Order Number 9521283

Structural and functional studies of the *Escherichia coli* cyclic AMP receptor protein

Jin, Ruzhong, Ph.D.

City University of New York, 1995

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

STRUCTURAL AND FUNCTIONAL STUDIES

OF THE *ESCHERICHIA COLI* CYCLIC AMP RECEPTOR PROTEIN

by

RUZHONG JIN

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy , The City University of New York

1995

This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/30/95
Date

Chairman of Examining Committee
Prof. Joseph S. Krakow, Hunter College

2/1/95
Date

Executive Officer
Dr. Richard L. Chappell

Prof. David Foster, Hunter College

Prof. Rivka Rudner, Hunter College

Prof. David Calhoun, City College

Prof. Michael Small, University of Medicine and Dentistry of New Jersey

Supervising Committee

The City University of New York

ABSTRACT

STRUCTURAL AND FUNCTIONAL STUDIES OF

THE *ESCHERICHIA COLI* CYCLIC AMP RECEPTOR PROTEIN

by

Ruzhong Jin

Adviser: Dr. Joseph S. Krakow

The *Escherichia coli* cAMP protein, which is referred to as CRP or CAP, is an allosteric transcriptional regulatory protein which regulates the transcription of more than 20 genes in *E. coli.* The relationship between CRP structure and how it mediates positive control of gene expression is investigated in this dissertation. One surface exposed loop of CRP located at amino acids 156-162 has been shown to be involved in contact with the α subunit of RNA polymerase and important in activating transcription. Insertion mutants were generated in this loop. All insertion mutants retained normal DNA binding activity and showed fairly high activity on the *gal* P1 (class II promoter, centered at -41.5) compared to wild type CRP. But several showed defective positive control activity on the *lac* P1 (class I promoter,

centered at -61.5). The results indicate that the conformation of the 156-162 loop of CRP is important in determining a functional interaction with RNA polymerase. Another surface exposed loop at around amino acid 52 is also considered to be a potential candidate for transcription activation. Using protein-protein photocrosslinking, the results showed that the 52-loop of CRP is in physical proximity to the σ subunit of RNA polymerase. The superactive CRP suppressor mutant, K52N, increased the efficiency of crosslinking, indicating that the K52N mutation improved the physical interaction between the CRP 52-loop and the σ subunit. As a control, the 156-162 loop was found to crosslink only with the RNA polymerase α subunit both on *lac* and *gal* promoters in the presence of cAMP, indicating that the 156-162 loop is close to the α subunit on the *gal* promoter as well as on the *lac* promoter.

Five conserved glycine residues at positions 33, 45, 56, 67 and 71 of CRP were substituted with alanine; G33C, G56C and G67C showed comparable *in vivo* activity on both *lac* and *gal* promoters with the wild type CRP. G71C showed only ~60% of the wild type activity and G45C was almost totally inactive. Western blotting indicated that G45C is a very unstable protein *in vivo* which may explain its low activity.

ACKNOWLEDGMENTS

I would like to express my appreciation and gratitude to my mentor, Dr. Joseph S. Krakow. His profound knowledge, keen scientific thinking, sincere scientific attitude, good sense of humor and constant support and encouragement made this work presented here possible. Special thanks are giving to Drs. Rivka Rudner, David Foster, David Calhoun, Michael Small, and Louis Levinger for their invaluable advice and encouragement.

I would like to thank Drs. Zhaohui Yang, Jianying Luo and Karim Sharif for teaching me the necessary techniques, invaluable discussion and generous help. I appreciate the opportunity to work together with those great fellows in the lab.

Thanks are due to the Biology Department faculties and staff who have provided much helpful assistance.

I thank my family members, especially my parents and grandparents, for their enormous love and encouragement throughout all my education. Without them, I could have never reached that far.

Finally, a special appreciation is given to my wife Xiaomei and my daughter Jennifer for their love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Ab: antibody

ApA: adenylyl (3', 5') adenosine

BSA: bovine serum albumin

BPM: benzophenone-4-maleimide

cAMP: 3', 5'-cyclic adenosine monophosphate

CAP: catabolite gene activator protein

CCpmelR: a synthetic *gal*-like promoter

CC+20pmelR: a synthetic *lac*-like promoter

CRP: cyclic AMP receptor protein

DNase I: deoxyribonuclease I

dNTP: deoxyribonucleoside triphosphate

DTT: dithiothreitol

EDTA: (ethylenedinitrilo) tetraacetic acid

gal P^+: wild type *gal* promoter

gal: galactose operon

kD: kilodalton

lac P^+: wild type *lac* promoter

lac: lactose operon

NFCM: non-fat Carnation milk

NTP: ribonucleoside triphosphate

$O^{met}GTP$: 3'-O-methylguanosine 5'-triphosphate

PAGE: polyacrylamide gel electrophoresis

PCR: polymerase chain reaction

PMSF: phenylmethanesulfonyl fluoride

RP_c: closed promoter complex

RP_o: open promoter complex

SDS: sodium dodecyl sulfate

TEMED: N, N, N', N'-tetramethylethylenediamine

Tris: Tris-(hydroxymethyl)aminomethane

UV: ultraviolet light

WT: wild type

INTRODUCTION

The *Escherichia coli* cAMP receptor protein (CRP, also referred to as the catabolite gene activator protein, CAP) is involved in the regulation of transcription of at least 25 genes. CRP is a dimeric protein of 45,000 daltons composed of two identical subunits of 209 amino acids (Anderson *et al.*, 1971; Aiba *et al.,* 1982; Cossart and Gicquel-Sanzey, 1982). The complete amino acid sequence has been deduced from the nucleotide sequence of the cloned CRP gene (Aiba *et al.,* 1982).

The crystal structure of CRP complexed with cAMP has been determined at 0.29 nm and further refined at 0.25 nm resolution (McKay and Steiz, 1981; Weber and Steiz, 1987). The CRP monomer has two structural and functional domains. The larger N-terminal domain contains extensive ß sheet structure forming the cAMP binding pocket, and is also involved in subunit-subunit interaction. The smaller C-terminal domain consists primarily of α helices involved in sequence specific DNA binding (Krakow and Pastan, 1973; McKay *et al.,* 1982), this helix-turn-helix DNA binding domain shows structural and sequence homologies with several other gene regulatory proteins such as the *cro,* cI and *lac* repressors (Steiz *et al.,* 1982; Matthews *et al.*, 1982; Sauer *et al.*, 1982; Weber *et al.*, 1982; Steiz and Weber, 1984). These two domains are connected by a "hinge" region which runs from residue 130 to 138. Although the crystal structure information for CRP is based on the CRP-$(cAMP)_2$ complex, it is believed that the physiologically active form of CRP is CRP-$(cAMP)_1$ (Takahashi *et al*,, 1989). The *in vivo* physiological

concentration range of cAMP is between 0 to 10 μM; the major complex of CRP and cAMP at cAMP concentration from 0 to 200 μM is CRP-$(cAMP)_1$ (Heyduk *et al.*, 1989).

The crystal structure of cAMP/CRP indicates that the cAMP is deeply buried within the CRP dimer in the cAMP binding pockets (Weber and Steitz, 1987). CRP is an allosteric regulatory protein. Binding of cAMP to the N-terminal domain would elicit conformational changes in the C-terminal domain (Eilen and Krakow, 1977; Eilen *et al.*, 1978). The earliest evidence for such a conformational change is that unliganded CRP is relatively resistant to attack by several proteases but is readily attacked in the presence of cAMP (Eilen *et al.*, 1978). Additional evidence for cAMP induced conformational change in CRP is provided by intersubunit crosslinking using dithionitrobenzoic acid (DTNB), this intersubunit crosslinking is cAMP dependent. (Eilen and Krakow, 1977). In the absence of cAMP, CRP only shows nonspecific DNA binding activity. On binding cAMP, CRP undergoes a conformational change which makes it capable of binding to specific DNA sites and activating transcription from the CRP-dependent promoters (Busby and Ebright, 1994). Although the major role of CRP is to activate transcription from most catabolite repressible genes, it can also act to repress some systems, such as the *crp* promoter (Aiba, 1983). In the *lac* (Malan and McClure, 1984) and *gal* (Musso *et al.*, 1977) operons, in which there are two overlapping promoters, the binding of the cAMP-CRP complex stimulates transcription from one promoter while inhibiting the other. In *ara*BAD (Lee *et al.*, 1981), *malK-malB* and *malE, F, G* (Chapon, 1982)

operons, additional regulatory proteins, such as Ara C or Mal T proteins, are also required for transcription activation.

cAMP independent mutants of CRP, termed CRP*, have been isolated. They do not require cAMP binding to activate transcription of those normally cAMP/CRP dependent promoters. One CRP* mutant, CRP*598 (R142H, A144T) was well characterized. The unliganded CRP*598 exhibits the behavior of a wild type CRP/cAMP complex (Ren *et al.*, 1988). In contrast to the wild type CRP, CRP* is sensitive to cleavage by proteases, including trypsin, chymotrypsin, subtilisin and *Staphylococcus aureus* V8 protease, in the absence of cAMP, which is a property of wild type CRP only in the presence of cAMP. The protease digestion patterns of CRP* were identical to those produced from the wild type cAMP/CRP complex (Ren *et al.,* 1988). This suggests that CRP* has a conformation similar to the cAMP/CRP complex. At high concentrations, the unliganded CRP*598 showed activity comparable to the wild type CRP in supporting abortive initiation on *lac* p1 by RNA polymerase. CRP*598 showed cooperativity with RNA polymerase for *lac* P1 DNA binding (Ren *et al.,* 1988). The two amino acid substitutions of CRP*598 reside near the "hinge" region of CRP (amino acid 130-138), which is thought to be responsible for the transmittance of conformational changes elicited by cAMP from the N-terminal domain to the C-terminal domain. It is suggested that the conformational changes of CRP upon binding of cAMP is caused by the inward motion between the C and D α-helices, releasing the F α-helix for DNA recognition (Kim *et al.*, 1992).

The consensus DNA binding site (Gunasekera *et al.,* 1992) for CRP is:

5'-AAATGTGATCTAGATCACATTT-3'

It is 22 base pairs long and exhibits perfect two fold sequence symmetry. The CRP-DNA complex is also symmetric; one subunit of CRP interacts with one half of the consensus DNA site, the other subunit interacts with the other half of the DNA site which has rotated to the same face of the DNA helix. The nucleotide sequence of the CRP binding site in *lac* P^+ is 5'-tAATGTGAgtTgctTCACtcat-3', where the lower case letters denote the bases different from those predicated from the consensus sequence. These differences in turn weaken the affinity of CRP binding to the *lac* promoter 450 times compared to the consensus sequence (Ebright *et al.*, 1989). The CRP binding site on the *gal* P1 promoter is 5'-AAgTGTGAcaTgGAatAaATTa-3' (de Crombrugghe *et al.*, 1984).

RNA transcription initiation involves three overall steps shown schematically below (Chamberlin, 1974; McClure, 1985):

$$\begin{array}{ccccccc} & & & & \text{NTPs} & & \\ R + P & \xleftrightarrow{K_b} & RP_c & \xrightarrow{K_f} & RP_o & \xrightarrow{\text{Promoter clearance}} & RNA \end{array}$$

First, RNA polymerase (R) binds to the DNA promoter (P), forming a binary competitor-sensitive complex, called the "closed promoter complex" (RP_c), the binding constant is Kb. This RP_c complex is then isomerized to a transcriptionally active complex with a rate constant K_f, this irreversiblly formed complex is called the "open promoter complex" (RP_o). RNA polymerase is then able to clear the promoter

in the presence of ribonucleoside triphosphates following the synthesis of a certain transcript length. Transcription of a CRP dependent promoter may be limited at different stages of the above scheme and CRP supports transcription by functioning at a certain stage where the RNA polymerase is trapped. For example, in the case of the *lac* promoter, without CRP, RNA polymerase binds to the upstream weaker promoter *lac* P2, and shows a very low level of transcription. In the presence of cAMP/CRP, the binding of RNA polymerase is shifted to the stronger *lac* P1 promoter and isomerization of RP_c to RP_o is accelerated (Straney *et al.*, 1989). In *gal* P1, CRP only accelerates the isomerization of the initial closed promoter complex to the open promoter complex without affecting the initial binding of RNA polymerase to the promoter (Herbert *et al.,* 1986). In the *mal* T promoter, CRP only supports the movement of RNA polymerase away from the already formed RP_o without helping the binding of RNA polymerase to the promoter nor the isomerization of the closed promoter complex to open promoter complex (Menendez *et al.*, 1987).

The binding of CRP sharply bends DNA to approximately 90° or more (Schultz *et al.,* 1991; Wu *et al.*, 1984, Liu-Johnson and Crothers, 1986). The bending results almost entirely from two 40° kinks that occur in the conserved TGTGA sequence, between TG/CA at positions 5 and 6 on each of the CRP half binding sites. The kinks derive from ionic and hydrogen bond interactions between residues of CRP and DNA bases and phosphates (Schultz *et al.,*1991). DNA bending does not seem to be sufficient for transcription activation (Bell *et al.*, 1990;

Zhou *et al.*, 1993). Instead, it is assumed that DNA bending facilitates additional protein-DNA or protein-protein interactions required for the initiation of transcription by RNA polymerase. Schultz *et al.* (1991) proposed that the DNA bending induced by CRP binding might facilitate the contact between the RNA polymerase and the upstream DNA sequence. Alternatively, DNA bending could play an energetic role essential for transcription initiation.

The location of the CRP binding sites upstream of the promoters varies for different catabolite-sensitive operons and changing the location of binding sites affects the activation efficiency. It is also considered that the angular orientation, rather than the actual distance of the binding site with respect to the promoter, is an important factor (Straney *et al.* 1989; Gaston *et al.* 1990). The operons which have been studied most extensively for CRP activation are the *lac* and *gal* systems, they have CRP binding sites at positions -61.5 (type I promoter) and -41.5 (type II promoter) respectively. How the different CRP-dependent promoters are activated by CRP is a question that has received much attention.

It has long been suspected that CRP activates transcription through direct contact with RNA polymerase. Perhaps this contact enhances the binding of RNA polymerase to DNA by providing a stabilizing force compensating for those provided by a non-consensus promoter DNA sequence. DNA bending upon binding CRP also makes it possible for RNA polymerase to contact upstream DNA sequences. Another possibility is that CRP induces favorable conformational changes in RNA polymerase. There is evidence indicating that direct CRP-RNA

polymerase contact plays an important role in at least *lac* promoter activation. DNase I footprinting demonstrated that CRP and RNA polymerase cooperatively bind to the *lac* promoter (Li and Krakow, 1985; Ren *et al.,* 1988; Straney *et al.,* 1989). Fluorescence polarization experiments showed that CRP and RNA polymerase interact in solution (dissociation constant, ~1 μM) (Pinkney *et al.*, 1988). CRP "positive control" mutants which affect transcription activation without altering specific DNA binding and bending have been isolated. They map on an exposed loop spanning amino acid residues 156 to 162 in the C-terminal domain (Bell *et al.*, 1990; Eschenlauer *et al.*, 1991; Zhou *et al.*, 1993).

C-terminal truncation mutants of the α subunit of RNA polymerase differently affect promoters activated by CRP depending on the promoter class (Igarashi *et al.,* 1991; Kolb *et al.,* 1993). When the CRP binding site is located at -61.5 (type I, e. g. *lac*), stimulation of activation by the C-terminal mutants of α is totally lost with little effect seen where the CRP binding site is located at the -41.5 position (type II, e. g. *gal*). These results suggest that CRP activates transcription through different mechanisms on different promoters. The distance between the CRP binding site and transcription initiation site may affect the way CRP interacts with RNA polymerase.

The loss of cooperative binding of CRP with RNA polymerase is observed in CRP positive control mutants (amino acid residues 156-162). All of these *pc* mutants are single amino acid substitutions or deletions. They can bind to the promoter DNA normally, and bend the DNA to the extent seen for the wild type, but

failed to support transcription initiation by RNA polymerase. *PC* mutants also lost their cooperative binding ability with RNA polymerase on the *lac* promoter. But one important fact needs to be noted, most of the *pc* mutants found in the 156-162 region of CRP (except for H159L and T158A) are *lac*$^-$ but *gal*$^+$, which means that they were unable to support *lac* transcription but were still able to support transcription from the *gal* promoter. Combining the information for these *pc* mutants together with α truncations, strongly indicates that CRP interacts with the RNA polymerase α subunit through its 156-162 loop and thereby supports transcription on the *lac* promoter. Recently, Ebright's group demonstrated site specific crosslinking through the CRP-161 position with the C-terminus of the RNA polymerase α subunit on the *lac* promoter (Chen *et al.,* 1994).

The nature of the contacts between CRP and RNA polymerase at class II promoters, where the CRP binding site is centered at position -41.5 upstream of the transcription initiation site, is not very clear. There are several arguments indicating that the contacts of CRP with RNA polymerase on class II promoters differ from those on class I promoters. Point mutations and deletions in the C-terminal region of the RNA polymerase α subunit prevent CRP from activating class I promoters but have very little effect on class II promoters (Igarashi and Ishihama, 1991; Zou *et al.*, 1992). C-terminal truncation of α subunit abolished the ability of RNA polymerase to cooperatively interact with CRP on *lac* P1 but not on *gal* P1 (Kolb *et al.*, 1993). One interesting *pc* mutant, H159L, which has a leucine substituted for the wild type histidine, leads to total inactivation for stimulation of *lac* and *gal* transcription (Bell

et al., 1990). Interestingly, mutations at position 52 (K52N and K52Q) which is in another exposed loop present in the N-terminal domain increase CRP activity and reverse the effect of H159L. However, the K52N-H159L double mutant only increased expression from type II promoters but not from type I promoters (Bell *et al.*, 1990). Since position 52 is situated far away from H159, it is very unlikely that the substitutions restored the correct contact at the 159 region. It is more reasonable to propose that the 52 site mutants somehow unmasked or created a second possible contact site for RNA polymerase. According to the crystallographic structure of the CRP-DNA complex (Schultz *et al.*, 1991), it is possible that the two surface exposed loops around residues 52 and 159 contain elements that are potential RNA polymerase contact sites. These two loops are located on adjacent faces of two different subunits. FNR, another regulatory protein of *E. coli*, shares structural similarity with CRP and activates the expression of a variety of anaerobic growth regulated operons. Because of the extreme similarities between these two proteins, it is probable that CRP and FNR stimulate transcription by a related mechanism. It has been shown that mutations at position 81-87 of FNR, which corresponds to and has high homology with position 52-58 of CRP (FNR has an extra N-terminal 29 amino acids compared with CRP), reduce expression from an FNR-dependent promoter without affecting binding to the DNA target (Williams *et al.*, 1991). This defines an activation region on FNR, homologous to the N-terminal exposed loop around position 52 of CRP, which might be the potential activation region involved in contact with RNA polymerase. Furthermore, FNR derivatives

which activate CRP-dependent genes in response to anoxia, and CRP derivatives which activate FNR-dependent genes in response to cAMP have been constructed by site specific mutagenesis in the corresponding DNA recognition helices (Spiro *et al.*, 1990; Spiro *et al.*, 1987). All of these findings indicate that the second surface exposed loop at around position 52 of CRP, might be another potential activating region. Recently, Busby's group further demonstrated that the effect of the K52N suppressor does not require the C-terminus of α. The K52N, K52E and E96G substitutions also improve the ability of CRP to activate transcription from CCpmeIR but not CC+20pmeIR (West, *et al.*, 1993). Another single amino acid substitution at the 58 position, E58H, isolated by random mutagenesis, caused a clear reduction in CRP dependent expression from CCpmeIR but not CC+20pmeIR. However, the effect of E58H also differs from one class II promoter to another, ranging from an 87% reduction in activity at *pmeIRcon*$_{-1}$ to a 33% reduction inactivity at *galP1p9Ap16C* (West, *et al.*, 1993).

It is generally assumed that the effects of positive control mutants are due to direct or indirect disruption of an indispensable interaction for activation. However, there are other possibilities; for example, the effects could in some cases be due to the creation of steric blockages. The C-terminus of α is known to be essential for transcription of Class I promoters but dispensable for Class II promoters. Interestingly, C-terminal truncation of α was found to suppress the positive control inability of *pc* mutants H159L and T158A on CCpmeIR (West, *et al.*, 1993). One possible explanation for this is that the 156-162 loop is not essential for

CCpmeIR activation, H159L and T158A substitutions may have created a clash with the C-terminus of α which could be relieved by the deletion. An alternative explanation is that α subunit interferes in some way with CRP dependent activation on CCpmeIR, and that interference is relieved by contact with the 156-162 loop of CRP. Mutations in this loop abolish this contact, but the effects of these mutants can be suppressed either directly or indirectly by the deletions in α.

Numerous studies of the *lac* promoter show that the surface exposed CRP loop consisting of amino acids from 156-162, known as Activating Region I, makes direct contact with a site on the C-terminus of the RNA polymerase α subunit. In a recent study, Busby and his collaborators found that α protects upstream sequences both in the binary RNA polymerase-*gal* P1 and ternary RNA polymerase-CRP-*gal* P1 complexes. In the latter complexes, α appears to make direct contact with the Activating Region I, spanning amino acids from 156-162, of CRP (Attey *et al.*, 1994). The probable mode of open promoter complex formation at *gal* P1 is that the Activating Region I of the upstream subunit of the CRP dimer makes contact with the α subunit of RNA polymerase while the downstream subunit is likely to make alternative contacts with other parts of RNA polymerase (West *et al.*, 1993; Attey *et al.*, 1994)

Several glycine residues are found to be conserved in CRP, FNR and another regulatory protein *fix*K in *Rhizobium meliloti* (Batut *et al.*, 1989). This infers a related role for these regulatory proteins. The conserved glycines are present in the N-terminal domain in CRP at positions 33, 45, 56, 67 and 71. Since glycine

residues allow high flexibility, they might act as swivel points when CRP undergoes conformational changes in response to cAMP binding. It would be interesting to see whether substitution of glycines with bulkier amino acids affects the activity of CRP.

CRP is one of the prime transcription factors serving as a model for the study of regulation at the transcriptional level. Although CRP has been studied extensively at biochemical, genetic and X-ray crystallographic levels, the relationship between the structure of the cAMP-CRP complex and the mechanism of transcription activation is still not completely understood, and this has been the major focus of investigation over the last decade. The study of cAMP induced transcription activation involves a variety of basic topics, such as protein-protein interaction, protein conformation, protein-DNA interaction, DNA conformation, *etc.* Recently, we have witnessed a major advance in understanding the direct interaction of CRP and RNA polymerase.

In this dissertation, some structural and functional studies of CRP regarding the mechanism of CRP dependent transcription activation are reported. More information on the Activation Region I of CRP is provided by using site-directed mutagenesis, and evidence of direct interaction between CRP and RNA polymerase holoenzyme was shown. The possible function of this protein-protein interaction is addressed.

Materials and Methods

1. Reagents:

Reagents were obtained as follows: cAMP, bovine serum albumin, Tris base, Tetrazolium Red, trypsin, subtilisin, chymotrypsin and *Staphylococcus aureus* V8 protease, pronase E, β-mercaptoethanol, Coomassie brilliant blue, bromophenol blue, PMSF, benzophenone-4-maleimide, Sigma; sodium dodecyl sulfate and TEMED, Bio-Rad; acrylamide and bisacrylamide, Serva; restriction endonucleases, DNA polymerase I Klenow fragment and Taq DNA polymerase, Boehringer Mannheim; T4 DNA ligase, BRL; DNase I, Promega; Sequenase Version 2.0, USB; Vent DNA polymerase, NEB; radioactive nucleotides and Ecolume, ICN; Sephacryl S-100, Sephacryl S-200, Sephacryl S-300, S-Sepharose Fast Flow, deoxyribonucleotides and ribonucleotides, O^{met} GTP, Pharmacia; ZetaChrom SP100 capsule, CUNO, Inc.; APDP, Pierce Chemical Co. ECL Western blotting analysis system, Amersham. Anti-σ monoclonal antibody 2D1 (Strickland *et al.,* 1985) was from Dr. Richard Burgess (University of Wisconsin); anti-CRP monoclonal antibodies 64D1 and 115D5 (Li and Krakow, 1985), anti-α monoclonal antibody 23C2 (Sharif *et al.*, 1994) and anti-α, β, β' monoclonal antibodies (Riftina *et al.*, 1988; Rockwell and Krakow, 1988) were prepared in this laboratory.

2. Bacterial Strains and Plasmids:

E. coli CJ236, a *dut⁻ ung⁻ E. coli* mutant used for the preparation of uracil-containing DNA, was purchased from IBI. The *crp⁻* strain, XE64.2, *crp⁻cya⁻* strain CA8445 and phagemid pXZCRP (Zhang *et al.*, 1992) were obtained from Dr. Richard Ebright (Rutgers University, New Jersey). *E. coli* KC1071 (Irwin and Ptashne, 1987) was obtained from Dr. Mark Ptashne (Harvard University). pHW104, a plasmid which has four segments of *lac* P^+ cloned under *Eco* RI sites, was obtained from Dr. Donald M. Crothers (Yale University). *E. coli* M182 (Δ*crp39,* Δ*lac*), plasmid pRW2, promoters CCpmel R and CC+20pmelR (Bell *et al.,* 1990) were obtained from Dr. Stephen Busby. (University of Birmingham, Birmingham, U.K.). pBR322 was purchased from Boehringer Mannheim.

3. Phosphorylation of Oligonucleotides:

The phosphorylation reaction mixtures (final volume 30 µl) contained: 100 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 7 mM DTT, 1 mM ATP, 1.7 µM oligonucleotide and 2 units of T4 polynucleotide kinase. The mixtures were incubated at 37°C for 30 minutes and then heated at 70°C for 10 minutes to inactivate the kinase. The phosphorylated oligonucleotides were stored at -20°C.

4. DNA Mutagenesis:

Site-directed Mutagenesis:

(A) Uracil-containing single-stranded DNA was prepared from pXZCRP with

co-infection of helper phage R408 in *E. coli* CJ236 as described in *Current Protocols in Molecular Biology* (p8.1.1-8.1.6) and used for mutagenesis. Oligonucleotide primers containing the desired mutations (Table I) were synthesized on an Applied Biosynthesis Model 380B DNA Synthesizer and purified using Nensorb columns. After phosphorylation with T4 polynucleotide kinase, the mutant oligonucleotides were annealed to the single-stranded phagemid DNA by heating at 70°C for 5 minutes and then incubating at 37° C for 30 minutes. Polymerization and ligation were carried out by addition of DNA polymerase I (Klenow fragment), T4 DNA ligase, dNTPs and incubated at 14°C for 16 hours. The mixtures were then used to transform *E. coli* XE64.2. Insertion mutants were selected by screening minipreps for the presence of plasmids containing a new *Bgl* II site and confirmed by DNA sequencing.

(B) Site specific mutagenesis by the use of polymerase chain reaction was carried out on the *crp* gene in pBR322 according to the method of Ito *et al.* (1991). Three common primers and primers specific for various mutations were used in the PCR reaction.

Random DNA Mutagenesis:

Random DNA mutagenesis was carried out according to the method of Ito *et al.* (1991) and Zhou *et al.* (1991) using PCR with *Taq* DNA polymerase in the presence of Mg^{++} and/or Mn^{++}.

5. Transformation of E. coli by Electroporation:

Electroporation of *E. coli* cells was done by using the Gene Pulser Transfection Apparatus (Bio-Rad). Cells were prepared according to *Current Protocols in Molecular Biology* (p1.8.4-1.8.5). The electroporation apparatus was set to 2.5 kV, 25 μF and the pulse controller to 200 ohms. 5 pg to 0.5 μg plasmid DNA were mixed with the prepared cells on ice and transfer to a 0.2 cm cuvette (Bio-Rad). Electroporation was carried out and 1 ml of LB medium was added afterwards. Cells were transferred into a sterile test tube and incubated at 37°C for 1 hour with constant shaking. Aliquots of the cell suspension were plated on different media where indicated.

6. DNA Sequencing:

The plasmids for double strand DNA sequencing were purified by using the QIAprep Spin Plasmid Kit (QIAGEN). The procedure for dideoxynucleotide DNA sequencing was carried out according to that described in the USB DNA sequencing kit.

7. Sugar Fermentation:

The ability of the CRP mutants to support sugar fermentation was tested on plates containing Difco Antibiotic Medium 2 plus 50 mg/L Tetrazolium Red, 100 μg/ml ampicillin and 10 g/L of the indicated sugar.

8. ß-galactosidase Assay:

ß-galactosidase assays were carried out in quadruplicate by the method of Miller (1972) using sodium dodecyl sulfate and chloroform. Cells were grown at 37°C overnight in LB medium supplemented with 100 µg/ml ampicillin and 100 µM isopropyl-β-D-thiogalactopyranoside. The overnight cultures were reinoculated into the same medium and grown at 37°C until they reached mid-log phase when they were used to determine ß-galactosidase activity.

9. Protein Purification:

Mutant and wild-type CRP were purified by the method of Eilen and Krakow (1977) with the following modifications. After cell lysis, polymin P addition and centrifugation, the supernant was adjusted to pH 6.5 with 1 M acetic acid and loaded onto a ZetaChrom SP-100 capsule (equilibrated with 50 mM sodium phosphate (pH 6.5), 0.1mM DTT, 0.1mM PMSF and 5% glycerol). After washing with the equilibration buffer, CRP was eluted with 50 mM sodium phosphate (pH 7.5), 0.5 M NaCl, 0.1 mM DTT, 0.1 mM PMSF and 5% glycerol.

RNA polymerase was isolated from *E. coli* K12 by a modification of the method of Burgess and Jendrisak (1975). Holoenzyme and core enzyme were separated by chromatography on denatured calf thymus DNA-agarose column (Lowe *et al.*, 1979). Protein concentrations were determined using the extinction coefficients: CRP, $E^{1\%}_{280nm}$ = 8.8 (Aiba and Krakow, 1981); RNA polymerase holoenzyme, $E^{1\%}_{280nm}$ = 6.7 (Levine *et al.*, 1980).

10. DNA Fragment Purification:

DNA promoter fragments were first excised from their vector plasmids by using the appropriate restriction endonucleases and then separated from the vector DNA by 0.8% agarose gel electrophoresis. The relevant DNA promoter fragment bands were excised from the gel and further purified using the Gene Clean kit (USB). The DNA fragments were resolved on a 0.8% agarose gel in TAE (40 mM Tris-acetate, 1 mM EDTA , pH 8.0) buffer. The promoter bands were cut out and weighed. 6 M NaI (4 times the volume of the gel slice) was added and the agarose pieces were melted by heating at 55°C for 5 minutes. Glassmilk was added into the high salt mixture and incubated for 10 minutes on ice. The DNA bound to glassmilk was spun down and washed with cold 50% ethanol three times, the DNA fragments were eluted with H_2O and were ready for use.

11. Abortive Initiation Assays:

Abortive initiation assays were performed by a modification (Ren *et al.*, 1988) of the method of Malan *et al.* (1984). Reaction mixtures (final vol. 50 µl) contained: 40 mM Tris-HCl, pH 8.0, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.1 mM cAMP, 5% glycerol, 2 nM RNA polymerase holoenzyme and the indicated amount of wild type or mutant CRPs. After preincubation at 37°C for 30 minutes, 0.5 mM ApA and 50 nM [^{3}H]UTP (360 cpm/pmol) were added as substrates. The reactions were allowed to proceed for another 15 minutes at 37°C and then terminated by addition of 10 µl of 0.5 M EDTA. The reaction products were resolved by paper chromatography in

WASP solvent (100 ml solvent, contains 18 ml of H_2O, 80 ml of saturated ammonium sulfate, pH 8.0, and 2 ml of isopropanol) (McClure *et al.*, 1978). After chromatography, the segment containing the product ApApUpU was cut out and the amount of ApApUpU synthesized was determined by counting the radioactivity of the appropriate segments in a scintillation counter.

12. In Vitro Run-Off Assays:

Runoff assays (final volume 35 µl) were carried out in 50 mM Tris-HCl (pH 7.8), 3 mM $Mg(OAc)_2$, 0.1 mM EDTA, 0.1 mM DTT, 25 µg/ml BSA, 50 mM NaCl, 0.3 pmol promoter DNA fragment, 1.5 pmol mutant or wild-type CRP and 3 pmol RNA polymerase holoenzyme. After incubation at 37°C for 1 hr, 15 µl of a mix containing 10 µg heparin, 0.5 mM each of ATP, GTP, CTP and 0.16 mM UTP + 2 µCi [α-^{32}P]UTP was added and incubated for another 5 min. Reaction mixtures were denatured and resolved by electrophoresis on a 8% urea polyacrylamide gel.

13. Labeling of DNA Fragments:

The [^{32}P]*lac* P$^+$ fragment was prepared by labeling the *lac* P$^+$ DNA fragment with [α-^{32}P]dATP using DNA polymerase I Klenow fragment. The reaction mixture (final vol. 50 µl) contained: 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT, 40 µCi [α-^{32}P]dATP (3000 mCi/mmol), 4 µg *lac* P$^+$ fragment and 5 units of DNA polymerase I Klenow fragment. After 15 minutes incubation at room temperature, 10 µl of 1 mM cold dATP was added and the incubation was chased

for an additional 15 minutes. For gel retardation assays, the labeled DNA fragments were ethanol precipitated and washed with 70% ethanol, and resuspended in TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA) buffer. For DNase I footprinting assays, the doubly end-labeled fragments were digested with *Pvu*II, the smaller fragment was removed by using spin-100 filters (Clontech), the larger fragment which has only the upper strands labeled were used for DNase I footprinting.

14. Gel Retardation Assays:

Gel retardation assays were carried out by the modified method of Garner and Revzin (1981). Reaction mixtures (final vol. 15 µl) contained: 20 mM Tris-HCl (pH 8.0), 100 mM KCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.4 mM cAMP, 2 nM [^{32}P]*lac* P^+ DNA fragment, 90 nM wild type or mutant CRPs, and 0.5 µg tRNA. The mixtures were incubated at room temperature for 15 minutes and gel shift was assayed by 7.5% native polyacrylamide gel electrophoresis with TBE (90 mM Tris-Borate, 2 mM EDTA, pH 8.0) buffer plus 0.1 mM cAMP followed by autoradiography.

15. DNase I Footprinting assays:

DNase I footprinting was performed essentially as given in Ren *et al.* (1988). Incubation reactions were carried out in 40 mM Tris-HCl (pH 8.0), 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT with 3 nM ^{32}P-labeled (upper strand) *lac* promoter fragment, the amount of wild-type or mutant CRP as indicated and 20 nM RNA polymerase. After incubation at 37°C for 30 min, 1 unit of DNase I was added to each reaction

and the incubation was continued for 30 sec. The reaction mixtures were denatured and run on a 8% denaturing sequencing gel according to Maxam and Gilbert (1980).

16. Proteolytic Cleavage of CRP:

Mixtures contained (final volume 50 µl): 10 mM Tris-HCl (pH 8.0), 20 mM NaCl, 0.1 mM DTT, 7.5 µg wild-type or mutant CRP plus 0.1 mM cAMP where indicated. Following addition of 75 ng trypsin or 38 ng subtilisin the mixtures were incubated for 20 minutes at 37°C. The reactions were terminated by addition of 2.5 µl of 20 mM phenylmethylsulfonyl fluoride. The resultant cleavage products were resolved by SDS-polyacrylamide gel electrophoresis (Laemmli, 1970) on a 15 % polyacrylamide gel with a 4.75% stacking gel.

17. Western Blot Assay:

Proteins were separated on a SDS polyacrylamide gel following the procedure of Laemmli (1970). The proteins were transferred at 2.5 mA/cm^2 for 45 min to a nitrocellulose membrane using a MiniBlot-SDE (Millipore). Identification of antigens on the nitrocellulose membrane was performed according to the method of Johnson *et al.* (1984) Due to the high molecular weight of ß, ß' and crossed-linked product, pronase was used to facilitate their transfer to the nitrocellulose membrane(Gibson, 1981).

18. Preparation of BPM-CRP:

Benzophenone-4-maleimide modified CRP (BPM-CRP) was prepared by a modification of the method of Tao *et al.* (1985). Mutant CRPs (1-2 mg/ml) in CRP storage buffer (10 mM Potassium Phosphate [pH 7.0], 1 mM EDTA, 0.2 M NaCl, without DTT) was modified by adding BPM from a 20 mM stock solution in dimethylformamide to a final molar ratio of BPM to CRP of 4. The reaction was allowed to proceed in the dark overnight on ice, followed by quenching with excess ß-mercaptoethanol and dialysis against CRP storage buffer (with 0.1 mM DTT).

19. Photocrosslinking Assays:

Photocrosslinking assays were carried out in a final volume of 50 µl of 50 mM Tris-Cl (pH 7.8), 3 mM $Mg(OAc)_2$, 0.1 mM EDTA, 0.1 mM DTT, 50 mM NaCl, 200 pmol BPM-CRP, 120 pmol promoter DNA fragment (as indicated), 80 pmol RNA polymerase holoenzyme with or without 0.1 mM cAMP. The mixture was incubated at 37°C for 30 min and subjected to UV illumination (>300 nm, on a Fotodyne UV box filtered by Pyrex glass) for 30 min at room temperature followed by DNase I digestion for 15 min at 37°C.

20. ECL Detection of Western Blots:

ECL Western blot detection system (Amersham) was used to develop the Western blots after UV photocrosslinking. The UV irradiated mixtures were separated by SDS polyacrylamide gel electrophoresis and transferred to a

nitrocellulose membrane. The membrane was cut into strips if necessary and blocked in TBS (20 mM Tris-HCl, pH 7.4, 0.9% NaCl) with 5% Non-fat Carnation Milk (NFCM, Carnation Company) at room temperature for 1 hour. The blocked filters were then incubated with 1 μg/ml of the appropriate primary antibodies in TBS buffer with 1% NFCM at 4°C overnight with gentle rotation. After three washes in washing buffer (TBS with 0.5% NFCM and 0.1%(V/V) Tween 80), the filter strips were incubated with goat anti-mouse IgG-horse radish peroxidase (1:20,000 dilution in TBS with 1% NFCM) for 1 hour at room temperature on a rocker. After washing several times with a total of 500 ml of buffer for 2 hours, the nitrocellulose filter stripes were immersed in a mixture of ECL detection solution 1 and 2 (1:1, V/V) for 1-2 minutes. The excess detection reagent was drained off the strips and the strips were lined up according the SDS-PAGE run on Whatman 3MM paper. Saran Wrap was used to wrap the membrane and the whole blot was exposed to a autoradiography film for 15 seconds to 5 minutes.

RESULTS

Characterization of CRP 156-162 loop:

A *Bgl* II restriction site AGATCT (-Arg-Ser-) was inserted at different positions by site-directed mutagenesis into the cloned *crp* gene under the control of the *crp* promoter on a phagemid. The insertion sites are illustrated in Table 1. These insertion mutants were screened by *Bgl* II digestion and confirmed by DNA sequencing. Each of the CRP insertion mutants was expressed in crp$^-$ *E. coli* XE64.2 and assayed for ability to ferment lactose, galactose, maltose and arabinose on Tetrazolium plates. The results (Table II) indicated that none of the insertion mutants showed any apparent impairment in the ability to ferment galactose or arabinose while all four were unable to ferment maltose. CRP-HrsP^{160} and CRP-GrsM^{163} were unable to support lactose fermentation while CRP-QrsP^{154} and CRP-PrsD^{161} retained partial activity.

To quantitate the *in vivo* effect of the insertion mutations on *lac* expression, ß-galactosidase assays were carried out. The results (Figure 3) show that CRP-HrsP^{160} and CRP-GrsM^{163} were severely affected in their ability to support *lac* expression. CRP-QrsP^{154} and CRP-PrsD^{161} retained about 40% of the wild-type CRP activity. The results obtained using the *lac* fermentation and ß-galactosidase assays are qualitatively consistent.

To determine whether the inability of the insertion mutants to support *lac* expression was due to a defect in DNA binding, assays for ß-galactosidase

expression were carried out in *E. coli* KC1071 which has the *lac* promoter region converted to a CRP-independent *lac* UV5 promoter and the *lac* operator changed to a symmetric high affinity CRP binding site (Irwin and Ptashne, 1987). CRP-QrsP^{154} and CRP-PrsD^{161} which retained partial activity in supporting ß-galactosidase expression in the wild-type *lac* system showed repression of ß-galactosidase synthesis in *E. coli* KC1071 in a range comparable to that of the wild-type CRP (Figure 4). CRP-HrsP^{160} and CRP-GrsM^{163} severely repressed ß-galactosidase synthesis indicating that these mutants had retained normal DNA binding activity. The ability of these mutants to bind *lac* promoter DNA *in vitro* is also shown in the DNA gel retardation assay (Figure 5).

The ability of the CRP mutants to support abortive initiation by RNA polymerase directed by *lac* P$^+$ is shown in Figure 6. The results indicate that CRP-PrsD^{161} and CRP-QrsP^{154} retained about 7% and 20% respectively of the activity of the wild-type CRP. CRP-HrsP^{160} and CRP-GrsM^{163} were essentially inactive consistent with the *in vivo* results. The CRP^{pc} mutants which have been mapped within the 156-162 loop are generally *lac*$^-$ and *gal*$^+$. The runoff assay shown in Figure 7 demonstrate that all four insertion mutants still retain *in vitro gal* activity. At the five-fold excess of CRP to *gal* DNA used the wild type CRP completely blocks transcription from the CRP-independent P2. The insertion mutants clearly support transcription from *gal* P1. The residual P2 activity may be due to the somewhat lower affinity for site specific DNA binding shown by these mutants.

DNase I footprinting (Figure 8) was carried out to define the effect on *lac* P$^+$

binding by the CRP insertion mutants. Since two of the mutants showed properties consonant with other CRP^{pc} mutants it was also of interest to determine whether they could act to cooperatively establish the formation of an open promoter complex with RNA polymerase. The results indicate that each of the insertion mutants was able to bind to the CRP site spanning -55 to -70 bp on the *lac* promoter. The two insertion mutants (CRP-$HrsP^{160}$ and CRP-$GrsM^{163}$) which showed a strong defect in positive control were able to bind to the CRP site with an affinity approximating that of the wild-type CRP (Figure 8, lanes c and j). RNA polymerase and wild-type CRP cooperatively interact to form an open complex at *lac* P1 (Ren *et al.*, 1988; Straney *et al.*, 1989). In addition to the large region protected against DNase I attack there is a region of enhanced cutting at around -23 bp (Figure 8, lane b). The results indicated that CRP-$QrsP^{154}$ and CRP-$PrsD^{161}$ only weakly interacted with RNA polymerase resulting in weak formation of an open complex. In contrast CRP-$HrsP^{160}$ and CRP-$GrsM^{163}$ did not support binding by RNA polymerase. The inability of CRP-$HrsP^{160}$ and CRP-$GrsM^{163}$ to activate transcription is due to a defect in interaction with RNA polymerase characteristic of positive control mutants.

Proteolytic sensitivity provides a convenient assay for determining the stability of unliganded CRP and the allosteric effect of cAMP. The results shown in Figure 9 indicate that while unliganded CRP is resistant to trypsin and subtilisin each of the insertion mutants was relatively sensitive to attack by these proteases. The mutants responded to cAMP binding by shifting into a trypsin-sensitive conformation resulting in core forms which differ from that of wild type CRP due the

presence of the arginine residue in the insertion site. The results indicate that the insertion of Ser-Arg at any of the four positions altered the conformation of the unliganded CRP mutants. Earlier studies on tryptic digestion of cAMP/CRP indicated that the initial cut is at Lys-166 or Arg-169 with subsequent digestion resulting in the accumulation of a resistant core terminating at Lys-130 (Angulo and Krakow, 1986). The additional trypsin-sensitive site resulting from the insertion of the arginine residue results in slightly smaller intermediate forms along with the limit Lys-130 core. The presence of cAMP is required to stabilize the Lys-130 core. CRP^{*} mutants which are intrinsically less stable in the absence of cAMP are further digested (Ren *et al.*, 1988) with trypsin cutting at Arg-115. The unliganded insertion mutants are also more sensitive to attack by subtilisin.

In order to select for mutants which can restore the stimulatory activity of $HrsP^{160}$ or $GrsM^{163}$ on the *lac* promoter, random mutagenesis was carried out by PCR with *Taq* DNA polymerase; transformation of *E. coli* XE64.2 was by electroporation. As potential candidates, white or light pink colonies (crp^{+}) were picked from lactose Tetrazolium plates. Thirty independent PCR reactions (Zhou *et al.*, 1991) were performed on each template containing the *crp* gene with either $HrsP^{160}$ or $GrsM^{163}$. A total of 20,000 mutagenized transformants was screened from each template, no revertants were found for the $HrsP^{160}$ insertion and five potential candidates were found for the $GrsM^{163}$ insertion. These mutant *crp* gene plasmids were isolated and transformed back into XE64.2 to rule out the possibility of individual host cell mutation. The entire CRP gene of these candidates was

sequenced and transcription activation was assessed by ß-galactosidase activity. Unfortunately, none of these mutants can totally restore activity compared to wild type, the most active suppressor mutant shows about 45% of the wild type activity. Four out of these five partial revertants have substitutions at amino acid position 161, two have glutamic acid and two have tyrosine substitutions for its original aspartic acid. The last mutant is a triple mutant with amino acid substitutions at amino acids 5, 111 and 122 (Q5R-D111G-R122C) (Table III).

Double mutants K52N-HrsP^{160} and K52N-GrsM ^{163} (with asparagine substituted at position 52 for lysine in addition to the insertions) were also made by site specific mutagenesis, but they were not able to restore transcription activation for these two insertion mutants on the *lac* promoter (type I).

To further characterize the effect of insertions at positions H-P^{160}, P-D^{161} and G-M^{163}, and to determine whether the effect is caused by the charge or size of the amino acid, other amino acid insertions :-Arg-Gly-, -Gly-Ser-, -Arg- and -Ser- were generated at those positions. ß-galactosidase activity assays were carried out to determine transcription activation by each mutant. The results shown in Figure 10 suggested that the property of the inserted amino acid as well as the steric effect are involved in the inactivation of CRP. Different amino acid insertions are effective at different positions (although these positions are quite close to and in the same exposed loop) to make CRP behave as a positive control mutant on the *lac* promoter. Generally, insertions at position H-P^{160} and G-M^{163} are more destructive than insertions at P-D^{161} position. At H-P^{160} position, only -Ser- and -Gly-Ser-

insertions retained relatively higher activity (~40% and 30% of the wild type respectively), others were all inactive on the *lac* promoter. It appears that the introduction of the positively charged amino acid, arginine, is very deleterious at that position. At G-M^{163} position, only the -Arg-Gly- insertion retained about 55% of the wild type activity and the other four were totally inactive. H-P^{160} is the least sensitive site for insertion, only -Gly-Ser- made the protein inactive, others all retained 50-70% of the activity. All these insertion mutation retained equally high DNA binding activity (Figure 11) and fairly high activity on *the gal* promoter compared to the wild type CRP (Figure 12).

It is believed that CRP stimulates transcription on a type I promoter by contact of its activating region at amino acids 156-162 with the C-terminal region of the α subunit of RNA polymerase (based on the results with CRP *pc* mutants and α truncation mutants) (Bell *et al.,* 1990; Eschenlauer *et al.*, 1991; Igarashi *et al.*, 1991; Kolb *et al.,* 1993; Zhou *et al.,* 1993). However, on type II promoters, this does not seem to be the case. C-terminal truncations in σ impair the ability of the RNA polymerase holoenzyme to respond to CRP activation of *gal* transcription, indicating that the RNA polymerase activation site for CRP for a type II promoter is located within the deleted C-terminal region of σ (Kumar *et al.*, 1994). But whether the activating region on CRP for a type II promoter remains the same as for a type I promoter is still unknown.

Characterization of the CRP 52 loop:

CRP stimulates transcription on a Class I promoter by contact of its activating region at the 156-162 loop with the C-terminal region of the α subunit of RNA polymerase. Another surface-exposed loop including CRP amino acid 52 is also considered a potential candidate for contacting RNA polymerase during Class II promoter activation (Williams *et al.*, 1991; West *et al.*, 1993). To directly study possible contact with RNA polymerase, site specific photocrosslinking was carried out using benzophenone-4-maleimide as a sulfhydryl specific photoreactive crosslinker.

The reactive cysteine-178 of CRP was first replaced with serine (Pendergrast *et al.*, 1992) and then a new cysteine was introduced into the CRP 52-loop at position 52 or 56. Another mutant was constructed having the K52N substitution in addition to its G56C substitution. The CRP mutants produced for photocrosslinking were C178S-G56C, C178S-K52N-G56C and C178S-K52C. Transcription activation was determined *in vivo* by assaying β-galactosidase activity. All three mutant CRPs retain the ability to activate transcription on both *lac* and *gal* type promoters (Figure 13). The mutant CRPs were purified and reacted with benzophenone-4-maleimide (BPM) to specifically modify the solvent accessible cysteine residue in CRP introduced by site directed mutagenesis into the 52-loop. The BPM-derivatized CRP mutants were used to test their ability to support single round run-off transcription, the results indicate that they were all able to transcribe both CCpmeIR and CC+20pmeIR equally well as compared to the wild type (Figure 14).

To assess the ability of the BPM-CRP to contact RNA polymerase, photocrosslinking was carried out on the (BPM-CRP)-DNA-RNA polymerase ternary complex. After UV irradiation, DNase I was added to digest the promoter DNA and the mixture was resolved on a 10% SDS polyacrylamide gel. Western blotting with monoclonal antibodies specific for CRP, α, β, β', and σ was carried out to detect possible crosslinked products. The results indicated that BPM-CRP crosslinked with only the σ^{70} subunit of RNA polymerase *via* the CRP N-terminal 52-loop. After photoactivation in the presence of cAMP and promoter DNA, a novel protein band with a molecular weight of approximately 100 kDa was observed on the Western blot which crossreacted with both the anti-CRP (64D1) and anti-σ (2D1) monoclonal antibodies (Figure 15). Several crosslinked bands which cluster at around 45 kDa were also detected. Since the 45 kDa bands are only recognized by anti-CRP antibodies, we believe they are the crosslinked products formed between CRP monomers. No crosslinking of CRP with α, β, or β' was detected (results not shown).

The photocrosslinking between the CRP 52-loop and the σ subunit of the RNA polymerase holoenzyme is strictly cAMP dependent and to lesser degree promoter DNA dependent. In the absence of cAMP, no crosslinked product of ~100 kDa was detectable, whereas the crosslinked CRP dimers at ~45 KDa were still present. In the presence of cAMP, all three CRP derivatives C178S-G56C, C178S-K52N-G56C and C178S-K52C crosslink with the σ subunit, but each of the BPM-derivatized CRP mutants show crosslinking to σ to varying extent on the different

DNA promoters (Figure 15). Overall, C178S-K52C showed the highest efficiency for crosslinking with σ^{70} on the *gal*-like CCpmelR. The relative efficiency of crosslinking of different mutants on different DNAs is summarized in Table IV. The crosslinking efficiencies of the CRP derivatives on *gal*-like or *lac*-like CCpmelR and CC+20pmelR with σ^{70} are better than on the actual *gal* and *lac* promoters, presumably because of the consensus CRP binding sites on these synthetic promoters (Figure 16).

In order to determine the specificity of the crosslinking of CRP 52-loop with σ^{70}, another CRP mutant C178S-T158C was made with a cysteine substitution at the 158 position in the surface exposed loop 156-162 present at the C-terminal domain of CRP. The same UV crosslinking reaction was carried out and the results are shown in Figure 17. C178S-T158C was found to crosslink with the α subunit of RNA polymerase consistent with previously published results (Chen *et al.*, 1994) while no crosslinking was detected between this mutant and the σ^{70} subunit. Ebright's group has shown that the 156 loop is close to the RNA polymerase α subunit on the *lac* promoter. In addition, my results indicated that this loop can crosslink with α on both *lac* and *gal* promoters and that this crosslinking is also strictly DNA dependent.

To strengthen the conclusion that the crosslinking of CRP with σ^{70} relies on the functional preinitiation complex, substrate nucleotides (ATP, CTP, UTP and O^{met}GTP) were incubated with the [C178S-K52N]-*gal* P^{+}-RNA polymerase ternary complex mixture before UV irradiation. After 5 minutes incubation at 37°C, a 10

nucleotide long transcript (terminated by the presence of $O^{met}GTP$) was allowed to form and then followed by UV irradiation as usual. The results indicated that following initiation of transcription and the clearance of the *gal* promoter, the crosslinked CRP-σ^{70} product disappeared or significantly decreased (Figure 18).

In order to identify whether the crosslinking of the CRP 52-loop with the σ subunit depends on the functional activating region I (156-162 loop), an amino acid substitution T158A (*pc*) was generated in addition to the 52-loop cysteine substitutions. Three novel CRP mutants made were C178S-G56C-T158A, C178S-K52N-G56C-T158A, and C178S-K52C-T158A. Photocrosslinking reactions were carried out by using these mutants. The results indicated that T158A substitution, which made CRP a positive control mutant (for *lac* expression), did not affect the contact between the CRP 52-loop and σ (Figure 19).

Characterization of the N-terminal glycine substitutions:

N-terminal glycine residues of CRP were found to be conserved in CRP, FNR of *E. coli* and another regulatory protein *fix* K in the *Rhizobium meliloti* (Batut *et al.*, 1989). These glycine residues are located at positions 33, 45, 56, 67 and 71 which are mainly in the turns or inside the β strands of the β sheets that form the cAMP binding pocket. These glycine residues were substituted with alanine by using PCR site directed mutagenesis, the sequence of the mutagenized gene was confirmed by dideoxynucleotide DNA sequencing. The transcriptional stimulation activity of these substitutions were measured *in vivo* on both *lac* P1 and *gal* P1 promoters by using β-galactosidase activity assays. The results indicate that G33A, G56A and G67A retained fairly high activity (>85%) on the *lac* promoter compared to the wild type. G71A is less active (~50%) and G45A is almost inactive, retaining only about 20-25% of the wild type activity (Figure 20A). The same results were observed on the *gal* promoter (Figure 20B). In *in vivo* DNA binding activity assays, G33A, G56A and G67A showed comparable affinity for their specific DNA binding site, but G71A showed ~60% and G45A only retained about 10% of the DNA binding activity (Figure 21). Western blotting assays were performed on the crude cell extracts of these substitution mutants, the results indicate that G45A is a very unstable protein, it only showed an extremely faint band after Western blotting which explains the low activity of this protein. The amount of G71A protein is somewhat quantitatively lower than the rest of the substitutions (Figure 22).

DISCUSSION

CRP mutants defective in transcription activation for class I promoters have been located within the CRP loop extending from amino acids 156 to 162. It has been proposed that this region which is exposed in the DNA-CRP complex is involved in contact with RNA polymerase (Reznikoff, 1992; Ebright, 1993). Ebright (Ebright, 1993; Zhou *et al.*, 1993) has proposed that Thr-158 and His-159 in CRP are the residues directly involved in RNA polymerase contact since they are exposed in the CRP-DNA complex. Our results indicate that the conformation as well as the particular amino acid side chains of the activation domain are important factors in activation. CRP-HrsP^{160} and CRP-GrsM ^{163} demonstrated to be CRP ^{pc} mutants have insertions which are adjacent to His-159 and Gly-162; the insertion of two amino acid residues would be expected to result in a locally altered conformation. The lesser effect on activation consequent to insertion of Arg-Ser seen for CRP-PrsD^{161} is somewhat puzzling but points out the importance of position for inactivation. The insertions do not act by replacing a critical amino acid residue as has been shown for other CRP^{pc} mutants but appear to act by displacing it. The complementary interaction between amino acid residues in CRP and RNA polymerase presumably involve a particular conformation for each partner. Insertion into sites within the activation loop of CRP can be expected to distort the conformation as well as the local amino acid composition and thus impair productive interaction.

The CRP^{pc} mutants resulting from single amino acid substitutions in the activation loop are defective in their ability to activate transcription from the *lac* promoter both *in vivo* and *in vitro* while retaining the ability to activate *in vitro gal* expression (Reznikoff, 1992) (except T158A and H159L).This property is also evinced by our insertion mutants.

CRP has evolved to be able to assume two distinct conformations. The unliganded form is more compact than the cAMP-bound form of CRP. The unliganded CRP is relatively resistant to attack by endoproteases (Eilen *et al.*, 1978; Ebright *et al.*, 1985) and carboxypeptidase Y (Yang and Krakow, 1991). The unliganded form of each of the four CRP -Arg-Ser- insertion mutants is relatively sensitive to attack by trypsin and subtilisin. The trypsin sensitivity is not merely due to the inserted arginine residue. There are several arginine and lysine residues present in the C-terminal domain of wild-type CRP which are not available to attack by trypsin in the absence of cAMP. The insertions presumably alter the interactions required to maintain the normal folded conformation of CRP.

Compelling evidence indicates that direct contact between RNA polymerase and CRP is required for transcription activation. In the case of Class I promoters such as *lac* P1, contact involves CRP loop-161 and the C-terminal region of the α subunit of RNA polymerase. In Class I promoters the DNA site for CRP is centered at position -61.5 or further upstream (Ishihama, 1988). The involvement of the α subunit in activation is not limited to CRP but also involves a number of other transcription factors which bind to Class I sites (Ishihama, 1993).

Class II promoters such as *gal* P1 (CRP binding site centered at position -41.5) are affected to varying extent by *crp*^{pc} mutations which result in loss of transcription activation from Class I promoters. West *et al.* (1993) showed that amino acid substitution in the 159-contact loop (H159L, T158A) reduced CRP-dependent activation according to the Class II promoter used. Earlier work indicated that mutation in the loop involving amino acid 52 could reactivate the *crp*^{pc} mutation at position 159 (Bell *et al.*, 1990). It was suggested by West *et al.* (1993) that mutation in the 52-loop could "improve or create a second 'patch' on CRP" which may contact RNA polymerase.

I have provided direct evidence for contact between the 52-loop and the σ subunit of RNA polymerase. Crosslinking is specific for the σ subunit and requires the cAMP bound conformer of CRP. Little crosslinking occurs in the absence of DNA. The mutants prepared for the present study show activity comparable to wild type CRP for *lac* expression. Two of the mutants (K52C and K52N-G56C) support two to three fold increased expression from a *gal*-like promoter. Since the photocrosslinker benzophenone-4-maleimide used in this study is a rather small molecule (about 9 Å in length), the 52 position loop of CRP and the σ^{70} subunit must be in very close proximity. Mutant CRPs which have cysteine substituted at different positions in the 52-loop showed different efficiencies for crosslinking with different promoters, suggesting that the structure of the CRP protein as well as the promoter DNA is very important for CRP-RNA polymerase contact. On the *gal* promoter, the ability of C178S-K52N-G56C to support *in vivo* transcription is almost three times

stronger than C178S-G56C, and C178S-K52N-G56C also showed higher efficiency of crosslinking to σ^{70} than C178S-G56C. This implies that stronger contact between CRP and σ^{70} through the 52-loop might account for the superactivity of C178S-K52N-G56C. This phenomenon was observed in another mutant, C178S-K52C, which also showed increased biological activity as well as crosslinking efficiency. Besides the K52N substitution, K52E and K52Q were also found to act as superactive mutants (West *et al.,* 1993). My K52C substitution was another new member of these superactive mutants, indicating that the removal of Lys-52 could dramatically increase the activity of CRP to activate the transcription from *gal* but not *lac* promoters.

Transcription factors can interact with RNA polymerase both in the presence and absence of promoter DNA (Ishihama, 1988). The contact site of CRP with RNA polymerase on Class I promoters has been located on the 159-loop of CRP and the C-terminus of the α subunit of RNA polymerase. While the C-terminus of α is indispensable for factor activation on Class I promoters, it does not appear to be essential for Class II promoter activation (Ishihama, 1993; Kolb *et al.*, 1993; West *et al.,* 1994). On Class II promoters like *gal* P1, where the CRP binding site is located at the -41.5 position it appears that protein-protein contact also plays an important role in formation of the CRP-RNA polymerase-DNA complex essential for transcription initiation. It has been demonstrated that mutations near the upstream end of region 4.2 (the recognition domain for the promoter -35 region) of σ^{70} render RNA polymerase inactive in response to PhoB (Makino *et al.,*1993). The C-terminal

deletions of the σ^{70} subunit of RNA polymerase render it incapable for stimulation by CRP on the CRP-dependent *gal* promoter and PhoB on the PhoB-dependent P*pstS* promoter (Kumar *et al.,* 1994). The latter results not only confirmed the previous result that the contact site for PhoB is located in the C-terminal region of σ^{70}, but also indicate that the contact site for CRP, at least on the *gal* promoter, is at the same C-terminal region of σ^{70}. Kumar *et al.* (1994) also proposed that the contact site for CRP for *gal* P1 activation could be located between amino acids 530-539 on σ^{70}. Our crosslinking results are consistent with the conclusion of Kumar *et al.* (1994) that CRP physically contacts σ^{70} . Whether the σ^{70} region comprising amino acids 530-539 actually is the site for contact with CRP remains to be directly determined.

In order to demonstrate that the formation of a functional preinitiation complex on the *gal* promoter is necessary for the crosslinking between CRP 52-loop and σ^{70} subunit of RNA polymerase, photocrosslinking assays were carried out after the initiation of transcription. This was first done by incubating the cAMP/CRP derivative C178S-K52N-G56C, *gal* p1 promoter fragment and RNA polymerase holoenzyme to form a functional ternary preinitiation complex. The ATP, CTP, UTP and O^{met}GTP mixture was added to the preinitiation complex as substrates for transcription. Since the initial 10 ribonucleotide sequence of the *gal* P1 transcript is AUACCAUAAG..., this would enable transcription to initiate but pause after incorporation of the first 10 ribonucleotides because of the addition of the chain terminating O^{met}GMP at the tenth position. Photo-irradiation was then carried out;

crosslinking between CRP and σ^{70} could be hardly detected and sometimes was undetectable after the initiation of transcription. This result strongly demonstrated that CRP- σ^{70} crosslinking relies on the formation of a functional transcription complex. Heparin, a polyanionic molecule which mimics DNA and might compete for the DNA binding protein, was added together with the substrate mix and the same result was observed. As a control, adding heparin to the preformed RPo, did not inhibit CRP-σ^{70} crosslinking. This indicated that once the open preinitiation complex was formed, the binding of CRP and RNA polymerase on promoter DNA was cooperative and strong enough not to be dissociated by heparin competition. Furthermore, CRP and σ^{70} remains in contact even after the formation of the open promoter complex and is released only after promoter clearance.

The surface exposed loop of CRP spanning amino acids 156-162 was demonstrated to be in close contact with the C-terminus of the α subunit of RNA polymerase on the *lac* promoter (Chen and Ebright, 1994). To further establish the specificity of the crosslinking of my N-terminal CRP 52-loop mutations with the σ^{70} subunit, a cysteine substitution at the 158 position was made (C178S-T158C). This mutant was used for crosslinking reactions using the same conditions as for the 52-loop substitutions. The crosslinking results on the *lac* promoter were the same as those reported by Chen and Ebright (1994); C178S-T158C crosslinked with the α subunit of RNA polymerase on the *lac* promoter, while no crosslinking with β, β', or σ was detected. This crosslinking was also cAMP and DNA dependent. Surprisingly, C178S-T158C also showed comparable efficiency for crosslinking with

the α subunit on *gal* as well as *lac* promoters. This was at first puzzling to us, but recently published evidence indicated that CRP contacts the α subunit of RNA polymerase on the *gal* P1 promoter as demonstrated by DNA footprinting (Attey *et al.* 1994). They suggested that α binds at the upstream end of both the binary RNA polymerase-*gal* P1 and ternary RNA polymerase-CRP-*gal* P1 complexes. In the ternary complex, α appears to make direct contact with activating region 1 (which includes the 156-162 loop) in CRP. Taking this into consideration, our results showing crosslinking of C178S-T158C with α on both *lac* and *gal* promoters not only clarified the specificity of crosslinking (52-loop with σ, and 158-loop with α), but also provided another line of direct evidence that the activating region 1 of CRP makes direct contact with α on both *lac* and *gal* promoters.

The comparable crosslinking efficiency of CRP 52-loop with σ, whether or not the CRP derivative has the *PC* substitution, T158A, suggests that the second contact of CRP 52 loop with the RNA polymerase σ subunit is independent of the first contact of the CRP 156-162 loop with the RNA polymerase α subunit. Heyduk *et al.* (1993) demonstrated that CRP interacts with RNA polymerase in solution in the absence of promoter DNA, and this interaction became undetectable when the *PC* mutation T158A is introduced into the 156-162 loop. It would be very interesting to see if the introduction of the superactive substitution K52N could increase the interaction between CRP and RNA polymerase in solution.

No direct evidence is available indicating that the 52-loop of CRP is essential for stimulating transcription of *gal* P1. It remains possible that the contact between

the CRP 52-loop and σ is gratuitous and not directly involved in transcription activation. It is also possible that more than one site on CRP may contact one or more RNA polymerase subunits for overall stimulating activity. So far, at least two of the RNA polymerase contact sites on CRP have been identified. West *et al.* (1993) suggest that neither the 52-loop nor the 159-loop is essential for activating transcription on Class II promoters, another activating region, perhaps located around E96, might also be involved in transcription stimulation.

Due to its ability to assume many backbone dihedral angles, glycine can play a role in conformational transitions. Glycine residues at positions 33, 45, 56, 67 and 71 are conserved in CRP, FNR and *fix* K (Batut *et al.*, 1989). The conservation of glycines infers a comparable role for glycines in these regulatory proteins. This feature is also important structurally because these glycine residues belong to turns or loops which require positive Φ dihedral angles in the CRP structure. These glycine residues are in the large N-terminal domain of CRP which functions in cAMP binding and subunit:subunit interaction. These glycines in CRP were individually changed to alanine by using site directed mutagenesis. Since these glycines are mainly in the turns or inside the β strands of the β sheets which form the cAMP binding pocket, it seemed likely that some of the substitutions could disrupt the normal conformation for cAMP binding or local β sheet structure of the N-terminal domain. The results of the alanine substitution indicate that glycine 45 is very important for maintaining the normal CRP structure, a single alanine substitution led to almost the total loss of the protein *in vivo.* Alanine substitutions

at 33, 56 and 67 do not seem to markedly affect DNA binding and activity suggesting that these positions are more flexible and can tolerate a somewhat more rigid amino acid while still maintaining the normal structure and function of CRP. Whether CRP can tolerate the substitution of other amino acids is still unknown.

Table I. Oligonucleotide primers containing the AGATCT (Arg-Ser) insertion site.

QrsP[154] -CTGGCAAAACAAAGATCTCCAGACGCTATGACT-

HrsP[160] -GCTATGACTCACAGATCTCCGGACGGTATGCAA-

PrsD[161] -ATGACTCACCCGAGATCTGACGGTATGCAAATC-

GrsM[163] -CACCCGGACGGTAGATCTATGCAAATCAAAATTACC-

Table II. Ability of mutant CRPs to support sugar fermentation.

	Growth on Tetrazolium Plates Containing			
	Lactose	Galactose	Maltose	Arabinose
E. coli XE (*crp*$^-$)				
CRP	++	++	++	++
CRP-PrsD^{161}	+	++	-	++
CRP-HrsP^{160}	-	++	-	++
CRP-QrsP^{154}	+	++	-	++
CRP-GrsM^{163}	-	++	-	++

* ++, active; +, partially active; -, inactive

Table III. Summary of the mutation positions and activities of the 5 partial suppressors of the dipeptide insertion mutant $GrsM^{163}$.

	Mutagenized amino acid positions	Codon change	Amino acid substitutions	Activity (%)
$GrsM^{163}$ (S-1)	161(Asp)	GAC → GAA	Asp → Glu	34%
$GrsM^{163}$ (S-2)	161(Asp)	GAC → TAC	Asp → Tyr	45%
	Inserted Arg	AGA → AGG	Arg → Arg	
$GrsM^{163}$ (S-3)	127(Thr)	ACT → ACA	Thr → Thr	38%
	161(Asp)	GAC → GAA	Asp → Glu	
$GrsM^{163}$ (S-4)	161(Asp)	GAC → TAC	Asp → Tyr	40%
$GrsM^{163}$ (S-5)	5(Gln)	CAA → CGA	Gln → Arg	36%
	111(Asp)	GAC → GGC	Asp → Gly	
	122(Arg)	CGT → TGT	Arg → Cys	

Table IV. Relative crosslinking efficiency of the CRP derivatives with sigma factor of RNA polymerase on type I and type II promoters

	C178S-G56C	C178S-K52N-G56C	C178S-K52C
Type I (*lac*-like) promoter	+	++	+++
Type II (*gal*-like) promoter	+++	++++	+++

Figure 1. The amino acid sequence and DNA sequence of the molecule of *Escherichia coli* cyclic AMP receptor protein (From Aiba *et al.*, 1982).

CRP

```
          12                  24                  36                  48                  60                  72                  84
ATG GTG CTT GGC AAA CCG CAA ACA GAC CCG ACT CTC GAA TGG TTC TTG TCT CAT TGC CAC ATT CAT AAG TAC CCA TCC AAG AGC
TAC CAC GAA CCG TTT GGC GTT TGT CTG GGC TGA GAG CTT ACC AAG AAC AGA GTA ACG GTG TAA GTA TTC ATG GGT AGG TTC TCG
Met Val Leu Gly Lys Pro Gln Thr Asp Pro Thr Leu Glu Trp Phe Leu Ser His Cys His Ile His Lys Tyr Pro Ser Lys Ser
     1                   4                   7                  10                  13                  16                  19                  22                  25

          96                 108                 120                 132                 144                 156                 168
ACG CTT ATT CAC CAG GGT GAA AAA GCG GAA ACG CTG TAC TAC ATC GTT AAA GGC TCT GTG GCA GTG CTG ATC AAA GAC GAA GAG
TGC GAA TAA GTG GTC CCA CTT TTT CGC CTT TGC GAC ATG ATG TAG CAA TTT CCG AGA CAC CGT CAC GAC TAG TTT CTG CTT CTC
Thr Leu Ile His Gln Gly Glu Lys Ala Glu Thr Leu Tyr Tyr Ile Val Lys Gly Ser Val Ala Val Leu Ile LYS Asp Glu Glu
28                  31                  34                  37                  40                  43                  46                  49                  52                  55

          180                 192                 204                 216                 228                 240                 252
GGT AAA GAA ATG ATC CTC TCC TAT CTG AAT CAG GGT GAT TTT ATT GGC GAA CTG GGC CTG TTT GAA GAG GGC CAG GAA CGT AGC
CCA TTT CTT TAC TAG GAG AGG ATA GAC TTA GTC CCA CTA AAA TAA CCG CTT GAC CCG GAC AAA CTT CTC CCG GTC CTT GCA TCG
Gly Lys Glu Met Ile Leu Ser Tyr Leu Asn Gln Gly Asp Phe Ile Gly Glu Leu Gly Leu Phe Glu Glu Gly Gln Glu Arg Ser
       58                  61                  64                  67                  70                  73                  76                  79                  82

          264                 276                 288                 300                 312                 324                 336
GCA TGG GTA CGT GCG AAA ACC GCC TGT GAA GTG GCT GAA ATT TCG TAC AAA AAA TTT CGC CAA TTG ATT CAG GTA AAC CCG GAC
CGT ACC CAT GCA CGC TTT TGG CGG ACA CTT CAC CGA CTT TAA AGC ATG TTT TTT AAA GCG GTT AAC TAA GTC CAT TTG GGC CTG
Ala Trp Val Arg Ala Lys Thr Ala Cys Glu Val Ala Glu Ile Ser Tyr Lys Lys Phe Arg Gln Leu Ile Gln Val Asn Pro Asp
    85                  88                  91                  94                  97                 100                 103                 106                 109

          348                 360                 372                 384                 396                 408                 420
ATT CTG ATG CGT TTG TCT GCA CAG ATG GCG CGT CGT CTG CAA GTC ACT TCA GAG AAA GTG GGC AAC CTG GCG TTC CTC GAC GTG
TAA GAC TAC GCA AAC AGA CGT GTC TAC CGC GCA GCA GAC GTT CAG TGA AGT CTC TTT CAC CCG TTG GAC CGC AAG GAG CTG CAC
Ile Leu Met Arg Leu Ser Ala Gln Met Ala Arg Arg Leu Gln Val Thr Ser Glu Lys Val Gly Asn Leu Ala Phe Leu Asp Val
112                 115                 118                 121                 124                 127                 130                 133                 136                 139

          432                 444                 456                 468                 480                 492                 504
ACG GGC CGC ATT GCA CAG ACT CTG CTG AAT CTG GCA AAA CAA CCA GAC GCT ATG ACT CAC CCG GAC GGT ATG CAA ATC AAA ATT
TGC CCG GCG TAA CGT GTC TGA GAC GAC TTA GAC CGT TTT GTT GGT CTG CGA TAC TGA GTG GGC CTG CCA TAC GTT TAG TTT TAA
Thr Gly Arg Ile Ala Gln Thr Leu Leu Asn Leu Ala Lys Gln Pro Asp Ala Met Thr His Pro Asp Gly Met Gln Ile Lys Ile
       142                 145                 148                 151                 154                 157                 160                 163                 166

          516                 528                 540                 552                 564                 576                 588
ACC CGT CAG GAA ATT GGT CAG ATT GTC GGC TGT TCT CGT GAA ACC GTG GGA CGC ATT CTG AAG ATG CTG GAA GAT CAG AAC CTG
TGG GCA GTC CTT TAA CCA GTC TAA CAG CCG ACA AGA GCA CTT TGG CAC CCT GCG TAA GAC TTC TAC GAC CTT CTA GTC TTG GAC
Thr Arg Gln Glu Ile Gly Gln Ile Val Gly Cys Ser Arg Glu Thr Val Gly Arg Ile Leu Lys Met Leu Glu Asp Gln Asn Leu
    169                 172                 175                 178                 181                 184                 187                 190                 193

          600                 612                 624
ATC TCC GCA CAC GGT AAA ACC ATC GTC GTT TAC GGC ACT CGT TAA
TAG AGG CGT GTG CCA TTT TGG TAG CAG CAA ATG CCG TGA GCA ATT
Ile Ser Ala His Gly Lys Thr Ile Val Val Tyr Gly Thr Arg
196                 199                 202                 205                 208
```

Figure 2. Drawing of the dimer of the cyclic AMP receptor protein of *E. Coli.* (From de Crombrugghe *et al.*, 1984) . The N-terminal domain consists of α -helix A, β-sheets 1 to 8 and α-helices B and C. The DNA binding C-terminal domain consists of α-helices D, E and F, and the residues connecting these secondary structures. The two F helices, which clearly protrude from the dimer, are thought to provide many of the interactions with DNA. All of the interactions between the two subunits are provided by the large N-terminal domain and the majority of these are provided by the two long C helices that lie together in the center of the dimer. The two subunits are not exactly related by a perfect dyad axis of symmetry.

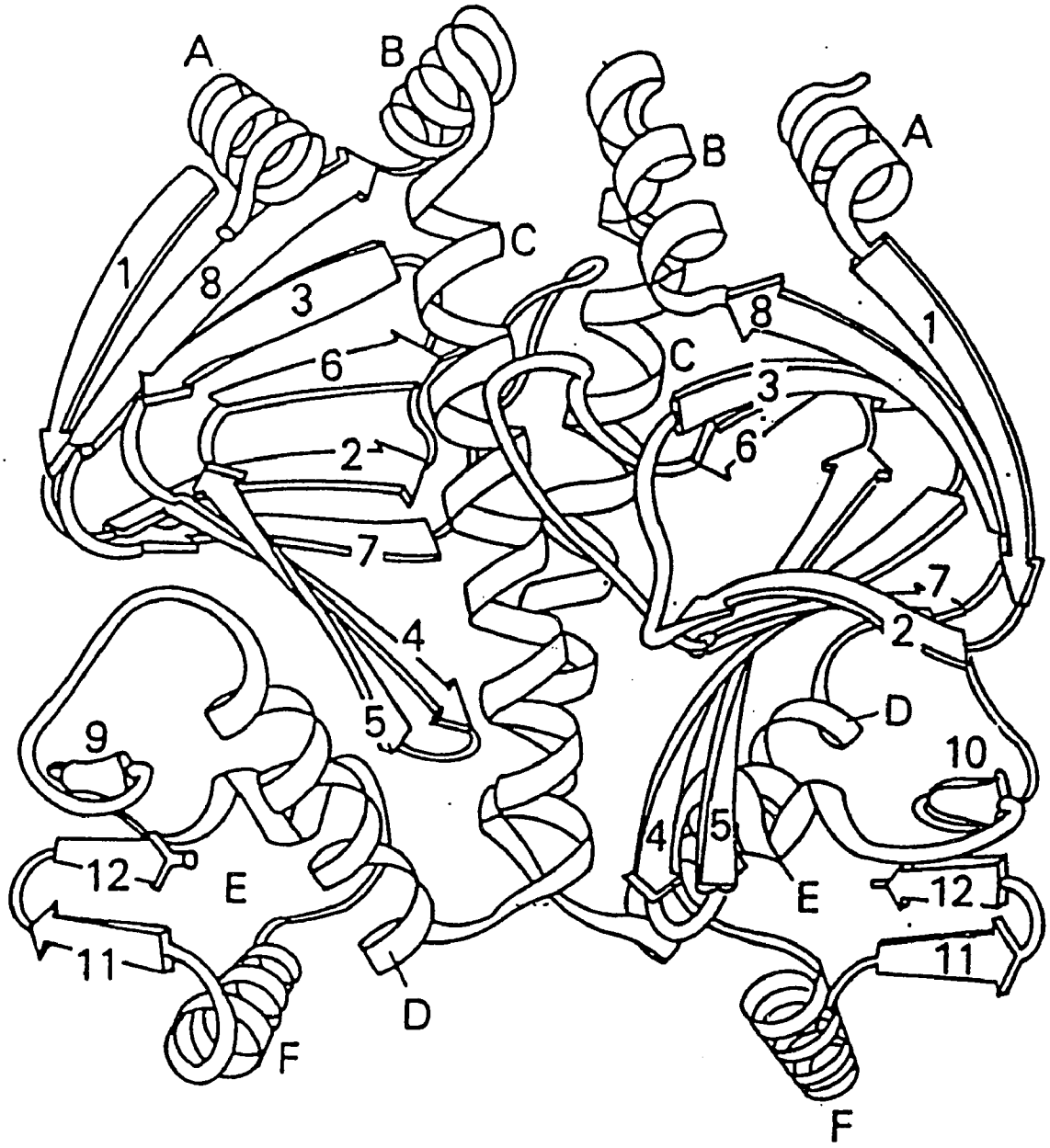

A
B
C
1
8
3
6
2
7
4
5
9
12
E
11
D
F
B
A
C
8
1
3
6
7
2
D
10
4
5
E
12
11
F

Figure 3. *In vivo* expression of the *lac* operon by the dipeptide insertion mutants. ß-galactosidase assays show the stimulating activity of dipeptide insertion mutant and wild-type CRPs which were expressed from pXZCRP variants in *E. coli* XE64.2 cells.

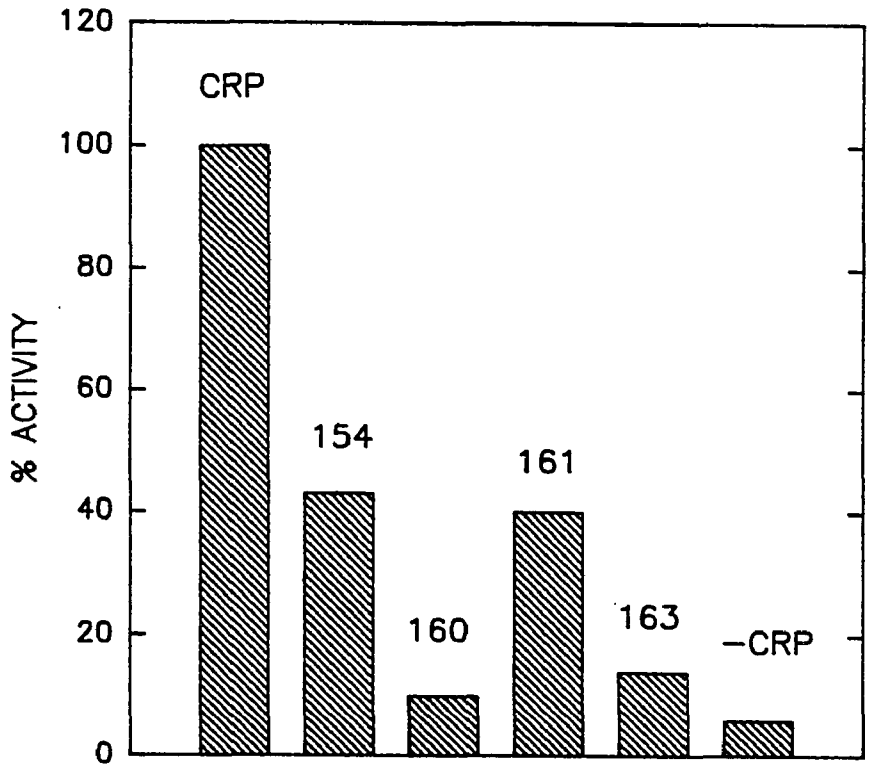

120
100
80
60
40
20
0
% ACTIVITY
CRP
154
160
161
163
−CRP

Figure 4. ß-galactosidase assays show the repressing ability of dipeptide insertion mutant and wild-type CRPs. Wild-type and mutant CRPs were expressed from pXZCRP variants in *E. coli* KC1071 cells.

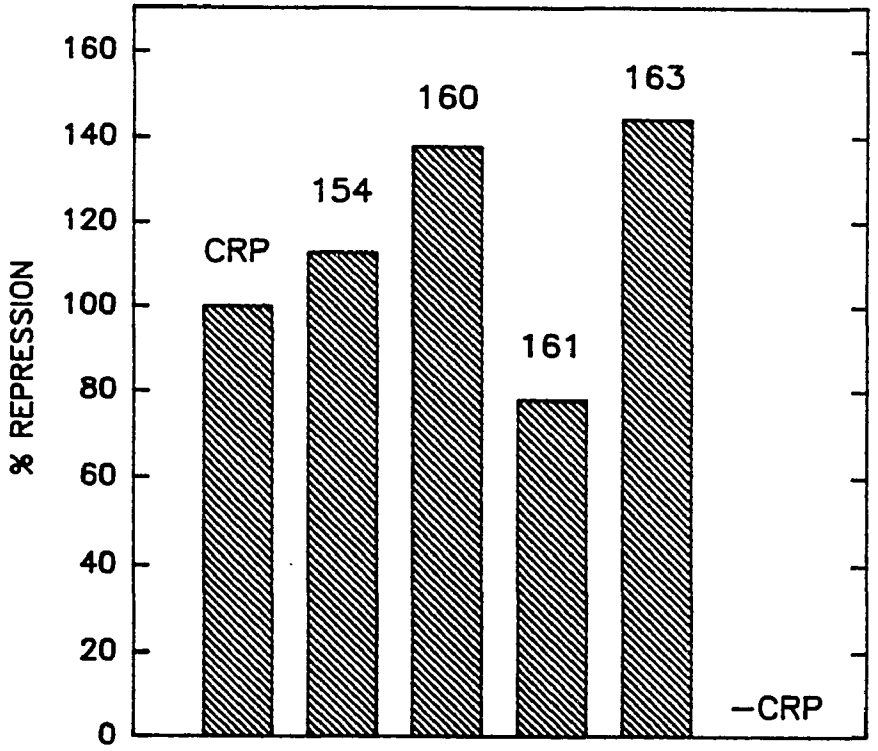

% REPRESSION
160
140
120
100
80
60
40
20
0
CRP
154
160
161
163
−CRP

Figure 5. Gel retardation assay show the specific DNA binding ability of dipeptide insertion mutant and wild-type CRPs. Conditions for the assays are described in "Materials and Methods" using 2 nM ^{32}P-labelled *lac* P $^{+}$ fragment and 90 nM mutant or wild-type CRP.

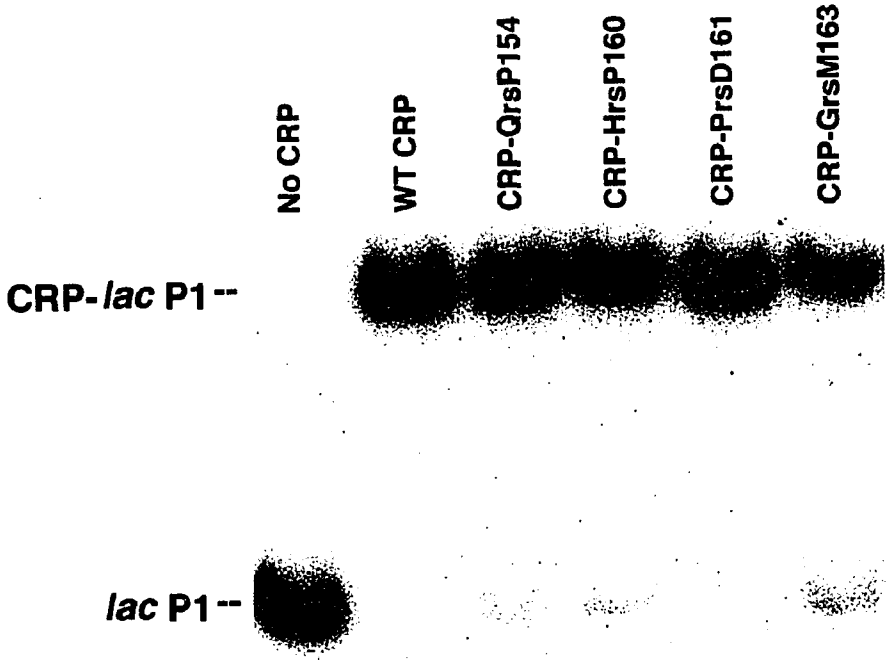

No CRP
WT CRP
CRP-QrsP154
CRP-HrsP160
CRP-PrsD161
CRP-GrsM163
CRP-lac P1--
lac P1--

Figure 6. Abortive initiation assay of dipeptide insertion mutants. The reaction were carried out as described in "Experimental Procedures" in 40 mM Tris-HCl, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT with 4 nM *lac* P^+ promoter. ApA and ^{3}H-UTP were used as substrates and incorporation of ^{3}H-UMP was determined. Wild-type CRP, ●-●; CRP-QrsP^{154}, ▽-▽; CRP-PrsD^{161} ▾-▾; CRP-GrsM^{163}, □-□; CRP-HrsP^{160}, ■-■.

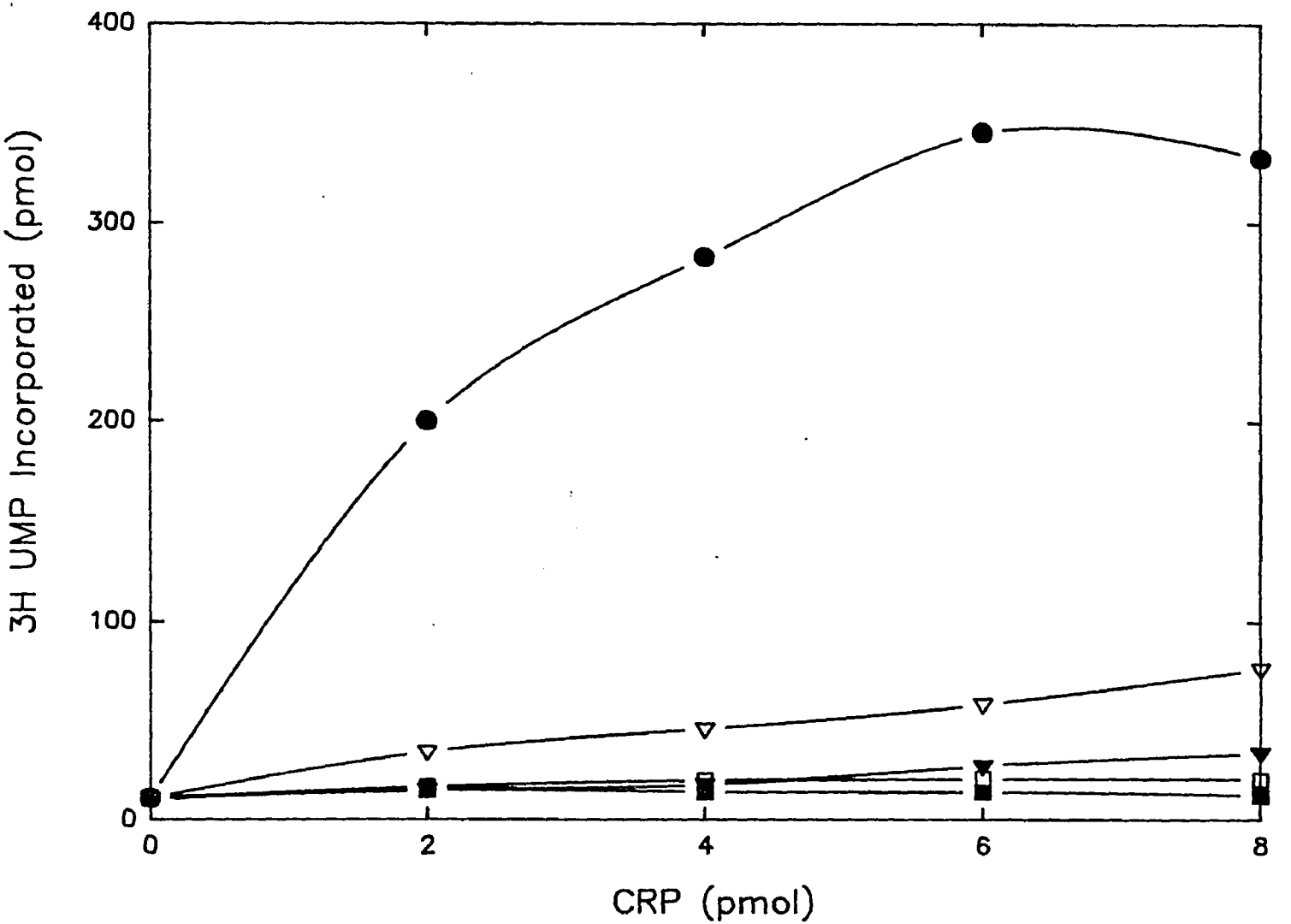

400
300
200
100
0
3H UMP Incorporated (pmol)
0
2
4
6
8
CRP (pmol)

Figure 7. Run-off assay showing the transcription activating activity of dipeptide insertion mutant CRPs on *gal* P^+ fragments. Conditions are described in "Experimental Procedures" with the *gal* P^+ fragment used as template. Lane a, wild-type CRP; lane b, CRP-PrsD^{161}; lane c, CRP-HrsP^{160}; lane d, CRP-QrsP^{154}; lane e, CRP-GrsM^{163}; lane f, no CRP.

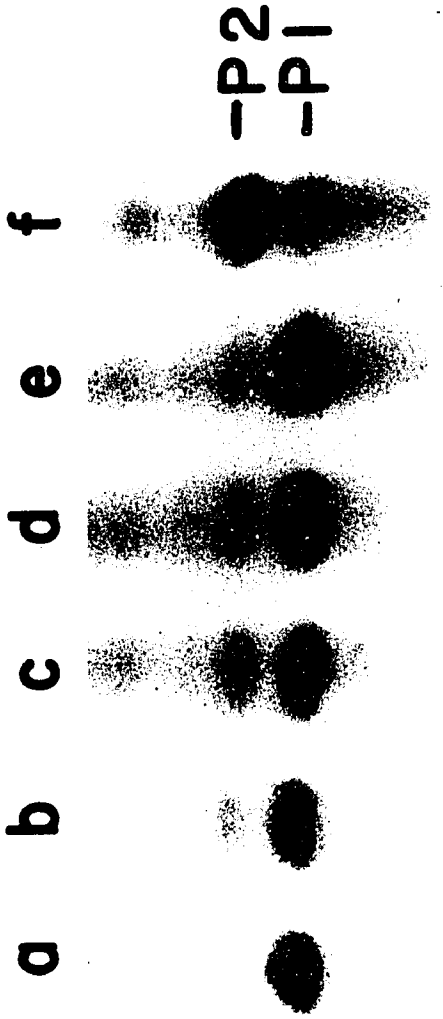

P2
P1
a
b
c
d
e
f

Figure 8. DNase I footprinting showing the ability of the dipeptide insertion mutant CRPs to interact with the CRP site on *lac* and with RNA polymerase. Conditions for footprinting are described in "Experimental Procedures". Lane a, wild-type CRP; lane b, wild-type CRP + RNA polymerase; lane c, CRP-PrsD^{161}; lane d, CRP-PrsD^{161} + RNA polymerase; lane e, CRP-HrsP^{160}; lane f, CRP-HrsP^{160} + RNA polymerase; lane g, CRP-QrsP ^{154} ; lane h, CRP-QrsP^{154} + RNA polymerase; lane i, CRP-GrsM^{163}; lane j, CRP-GrsM^{163} + RNA polymerase; lane k, omit CRP and RNA polymerase; lane l, RNA polymerase.

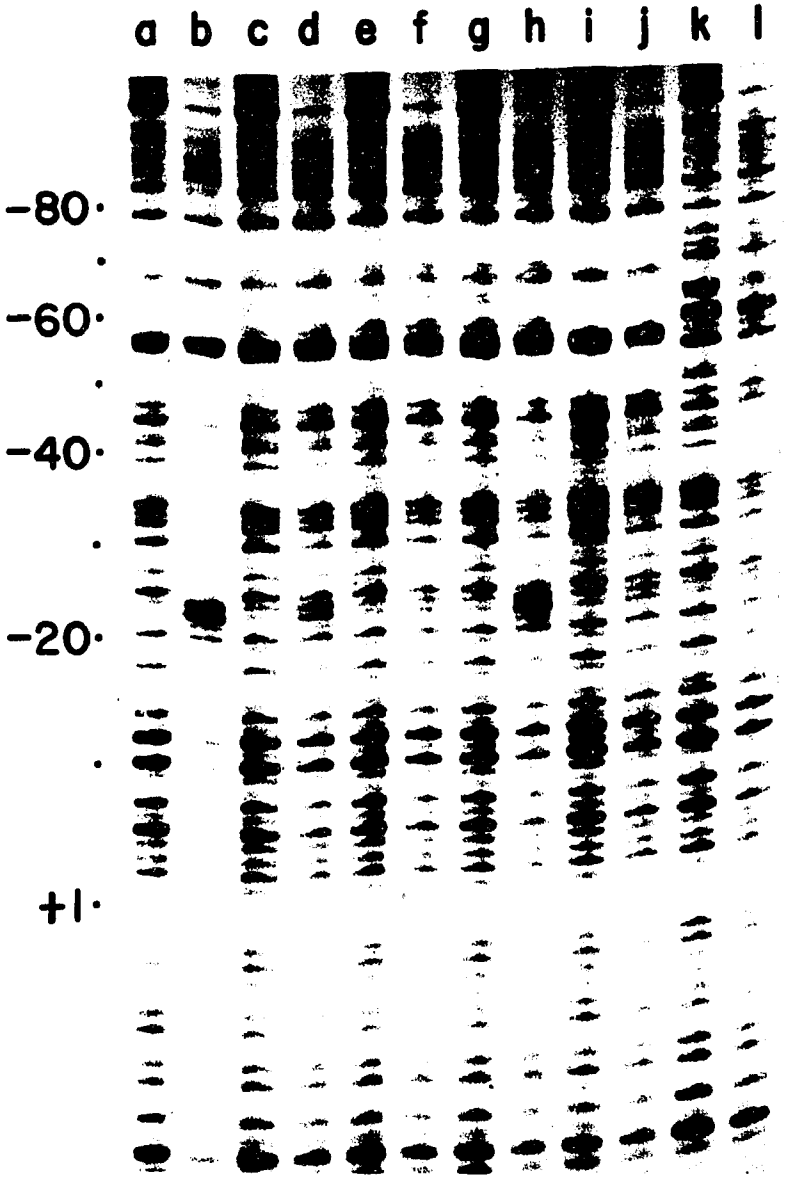

a b c d e f g h i j k l
-80
-60
-40
-20
+1

Figure 9A. Trypsin digestion patterns of dipeptide insertion mutant and wild-type CRPs. Conditions are described under "Experimental Procedures". Panel A: trypsin digestion; Panel B: subtilisin digestion. Lane a, wild-type CRP; lane b, CRP-PrsD^{161}; lane c, CRP-HrsP ^{160}; lane d, CRP-QrsP ^{154}; e, CRP-GrsM^{163}; lanes f-j contain 0.1 mM cAMP: lane f, wild-type CRP; lane g, CRP-PrsD^{161}; lane h, CRP-HrsP^{160}; lane i, CRP-QrsP^{154}; j, CRP-GrsM^{163}.

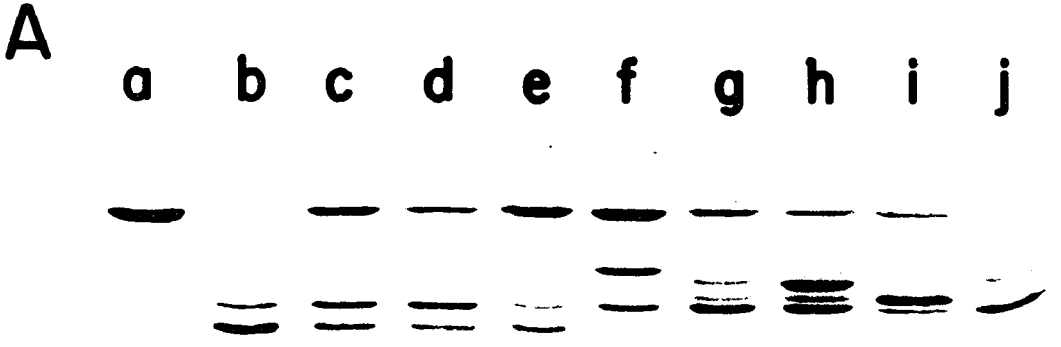

A
a b c d e f g h i j

Figure 9B. Subtilisin digestion patterns of dipeptide insertion mutant and wild-type CRPs. Conditions are described under "Experimental Procedures". Panel A: trypsin digestion; Panel B: subtilisin digestion. Lane a, wild-type CRP; lane b, CRP-PrsD^{161}; lane c, CRP-HrsP ^{160}; lane d, CRP-QrsP ^{154}; e, CRP-GrsM^{163}; lanes f-j contain 0.1 mM cAMP: lane f, wild-type CRP; lane g, CRP-PrsD^{161}; lane h, CRP-HrsP^{160}; lane i, CRP-QrsP^{154}; j, CRP-GrsM^{163}.

B
a b c d e f g h i j

Figure 10. *In vivo* β-galactosidase activity assays show the stimulating activity of the 156-162 loop insertion mutant CRPs on *lac* P^+ promoter. These mutant and wild-type CRPs were expressed from pPBRCRP variants in *E. coli* XE64.2 cells.

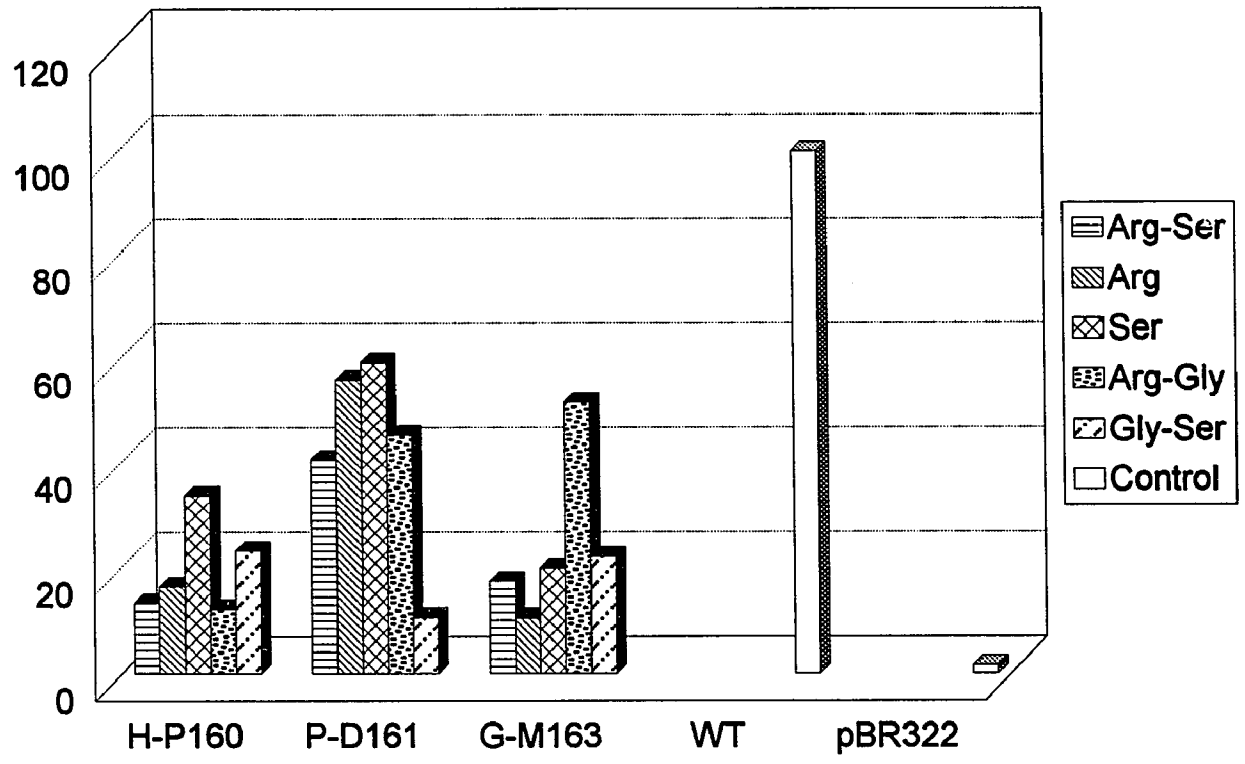

120
100
80
60
40
20
0
H-P160
P-D161
G-M163
WT
pBR322
Arg-Ser
Arg
Ser
Arg-Gly
Gly-Ser
Control
(% of activity on lacP1)

Figure 11. *In vivo* β-galactosidase activity assays showing the repressing ability of the 156-162 loop insertion mutant CRPs. These mutant and wild-type CRPs were expressed from pPBRCRP variants in *E. coli* KC 1071 cells.

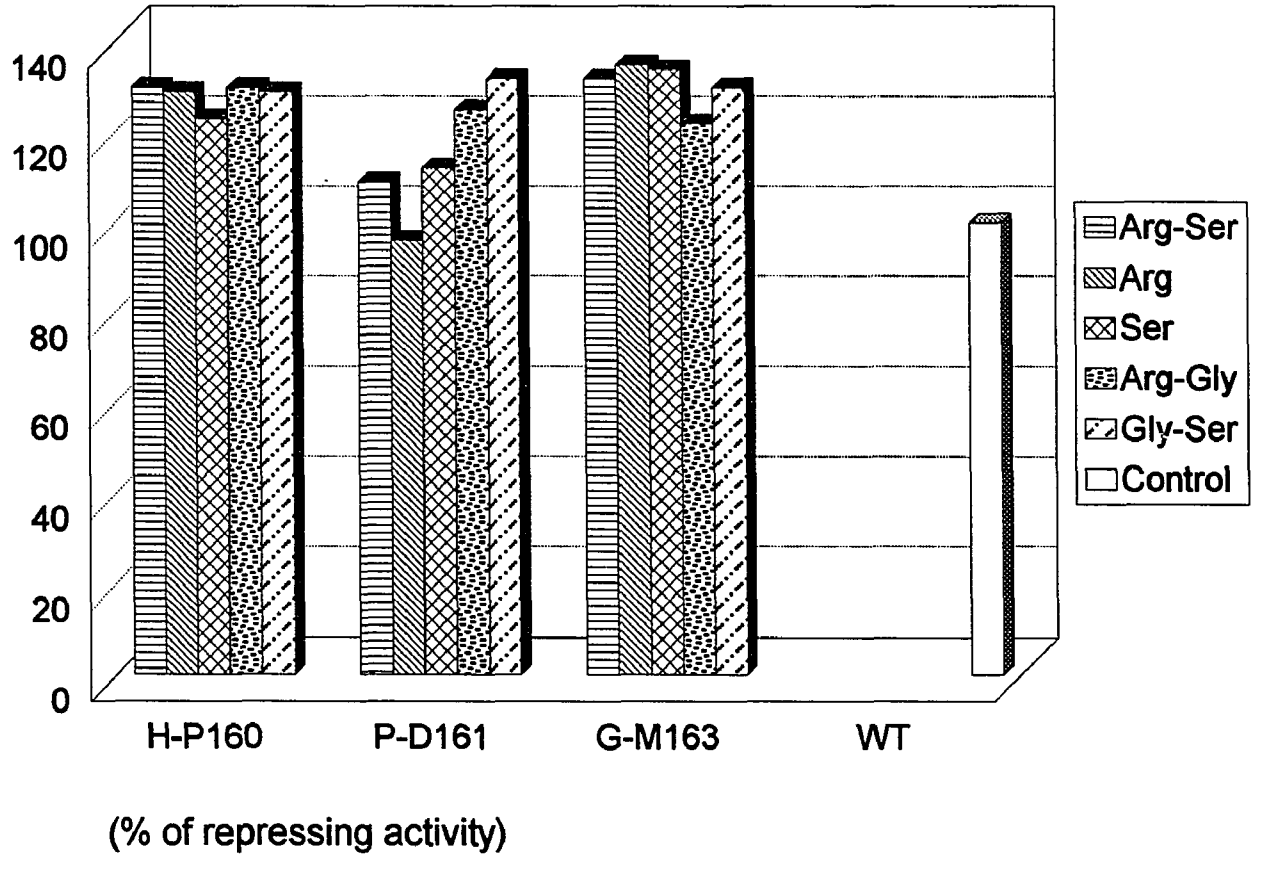
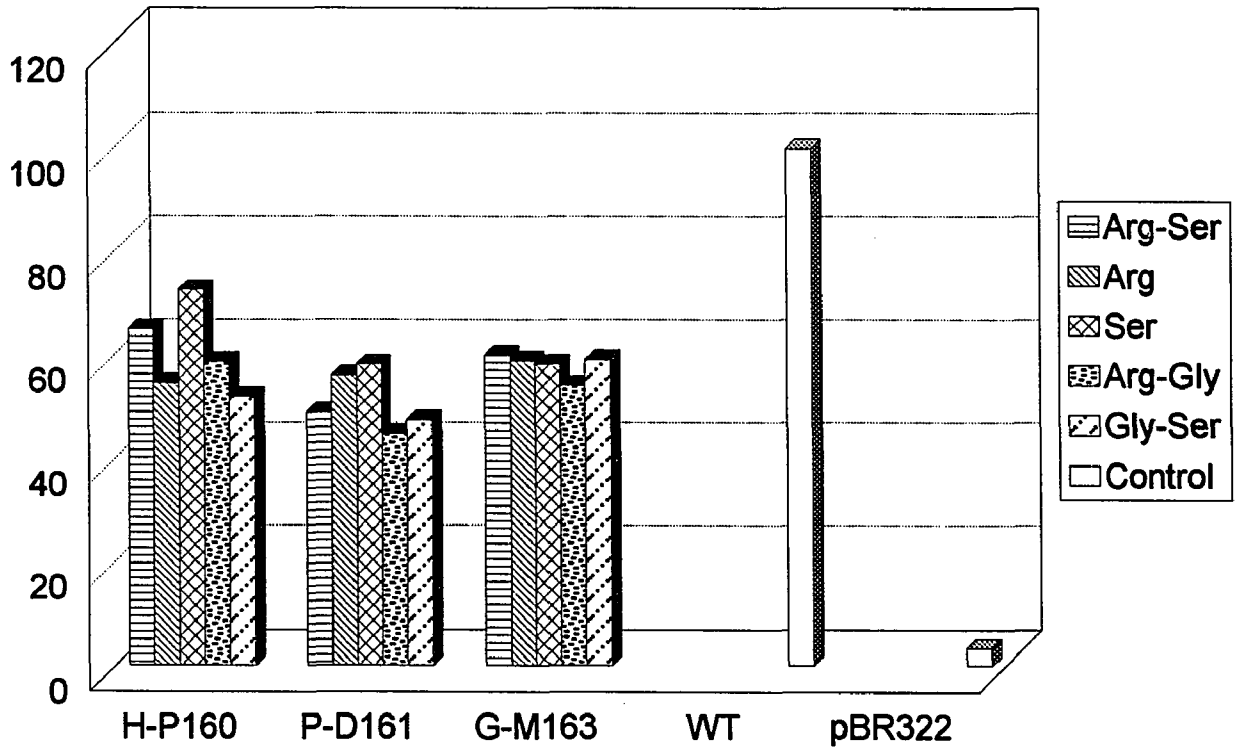

140
120
100
80
60
40
20
0
H-P160
P-D161
G-M163
WT
Arg-Ser
Arg
Ser
Arg-Gly
Gly-Ser
Control
(% of repressing activity)

Figure 12. *In vivo* β-galactosidase activity assays showing the stimulating activity of the 156-162 loop insertion mutant CRPs on *gal* P^+ promoter. These mutant and wild-type CRPs were expressed from pPBRCRP variants in *E. coli* M182 cells (Δ*crp39,Δlac*) with the co-transformation of *gal* P1 promoter in pRW2.



(% of activity on gal P1)

Figure 13 (A). *In vivo* expression of the *lac* operon stimulated by the 52-loop substitution mutant and wild-type CRPs. β-galactosidase assays showing the stimulating activity of mutant and wild-type CRP which were expressed from pPBRCRP variants in *E. coli* XE64.2 cells.

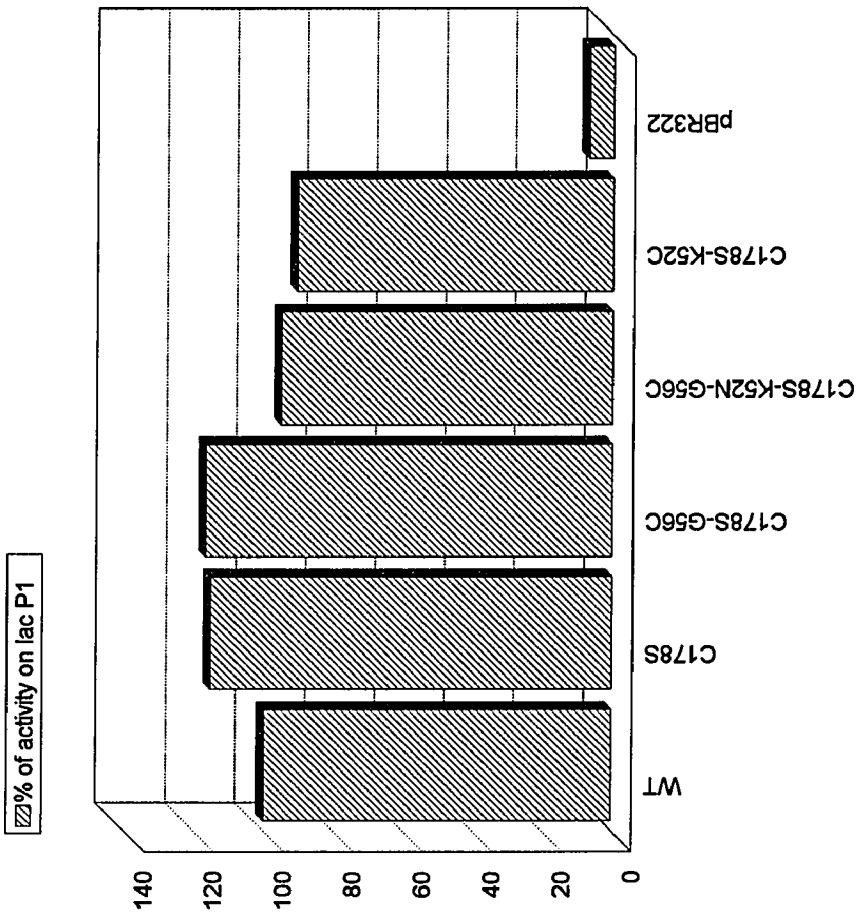

% of activity on lac P1
140
120
100
80
60
40
20
0
WT
C178S
C178S-G56C
C178S-K52N-G56C
C178S-K52C
pBR322

Figure 13 (B). *In vivo* expression of the *gal*-like class II promoter CCpmel R stimulated by the 52-loop substitution mutant and wild-type CRPs. β-galactosidase assays showing the stimulating activity of mutant and wild-type CRP which were expressed from pPBRCRP variants in *E. coli* M182 cells (*Δcrp39,Δlac*) with the co-transformation of CCpmel R promoter in pRW2.

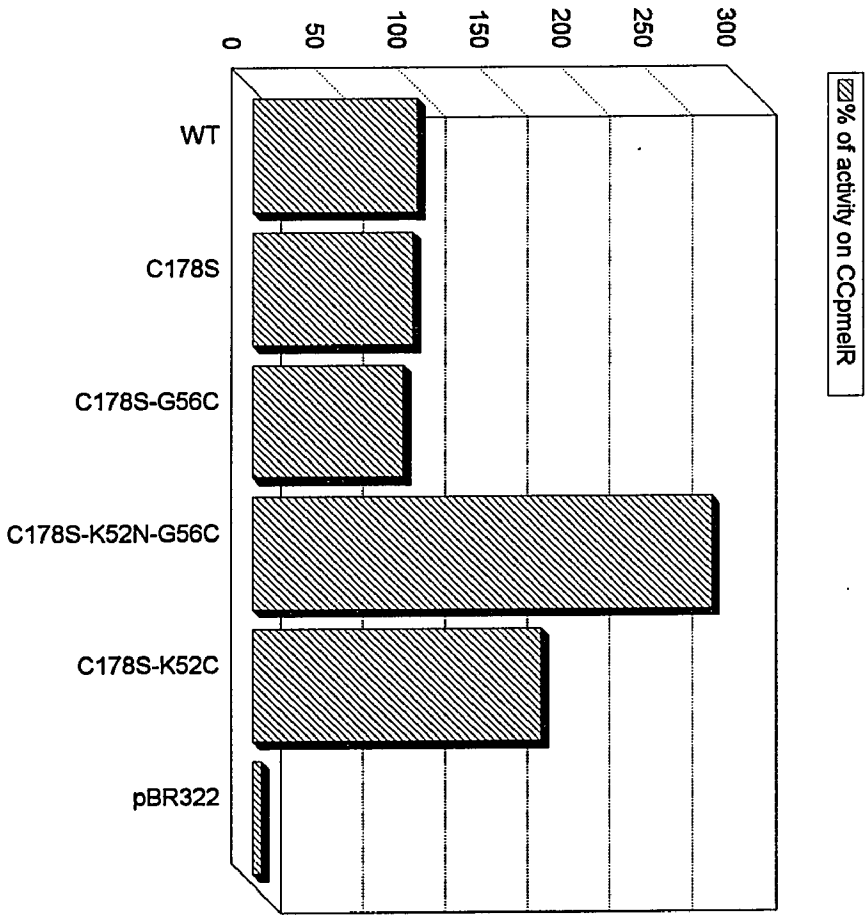

0
50
100
150
200
250
300
WT
C178S
C178S-G56C
C178S-K52N-G56C
C178S-K52C
pBR322
% of activity on CCpmelR

Figure 14. Single round run-off transcription assay show the ability of the BPM-derivatized CRP mutants to transcribe CCpmelR and CC+20pmelR promoters. The condition of the assay is described in "Materials and Methods". Runoff assays (final volume 35 µl) were carried out in 50 mM Tris-HCl (pH 7.8), 3 mM Mg acetate, 0.1 mM EDTA, 0.1 mM DTT, 25 µg/ml BSA, 50 mM NaCl, 0.5 pmol *lac* P^+ promoter fragment plus 0.5 pmol either CCpmelR (Panel A) or CC+20pmelR (Panel B) fragments, 4 pmol wild-type or BPM-derivatized CRP and 2 pmol RNA polymerase. After incubation at 37°C for 1 hr, 15 µl of a mix containing 10 µg heparin, 0.5 mM each of ATP, GTP, CTP and 0.16 mM UTP + 2 µCi ^{32}P-UTP was added and incubated for another 5 min. Reaction mixtures were denatured and resolved by electrophoresis on a 8% urea polyacrylamide gel.

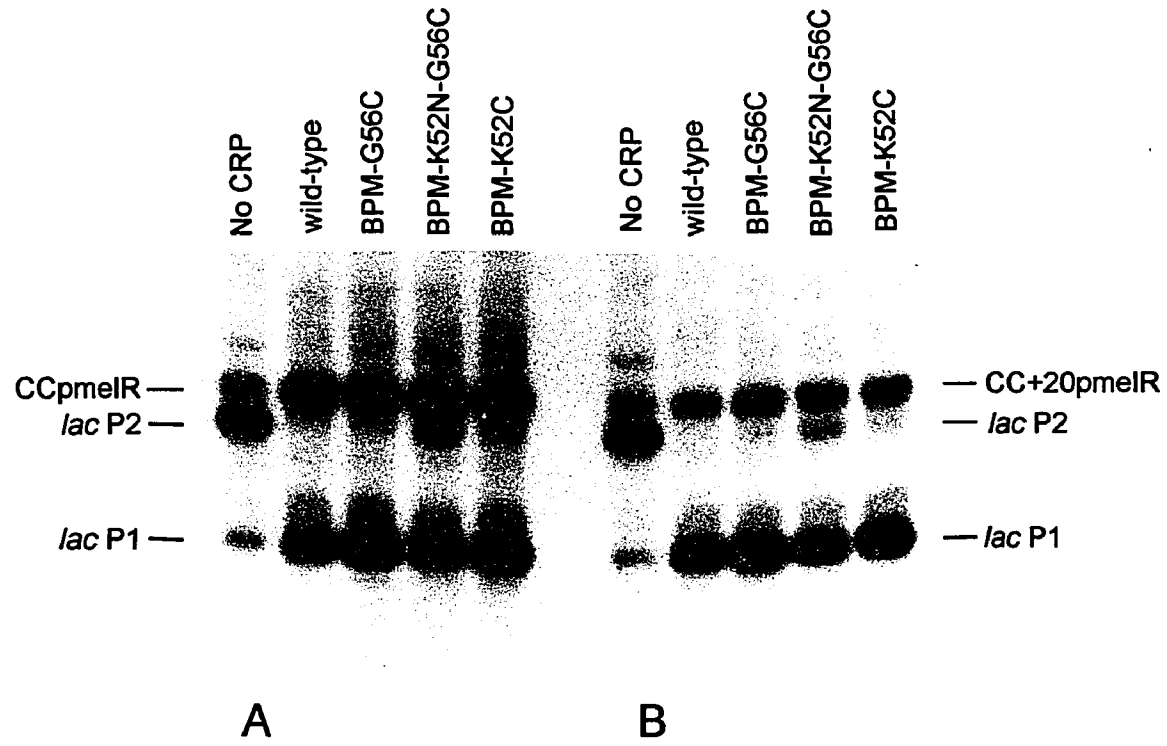

No CRP
wild-type
BPM-G56C
BPM-K52N-G56C
BPM-K52C
No CRP
wild-type
BPM-G56C
BPM-K52N-G56C
BPM-K52C
CCpmeIR
lac P2
lac P1
CC+20pmeIR
lac P2
lac P1
A
B

Figure 15 (A). Western blotting experiments showing the photocrosslinking results of BPM-C178S-G56C on different DNA fragments. Conditions for the assys are described in "Materials and Methods".

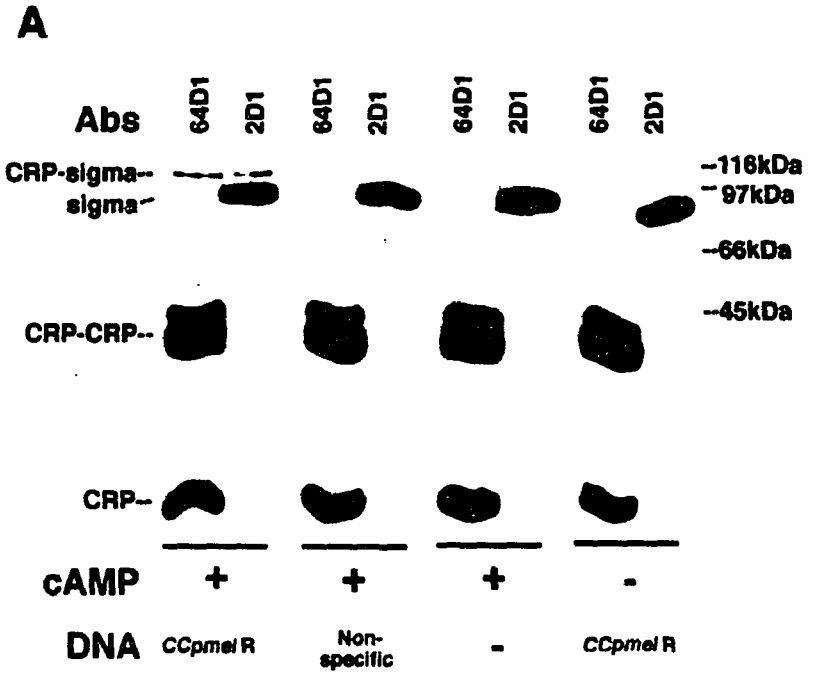

A
Abs
64D1
2D1
64D1
2D1
64D1
2D1
64D1
2D1
CRP-sigma--
sigma-
--116kDa
-97kDa
--66kDa
--45kDa
CRP-CRP--
CRP--
cAMP
+
+
+
-
DNA
CCpmel R
Non-specific
-
CCpmel R

Figure 15 (B). Western blotting experiments showing the photocrosslinking results of BPM-C178S-K52N-G56C on different DNA fragments. Conditions for the assys are described in "Materials and Methods".

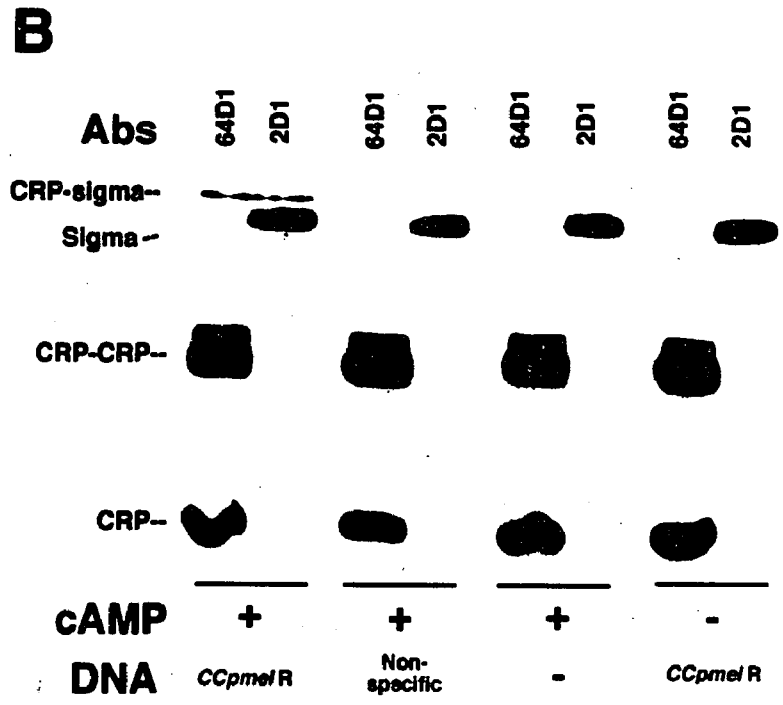

B
Abs
64D1
2D1
64D1
2D1
64D1
2D1
64D1
2D1
CRP-sigma--
Sigma--
CRP-CRP--
CRP--
cAMP
+
+
+
-
DNA
CCpmel R
Non-specific
-
CCpmel R

Figure 15 (C). Western blotting experiments showing the photocrosslinking results of BPM-C178S-K52C on different DNA fragments. Conditions for the assys are described in "Materials and Methods".

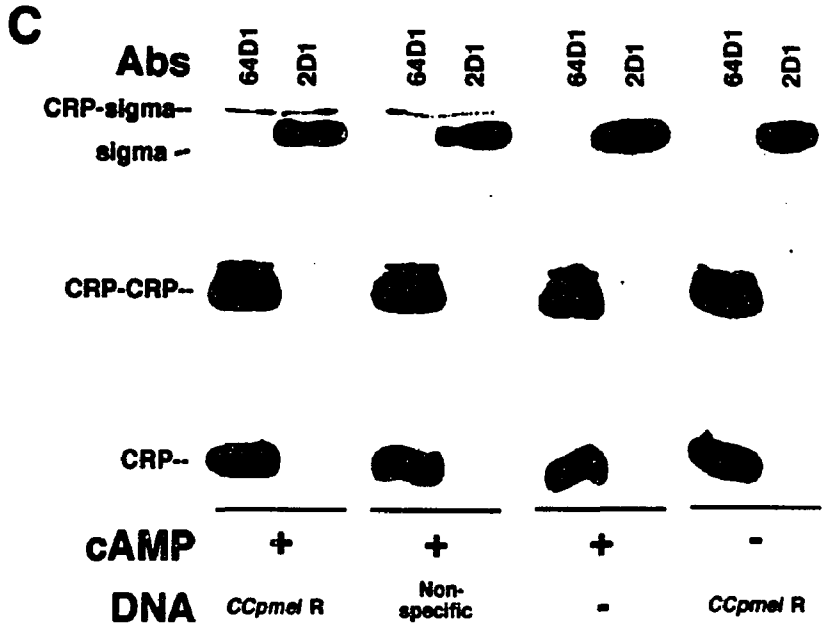

C
Abs
64D1
2D1
64D1
2D1
64D1
2D1
64D1
2D1
CRP-sigma--
sigma --
CRP-CRP--
CRP--
cAMP
+
+
+
-
DNA
CCpmel R
Non-specific
-
CCpmel R

Figure 16. Western blotting experiment showing the photocrosslinking result of BPM-C178S-K52N-G56C on different DNA promoters. Condition for the assay is described in "Materials and Methods". Promoters CCpmeIR, CC+20pmeIR, *gal* P^+ and *lac* P^+ are used in this experiment.

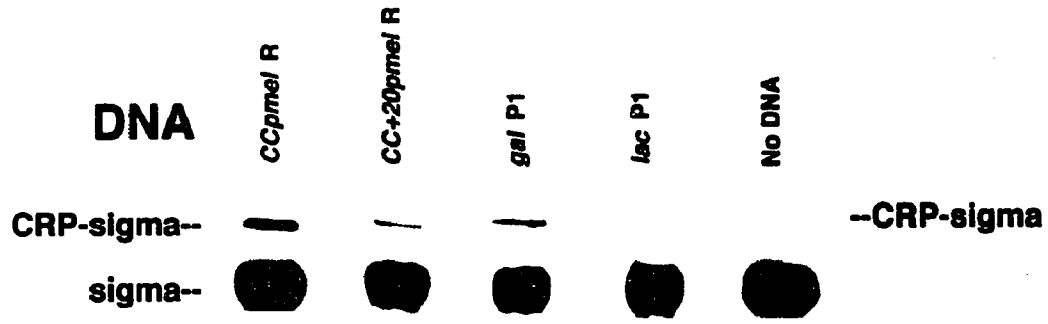

DNA
CCpmel R
CC+20pmel R
gal P1
lac P1
No DNA
CRP-sigma--
sigma--
--CRP-sigma

Figure 17. Western blotting experimrnt showing the photocrosslinking result of BPM-C178S-T158C on *gal* P^+ and *lac* P^+ promoters. Condition for the assay is described in "Materials and Methods". Monoclonal anti-CRP antibody 64D1, anti-σ antibody 2D1 and anti-α antibody 23C2 were used in the experiment where indicated.

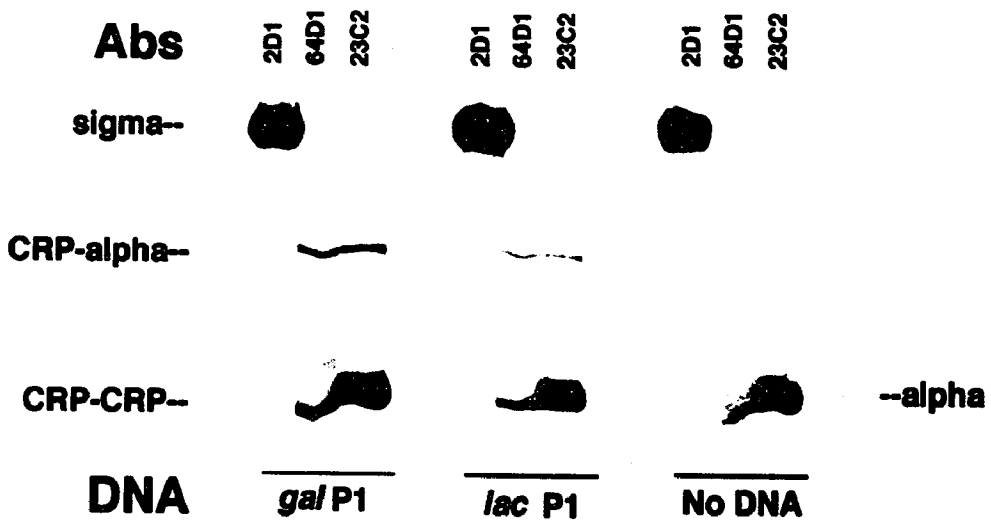

Abs
2D1
64D1
23C2
2D1
64D1
23C2
2D1
64D1
23C2
sigma--
CRP-alpha--
CRP-CRP--
--alpha
DNA
gal P1
lac P1
No DNA

Figure 18. Western blotting experiment showing the photocrosslinking result of BPM-C178S-K52N-G56C on *gal* P^+ before and after the clearance of the promoter. Condition for the assay is described in "Materials and Methods". Nucleotide substrate mix (ATP, CTP, UTP and $O^{met}GTP$) and/or heparin were added to the CRP-*gal* P^+-RNA polymerase ternary complex mixture before UV irradiation.

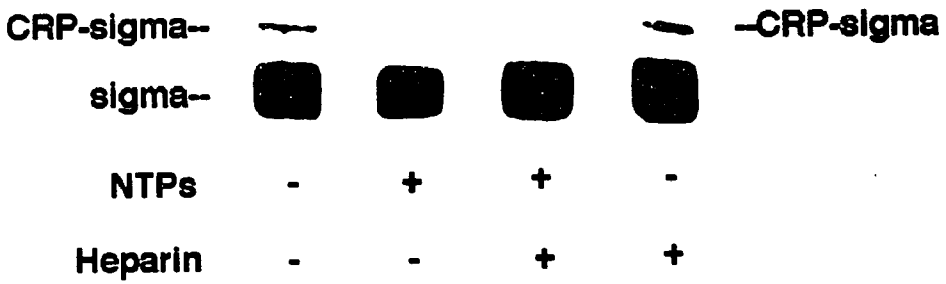

CRP-sigma--
sigma--
NTPs - + + -
Heparin - - + +
--CRP-sigma

Figure 19. Western blotting experiment showing the photocrosslinking result of BPM-C178S-K52N-G56C and BPM-C178S-K52N-G56C-T158A on *gal* P^+ and *lac* P^+ promoters. Condition for the assay is described in "Materials and Methods".

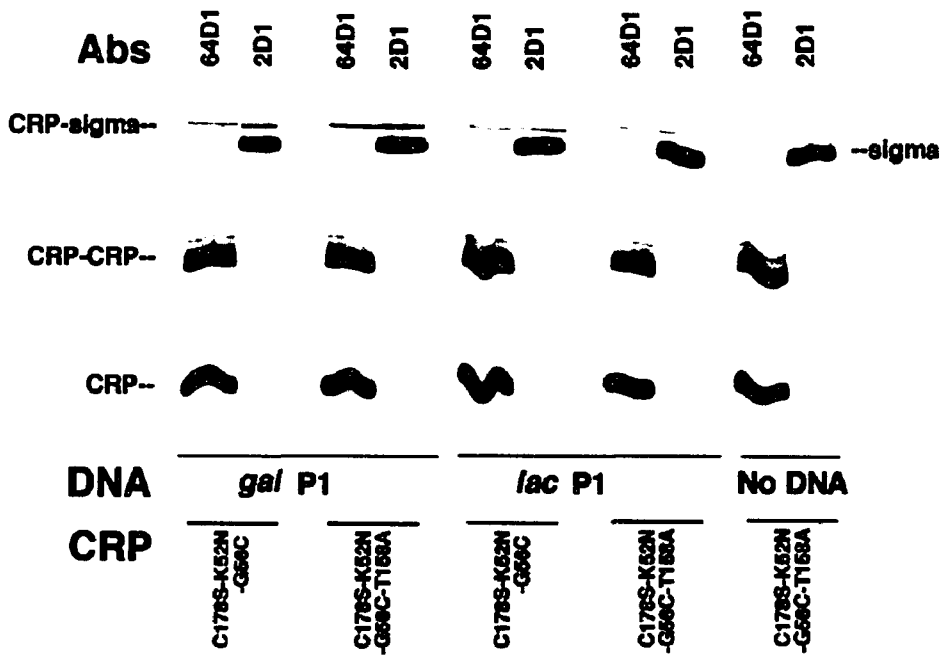

Abs
64D1
2D1
64D1
2D1
64D1
2D1
64D1
2D1
64D1
2D1
CRP-sigma--
--sigma
CRP-CRP--
CRP--
DNA
gal P1
lac P1
No DNA
CRP
C178S-K52N -G56C
C178S-K52N -G56C-T158A
C178S-K52N -G56C
C178S-K52N -G56C-T158A
C178S-K52N -G56C-T158A

Figure 20 (A). *In vivo* expression of the *lac* operon by the CRP N-terminal glycine substitution mutants. ß-galactosidase assays showing the stimulating activity of N-terminal glycine substitution mutant and wild-type CRPs which were expressed from pXZCRP variants in *E. coli* XE64.2 cells.

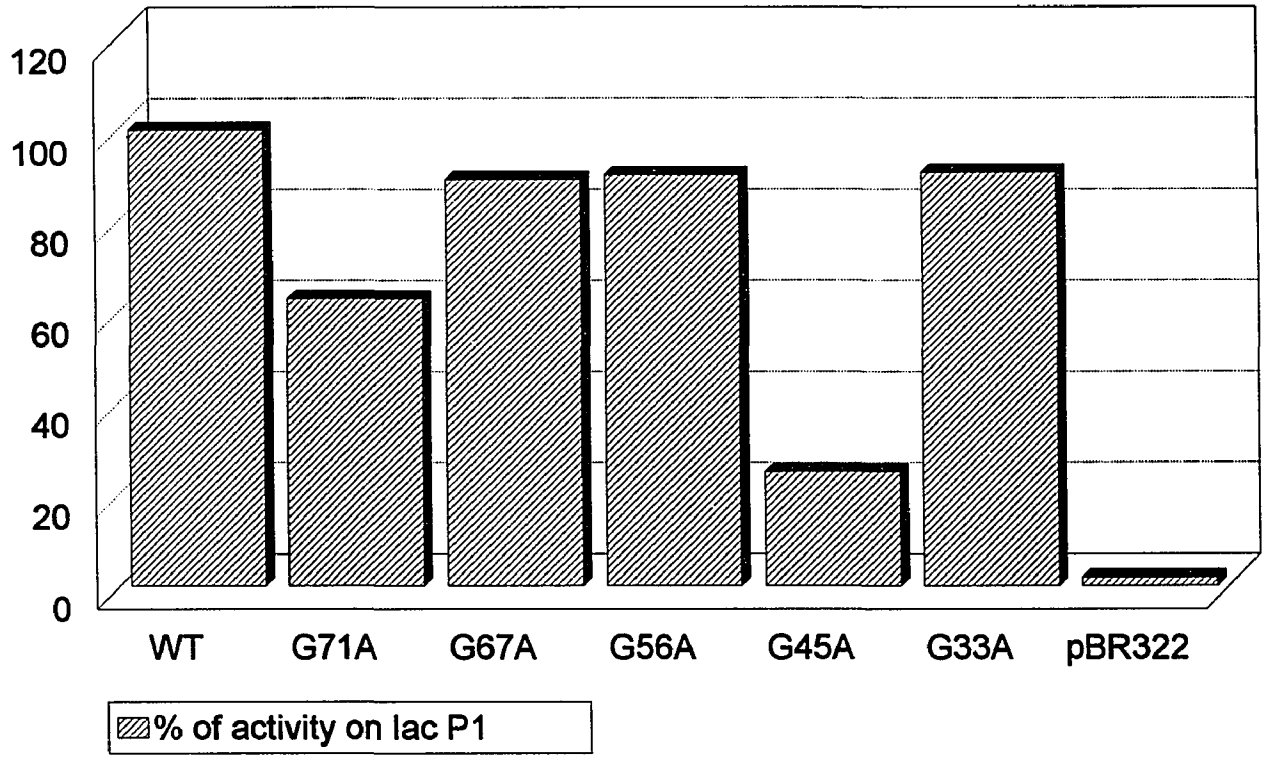

120
100
80
60
40
20
0
WT
G71A
G67A
G56A
G45A
G33A
pBR322
% of activity on lac P1

Figure 20 (B). *In vivo* expression of the *gal* operon by the CRP N-terminal glycine substitution mutants. ß-galactosidase assays showing the stimulating activity of N-terminal glycine substitution mutant and wild-type CRPs which were expressed from pXZCRP variants in *E. coli* M182 cells (Δ*crp39,Δlac*) with the co-transformation of *gal* P1 promoter in pRW2.

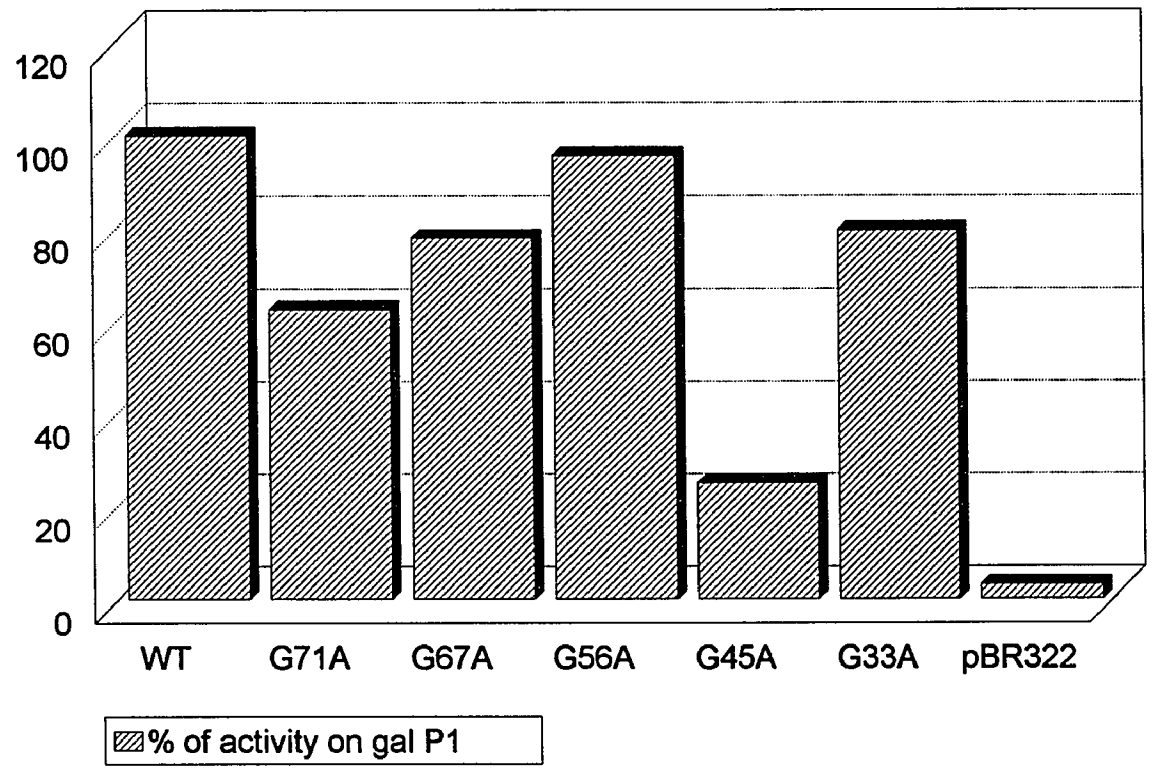

120
100
80
60
40
20
0
WT
G71A
G67A
G56A
G45A
G33A
pBR322
% of activity on gal P1

Figure 21. ß-galactosidase assays showing the repressing ability of N-terminal glycine substitution mutant and wild-type CRPs. Wild-type and mutant CRPs were expressed from pXZCRP variants in *E. coli* KC1071 cells.

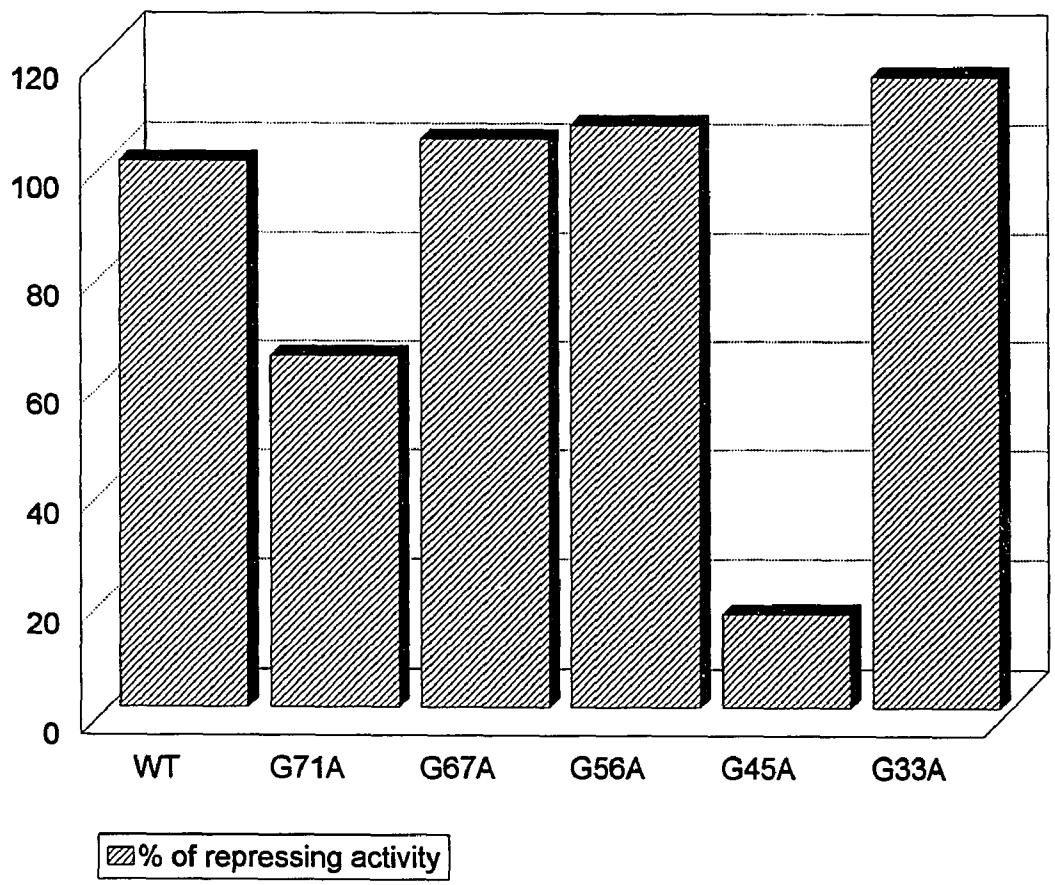
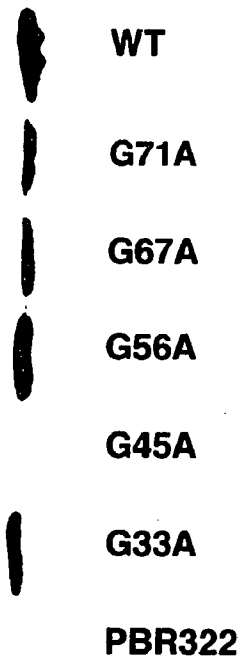

120
100
80
60
40
20
0
WT
G71A
G67A
G56A
G45A
G33A
% of repressing activity

Figure 22. Western blotting experiment showing the amount of CRP N-terminal glycine substitution mutants in the cell crude extracts. The cell crude extracts were prepared by boiling cells in 1X SDS sample buffer (50 mM Tris-HCl, pH 6.8, 10% (v/v) β-mercaptoethanol, 10% (v/v) glycerol, 1% SDS) for 5 minutes. Cell debris was spun down and the supernatants were resolved on a 15% SDS-PAGE followed by Western blotting.


WT
G71A
G67A
G56A
G45A
G33A
PBR322

BIBLIOGRAPHY

Aiba, H. (1982) Autoregulation of the *E. coli crp* gene: CRP is a transcriptional repressor for its own gene. *Cell* **32**, 141-149

Aiba, H., Fujimoto, S., and Ozaki, N. (1982) Molecular cloning and nucleotide sequencing of the gene for the *E. coli* cAMP receptor protein. *Nucleic Acids Res.* **10**, 1345-1361

Aiba, H., Nakamura, T., Mitani, H., and Mori, H. (1985) Mutations that alter the allosteric nature of cAMP receptor protein of *Escherichia coli. EMBO J.* **4,** 3329-3332

Anderson, W. B., Schneider, A. B., Emmer, M., Perlman, R. L., and Pastan, I. (1971) Purification and properities of the cyclic adenosine 3', 5'-monophosphate receptor protein which mediates cyclic adenosine 3', 5'-monophosphate-dependent gene transcription in *E. coli. J. Biol. Chem.* **246,** 5929-5937

Angulo, J., and Krakow, J. S. (1986) Sensitization of the *E. coli* cyclic AMP receptor protein to trypsin cleavage by polydeoxynucleotides and polyribonucleotides. *J. Biol. Chem.* **261**, 11315-11319

Attey, A., Belyaeva, T., Savery, N., Hoggett, J., Fujita, N., Ishihama, A. and Busby, S. (1994) Interactions between the cyclic AMP receptor protein and the alpha subunit of RNA polymerase at *Escherichia coli* galactose operon *P1* promoter. *Nucleic Acids Res.* **22**, 4375-4380

Batut, J., Daveran-Mingot, M.-L., David, M., Jacobs, J., Garnerone, A. M. and Kahn, D. (1989) *fixK*, a gene homologous with *fnr* and *crp* from *Escherichia coli,* regulates nitrogen fixation genes both positively and negatively in *Rhizobium meliloti. J. EMBO* **8**, 1279-1286

Bell, A. and Busby, S. (1994) Location and orientation of an activating region in the *Escherichia coli* transcription factor, FNR. *Molecular Microbiology* **11**, 383-390

Bell, A., Gaston, K., Williams, R., Chapman, K., Kolb, A., Buc, H., Minchin, S., Williams, J., and Busby, S. (1990) Mutations that alter the ability of the *Escherichia coli* cyclic AMP receptor protein to activate transcription. *Nucl. Acids Res.* **18**, 7243-7250

Botsford, J. L., and Harman, J. G. (1992) Cyclic AMP in prokaryotes. *Microbiol.*

Revs. **56**,100-122

Burgess, R. R., and Jendrisak, J. J. (1975) A procedure for the rapid large-scale purification of *E. coli* DNA dependent RNA polymerase involving polymin P precipitation and DNA-cellulose chromatography. *Biochemistry* **14**, 4634-4638

Busby, S., and Ebright, R. H. (1994) Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell*, **79**, 743-746

Busby, S., West, D., Lawes, M., Webster, C., Ishihama, A., and Kolb, A. (1994) Transcription activation by the *Escherichia coli* cyclic AMP receptor protein. Receptors bound in tandem at promoters can interact synergistically. *J. Mol. Biol.* **241**, 341-352

Chapon, C. (1982) Role of the catabolite activator protein in the maltose regulation of *E. coli. J. Bacteriol.* **150,** 722-729

Chen, Y., Ebright, Y. W., and Ebright, R. H. (1994) Identification of the target of a transcription activator protein by protein-protein photocrosslinking. *Science* **265**, 90-92

Cossart, P., and Gicquel-Sanzey, B. (1982) Cloning and sequence of the CRP gene of *Escherichia coli. Nucl. Acids Res.* **10**, 1363-1378

de Crombrugghe, B., Busby, S., and Buc, H. (1984) Cyclic AMP receptor protein: role in transcription activation. *Science* **224**, 831-838

Dombroski, A. J., Walter, W. A., Record, M. T., Siegele, D. A., and Gross, C. A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor σ^{70} exhibit specificity of binding to promoter DNA. *Cell* **70**, 501-512

Ebright, R. H.(1993) Transcription activation at class I CAP-dependent promoters. *Molec. Microbiol.* **8**, 797-802

Ebright, R. H., Ebright, Y. W., and Gunaskera, A. (1989) Consensus DNA site for the *Escherichia coli* catabolite gene activator protein (CAP): CAP exhibits a 450-fold higher affinity for the consensus DNA site than for the *E. coli lac* DNA site. *Nucl. Acids Res.* **17**, 10295-10305

Ebright, R. H., LeGrice, S. F. J., Miller, J. P., and Krakow, J. S. (1985) Analogs of cyclic AMP that elicit the biochemically defined conformational change in catabolite gene activator protein (CAP) but do not stimulate binding to DNA. *J. Mol. Biol.* **182**, 91-107

Eilen, E., and Krakow, J. S. (1977) Cyclic AMP-mediated intersubunit disulfide crosslinking of the cyclic AMP receptor protein of *Escherichia coli*. *J. Mol. Biol.* **114**, 47-60

Eilen, E., Pampeno, C., and Krakow, J. S. (1978) Production and properties of alpha core derived from adenosine monophosphate receptor protein of *E. coli*. *Biochemistry* **17**, 2469-2473

Eschenlauer, A. C., and Reznikoff, W. S. (1991) *Escherichia coli* catabolite gene activator mutants defective in positive control of *lac* operon transcription. *J. Bacteriol.* **173**, 5024-5029

Gaston, K., Bell, A., Buc, H., and Busby, S., (1990) Stringent spacing requirements for transcription activation by CRP. *Cell* **62**, 733-743

Gronenborn, A., Sandulache, R., Gartner, S., and Clore, G. M. (1988) Proton nuclear magnetic resonance studies on cyclic nucleotide binding to the *E. coli* adenosine cyclic 3', 5'-phosphate receptor protein. *J. Biochem.* **253**, 801-807

Herbert, M., Kolb, A. and Busby, S. (1986) Overlapping promoters and their control in *Escherichia Coli:* the *gal* case. *Proc. Natl. Acad. Sci. U.S.A.* **83,** 2807-2811

Heyduk, T., Lee, J. C., Ebright, Y. W., Blatter, E. E., Zhou, Y., and Ebright, R. H. (1993) CAP interacts with RNA polymerase in solution in the absence of promoter DNA. *Nature* **364**, 548-549

Igarashi, K. and Ishihama, A. (1991) Bipartite functional map of the *E. coli* RNA polymerase α subunit: involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell* **65,** 1015-1022

Irwin, N., and Ptashne, M., (1987) Mutants of the catabolite activator protein of *Escherichia coli* that are specifically deficient in the gene-activation function. *Proc. Natl. Acad. Sci. USA* **84**, 8315-8319

Ishihama, A., (1993) Protein-protein communication within the transcription apparatus. *J. Bacteriol.*, **175**, 2483-2489

Ito, W., Ishiguro, H. and Kurosawa, Y., (1991) A general method for introducing a series of mutations into cloned DNA using the polymerase chain reaction. *Gene* **102**, 67-70

Kolb, A., Igarashi, K., Ishihama, A., Lavigne, M., Buckle, M., and Buc, H. (1993) *E. coli* RNA polymerase, deleted in the C-terminal part of its α-subunit, interacts

differently with the cAMP-CRP complex at the *lac*P1 and at the *gal*P1 promoter. *Nucl. Acids Res.* **21,** 319-326

Krakow, J. S. and Pastan, I. (1973) Cyclic adenosine monophosphate receptor: loss of cAMP-dependent DNA binding activity after proteolysis in the presence of cyclic adenosine monophosphate. *Proc. Natl. Acad. Sci. USA* **70,** 2529-2533

Kumar, A., Grimes, B., Fujita, N., Makino, K., Malloch, R. A., Hayward, R. S., Ishihama, A., (1994) Role of the Sigma^{70} subunit of *Escherichia coli* RNA polymerase in transcription activation. *J. Mol. Biol.* **235**, 405-413

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685

Landick, R., Colwell, A., and Stewart, J. (1990) Insertional mutagenesis of a plasmid-borne *Escherichia coli rpoB* gene reveals alterations that inhibit β-subunit assembly into RNA polymerase. *J. Bacteriol.* **172**, 2844-2854

Lee, N. L., Gielow, W. O. and Wallace, R. G. (1981) Mechanism of *ara*C autoregulation and the domains of two overlapping promoters P_c and P_{BAD} in the L-arabinose regulatory region of *E. coli. Proc. Natl. Acad. Sci. USA* **78**, 752-756

Levine, B. J., Orphanos, P. D., Fischmann, B. S., and Beychok, S. (1980) Physicochemical properties and interactions of *E. coli* RNA polymerase holoenzyme, core enzyme, subunits and subassembly alpha$_2$beta. *Biochemistry* **19,** 4808-4814

Li, X.-M. and Krakow, J. S. (1985) Characterization of nine monoclonal antibodies aganist the *E. coli* cyclic AMP receptor protein. *J. Biol. Chem.* **260**, 4378-4383

Makino, K., Amemura, M., Kim, S. K., Nakata, A., and Shinagawa, H. (1993) Role of the σ^{70} subunit of RNA polymerase in transcriptional activation by activator protein PhoB in *Escherichia coli. Genes Dev.* **7,** 149-160

Malan, T. P., Kolb. A., Buc, H., and McClure, W. R. (1984) Mechanism of CRP-cAMP activation of *lac* operon transcription initiation: activation of P1 promoter. *J. Mol. Biol.* **180**, 881-909

Malan, T. P. and McClure, W. R. (1984) Dual promoter control of the *E. coli* lactose operon. *Cell* **39,** 173-180

Matthews, K. S. (1992) DNA looping. *Microbiological Reviews* **56**, 123-136

Maxam, A., and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499-560

McClure, W. R. (1985) Mechanism and control of the transcription initiation in prokaryotes. *Ann. Rev. Biochem.* **1985,** 171-204

Mckay, D. And Steitz, T. (1981) Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* (London) **290**, 744-749

McKay, D. B., Weber, I. T., and Steitz, T. A. (1982). Structure of catabolite gene activator protein at 2.9 Å. *J. Biol. Chem.* **257,** 9518-9524

Menendez, M., Kolb, A. And Buc, H. (1987) A new target for CRP action at the *malT* promoter. *EMBO J.* **6,** 4227-4234

Miller, J. (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Moore, J., Kantorow, M., Vanderzwaag, D., and McKenney, K. (1992) *Escherichia coli* cyclic AMP receptor protein mutants provide evidence for ligand contacts important in activation. *J. Bacteriol.* **174**, 8030-8035

Musso, R. E., DiLauro, R., Adhya, S., and de Crombrugghe, B. (1977) Dual control for transcription of the *galactose* operon by cyclic AMP and its receptor protein at two interspersed promoters. *Cell* **12**, 847-854

Pendegrast, P. S., Chen, Y., Ebright, Y. W., and Ebright, R. H. (1992) Determination of the orientation of a DNA binding motif in a protein-DNA complex by photocrosslinking. *Proc. Natl. Acad. Sci. USA* **89**, 10287-10291

Pinkey, M. and Hoggett, J. G. (1988) Binding of the cAMP receptor protein of *Escherichia coli* to RNA polymerase. *J. Biochem.* **250,** 897-902

Ren, Y. L., Garges, S., Adhya, S. and Krakow, J. S. (1988) Cooperative DNA binding of heterologous proteins: Evidence for contact between the cyclic AMP receptor and RNA polymerase. *Proc. Natl. Acad. Sci. USA* **85**, 4138-4142

Ren, Y. L., Garges, S., Adhya, S., and Krakow, J. S. (1990) Characterization of the binding of cAMP and cGMP to the CRP*598 mutant of the *E. coli* cAMP receptor protein. *Nucl. Acids Res.* **18**, 5127-5132

Reznikoff, W. S. (1992) Catabolite gene activator protein activation of *lac* transcription. *J. Bacteriol.* **174**, 655-658

Ross, W., Gosink, K. K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R. L. (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* **262,** 1407-1413

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

Sharif, K. A., Fujita, N., Jin, R., Igarashi, K., Ishihama, A., and Krakow, J. S. (1994) Epitope mapping and functional characterization of monoclonal antibodies specific for the α subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **269,** 23655-23660

Spiro, S. and Guest, J. R. (1987) Regulation and over-expression of the *fnr* gene of *Escherichia coli. J.Gen. Microbiol.* **133,** 3279-3288

Spiro, S. and Guest, J. R. (1987) Activation of the*lac* operon of *Escherichia coli* by a mutant FNR protein. *Mol. Microbiol.* **1**, 53-58

Spiro, S. and Guest, J. R. (1990) FNR and its role in oxygen-regulated gene expression in *Escherichia coli. FEMS Microbiol. Rev.* **75,** 399-428

Straney, D. C., Straney, S. M., and Crothers, D. M. (1989) Synergy between *Escherichia coli* CAP protein and RNA polymerase in the *lac* promoter open complex. *J. Mol. Biol.* **206**, 41-57

Starzyk, R., Burbaum, J. J., and Schimmel, P. (1989) Insertion of new sequences into the catalytic domain of an enzyme. *Biochemistry* **28**, 8479-8484

Strickland, M. S., Thompson, N. E., and Burgess, R. R. (1988) Structure and function of the σ-70 subunit of *Escherichia coli* RNA polymerase . Monoclonal antibodies: localization of epitopes by peptide mapping and effects of transcription. *Biochemistry* **27**, 5755-5762

Taniguchi, T., O'neill, M., and De Crombrugghe, B. (1979) Interaction site of *Escherichia coli* cyclic AMP receptor protein on DNA of galactose operon promoters. *Proc. Natl. Acda. Sci. USA* **76**, 5090-5094

Tao, T., Lamkin, M., and Scheiner, C. J. (1985) The conformation of the C-terminal region of actin: a site specific photocrosslinking study using benzophenone-4-maleimide. *Arch. Biochem. Biophys.* **240**, 627-634

Ushida, C., and Aiba, H. (1990) Helical phase dependent action of CRP: effect of the distance between the CRP site and the -35 region on promoter activity. *Nucl.*

Acids Res. **18**, 6325-6330

Weber, I. T., and Steitz, T. A. (1987) Structure of complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. *J. Mol. Biol.* **198**, 311-326

West, D., Williams, R., Rhodius, V., Bell, A., Sharma, N., Zou, C., Fujita, N., Ishihama, A., and Busby, S. (1993) Interactions between the *Escherichia coli* cyclic AMP receptor protein and RNA polymerase at Class II promoters. *Molecular Microbiology* **10**, 789-797

Williams, R., Bell, A., Sims, G., and Busby, S. (1991) The role of two surface exposed loops in transcription activation by the *Escherichia coli* CRP and FNR proteins. *Nucl. Acids Res.* **19**, 6705-6712

Wu, H.-M., and Crothers, D. M. (1984) The locus of sequence-directed and protein induced DNA bending. *Nature* (London) **308,** 509-513

Yang, Z., and Krakow, J. S. (1991) Characterization of the CRP^{CY} core formed after treatment with carboxypeptidase Y. *Nucl. Acids Res.* **19**, 4253-4257

Zhang, X., and Ebright, R. (1990) Identification of a contact between arginine-180 of the catabolite gene activator protein (CAP) and base pair 5 of the DNA site in the CAP-DNA complex. *Proc. Natl. Acad. Sci. USA* **87**, 4717-4721

Zhang, X., Zhou, Y., Ebright, Y., and Ebright, R. (1992) Catabolite gene activator protein (CAP) is not an "acidic activating region" transcription activator protein. *J. Biol. Chem.* **267**, 8136-8139

Zhou, Y., Zhang, X., and Ebright, R. H. (1991) Random mutagenesis of gene-sized DNA molecules by use of PCR with Taq DNA polymerase. *Nucl. Acids Res.* **19**, 6052

Zhou, Y., Zhang, X., and Ebright, R. H. (1993) Idenfication of the activating region of catabolite gene activator protein (CAP): Isolation and characterization of CAP specifically defective in transcription activation. *Proc. Natl. Acad. Sci. USA* **90**, 6081-6085