visualize the homology with *deoP2* and *cytRP*. In *deoP2*, the substitutions interfering with binding of CytR are indicated. In *cddP*, cAMP-CRP binds to CRP-1 and CRP-3 in presence of CytR, and CytR binds between these two sites, the mutations that interfere with binding of CytR are indicated (7). In *cytRP*, cAMP-CRP binds to only one CRP site (CRP) and CytR binds immediately downstream from the CRP site (8). In *tsxP2*, cAMP-CRP binds to two sites CRP-1 and CRP-2 in absence of CytR, and CytR binds to a site that overlaps extensively with CRP-1 in absence of cAMP-CRP; the structure of the combined complex is not known in details (9). Below a simple consensus based on the number of occurrences of each base at each position is presented, the suffix indicates the number of occurrences of the most frequent base. In the symmetry consensus, the simple consensus is converted to a palindrome based on the most frequent base at each symmetrical position, i.e. in the right half site, the base at the first position is not conserved, however, an A is found at the symmetrical position in the left half-site in all four sites, hence, the first position in the right half site is a T in the symmetry consensus.

footprint in which the protected regions were separated by approximately 10 bp and located between the regions protected by the two cAMP-CRP complexes. Displaying the protein contacts to the backbone on a DNA helix with 10.5 bp per turn reveals that all contacts are positioned on the same side of the helix and span the minor groove (Fig. 5). These results strongly indicate that the repression complex adapts a structure in which the regulatory region of *deoP2* is wrapped around a complex of two cAMP-CRP complexes and one CytR molecule. This picture also highlights that the mutations that directly affect CytR binding to *deoP2* are positioned adjacent to and partially overlapping the regions protected from attack by hydroxyl radicals by CytR. The assessment of the CytR–DNA backbone contacts, together with the genetically defined binding sequence, therefore, suggests that CytR lies across two adjacent minor grooves and recognizes its binding site by contacting bases in two neighbouring major grooves.

The DNA segment in which the five mutations that interfere with CytR–DNA interactions map may be viewed as an imperfect inverted repeat containing the sequence 5'TGCAA on the top strand and 5'TACAA on the bottom strand (Fig. 1). However, two five bp boxes (5'TGCAA and 5'TGTAA on the top strand) oriented in a directly repeated configuration are also contained in this DNA segment (Fig. 1). At present it is not possible to draw any definite conclusions about the binding mode of CytR. However, several lines of evidence implicate that CytR is an oligomer with rotational symmetry that contacts the DNA via an α -helix-turn- α -helix motif. CytR has extensive homology with well-characterized DNA binding proteins like LacI and GalR that contact palindromic sites via two identical subunits (28–31), and specifically CytR contains a region exhibiting a high level

of homology with the α -helix-turn- α -helix motif found in the DNA binding domain of several well-characterized DNA binding proteins (32). Amino acid changes in this region of CytR abolish its repressor function and the mutant proteins are *trans*-dominant (33; Søgaard-Andersen *et al.*, unpubl.). Furthermore, CytR exists as an oligomer in solution (33). Finally, CytR would be expected to make identical contacts to the two cAMP-CRP complexes in *deoP2*. As the arrangement of the repeats in a protein binding site is thought to reflect the structure of the active form of the protein, the most likely binding model is that CytR interacts with an inverted repeat.

Comparison of the CytR recognition motif in *deoP2* with known CytR binding sites reveals that the pentamer motif 5'TGCAA or degenerate versions of this sequence are conserved among all sites (Fig. 6). However, in order to obtain maximal homology with the inverted repeat in *deoP2*, the spacer region between the half-sites must be allowed to vary between 2 bp (in *deoP2* and *tsxP2*) and 3 bp (in *cddP* and *cytRP*) (Fig. 6). There are precedents for such moderate variations in spacer length differences between operator half-sites. The natural *lac* operator is apparently flawed by an incorrect spacing between the operator half-sites with respect to maximal repressor binding (34, 35), also spacer length differences have been suggested between the half-sites in CRP binding sites (36, 37). Finally, two additional lines of evidence strongly emphasize the importance of the inverted repeat for sequence specific binding of CytR. First, mutations in the 5'TGCAA motif in *cddP* reduces binding of CytR (7) (Fig. 6); and, second, the CytR–DNA backbone contacts in *cytRP* are arranged around the inverted repeat in this promoter as they are in *deoP2* (8). Based on the above considerations, a consensus CytR recognition sequence might be

proposed (denoted 'simple consensus' in Fig. 6). This consensus may be refined by imposing a symmetry constraint to give the 'symmetry consensus' (Fig. 6). From the alignment in Fig. 6 it also appears that the CytR recognition sequence is positioned asymmetrically with respect to the two CRP sites in *deoP2* and *cddP*. Furthermore, the CytR recognition sequence is located 2 bp closer to the nearest CRP site in *cddP* and *cytRP* than in *deoP2*. It is still too early to evaluate the importance of these variations, however, they may indicate that CytR is a flexible protein. A simple model based on the homology of CytR with the LacI protein might be that CytR contains an NH₂-terminal DNA binding domain (equivalent to the LacI headpiece (38)) that is connected with the dimerization domain of the protein through a flexible hinge region as was recently described for the $\alpha 2$ protein in yeast (39, 40). Hence, the two parts of the protein could move independently with respect to each other providing CytR with the possibility to interact with binding sites with different structures and cAMP-CRP complexes at varying distances.

ACKNOWLEDGEMENTS

We thank Pia Hovendal and Marianne Hald for expert technical assistance. This work was supported by The Danish Center for Microbiology and The Carlsberg Foundation. L.S.-A. was supported by a post-doctoral fellowship from The Danish Natural Science Research Council.

REFERENCES

- Hammer-Jespersen, K. (1983) In Munch-Pedersen, A. (ed.), *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms*. Academic Press, London, pp.203–258.
- Søgaard-Andersen, L., Pedersen, H., Holst, B. and Valentin-Hansen, P. (1991) *Mol. Microbiol.* **5**, 969–975.
- Søgaard-Andersen, L., Martinussen, J., Møllegaard, N.E., Douthwaite, S.R. and Valentin-Hansen, P. (1990) *J. Bacteriol.* **172**, 5706–5713.
- Søgaard-Andersen, L., Møllegaard, N.E., Douthwaite, S.R. and Valentin-Hansen, P. (1990) *Mol. Microbiol.* **4**, 1595–1601.
- Søgaard-Andersen, L. and Valentin-Hansen, P. (1991) *Mol. Gen. Genet.* **231**, 76–80.
- Pedersen, H., Søgaard-Andersen, L., Holst, B. and Valentin-Hansen, P. (1991) *J. Biol. Chem.* **266**, 17804–17808.
- Holst, B., Søgaard-Andersen, L., Pedersen, H. and Valentin-Hansen, P. (1992) *EMBO J.* **11**, 3635–3643.
- Pedersen, H., Søgaard-Andersen, L., Holst, B., Gerlach, P., Bremer, E. and Valentin-Hansen, P. (1992) *J. Mol. Biol.* **227**, In press.
- Gerlach, P., Søgaard-Andersen, L., Pedersen, H., Valentin-Hansen, P. and Bremer, E. (1991) *J. Bacteriol.* **173**, 5419–5430.
- Søgaard-Andersen, L., Mironov, A.S., Pedersen, H., Sukhodelts, V.V. and Valentin-Hansen, P. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4921–4925.
- Valentin-Hansen, P. (1982) *EMBO J.* **1**, 1049–1054.
- de Crombrughe, B., Busby, S. and Buc, H. (1984) *Science* **224**, 831–838.
- Maniatis, T., Fritsch, E.T. and Sambrook, J. (1982) *Molecular Cloning. A laboratory manual*. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.
- Valentin-Hansen, P., Hammer-Jespersen, K., Boetius, F. and Svendsen, I. (1984) *EMBO J.* **3**, 179–183.
- Valentin-Hansen, P., Svenningsen, A., Munch-Pedersen, A. and Hammer-Jespersen, K. (1978) *Mol. Gen. Genet.* **159**, 191–202.
- Valentin-Hansen, P., Albrechtsen, B. and Løve-Larsen, J.E. (1986) *EMBO J.* **5**, 2015–2021.
- Vieira, J. and Messing, J. (1982) *Gene* **19**, 259–268.
- Myers, R.M., Lerman, L.S. and Maniatis, T. (1985) *Science*, **229**, 242–247.
- Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103–119.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci., USA* **74**, 5463–5467.
- Taylor, J.W., Ott, J. and Eckstein, F. (1985) *Nucl. Acids. Res.* **13**, 8765–8785.
- Ghosaini, L.R., Brown, A.M. and Sturtevant, J.M. (1988) *Biochemistry* **27**, 5257–5261.
- Tullius, T.D. and Dombroski, B.A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5469–5473.
- O'Halloran, T.V., Frantz, B., Shin, M.K., Ralston, D.M. and Wright, J.G. (1989) *Cell* **56**, 119–129.
- Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Dandanell, G., Valentin-Hansen, P., Lve-Larsen, J.E. and Hammer-Jespersen, K. (1987) *Nature* **325**, 823–826.
- Schultz, S.C., Shields, G.C. and Steitz, T.A. (1991) *Science* **253**, 1001–1007.
- Valentin-Hansen, P., Lve-Larsen, J.E., Hjrur, P., Short, S.A. and Barbier, C.S. (1986) *Nucl. Acids. Res.* **14**, 2215–2228.
- Gilbert, W. and Müller-Hill, B. (1967) *Proc. Natl. Acad. Sci. USA* **58**, 2415–2419.
- DiLauro, R., Taniguchi, T., Musso, R. and de Crombrughe, B. (1979) *Nature* **279**, 94–500.
- Majumdar, A. and Adhya, S. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6100–6104.
- Aggarwal, A.K. and Harrison, S.C. (1990) *A. Rev. Biochem.* **59**, 933–969.
- Barbier, C.S. and Short, S.A. (1992) *J. Bacteriol.* **174**, 2881–2890.
- Sadler, J.R., Sasmor, H. and Betz, J.L. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6785–6789.
- Simons, A., Tils, D., Wilcken-Bergmann, B.v. and Müller-Hill, B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1624–1628.
- Berg, O.G. and Hoppel, P.H.v. (1988) *J. Mol. Biol.* **200**, 709–723.
- Barber, A.M. and Zhurkin, V.B. (1990) *J. Biomol. Struct. Dyn.* **8**, 213–232.
- Adler, K., Beyreuther, K., Fanning, E., Geisler, N., Gronenborn, B., Klemm, A., Müller-Hill, B., Pfahl, M. and Schmotz, A. (1972) *Nature* **237**, 322–327.
- Sauer, R.T., Smith, D.L. and Johnson, A.D. (1988) *Genes Dev.* **2**, 807–816.
- Smith, D.L. and Johnson, A.D. (1992) *Cell* **68**, 133–142.