The EMBO Journal Vol.1 No.5 pp.591 – 595, 1982

Homologies between different procaryotic DNA-binding regulatory proteins and between their sites of action

B. Gicquel-Sanzey* and P. Cossart[1]

Unité de Biochimie des Régulations Cellulaires, et [1]Unité de Biochimie Cellulaire, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

Communicated by J. Janin
Received on 17 May 1982

Comparison of the amino acid sequences of 13 procaryotic regulatory proteins, including the products of genes *crp* (catabolite activator protein; CAP), *lac*I, *gal*R, *lex*A, *lys*R, *ara*C, *trp*R, and *tnp*R of *Escherichia coli*, of genes *c*I, *c*II and *cro* of phage λ, *cro* of phage 434, and *c*2 of phage P22, has revealed two regions of homology. The sites of action of these proteins also share common features in their DNA sequence. Taking into account the models proposed for the λ repressors, *cro* and *c*I, and for CAP, a general type of DNA-protein interaction is suggested.

Key words: DNA-binding regulatory proteins/DNA sites/ homologies

Introduction

In bacteria, the regulation of transcription initiation involves the interaction of repressor or activator proteins with specific sites on the DNA. Several of these sites have been sequenced, including the regulatory regions of the *lac* operon (Dickson *et al.*, 1975), the *gal* operon (Musso *et al.*, 1974, 1977b; Sklar, 1977), the *ara* operon (Greenfield *et al.*, 1978; Smith and Schleif, 1978), *lex*A (Miki *et al.*, 1982; Horii *et al.*, 1981, Markham *et al.*, 1981), *rec*A (Horii *et al.*, 1980), and *uvr*B (Van den Berg *et al.*, 1981; Sancar *et al.*, 1982). The sequences of the bacteriocin promoters of CloDF13 and ColE1 (Van den Elzen, 1982), the regulatory region of the transposon tn3 (Chou *et al.*, 1979; Heffron *et al.*, 1979), that of the *trp* operon (Bennett *et al.*, 1978), *aro*H (Zurawski, 1981), *trp*R (Gunsalus and Yanofsky, 1980; Singleton *et al.*, 1980), and *lys*R (P. Stragier, personal communication) are also known. In phage λ, some regulatory regions have also been sequenced such as the left and right operators (Maniatis *et al.*, 1975a; 1975b) and promoters P_i (Abraham *et al.*, 1980) and P_E (Schwarz *et al.*, 1978). Nucleotides presumed to make contacts with the regulatory proteins have been identified in a few instances: the *lac* operator (Goeddel *et al.*, 1978; Schmitz and Galas, 1979), the *lac* promoter (Majors, 1975; Simpson, 1980), the *ara* regulatory region (Ogden *et al.*, 1980 Lee *et al.*, 1981), the *lex*A and *rec*A regulatory regions (Brent and Ptashne, 1981; Little *et al.*, 1981), the *gal* regulatory regions (Taniguchi *et al.*, 1979), the *trp, aro*H, and *trp*R regulatory regions (Oppenheim *et al.*, 1980; Gunsalus and Yanofsky, 1980), the pBR-P4 promoter of pBR322 (Queen and Rosenberg, 1981), and the *lys*R regulatory region (P. Stragier, personal communication), in the left and right operators (Ptashne *et al.*, 1976; Humayun *et al.*, 1977), and the two promoters P_i and P_E (Ho and Rosenberg, 1982) of phage λ.

*To whom reprint requests should be sent.

However, much less is known about the DNA binding sites on the regulatory proteins. The only relevant data have been the identification of large domains responsible for DNA binding, in one of the λ repressors, the product of gene *c*I (Pabo *et al.*, 1979), the *lac* repressor (reviewed by Beyreuther, 1978; Miller, 1978) and in the catabolite activator protein (CAP) (Krakow and Pastan, 1973; Aiba and Krakow, 1981) of *Escherichia coli*. To study the DNA-CAP interaction, we (Cossart and Gicquel-Sanzey, 1982) and others (Aiba *et al.*, 1982) have recently sequenced the *crp* gene of *E. coli*. Here we compare: (i) the primary structure of CAP with that of other DNA-binding regulatory proteins; and (ii) the sites recognized by CAP and the other regulatory proteins. The existence of homologies at these two levels suggests a common mechanism for DNA-protein interaction.

Results

Comparison of the primary structures of several DNA-binding regulatory proteins

Figure 1 shows the homologies in the amino acid sequences of 13 procaryotic DNA binding regulatory proteins: repressors such as *trp*R, *lac*I, *gal*R, *ara*C, *lex*A, *tnp*R, and *lys*R of *E. coli*, and *cro* of phage λ, *cro* of phage 434, and *c*2 of phage P22, or activators such as CAP of *E. coli*, *c*I and *c*II of λ (for references, see Figure 1). (The proteins encoded by genes *trp*R, *lac*I, *gal*R, *ara*C, *lex*A, *tnp*R, *lys*R of *E. coli*, *c*I, *c*II, and *cro* and λ, *cro* of 434, *c*2 of P22 are designated by the name of their structural gene. The *crp* gene product is called CAP.)

A first set of homologies, that we have already pointed out (Cossart and Gicquel-Sanzey, 1982) and designate here Region 1, reveals in the 13 proteins, the presence of a rather conserved sequence ^{1}V/L/A.^{2}G.^{3}V.^{4}S.^{5}Q.^{6}S.^{7}T.^{8}V/I.^{9}S/G.^{10}R.–.^{12}V.^{13}N. The number of proteins containing these conserved residues in the exact position is nine for ^{1}V/L/A, nine for ^{2}G, six for ^{3}V, five for ^{4}S, four for ^{5}Q, three for ^{6}S, eight for ^{7}T, 13 for ^{8}V/I, eight for ^{9}S/G, nine for ^{10}R, three for ^{12}V, four for ^{13}N. The most homologous segment is the ^{7}T.^{8}V/I.^{9}S/G.^{10}R tetrapeptide totally conserved in CAP, *lac*I, *gal*R, *ara*C, *lys*R and, to a lesser extent, in *trp*R, *lex*A, and *tnp*R, as already mentioned (Cossart and Gicquel-Sanzey, 1982). The homologies between the amino-terminal sequence of *lac*I and those of *cro* of λ and *cro* of 434, *c*I and *c*II of λ and *c*2 of P22 have also been recently reported (Matthews *et al.*, 1982).

We have identified a second region of homology that we call Region 2 (Figure 1b). This region appears in CAP, *trp*R, *lac*I, *gal*R, *lex*A of *E. coli*, *c*I of λ and *c*2 of P22. Homologies between *lac*I, *gal*R of *E. coli*, *c*I of λ and *c*2 of P22 in this region have already been reported (Wilcken-Bergman and Müller-Hill, 1982). To maximize the homology between these and the other DNA-binding proteins in this region, we have not aligned the sequences in exactly the same way as these authors (they align ^{50}N of *c*2 with ^{61}N of λ*c*I, ^{48}N of *gal*R, ^{50}N of *lac*I). In this Region 2, the peptide of *lac*I (residues 44–64) which corresponds to peptide 53–73 of *c*I

© IRL Press Limited, Oxford, England. 0261-4189/82/0105-0591$2.00/0.

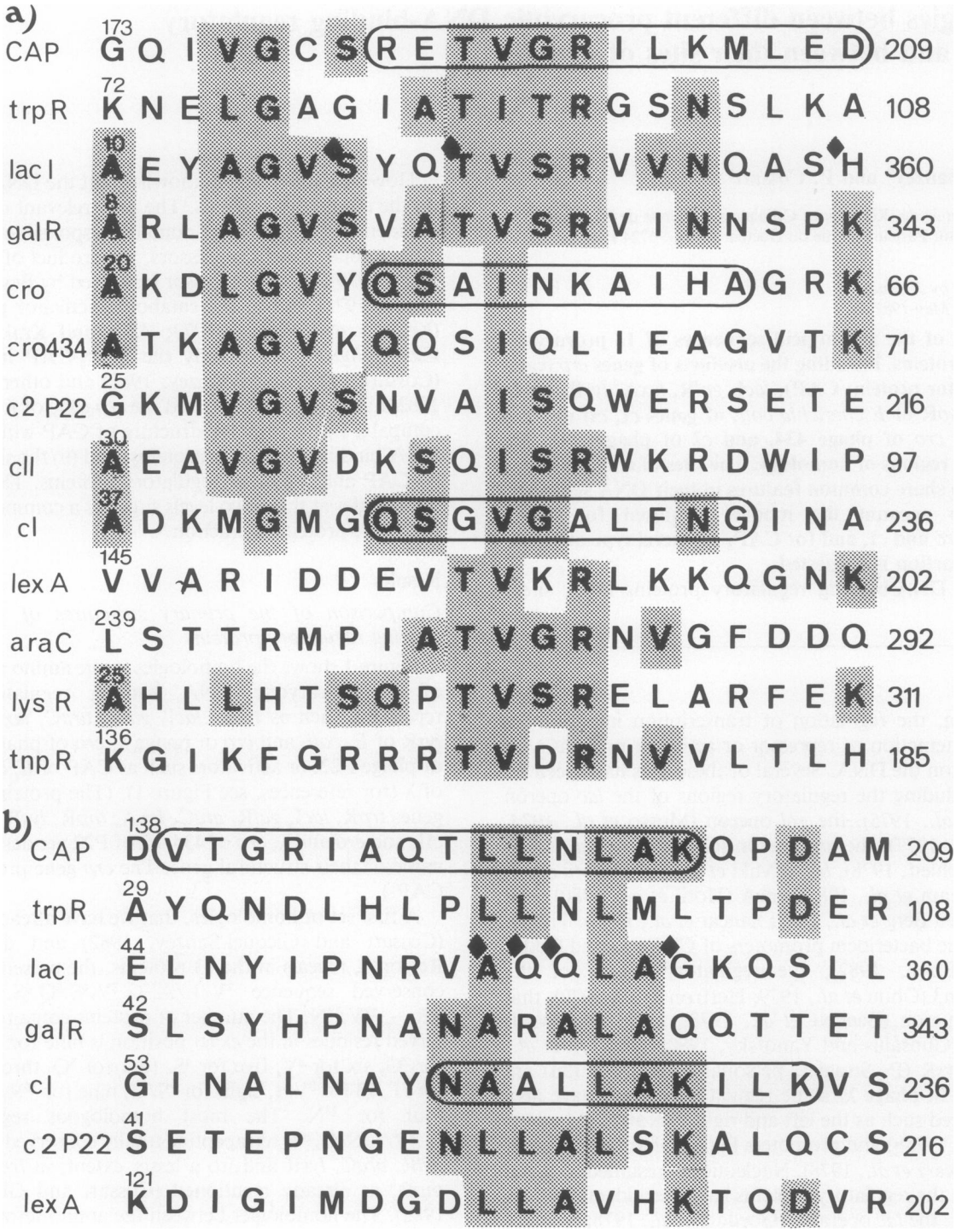

Fig. 1. Homology between the amino acid sequences of 13 DNA-binding regulatory proteins. CAP (Cossart and Gicquel-Sanzey, 1982; Aiba *et al.*, 1982), *trp*R (Gunsalus and Yanofsky, 1980), *lac*I (Farabaugh, 1978), *gal*R (Wilcken-Bergman and Müller-Hill, 1982), *ara*C (Miyada *et al.*, 1980; Wallace *et al.*, 1980; Stoner and Schleif, 1982), *lex*A (Horii *et al.*, 1981; Markham *et al.*, 1981), *tnp*R (Chou *et al.*, 1979; Heffron *et al.*, 1979), *lys*R (P. Stragier, personal communication) of *E. coli*, cI (Sauer and Anderegg, 1978), cII (Schwarz *et al.*, 1978), *cro* (Schwarz *et al.*, 1978) of λ, *cro* (Grosscheld and Schwarz, 1979) of 434 and *c2* of P22 (Sauer *et al.*, 1981) (**a**) Region 1, (**b**) Region 2. The number indicated at the right of the sequences is the total number of codons in the gene coding for the regulatory protein. For each protein, only part of the sequence is shown. The number above the first amino acid indicates the position of this residue in the sequence. The sequences are aligned in such a way that the conserved (present at least three times) residues (shaded), fall on the same vertical line. Residues known to belong to α helices in the tertiary structure are boxed. ♦: corresponds to i^{-D} mutations (see text). The one letter amino acid abbreviations are given in Dayhoff (1978)

and peptide 41 – 64 of *c2* is not the peptide starting at position 26, reported by Matthews *et al.* (1982). Our alignment allows a better correspondence between the seven proteins listed in Figure 2b.

Both regions of homology are located in the known DNA-binding domains of *lac*I (reviewed by Beyreuther, 1978; Miller, 1978), cI (Pabo *et al.*, 1979) and CAP (Krakow and Pastan, 1973; Aiba and Krakow, 1981). Moreover, i^{-D}

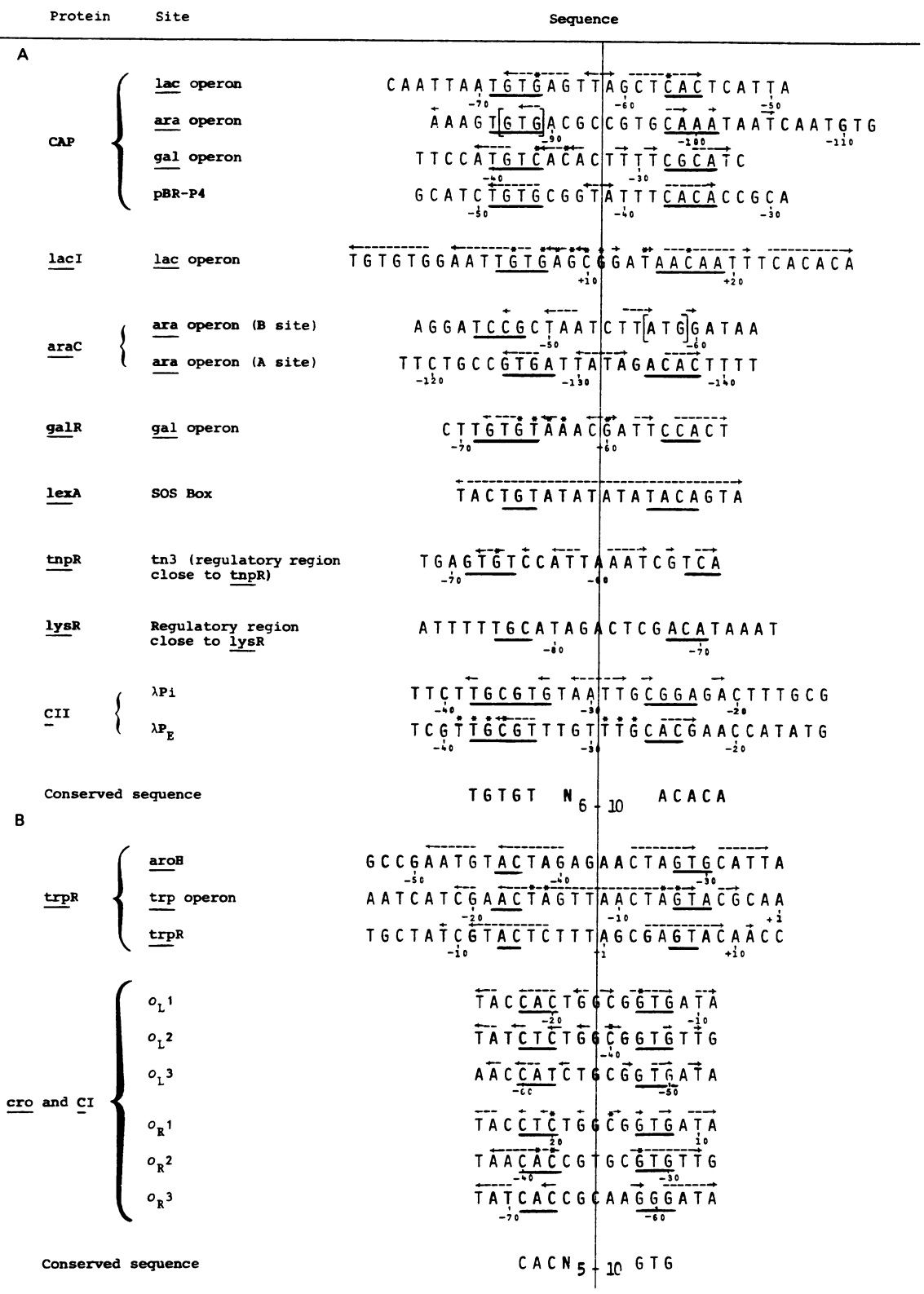

Fig. 2. Comparison of the sites recognized by the 13 regulatory proteins. The different regulatory proteins are indicated in the first column and their sites of action in the second column. The sequences of these sites are shown in the last column. They are aligned in such a way that all the centers of symmetry are on the same vertical line. Symmetrical regions are indicated by a dashed line above the sequence. Nucleotides corresponding to the conserved sequences are underlined. Asterisks and sequences between brackets indicate that the corresponding position was shown to be the site of mutations preventing the binding (or the action) of the regulatory protein: asterisks correspond to ponctual mutations in *lac* (Gilbert *et al.*, 1975; Dickson *et al.*, 1977), in *gal* (Di Lauro *et al.*, 1979; Busby *et al.*, 1982), *tn3* (Chou *et al.*, 1979), λP_E (Ho and Rosenberg, 1982), *trp* (Bennett and Yanofsky, 1978), or λ operators (reviewed in Ptashne, 1978) and brackets correspond to deletions in *ara* (Miyada *et al.*, 1982). The numbers below the nucleotide sequence indicate the coordinates with respect to the *in vitro* transcription initiation site except for *lys*R and *tnp*R binding sites and λ operators; in these cases the coordinates are indicated with respect to the translational start codon. (**A**) corresponds to the sites containing the conserved sequence TGTGT N_{6-10} ACACA (N, being any nucleotide); (**B**) corresponds to the sites containing the conserved sequences CAC N_{5-10} GTG.

mutations which impair the binding of the *lac* repressor to the operator are clustered in these two regions (reviewed by Müller-Hill, 1975). Some of these mutations have been sequenced and correspond to a change of ^{16}Serine (S) to Proline, of ^{19}Threonine (T) to Alanine, ^{29}Histidine (H) to Tyrosine (Region 1) and of ^{53}Alanine (A) to Valine or Threonine, of ^{54}Glutamine (Q) to Tyrosine or Lysine, of ^{55}Glutamine (Q) to Tyrosine or Lysine, and of ^{58}Glycine (G) to Aspartic acid or Serine (Region 2). In addition, the tertiary structure of *cro* (Anderson *et al.*, 1981), CAP (McKay *et al.*, 1981), and the head piece of *c*I (Pabo and Lewis, 1982) have been determined recently and models for DNA-protein interaction have been proposed. In each case, a pair of 2-fold related α-helices of the protein comes in contact with two successive major grooves of the DNA: these are helix α3 of *cro* (residues 27–36), helix F of CAP (residues 181–194, T. Steitz, personal communication), and helix 3 of *c*I (residues 44–53), these three helices (boxed in Figure 1a) correspond to the strongest zone of homology in Region 1.

These findings suggest that for all proteins listed in Figure 1, the homologous sequences of Region 1 (and to some extent those of Region 2) could play an important role in the interaction with DNA. It is interesting to note that Region 2 corresponds to helix D of CAP and helix 4 of *c*I, both of which belong to the DNA binding domain of the corresponding protein.

Since we find homologies between the DNA-binding domains of the different regulatory proteins, we expect to find also some analogies between their sites of action. This appears to be the case.

Comparison of the specific DNA-binding sites of several regulatory proteins

Since all the regulatory proteins so far studied have been reported to be dimers such as CAP (Eilen and Krakow, 1977), *ara*C (Steffen and Schleif, 1977), *c*I (Pirotta *et al.*, 1970), *cro* of λ (Takeda *et al.*. 1977), or tetramers such as *lac*I (Riggs and Bourgeois, 1968), *trp*R (Gunsalus and Yanofsky, 1980), and *c*II (Ho and Rosenberg, 1982), one could expect to find symmetries in their sites of action. The binding sites of *lac*I, *ara*C, *lex*A, CAP, *trp*R, *gal*R, *tnp*R, *c*I, and *c*II and *cro* of λ, have been localized by DNA binding experiments, *in vitro* protection against DNase I, restriction enzymes, or alkylating agents (for references, see first paragraph) and/or location of mutations which impair the binding of these regions with the corresponding proteins (for references, see Figure 2). All these sites are presented in Figure 2. As reported for all of them (except P_i, P_E, and the *lys*R site of action), these sites (~25 bp long) have a 2-fold rotational symmetry. From the comparison of these sites, two consensus sequences can be drawn which are:

$$\begin{array}{l} \text{TGTGT } N_{6-10} \text{ ACACA (Figure 2a) or} \\ \text{ACACA} \qquad\quad \text{TGTGT} \end{array}$$

$$\begin{array}{l} \text{CAC } N_{5-10} \text{ TGT (Figure 2b)} \\ \text{GTG} \qquad\quad \text{ACA} \end{array}$$

with an axis of symmetry in the center of the region. Most of the mutations preventing the binding of the corresponding regulatory proteins are located in the conserved sequences. Some of the regulatory proteins are known to act on several sites which have been reported to share homologies. It is the case for CAP ("CAP sites") (Queen and Rosenberg, 1981; O'Neill *et al.*, 1981), *lex*A ("SOS boxes") (Little *et al.*, 1981), *trpR* (Gunsalus and Yanofsky, 1982) and *c*II (Ho and Rosenberg, 1982). However, homologies between all the sites had not been pointed out.

From these findings, one could predict that a given regulatory protein could also interact with the binding site of another regulatory protein. In fact, it has been reported that, in the *lac* and *ara* operons which have a CAP-binding site, CAP is able to bind, albeit with a lower affinity, to the operator site (Schmitz, 1981; Ogden *et al.*, 1980).

Discussion

We have shown two regions of homology (Regions 1 and 2, Figure 1) in the amino acid sequence of 13 DNA-binding proteins: CAP, *trp*R, *tnp*R, *lex*A, *ara*C, *lys*R, *lac*I, *gal*R of *E. coli*, *c*I, *c*II, *cro* of λ1, *c*2 of P22, and *cro* of 434. Region 1 has a conserved sequence V/L/A.G.V.S.Q.S.T.V/I.S/G.R.–. V.N. Region 2 contains the conserved sequence L/A.L.A.L. A.K.

From the three dimensional structure of *cro*, CAP, and the head piece of *c*I, different models for DNA-protein interaction have been proposed, all of them involving α-helices: helix F of CAP, helix α3 of *cro*, and helix 3 of *c*I. These helices correspond to the homologous sequences of Region 1.

Comparison of the different sites of action of the regulatory proteins have shown that they all present a 2-fold rotational symmetry and that most of them contain a consensus sequence $\begin{array}{l} \text{TGTGT } N_{6-10} \text{ ACACA} \\ \text{ACACA} \qquad\quad \text{TGTGT} \end{array}$ or $\begin{array}{l} \text{CAC } N_{5-10} \text{ TGT}. \\ \text{GTG} \qquad\quad \text{ACA} \end{array}$ It is interesting to note that the distance between the conserved nucleotides is ~10 bp which corresponds to one turn of a helix.

These findings, and the observation that the regulatory proteins are multimeric, suggest a general type of DNA-regulatory protein interaction, whether the protein be an activator or a repressor: two α-helices of two monomers would make contacts with a symmetrical sequence, by interacting with DNA in two sucessive major grooves of a right-handed B DNA.

The mode of action of CAP, which in most cases is an activator, could be similar to that of repressors: it has, in fact, been reported that CAP inhibits the transcription initiated at P_2, one of the two promoters of the *gal* operon by binding in the "–35" region of this promoter (Musso *et al.*, 1977a). In *lac*, it has also been recently shown (Malan, 1981), that transcription can be initiated *in vitro* not only at the previously determined promoter (P_1) but also at a second promoter (P_2) located 22 bp upstream from P_1. These two promoters would compete for the binding of the RNA polymerase, transcription initiation being more efficient from P_1 than from P_2. According to this model, CAP would repress RNA polymerase binding on P_2 by interacting with its "–35" region, and would be properly positioned to activate promoter P_1.

Acknowledgements

We thank T. Steitz and P. Stragier for communicating results before publication, A. Ullman for stimulating discussions and help with the manuscript, G. Cohen and J. Janin for interest in our study and encouragement. We are grateful to S. Busby for critical reading of the manuscript. We are greatly indebted to M. Ferrand for her excellent secretarial help. This research was supported by grants from the Centre National de la Recherche Scientifique (LA No. 270; ATP No. AI-5022; GR No. 30), Délégation Générale à la Recherche Scientifique et Technique (No. 81F1202), Institut National de la Santé et de la Recherche Médicale (No. 811026), and Fondation pour la Recherche Médicale Francaise.

References

Abraham,J., Mascarenas,D., Fisher,R., Benedik,M., Campbell,A., and Echols,H. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 2477-2481.

Aiba,H., Fujimoto,S., and Ozaki,N. (1982) *Nucleic Acids Res.,* **10**, 1345-1361.

Aiba,H., and Krakow,J.S. (1981) *Biochemistry (Wash.),* **20**, 4774-4780.

Anderson,W.F., Ohlendorf,D.H., Takeda,Y., and Matthews,B.W. (1981) *Nature,* **290**, 754-758.

Bennett,G.N., and Yanofsky,C. (1978) *J. Mol. Biol.,* **121**, 179-192.

Bennett,G.N., Schweingruber,M.E., Brown,K.D., Squires,C., and Yanofsky,C. (1978) *J. Mol. Biol.,* **121**, 113-137.

Beyreuther,K. (1978) in Miller,J.H., and Reznikoff,W. (eds.), *The Operon,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 123-154.

Brent,R., and Ptashne,M. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 4204-4208.

Busby,S., Aiba,H., and De Crombrugghe,B. (1982) *J. Mol. Biol.,* **154**, 211-228.

Chou,J., Casadaban,M., Lemaux,P., and Cohen,S.N. (1979) *Proc. Natl. Acad. Sci. USA,* **76**, 4020-4024.

Cossart,P., and Gicquel-Sanzey,B. (1982) *Nucleic Acids Res.,* **10**, 1363-1378.

Dayhoff,M.O. (1978) *Atlas of Protein Sequence and Structure,* published by National Biochemical Research Foundation, Silver Spring, MD.

Dickson,R.C., Abelson,J., Barnes,W.M., and Reznikoff,W.S. (1975) *Science (Wash.),* **187**, 27-35.

Dickson,R.C., Abelson,J., Johnson,R., Reznikoff,W.S., and Barnes,W.M. (1977) *J. Mol. Biol.,* **111**, 65-75.

Di Lauro,R., Taniguchi,T., Musso,R., and De Crombrugghe,B. (1979) *Nature,* **279**, 494-500.

Eilen,E., and Krakow,J.S. (1977) *J. Mol. Biol.,* **114**, 47-60.

Farabaugh,P. (1978) *Nature,* **274**, 765-769.

Gilbert,W., Gralla,J., Majors,J., and Maxam,A. (1975) in Sund,H., and Blauer,G. (eds.), *Protein-Ligand Interactions,* Walter de Gruyter, Berlin, pp. 193-206.

Goeddel,D.V., Yansura,D.G., and Caruthers,M.H. (1978) *Proc. Natl. Acad. Sci. USA,* **75**, 3578-3582.

Greenfield,L., Boone,T., and Wilcox,G. (1978) *Proc. Natl. Acad. Sci. USA,* **75**, 4724-4728.

Grosscheld,R., and Schwarz,E. (1979) *Nucleic Acids Res.,* **6**, 867-881.

Gunsalus,R.P., and Yanofsky,C. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 7117-7121.

Heffron,F., McCarthy,B., Ohtsubo,H., and Ohtsubo,E. (1979) *Cell,* **18**, 1153-1163.

Ho,Y., and Rosenberg,M. (1982) *Ann. Microbiol. (Inst. Pasteur),* **133A**, 215-218.

Horii,T., Ogawa,T., and Ogawa,H. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 313-317.

Horii,T., Ogawa,T., and Ogawa,H. (1981) *Cell,* **23**, 689-697.

Humayun,Z., Kleid,D., and Ptashne,M. (1977) *Nucleic Acids Res.,* **4**, 1595-1607.

Krakow,J.S., and Pastan,I. (1973) *Proc. Natl. Acad. Sci. USA,* **70**, 2529-2533.

Lee,N.L., Gielow,W.O., and Wallace,R.G. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 752-756.

Little,J.W., Mount,D.W., and Yanisch-Perron,C. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 4199-4203.

McKay,D.B., and Steitz,T.A. (1981) *Nature,* **290**, 744-749.

Majors,J. (1975) *Nature,* **256**, 672-674.

Malan,P.M.,Jr. (1981) Ph.D.Thesis, Harvard University, USA.

Maniatis,T., Jeffrey,A., and Kleid,D.G. (1975a) *Proc. Natl. Acad. Sci. USA,* **72**, 1184-1188.

Maniatis,T., Ptashne,M., Backman,K., Kleid,D., Flashman,S., Jeffrey,A., and Maurer,R. (1975b) *Cell,* **5**, 109-113.

Markham,B., Little,J., and Mount,D. (1981) *Nucleic Acids Res.,* **9**, 4149-4161.

Matthews,B.W., Ohlendorf,D.H., Anderson,W.F., and Takeda,Y. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 1428-1432.

Miki,T., Ebina,Y., Kishi,F., and Nakazawa,A. (1981) *Nucleic Acids Res.,* **9**, 529-543.

Miller,J. (1978) in Miller,J., and Reznikoff,W. (eds.), *The Operon,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 31-38.

Miyada,C., Horwitz,A., Cass,L.G., Timko,J., and Wilcox,G. (1980) *Nucleic Acids Res.,* **8**, 5267-5274.

Miyada,C., Soberon,X., Itakura,K., and Wilcox,G. (1982) *Gene,* **17**, 167-177.

Müller-Hill,B. (1975) *Prog. Biophys. Mol. Biol.,* **30**, 227-252.

Musso,R.E., De Crombrugghe,B., Pastan,I., Sklar,J., Yot,P., and Weissman,S. (1974) *Proc. Natl. Acad. Sci. USA,* **71**, 4940-4944.

Musso,R.E., Di Lauro,R., Adhya,S., and De Crombrugghe,B. (1977a) *Cell,* **12**, 847-854.

Musso,R.E., Di Lauro,R., Rosenberg,M., and De Crombrugghe,B. (1977b) *Proc. Natl. Acad. Sci. USA,* **74**, 106-110.

Ogden,S., Haggerty,D., Stoner,C., Kolodrubetz,D., and Schleif,R. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 3346-3350.

O'Neill,M., Amass,K., and De Crombrugghe,B. (1981) *Proc. Natl. Acad. Sci. USA,* **78**, 2213-2217.

Oppenheim,D.S., Bennett,G.N., and Yanofsky,C. (1980) *J. Mol. Biol.,* **144**, 133-142.

Pabo,C.O., and Lewis,M. (1982) *Nature,* in press.

Pabo,C.O., Saver,R.T., Sturtevant,J.N., and Ptashne,M. (1979) *Proc. Natl. Acad. Sci. USA,* **76**, 1608-1612.

Pirotta,V., Chadwick,P., and Ptashne,M. (1970) *Nature,* **277**, 41-44.

Ptashne,M. (1978) in Miller,J., and Reznikoff,W. (eds.), *The Operon,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 325-343.

Ptashne,M., Backman,K., Humayun,M.Z., Jeffrey,A., Maurer,R., Meyer, B., and Sauer,R.T. (1976) *Science (Wash.),* **194**, 156-161.

Queen,C., and Rosenberg,M. (1981) *Nucleic Acids Res.,* **9**, 3365-3377.

Riggs,A.D., and Bourgeois,S. (1968) *J. Mol. Biol.,* **34**, 361-364.

Sancar,G.B., Sancor,A., Little,J.W., and Rupp,W.D. (1982) *Cell,* **28**, 523-530.

Sauer,R.T., and Anderegg,R. (1978) *Biochemistry (Wash.),* **17**, 1092-1100.

Sauer,R.T., Pan,J., Hopper,P., Hehir,K., Brown,J., and Poteete,A.R. (1981) *Biochemistry (Wash.),* **20**, 3591-3598.

Schmitz,A. (1981) *Nucleic Acids Res.,* **9**, 277-291.

Schmitz,A., and Galas,D.J. (1979) *Nucleic Acids Res.,* **6**, 111-137.

Schwarz,E., Scherer,G., Hobom,G., and Kössel,M. (1978) *Nature,* **272**, 410-414.

Simpson,R.B. (1980) *Nucleic Acids Res.,* **8**, 759-766.

Singleton,C.K., Roeder,W.D., Bogosian,G., Somerville,R.L., and Weith, H.L. (1980) *Nucleic Acids Res.,* **8**, 1551-1560.

Sklar,J., Weissman,S., Musso,R., Di Lauro,R., and De Crombrugghe,B. (1977) *J. Biol. Chem.,* **252**, 3538-3547.

Smith,B.R., and Schleif,R. (1978) *J. Biol. Chem.,* **253**, 6931-6933.

Steffen,D., and Schleif,R. (1977) *Mol. Gen. Genet.,* **157**, 333-339.

Stoner,C.M., and Schleif,R. (1982) *J. Mol. Biol.,* **154**, 649-652.

Takeda,Y., Folkmanis,A., and Echols,H. (1977) *J. Biol. Chem.,* **252**, 6177-6183.

Taniguchi,T., O'Neill,M., and De Crombrugghe,B. (1979) *Proc. Natl. Acad. Sci. USA,* **76**, 5090-5094.

Van den Berg,E., Zwetsloot,J., Noordermeer,I., Panekoek,H., Dekker,B., Dijkema,R., and Van Ormondt,H. (1981) *Nucleic Acids Res.,* **9**, 5623-5643.

Van den Elzen,P.J.M., Maat,J., Walters,H.H.B., Veltkamp,E., and Nijkamp,H.J.J. (1982) *Nucleic Acids Res.,* **10**, 1913-1929.

Wallace,R.G., Lee,N., and Fowler,R.A. (1980) *Gene,* **12**, 179-190.

Wilcken-Bergman,B., and Müller-Hill,B. (1982) *Proc. Natl. Acad. Sci. USA,* **79**, 2427-2431.

Zurawski,G., Gunsalus,R.P., Brown,K.D., and Yanofsky,C. (1981) *J. Mol. Biol.,* **145**, 47-73.