Cell, Vol. 89, 1101–1109, June 27, 1997, Copyright ©1997 by Cell Press

Protein–Protein Communication: Structural Model of the Repression Complex Formed by CytR and the Global Regulator CRP

Birgitte H. Kallipolitis, Mads Nørregaard-Madsen, and Poul Valentin-Hansen
Department of Molecular Biology
Odense University
Campusvej 55
DK-5230 Odense M
Denmark

Summary

The cAMP receptor protein (CRP) and the LacI-related CytR antiactivator bind cooperatively to adjacent DNA sites at or near promoters, an interaction that involves direct protein contacts. Here, we identify a collection of amino acid substitutions in CytR that reestablish protein–protein communication to mutant CRP proteins specifically defective in cooperative binding with wild-type CytR. To assess the location and spatial arrangement of these substitutions, we built a three-dimensional model of CytR based on the recent X-ray structure of the highly homologous PurR repressor bound to DNA. This approach enables us to specify the patch on CytR's surface that contacts CRP. Furthermore, our results permit the construction of a three-dimensional structure of the higher order nucleoprotein complex formed by CytR and CRP.

Introduction

Protein–protein interactions are intrinsic to most cellular processes. Consequently, it is of great importance to understand how proteins interact to perform their biological functions. Recent studies of well-characterized complexes formed by heterologous pairs of proteins have considerably extended our understanding of the nature of protein-contact interfaces and the molecular principles governing protein–protein recognition (for review, see Davies and Cohen, 1996; Jones and Thornton, 1996; Wells, 1996). In particular, it is revealed that distinct protein partners such as antibody–antigen, enzyme–inhibitor, and receptor–hormone share related structural features and common molecular recognition principles for complex formation. In this work, we have explored the molecular architecture of a gene regulatory complex that is detectable only when the interacting partners are appropriately positioned on the DNA helix, i.e., a nucleoprotein complex that is held together by multiple specific protein–DNA and protein–protein interactions.

The system employed is a relatively simple bacterial gene regulatory circuit consisting of two well-characterized proteins, the global activator CRP (also referred to as the catabolite gene activator protein, CAP) and the CytR antiactivator, which act in concert to turn off transcription initiation (for review, see Valentin-Hansen et al., 1996). CytR belongs to the LacI-family of regulatory proteins (Weickert and Adhya, 1992) and exhibits a high level of sequence similarity to the PurR repressor. However, CytR cannot act on its own: efficient DNA binding of CytR relies on cobinding with CRP. In a highly cooperative fashion, the presence of one protein at a DNA site increases the affinity of the other protein for an adjacent binding site (Gerlach et al., 1991; Pedersen et al., 1991, 1992, 1995; Holst et al., 1992). The synergistic binding of CytR and CRP relies on direct protein–protein interactions (Søgaard-Andersen et al., 1991; Søgaard-Andersen and Valentin-Hansen, 1993). Two types of repression complexes can be formed with protein stoichiometries of 1:1 or 1:2. In the latter complex, the dimeric CytR protein is sandwiched between tandem DNA-bound CRP molecules. Structural analysis of these nucleoprotein complexes is greatly facilitated by the knowledge of the three-dimensional structure of CRP (Schultz et al., 1991) and the two CytR-related proteins PurR (Schumacher et al., 1994) and LacI (Lewis et al., 1996), bound to their respective DNA targets. Accordingly, this system is particularly well suited to study how proteins interact on DNA to form multiprotein complexes.

In previous work, the epitope on CRP involved in direct protein contact to CytR has been determined (Søgaard-Andersen et al., 1991). Here, we have used two mutagenesis strategies, molecular modeling, and in vitro binding assays to define the cognate site in CytR. Our results provide information regarding CytR's contact interface to CRP and the chemical nature of the contact residues in CytR that provide most of the binding energy. Using this information, it is possible to propose a three-dimensional model of the higher order nucleoprotein complex formed by CytR and CRP. In this model, the orientation of CytR with respect to the DNA helix is orthogonal to that seen for PurR and LacI bound to DNA.

Results

Because CRP is absolutely required for the action of CytR, it is not feasible to set up a genetic screen that permits the isolation of repressor mutants specifically defective in cooperative binding with CRP. Therefore, we designed a different kind of screen that allowed us to identify residues in CytR that closely approach CRP in repression complexes. To this end, we took advantage of mutant CRP proteins, designated CRP^{nc} for negative-control defective, that interact with the DNA and activate transcription in a manner indistinguishable from wild-type (wt) CRP but fail to support CytR repression (Søgaard-Andersen et al., 1991). In the three-dimensional structure of the CRP–DNA complex, all of the substituted amino acids conferring a CRP^{nc} phenotype are in close proximity on the surface of the large domain of CRP (i.e., on the face opposite the DNA-binding surface of CRP; Figure 1). Three of these amino acids have solvent exposed side chains and, therefore, are candidates to be involved in direct contact with CytR (i.e., amino acids H17, V108, and P110). The basic principle of the screen was to search for suppressor mutations in *cytR* that partially or fully restore repression in the

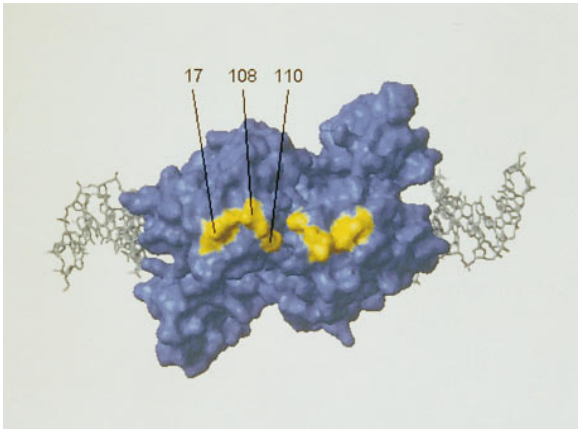

Figure 1. Contact-Surface Representation of the cAMP–CRP Dimer Complexed with DNA

The amino acids H17, V108, and P110, at which substitutions confer a CRP^{nc} phenotype, are shown in yellow.

presence of CRP^{nc} proteins. The functional properties of such compensating mutations can give valuable information on protein–protein interactions. Thus, CytR suppressors may reestablish repression in one of two ways. First, the mutants may create a productive interaction to CRP^{nc} proteins that resembles the contact between the wt components. This type of suppression is generally allele specific. Second, the mutants may provide an independent activity that compensates for the defects in the CRP^{nc} proteins. This type of suppression is not generally allele specific but can provide an indication of the surface of CytR that is in close proximity to CRP in the repression complexes. A prerequisite for this, however, is that the substituted residues are capable of forming new contacts to CRP (i.e., surface exposed).

To facilitate screening, we employed the strain SØ1328KΔcrp (Δ*crp*, Δ*cytR*) carrying a chromosomal fusion between the CRP–CytR-regulated *cdd* promoter and *lacZYA*. On indicator plates, this strain displays a Lac$^-$ phenotype in the presence of *cytR-wt* plus *crp-wt* and a Lac$^+$ phenotype in the presence of *cytR-wt* and the *crp*^{nc} alleles. Thus, in a *crp*^{nc} background the *cytR* alleles that result in enhanced CytR regulation could be readily identified by the Lac phenotype. To facilitate mutagenesis and biochemical analyses of mutant *cytR* genes, we used the *cytR* plasmid pBHK231 carrying a number of unique restriction sites in the structural gene for the repressor (see Experimental Procedures). CytR is slightly overproduced in cells harboring this plasmid, and consequently, repression is enhanced 3- to 4-fold relative to cells that bear a chromosome-encoded *cytR* gene.

Three mutant CytR libraries (pNK0.1, pKB0.6, and pKB1.3) were generated by PCR-based mutagenesis methods, which targeted either the central or C-terminal regions of *cytR*, encoding amino acids 59–270 and amino acids 271–341, respectively. It has previously been shown that a CytR protein deleted for its DNA-binding motif (amino acids 1–57) is still able to repress transcription in vivo and to bind cooperatively with CRP in vitro (Søgaard-Andersen and Valentin-Hansen, 1993; L. M. Lambertsen and L. Søgaard-Andersen, unpublished data). Therefore, none of the libraries employed here carry substitutions in the N-terminal DNA-binding region of *cytR*. Using the three libraries and the tester strain harboring either *crp17R* or *crp108A* (both *crp*^{nc} mutants), we obtained 10 different CytR candidates bearing single amino acid substitutions (Table 1). Five of the mutants (A156V, H165L, F227S, Q298L, and Q314K), however, improve repression only minimally (less than 2-fold relative to CytR-wt).

Allele Specificity of CytR Suppressor Mutants

The remaining five mutants were assayed for their abilities to respond to various *crp*^{nc} alleles. The results of these experiments are summarized in Table 2 and Figure 2. The CytR mutants substituted at residues 171, 322, and 329 respond to all CRP^{nc} mutants and improve repression from ~2- to ~7-fold relative to CytR-wt. The fifth mutant (C213R), however, only improves CytR regulation significantly in cells harboring the CRP-17R protein (4-fold). In addition, all five suppressors exhibited repression values comparable to CytR-wt in the presence of the CRP-wt protein (150-fold; Table 2).

The competence of the mutant CytR proteins to support repression in concert with CRP-wt was further tested in the *cddp* derivative *cddΔ79* (Holst et al., 1992). This promoter, which only bears a single DNA site for CytR and CRP, respectively, is partially repressed by the wt proteins (Table 3. The results established that all five suppressors were slightly more active in transcription repression of *cddΔ79* (improving repression 1.1- to 1.5-fold relative to CytR-wt).

Table 1. Sequences of *cytR* Suppressor Mutants

Amino Acid Substitution	Codon Substitution	Number of Isolates	*crp* Allele Present	Mutant *cytR* Library
A156V	GCA→GTA	1	*crp108A*	pNK0.1
H165L	CAT→CTT	3	*crp108A*	pNK0.1
A171T	GGG→AGC	1	*crp17R*	pNK0.1
C213R	TGC→CGT	1	*crp17R*	pNK0.1
F227S	TTC→TCC	3	*crp108A*	pNK0.1
Q298L	CAG→CTG	2	*crp108A*	pNK0.6
		1	*crp108A*	pKB1.3
Q314K	CAA→AAA	2	*crp108A*	pKB0.6
S322R	AGT→AGA	3	*crp108A*	pKB0.6
		1	*crp108A*	pKB1.3
		1	*crp17R*	pKB0.6
C329S	TGC→AGC	1	*crp108A*	pKB1.3
C329Y	TGC→TAC	2	*crp108A*	pKB0.6

Table 2. Allele Specificity of CytR Suppressor Mutants

CytR Proteins[a]	CRP Proteins[a]			
	CRP-wt	H17R	V108A	P110S
None	4.4[b]	3.9	5.0	4.5
CytR-wt	0.03 (150)[c]	0.41 (9.5)	0.50 (10)	0.30 (17)
A171T	0.03 (150)	0.14 (28)	0.12 (42)	0.11 (41)
C213R	0.03 (150)	0.10 (39)	0.46 (11)	0.21 (21)
S322R	0.03 (150)	0.09 (43)	0.09 (56)	0.04 (113)
C329Y	0.03 (150)	0.17 (23)	0.17 (29)	0.08 (56)
C329S	0.03 (150)	0.15 (26)	0.24 (21)	0.17 (26)

[a] CRP and CytR proteins were expressed from p31 and pBHK plasmids, respectively, in strain SØ1328KΔcrp.
[b] Specific β-galactosidase activities.
[c] Fold of CytR regulation.

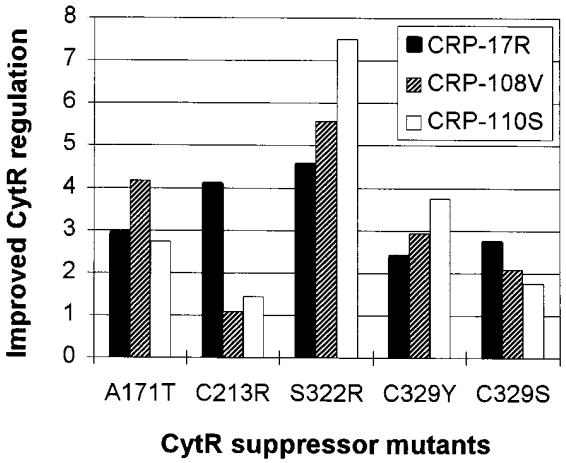

Figure 2. The Improved CytR Regulation Exhibited by CytR Suppressor Mutants at the *cdd* Promoter

In the presence of a given CRP^{nc} mutant, the improved CytR regulation was calculated as: (repression, CytR mutant/repression, CytR-wt) (see Table 2).

Taken together, the suppression analysis shows that four of the CytR mutants exhibit a broad, rather nonspecific suppression pattern (e.g., A171T, S322R, C329S, and C329Y), whereas the fifth mutant (C213R) shows a more narrow range of suppression. It is noteworthy that a *cdd* promoter repressed by any of the five suppressors can be fully induced by addition of cytidine (data not shown). Thus, the mutant proteins retain the ability to bind and respond to the natural inducer, suggesting that the overall protein structure is preserved.

Homology Modeling of the Three-Dimensional Structure of CytR

In order to assess the location and spatial arrangement of the substituted residues in the suppressors, we constructed a three-dimensional model of CytR by homology modeling based on the crystallographic structure of the PurR–DNA complex (Schumacher et al., 1994; Brookhaven accession number 1PNR). Within the LacI family of repressors, CytR and PurR are closely related. Amino acid sequences of the two proteins are 37% identical, with two single-residue gaps and one two-residue gap (see Experimental Procedures and Figure 6). Moreover, the predicted secondary structure of CytR matches the known secondary structure of PurR (Rost and Sander, 1993). Using homology modeling (Greer, 1991) within the WHAT IF suite of programs, we created a structure of CytR (residues 11–341; CytR contains 10 N-terminal amino acids for which there is no corresponding structure in PurR). Once all amino acids in CytR were defined, the model was subjected to multiple cycles of energy minimization (using GROMOS; van Gunsteren and Berendsen, 1977) until there was no further improvement, and the PROCHECK package (Laskowski et al., 1993) was used to evaluate the overall stereochemical quality of the CytR model (see Experimental Procedures). The backbone of PurR and the CytR model can be superimposed with a rmsd of 0.37 Å. A ribbon diagram of the CytR dimer is presented in Figure 3a.

The amino acids in CytR at which substitutions cause suppression of the CRP^{nc} proteins are positioned in three discrete segments on the surface of the core domain of the proposed model of CytR (Figure 3). We refer to these segments as patches 1, 2, and 3, respectively. Amino acid 213 is located on the COOH-terminal subdomain of the core, and the side chain is exposed to the solvent (patch 1). Amino acids 171 and 329 are partially buried and situated close to both the hinge region and the inducer pocket (patch 2). Finally, the side chain of amino acid 322 presumably is solvent exposed and resides in the NH_2-terminal subdomain of the core (patch 3).

Alanine-Scanning Mutagenesis

The structural model of CytR now places us in a position to analyze in further detail the importance of individual residues in or close to the three patches for cooperative binding with CRP. To this end, we performed alanine scanning. Alanine substitution eliminates all side-chain atoms beyond the β carbon and thereby permits identification of individual side-chain atoms involved in protein–protein interaction (Cunningham and Wells, 1989). Our homology modeling indicates that 6 amino acids at patch 1 (N177, Y180, K185, R211, R212, C213), 4 amino acids at patch 2 (T170, D174, D328, E330), 4 amino acids at the border between patches 1 and 2 (L169, F173, Q208, E181), and 10 amino acids at patch 3 (R144, E158, L159, E160, V164, H319, C320, S322, S324, R325) all

Table 3. CytR Regulation at the *cddΔ79*[a] Promoter in the Presence of Different CytR Proteins

CytR Proteins[b]	β-galactosidase Activity	Fold of CytR Regulation
None	2.1	—
CytR-wt	0.15	14
A171T	0.13	16
C213R	0.12	18
S322R	0.10	21
C329Y	0.12	18
C329S	0.12	18

[a] The *cddΔ79* promoter fused to *lacZ* was encoded from plasmid pGB2Δ79 in SØ928K (*crp*$^+$, cytR$^-$).
[b] CytR proteins were expressed from pBHK derivatives.

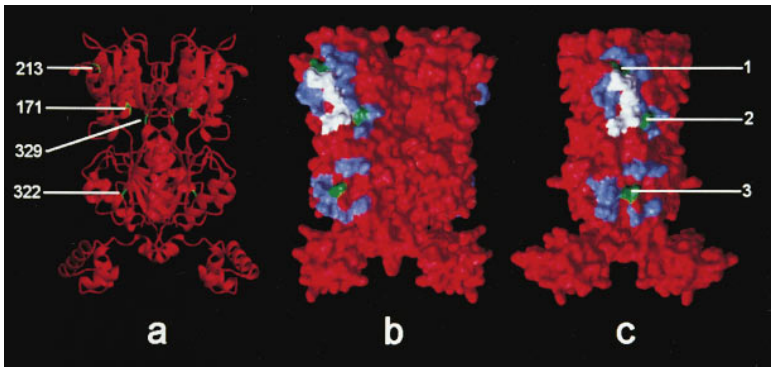

Figure 3. Three-Dimensional Model of CytR

(a) Ribbon diagram of the CytR dimer. The N-terminal DNA-binding domain of CytR can be divided into two important regions. The first contains the helix-turn-helix (HTH) motif, corresponding to helices 1 and 2, which is stabilized by helix 3. This structural element is followed by a linker region that connects the HTH motif to the core domain of the protein. In PurR and LacI, the linker is disordered in the absence of DNA. Upon interaction with specific operator DNA, this region undergoes induced fit, forming a stable DNA-binding unit consisting of two α helices (Nagadoi et al., 1995; Spronk et al., 1996). In CytR, however, the linker is likely to be unstructured also in the DNA-bound state (B. H. K., unpublished data). Each subunit of the dimeric core domain consists of two topologically similar subdomains: the NH_2 subdomain, which is directly attached via the linker region to the DNA-binding domain, and the COOH subdomain. The subdomains are linked by a hinge region, composed of three noncontinuous peptide crossovers, and the inducer-binding pocket is situated between the two subdomains. Amino acid residues 171, 213, 322, and 329, at which substitutions confer a CytR suppressor phenotype, are indicated in green.

(b and c) Contact-surface representation of CytR. Patches 1–3, defined by the CytR suppressor mutants, are shown in green. Amino acid residues 169, 173, and 212, at which alanine substitutions result in CytR mutants defective in repression, are shown in white. The alanine-substituted residues, which are not critical for repression, are indicated in blue. Note that alanine-scanning mutagenesis does not probe backbone nor β carbon atoms. Thus, the importance of several surface-exposed glycine residues in patch 1 (G205 and G214) and patch 3 (G321 and G323) cannot be determined. The figure was created using MOLMOL (Koradi et al., 1996).

have side chains beyond the β carbon that are clearly accessible on the surface of CytR. We reason that these surface-exposed amino acids are candidates to make side-chain interactions with CRP. It should be noted that except for A157, which corresponds to V164 in CytR, all of the analogous residues in PurR are surface exposed as well.

Each of the amino acids was substituted by alanine. Subsequently, the CytR mutants were tested for their ability to support transcription repression in concert with CRP-wt at the *cdd* promoter. The results are presented in Table 4. Alanine substitution of L169 and F173 at the border between patches 1 and 2 results in a dramatic reduction in transcription repression (50- and 65-fold, respectively); alanine substitution of R212 at patch 1 has a smaller effect (7.5-fold). Finally, alanine substitution of T170 at patch 2 has a small but significant effect on regulation (4-fold). No other amino acid tested results in a $>$2-fold defect. These results indicate that amino acids 169, 173, and 212 of CytR are essential for transcription repression at the *cdd* promoter and that a continuos region extending from patch 1 to patch 2 constitutes the functional epitope of CytR.

Analysis of CytR Mutants Defective in Repression

In order to confirm that the CytR proteins substituted by alanine at position 169, 173, or 212 are specifically defective in cooperative binding, we performed a number of tests. First, dissociation half-time experiments were carried out to investigate protein–DNA interactions. The results for CytR-wt and the CytR-173A derivative are presented in Figure 4. When incubated with a consensus CytR operator, both proteins form DNA complexes with identical dissociation half-time ($\sim$1.5 min). Therefore, the repression defect of CytR-F173A cannot be attributed to a reduced DNA-binding affinity. In contrast, the mutant protein is defective in formation of stable repression complexes when incubated with cAMP–CRP and a *deoP2* fragment containing an optimal CytR operator. The half-life of the CytR-173A/cAMP–CRP/DNA complex is $\sim$20 min, whereas the half-life of the wt-repression complex is several hours. Similar binding features were observed for CytR-169A and CytR-212A (data not shown); half-lives of repression complexes containing CytR-169A or CytR-212A are $\sim$20 and $\sim$75 min, respectively. Thus, the competence of CytR-wt and the mutant proteins to repress transcription in concert with CRP in vivo correlates with the stability of the repression complexes in vitro. Furthermore, in vivo experiments revealed that the three mutant proteins are transdominant and that the residual repression activity can be induced by cytidine. From this, we infer that other functions, including dimerization and inducer binding, are not grossly changed in the three CytR derivatives. We conclude that CytR-169A, CytR-173A, and CytR-212A are specifically defective in protein–protein interaction to CRP.

Modeling of the CytR–CRP Repression Complex

By combining the structural information on CytR and CRP, the genetic data on protein–protein interaction, and the DNA-binding features of the two proteins (Pedersen et al., 1992, 1995; Møllegaard et al., 1993; Søgaard-Andersen and Valentin-Hansen, 1993; Pedersen and Valentin-Hansen, 1997), we constructed a three-dimensional model of the multiprotein repression complex formed by CytR and CRP (Figure 5). The starting point for this work was the X-ray structure of two cAMP–CRP molecules situated on the DNA helix with a center-to-center distance of 52 bp. To facilitate modeling, we employed the CytR half-operator settings present in the *nupG* promoter (two inverted octamer boxes separated by 9 bp; Pedersen et al., 1995; Pedersen and Valentin-Hansen, 1997). This results in a roughly symmetrical arrangement of the operator sites for CRP and CytR. Finally, we oriented the CytR dimer between the two CRP molecules so that the functional epitopes on CytR and the adjacent CRP subunits are facing each other and that the helix-turn-helix (HTH) motif of each CytR monomer contacts half-operators in the major grooves in a manner very

Table 4. Alanine-Scanning Mutagenesis of CytR: Effects of Alanine Substitution on Transcription Repression at the *cdd* Promoter

CytR Protein[a]	β-galactosidase Activity	Effect on CytR Regulation[b]
None	4.3	—
CytR-wt	0.02	1.0
R144A	0.02	1.0
E158A	0.02	1.0
L159A	0.04	2.0
E158A, E160A	0.02	1.0
V164A	0.03	1.5
L169A	1.0	**50**
T170A	0.08	**4.0**
F173A	1.4	**65**
D174A	0.02	1.0
N177A	0.02	1.0
Y180A	0.04	2.0
E181A	0.03	1.5
K185A	0.02	1.0
Q208A	0.03	1.5
R211A	0.02	1.0
R212A	0.15	**7.5**
C213A	0.03	1.5
H319A	0.02	1.0
H319A, V320A	0.03	1.5
S322A	0.03	1.5
S324A	0.04	2.0
R325A	0.03	1.5
D328A	0.03	1.5
D328A, E330A	0.03	1.5

[a] CytR proteins were encoded from pBHK derivatives in SØ1328K.
[b] Enzyme activity, mutant CytR/enzyme activity, CytR-wt.

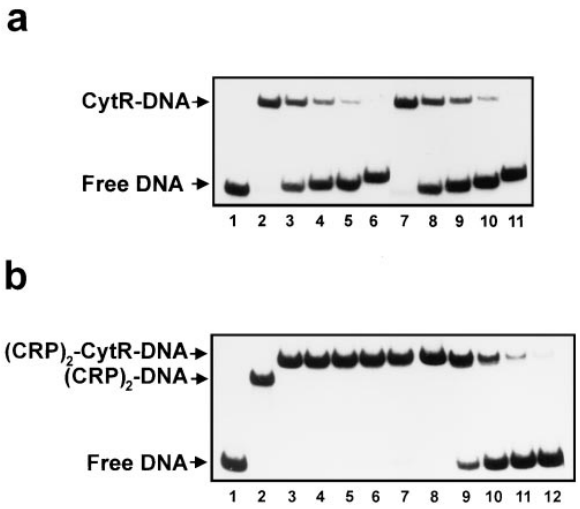

Figure 4. DNA Binding of CytR Protein Derivatives without or with CRP-wt

(a) Half-life experiments for CytR–DNA complexes containing CytR-wt (lanes 2–6) or CytR-173A (lanes 7–11) at 28°C. CytR proteins were bound to labeled DNA fragments containing a consensus CytR target. The dissociation reaction was started by adding unlabeled competitor DNA carrying the consensus target and stopped by loading onto a polyacrylamid gel. Lane 1, free DNA. Dissociation times: lanes 2 and 7, 0 min; lanes 3 and 8, 1 min; lanes 4 and 9, 2 min; lanes 5 and 10, 3 min; lanes 6 and 11, 5 min.

(b) Half-life experiments for $(CRP)_2$–CytR–DNA complexes containing CytR-wt (lanes 3–7) or CytR-173A (lanes 8–12) at 37°C. CRP-wt and CytR proteins were bound to labeled *deoP2*-derived DNA fragments containing a consensus CytR-binding target. In the repression complex formed, the arrangement of the CytR-binding target is homologous to the *cdd* promoter (Pedersen and Valentin-Hansen, 1997). Lane 1, free DNA; lane 2, DNA + CRP. Dissociation times: lanes 3 and 8, 0 min; lanes 4 and 9, 15 min; lanes 5 and 10, 30 min; lanes 6 and 11, 45 min; lanes 7 and 12, 60 min.

similar to PurR. To make the latter point feasible, we stretched out the interdomain linker connecting the HTH motif and the core domain of CytR. In this context, we note that there is an important functional difference between PurR and CytR. PurR carries two specific sets of DNA-binding motifs. One set consists of the HTH motifs, which interact with half-operators in DNA major grooves, and the other consists of two hinge helices that bind at the center of the operator and wedge open the DNA minor groove (Schumacher et al., 1994; see also legend to Figure 3). Thus, DNA recognition sequences for the dimeric PurR protein contain three types of information: the sequence of half-operators, the spacing of the two half-operators, and the sequence of the central base pairs of the operator. In contrast, CytR most likely bears only one specific DNA-binding motif, an HTH DNA-binding domain structurally similar to that of PurR, and is a highly adaptable DNA binder that can interact with various settings of half-operators (Pedersen and Valentin-Hansen, 1997). We envisage that the observed plasticity is provided by an unstructured and flexible interdomain linker. This would provide the DNA-binding domains with a considerable freedom of movement and allow interaction with half-operators in different orientations and with different length of spacer regions.

Discussion

In previous work, we have defined the region on CRP that is essential for protein–protein communication to CytR. The region consists of two discontinuous surface-exposed α helices (helices A and B). In the structure of CRP bound to DNA, the two α helices are situated in close proximity in a prominently accessible portion of CRP (i.e., on the face opposite to the DNA-binding surface of CRP; Figure 1). Here, we have combined genetic, biochemical, and homology-modeling analyses to define the cognate contact site in CytR and to explore the molecular architecture of the CytR–CRP nucleoprotein complex.

In the first approach, random mutagenesis of CytR-wt and screening for derivatives that strengthen repression in the presence of CRP^{nc} proteins yielded five nonallele-specific CytR suppressors carrying single amino acid substitutions at position 171, 213, 322, or 329. Homology modeling indicates that the substituted residues are located on a single face of the subunits of CytR but clearly separated from each other (patches 1–3; Figure 3). Furthermore, the side chain of amino acids 213 (patch 1) and 322 (patch 3) is exposed to the solvent, whereas amino acids 171 and 329 (patch 2) are partially buried in the proposed CytR structure. Finally, alanine scanning of the three patches unambiguously established that the critical residues for protein–protein communication to CRP are located in a region extending from patch 1 to patch 2 (just adjacent to amino acids 213 and 171).

The Chemical and Structural Nature of the Targets

Alanine substitution of three amino acids within CytR caused a substantial reduction of repression in vivo and

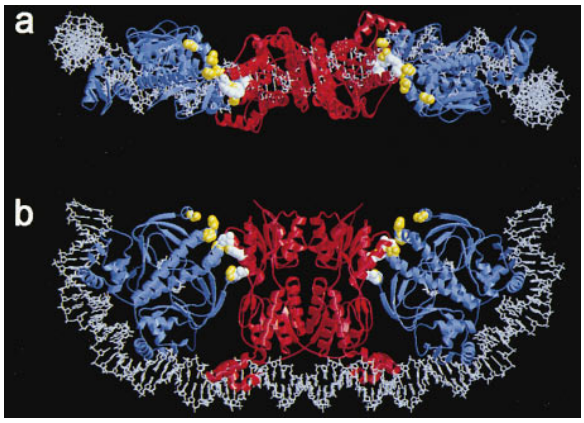

Figure 5. Structural Model of the CytR–CRP–DNA Repression Complex

(a) Top-view. (b) Side-view. The CytR dimer (red) is oriented between the two CRP molecules (blue) so that the repression patch on CRP defined by nc mutations (shown in yellow) and the repression patch on CytR defined by alanine substitutions (shown in white) are head to head. The figure was created using MOLSCRIPT (Kraulis, 1991) and Raster3D (Bacon and Anderson, 1988; Merritt and Murphy, 1994).

```
                .         .         .         .         .
  1 ........MATIKDVAKRANVSTTTVSHVINKTRFVAEETRNAVWAAIKE  42
            ||:|||| :|.|||.|||:.: ..  |.:.||| |  | :|
  1 MKAKKQETAAIMKDVALKAKVSTATVSRALMNPDKVSQATRNRVEKAARE  50
                .         .         .         .         .
 43 LHYSPSAVARSLKVNHTKSIGLLATSSEAAYFAEIIEAVEKNCFQLGYTL  92
    : | | :::|.:| |...| ::... .::|.||| ::| .. ..|| :
 51 VGYLPQPMGRNVKRNESRTILVIVPDICDPFFSEIIRGIEVTAANHGYLV 100
                .         .         .         .         .
 93 ILGNAWNNLEKQRAYLSMMAQKRVDGLLVMCSEYPEPLLAMLEEYRHIPM 142
    ::|:. :. :.::.::.:: |.:||:|::.| .  |: | :|| |::|
101 LIGDCAHQNQQEKTFIDLIITKQIDGMLLLGSRL..PFDASIEEQRNLPP 148
                .         .         .         .         .
143 VVMDWGEAKADF.TDAVIDNAFEGGYMAGRYLIERGHREIGVIPGPLERN 191
    :||. : |..     . ||| :.::: |..||.|.||: ||.|:|| |.
149 MVMANEFAPELELPTVHIDN.LTAAFDAVNYLYEQGHKRIGCIAGPEEMP 197
                .         .         .         .         .
192 TGAGRLAGFMKAMEEAMIKVPESWIVQGDFEPESGYRAMQQILSQPHRPT 241
     . ||.|::.|:  . |.|.. :|..|||. |.| :||||:|. |:.||
198 LCHYRLQGYVQALRRCGIMVDPQYIARGDFTFEAGSKAMQQLLDLPQPPT 247
                .         .         .         .         .
242 AVFCGGDIMAMGALCAADEMGLRVPQDVSLIGYDNVRNARYFTPALTTIH 291
    |||| :|:||:|||:.|.  ||:||:|:|:||:||:  ..: .|:||||
248 AVFCHSDVMALGALSQAKRQGLKVPEDLSIIGFDNIDLTQFCDPPLTTIA 297
                .         .         .
292 QPKDSLGETAFNMLLDRIVNKREEPQSIEVHPRLIERRSVADGPFRDYRR 341
    ||: .:| .|: :|||.: ..: :..|  :.. || | |. . |
298 QPRYEIGREAMLLLLDQMQGQHVGSGSRLMDCELIIRGSTRALP..... 341
```

Figure 6. Amino Acid Sequence Alignment of PurR (Upper) and CytR (Lower)

in the half-life of repression complexes in vitro. By far the greatest defect in repression occurred on substituting two hydrophobic residues (L169 and F173, 50- and 65-fold, respectively), whereas a more moderate reduction was seen by substituting a positively charged residue (R212, 7.5-fold). In addition, a substitution of a neighboring polar residue (T170) for alanine resulted in a minor defect in regulation (4-fold). Alanine substitution of no other amino acid tested resulted in a >2-fold defect. On the basis of transdominans assays, induction experiments, and in vitro DNA-binding experiments, we conclude that CytR proteins carrying alanine substitutions of L169, F173, or R212 are specifically defective in protein–protein interaction to CRP.

In the model of CytR, the three essential residues, as well as T170, map to a contiguous patch on the outer surface of the C-terminal subdomain (Figure 3). When compared with the contact site on CRP, a striking complementarity is revealed: two critical hydrophobic residues on CRP (V108 and P110; K. Lederballe, L. Søgaard-Andersen, and P. V.-H., unpublished data) match the most important residues in the functional epitope of CytR (L169 and F173). We suggest that these hydrophobic residues may be in direct contact in the repression complex and that hydrophobicity is the major factor in stabilizing protein–protein interaction between CytR and CRP. Moreover, we note that each functional epitope is composed of determinants located in two discontinuous polypeptide segments and that molecular modeling of repression complexes suggests that the protein interface or structural epitopes are large (see below).

Our analyses support the view that distinct partners use common themes in protein–protein recognition. Thus, structural and genetic studies of well-characterized interfaces in enzyme–inhibitor complexes (Schreiber and Fersht, 1995), antibody–antigen complexes (Jin et al., 1992; Kelley and O'Connell, 1993; Davies and Cohen, 1996), and receptor–hormone complexes (for review, see Wells, 1996) have revealed the following general features of the recognition sites: (i) the functional epitope is much smaller than the structural epitope and is generally located in discontinuous structures that make up a highly accessible patch in the folded polypeptide chain (e.g., the integral part of the structure), and (ii) the binding determinants are dominated by a small number of side chains in each binding site and are typically nonpolar or charged in character. In particular, the contact determinants defined for CytR and CRP resemble the functional epitopes of the human growth hormone and its receptor, consisting of central crucial hydrophobic residues surrounded by less important polar and charged groups (Clackson and Wells, 1995).

Features of the CytR–CRP Repression Complex

The present work further permitted the construction of a model of the entire repression complex, in which CytR is sandwiched between tandem DNA-bound CRP complexes (Figure 5). Three features of the architectural setup are worth noting. First, the orientation of the core domain of CytR with respect to the DNA helix is orthogonal to that seen for PurR bound to DNA (Schumacher et al., 1994). Second, in contrast to PurR, CytR does not induce a significant bend or kink in the DNA upon binding (H. Pedersen and P. V.-H., unpublished data). Third, a spacefill model of the nucleoprotein complex indicates that the protein interface between CRP and CytR is considerably larger (~11 × 35 Å) than the functional region ranging from patch 1 to patch 2 (~10 × 15 Å).

We can now consider the possible nature of the CytR suppressor mutants. As shown here, residues 213 and 322 of CytR are expected to be exposed to the solvent in the unbound state but are in very close proximity to CRP in the repression complexes. Substitution of C213

and S322 (small, uncharged polar groups) with a larger positively charged residue (R) resulted in suppressors that strengthen repression in the presence of both wt and nc CRP proteins. Therefore, it appears likely that substitution by arginine at these two positions improves repression by creating nonnative favorable interactions with CRP. The suppressors carrying substitutions in the hinge region connecting the two subdomains of the core (A171T, C329S, and C329Y) bear replacements of partially buried amino acid side chains. We infer that the substitutions in the hinge region affect repression indirectly by stabilizing a CytR conformation that is favorable for interacting with CRP.

Experimental Procedures

Enzymes used for DNA manipulations and Taq polymerase for PCR were purchased from Boehringer Mannheim. ^{32}P-labeled nucleotides were obtained from New England Nuclear-DuPont. Isolation of plasmid DNA, cloning, and cell transformation were carried out as previously described (Maniatis et al., 1989).

Bacterial Strains

E. coli strains used were as follows: SØ928K (Δ*deo*, Δ*lac*, Δ*cytR::Kan*^{r}); SØ1328K (as SØ928K but *cdd::*Δ*Mu::*S*λRp*1[209] by P1 transduction); SØ1328KΔcrp (as SØ1328K but Δ[*crp*]*96 zhd-732*::Tn10 by P1 transduction) (Søgaard-Andersen and Valentin-Hansen, 1993).

Plasmids

Plasmids used were p31 derivatives (Kan^R) encoding CRP^{wt} or CRP^{nc} proteins (Søgaard-Andersen et al., 1991).

To create the *cytR* plasmid pBHK231 (Ap^R), unique restriction sites for SnaBI, NcoI, and KpnI were introduced in *cytR* at bp positions −5, +172, and +812, respectively, by site-directed mutagenesis. First, an M13mp19 derivative (Yanisch-Perron et al., 1985) containing *cytR* sequences (EcoRI-SmaI fragment) from −247 to +145 (Pedersen et al., 1991) and an M13mp19 derivative containing *cytR* sequences from +145 to +1070 (SmaI-AflII fragment from pVH002; Valentin-Hansen et al., 1986) were subjected to site-directed mutagenesis (Taylor et al., 1985). None of the mutations lead to changes in the amino acid sequence. The mutated *cytR* fragments were joined by cloning into pUC13 cut with EcoRI and SmaI. An EcoRI-BamHI fragment containing the entire mutated *cytR* was then cloned into the R1 low copy number plasmid pJM241 cut with EcoRI and BamHI. Prior to this, a ClaI site in pJM241 had been filled in and thereby eliminated, resulting in a unique restriction site for ClaI in pBHK231 at bp position +347 in *cytR*. Finally, the 250 bp BglII-BglII fragment, carrying a part of *copB*, was deleted. Thereby, pBHK derivatives have a copy number of 6–8 per genome equivalent (Løve Larsen et al., 1984). To construct pBHK230 (Δ*cytR*), pBHK231 was restricted with BamHI and EcoRI, and after filling in the protruding ends the vector fragment was circularized.

Site-directed mutagenesis was performed by PCR. Mutagenized ClaI-BamHI fragments were cloned into ClaI-BamHI-restricted pBHK231.

pGB2Δ79 (Str^R) was constructed by cloning an EcoRI-SalI fragment containing a translational fusion between *cdd*Δ*79* and *lacZ* (Holst et al., 1992) into EcoRI-SalI-restricted pGB2 (Churchward et al., 1984).

Construction of Mutant *cytR* Libraries

Random mutagenesis was performed by error-prone PCR amplification (Leung et al., 1989). The entire *cytR* segment of pBHK231 was subjected to 25 cycles of PCR amplification using the two primers ocytR13 (5′-GATGTAGTACGCCTGACG-3′) and ocytR25 (5′-GCCC GGATCCCCTTAAGCGG-3′). The forward primer, ocytR13, and the reverse primer, ocytR25, flank the *cytR* gene in pBHK231. In the presence of 0.5 mM $MnCl_2$ and a biased nucleotide pool ([dGTP] = 0.2 mM, [dATP] = 0.2 mM, [dTTP] = 1 mM, [dCTP] = 1 mM), mutations were observed at a frequency of 1.3 per 100 base pairs. In the presence of 0.5 mM $MnCl_2$ and [dNTP] = 1 mM, mutations were observed at a rate of 0.6 per 100 bp. At standard PCR, the frequency of mutations was 0.1 per 100 bp.

Three different sets of mutant *cytR* libraries were generated: pKB0.6 and pKB1.3, carrying mutations in the C-terminal part of *cytR*, and pNK0.1, carrying mutations in the central part of *cytR*. pKB0.6 and pKB1.3 were constructed by joining KpnI-BamHI-restricted PCR fragments with purified KpnI-BamHI fragments of pBHK231. Correspondingly, pNK0.1 was constructed by employing NcoI-KpnI-restricted PCR and vector fragments.

β-galactosidase Assay

Cells were grown at 35°C in AB minimal medium (Clarke and Måløe, 1967) supplemented with 0.5% glycerol, 0.05% casamino acids, and antibiotics (for pBHK-derivatives, 30 μg/ml ampicillin; for pGB2Δ79, 200 μg/ml spectinomycin). For induction experiments, cytidine was added to a final concentration of 1 mM. Enzyme levels were determined as described by Miller (1972), and amounts of β-galactosidase are expressed as OD_{420}/OD_{450}/min.

Preparation of Protein Extracts

CRP protein was purified as previously described (Ghosaini et al., 1988). CytR protein derivatives were purified as described by Pedersen et al. (1991).

Gel-Retardation Assays

Radiolabeled DNA fragments and competitor DNA-containing consensus operators for CytR (CytR alone, AATGTAAC-GC-GTTGCATT; CytR plus cAMP–CRP, GTTGCATT-C-GTTGCACG) were produced by PCR amplification of the two pUC13 derivatives A12–30 and C09–20 (Pedersen and Valentin-Hansen, 1997) using the deoprim1 (5′-CCGAATTCCCTTTGAAAGTGAATTATTTGAACCAGATCGCA-3′) and deoprim2 (5′-GAGATCCGGAACACACTTCGATACACA-3′) primers. Radiolabeled PCR products were purified from 6% polyacrylamide gels. Gel-retardation experiments were performed according to Pedersen et al. (1991). The final concentration of the labeled fragments was 0.5 nM. In the absence of cAMP–CRP, the CytR proteins were added to final concentrations of 1.5 nM. In the presence of CRP (10 nM), CytR proteins were added to final concentrations of 3 nM (CytR-169A and CytR-173A) or 1 nM (CytR-212A and CytR-wt). At various times, the dissociation time course was started by adding a large molar excess (more than 100-fold) of competitor DNA. The samples were loaded on a 6% polyacrylamide gel with current on (200 V). Band intensities were measured by using a Molecular Dynamics PhosphorImager.

Molecular Modeling

A three-dimensional structure model of CytR was constructed using a homology-modeling approach (Greer, 1991). The modeling is based on the coordinates of the PurR repressor crystal solved to 2.7 Å resolution (Schumacher et al., 1994; Brookhaven accession number 1PNR). The WHAT IF program suite (Vriend, 1990) was used to assign homologous residues onto the known PurR structure according to the amino acid alignment in Figure 6. For regions that are nonhomologous, the structure was modeled using the DGLOOP feature of WHAT IF. Once the initial model was built, energy minimization was performed with the GROMOS program (van Gunsteren and Berendsen, 1977) to ensure that bond lengths, bond angles, dihedral angles, etc., were within physically meaningful values. The PROCHECK package (Laskowski et al., 1993) was used to evaluate and ensure the overall stereochemical quality of the model. The stereochemical analysis was performed relative to structures at 2.7 Å resolution, which is the resolution of the PurR crystal coordinates. For the main-chain parameters, it was found that >90% of residues in a Ramachandran plot are in the most favored regions, and peptide bond planarity, number of bad nonbonded interactions, and β carbon tetrahedral distortion, as well as hydrogen bond energies, are within or better than the mean of structures at 2.7 Å resolution. For side-chain parameters, it was found that chi angles are within or better than the mean of structures at 2.7 Å resolution. All calculations was performed on a Silicon Graphics Indy R4000.

Acknowledgments

This work was supported by the Danish Center for Interaction, Structure, Function and Engineering of Macromolecules. We thank R. Bywater for advice in choosing software for protein homology modeling and H. Pedersen for comments on the manuscript.

Received February 3, 1997; revised May 12, 1997.

References

Bacon, D.J., and Anderson, W.F. (1988). A fast algorithm for rendering space-filling molecule pictures. J. Mol. Graph. *6*, 219–220.

Churchward, G., Belin, D., and Nagamine, Y. (1984). A pSC101-derived plasmid which shows no sequence homology to other commonly used cloning vectors. Gene *31*, 165–171.

Clackson, T., and Wells, J.A. (1995). A hot spot of binding energy in a hormone-receptor interface. Science *267*, 383–386.

Clarke, J.D., and Måløe, O. (1967). DNA replication and the division cycle in Escherichia coli. J. Mol. Biol. *23*, 99–112.

Cunningham, B.C., and Wells, J.A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science *244*, 1081–1085.

Davies, D.R., and Cohen, G.H. (1996). Interactions of protein antigens with antibodies. Proc. Natl. Acad. Sci. USA *93*, 7–12.

Gerlach, P., Søgaard-Andersen, L., Pedersen, H., Martinussen, J., Valentin-Hansen, P., and Bremer, E. (1991). The cyclic AMP (cAMP)-cAMP receptor protein complex functions both as an activator and as a corepressor at the *tsx-p2* promoter of Escherichia coli K-12. J. Bacteriol. *173*, 5419–5430.

Ghosaini, L.R., Brown, A.M., and Sturtevant, J.M. (1988). Scanning calorimetric study of the thermal unfolding of catabolite activator protein from Escherichia coli in the absence and presence of cyclic mononucleotides. Biochemistry *27*, 5257–5261.

Greer, J. (1991). Comparative modeling of homologous proteins. Meth. Enzymol. *202*, 239–252.

Holst, B., Søgaard-Andersen, L., Pedersen, H., and Valentin-Hansen, P. (1992). The cAMP-CRP/CytR nucleoprotein complex in Escherichia coli: two pairs of closely linked binding sites for the cAMP-CRP activator complex are involved in combinatorial regulation of the *cdd* promoter. EMBO J. *11*, 3635–3643.

Jin, L., Fendly, B.M., and Wells, J.A. (1992). High resolution functional analysis of antibody-antigen interactions. J. Mol. Biol. *226*, 851–865.

Jones, S., and Thornton, J.M. (1996). Principles of protein–protein interactions. Proc. Natl. Acad. Sci. USA *93*, 13–20.

Kelley, R.F., and O'Connell, M.P. (1993). Thermodynamic analysis of an antibody functional epitope. Biochemistry *32*, 6828–6835.

Koradi, R., Billetter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. *14*, 51–55.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. J. Appl. Cryst. *24*, 946–950.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. *26*, 283–291.

Leung, D.W., Chen, E., and Goeddel, D.V. (1989). A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique *1*, 11–15.

Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G., and Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science *271*, 1247–1254.

Løve Larsen, J.E., Gerdes, K., Light, J., and Molin, S. (1984). Low-copy-number plasmid-cloning vectors amplifiable by derespression of an inserted foreign promoter. Gene *28*, 45–54.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1989). Molecular Cloning: A Laboratory Manual, 2nd Ed. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Merritt, E.A., and Murphy, M.E.P. (1994). Raster3D Version 2.0—a program for photorealistic molecular graphics. Acta Cryst. *D50*, 869–873.

Miller, J.H. (1972). Experiments in Molecular Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Møllegaard, N.E., Rasmussen, P.B., Valentin-Hansen, P., and Nielsen, P.E. (1993). Characterization of promoter recognition complexes formed by CRP and CytR for repression and by CRP and RNA polymerase for activation of transcription on the Escherichia coli *deoP2* promoter. J. Biol. Chem. *268*, 17471–17477.

Nagadoi, A., Morikawa, S., Nakamura, H., Enari, M., Kobayashi, K., Yamamoto, H., Sampei, G., Mizobuchi, K., Schumacher, M.A., Brennan, R.G., and Nishimura, Y. (1995). Structural comparison of the free and DNA-bound forms of the purine repressor DNA-binding domain. Structure *3*, 1217–1224.

Pedersen, H., and Valentin-Hansen, P. (1997). Protein-induced fit: the CRP activator protein changes sequence-specific DNA recognition by the CytR repressor, a highly flexible LacI member. EMBO J. *16*, 2108–2118.

Pedersen, H., Søgaard-Andersen, L., Holst, B., and Valentin-Hansen, P. (1991). Heterologous cooperativity in Escherichia coli: the CytR repressor both contacts DNA and the cAMP receptor protein when binding to the *deoP2* promoter. J. Biol. Chem. *266*, 17804–17808.

Pedersen, H., Søgaard-Andersen, L., Holst, B., Gerlach, P., Bremer, E., and Valentin-Hansen, P. (1992). cAMP-CRP activator complex and the CytR repressor protein bind co-operatively to the *cytRP* promoter in Escherichia coli and CytR antagonizes the cAMP-CRP-induced DNA bend. J. Mol. Biol. *227*, 396–406.

Pedersen, H., Dall, J., Dandanell, G., and Valentin-Hansen, P. (1995). Gene regulatory modules in Escherichia coli: nucleoprotein complexes formed by cAMP-CRP and CytR at the *nupG* promoter. Mol. Microbiol. *17*, 843–853.

Rost, B., and Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. J. Mol. Biol. *232*, 584–599.

Schreiber, G., and Fersht, A.R. (1995). Energetics of protein–protein interactions: analysis of the Barnase-Barstar interface by single mutations and double mutant cycles. J. Mol. Biol. *248*, 478–486.

Schultz, S.C., Shields, G.C., and Steitz, T.A. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90°. Science *253*, 1001–1007.

Schumacher, M.A., Choi, K.Y., Zalkin, H., and Brennan, R.G. (1994). Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α helices. Science *266*, 763–770.

Søgaard-Andersen, L., and Valentin-Hansen, P. (1993). Protein–protein interactions in gene regulation: the cAMP–CRP complex sets the specificity of a second DNA-binding protein, the CytR repressor. Cell *75*, 557–566.

Søgaard-Andersen, L., Mironov, A.S., Pedersen, H., Sukhodelets, V.V., and Valentin-Hansen, P. (1991). Single amino acid substitutions in the cAMP receptor protein specifically abolish regulation by the CytR repressor in Escherichia coli. Proc. Natl. Acad. Sci. USA *88*, 4921–4925.

Spronk, C.A.E.M., Slijper, M., van Boom, J.H., Kaptein, R., and Boelens, R. (1996). Formation of the hinge helix in the *lac* repressor is induced upon binding to the *lac* operator. Nature Structural Biol. *3*, 916–919.

Taylor, J.W., Ott, J., and Eckstein, F. (1985). The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorotioate-modified DNA. Nucleic Acids Res. *13*, 8765–8785.

Valentin-Hansen, P., Løve Larsen, J.E., Højrup, P., Short, S.A., and Barbier, C.S. (1986). Nucleotide sequence of the CytR regulatory gene of Escherichia coli K-12. Nucleic Acids Res. *14*, 2215–2228.

Valentin-Hansen, P., Søgaard-Andersen, L., and Pedersen, H. (1996). A flexible partnership: the CytR anti-activator and the cAMP-CRP activator protein, comrades in transcription control. Mol. Microbiol. *20*, 461–466.

van Gunsteren, W.F., and Berendsen, H.J.C. (1977). Algorithms for macromolecular dynamics and constraint dynamics. Mol. Phys. *34*, 1311–1327.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. *8*, 52–56.

Weickert, M.J., and Adhya, S. (1992). A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem. *267*, 15869–15874.

Wells, J.A. (1996). Binding in the growth hormone receptor complex. Proc. Natl. Acad. Sci. USA *93*, 1–6.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and puC19 vectors. Gene *33*, 103–119.