Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 7679–7683, December 1982
Biochemistry

The cAMP-binding domains of the regulatory subunit of cAMP-dependent protein kinase and the catabolite gene activator protein are homologous

(allosteric protein/evolution/gene regulatory protein/protein structure/DNA binding protein)

I. T. WEBER*, K. TAKIO†, K. TITANI†, AND T. A. STEITZ*‡

*Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511; and †Howard Hughes Medical Institute Laboratory at the Department of Biochemistry, SJ-70, University of Washington, Seattle, Washington 98195

Communicated by Edwin G. Krebs, September 20, 1982

ABSTRACT **Comparison of the recently determined amino acid sequences of the regulatory subunit of cAMP-dependent protein kinase (R_{II}) from bovine cardiac muscle and the *Escherichia coli* catabolite gene activator protein (CAP) shows significant homology. This homology extends over most of the amino-terminal domain in CAP and is particularly good for the region of the β-roll structure. The R_{II} sequence contains two adjacent and internally homologous regions, both of which have high resemblance to the cAMP-binding domain in CAP. This suggests that the protein kinase regulatory subunit contains two cAMP-binding domains in the carboxyl-terminal region, each having a β-roll structure similar to that in CAP. The cAMP molecule is expected to bind to the R_{II} within a pocket formed by residues from the β-roll, as is the case with CAP. One cAMP molecule would interact with residues from about 163 to 220, and the other cAMP would interact with amino acids in the stretch 285–350 of the R_{II} protein kinase sequence. As the carboxyl-terminal domain of CAP shows homologies to the DNA-binding domains of other regulatory proteins, the protein appears to be of modular construction: a DNA-binding domain joined to a cAMP-binding domain.**

Takio *et al.* (1) have recently determined the amino acid sequence of the regulatory subunit of cAMP-dependent protein kinase type II (R_{II}) from bovine heart muscle. The protein kinase catalyzes the transfer of phosphate from ATP to various protein substrates (2). It functions in the hormone-mediated regulation of cellular functions such as glycolysis in eukaryotes with cAMP acting as a "second messenger." The inactive holoenzyme consists of two regulatory and two catalytic subunits. In the presence of cAMP, the protein dissociates into two active catalytic subunits and a dimer of two regulatory subunits that is complexed with four cAMP molecules (3–5). R_{II} appears to consist of three domains (4) and contains a region very sensitive to proteolytic attack. This protease-sensitive region separates the smaller amino-terminal domain and the larger carboxyl-terminal domain, contains the "autophosphorylation" site (serine-95) (6), and probably interacts with the catalytic subunit within the holoenzyme. The amino-terminal fragment of limited proteolysis still forms a dimer (7) while the carboxyl-terminal fragment, which appears to contain internal sequence homology between residues 135–256 and 257–400, may bind one cAMP molecule within each of these regions (1).

McKay and Steitz (8) have solved the crystal structure of the *Escherichia coli* catabolite gene activator protein (CAP), otherwise known as the cAMP receptor protein, at 2.9-Å resolution. CAP binds to specific DNA sequences in the presence of cAMP (9, 10) and regulates transcription of several operons such as *lac* (11), *gal* (12), and *ara* C (13). In the absence of cAMP, CAP only binds nonspecifically to DNA. Thus it is also a cAMP-dependent regulatory protein but it acts at the level of transcription. Crystals of CAP contain one molecule of cAMP per subunit of protein dimer. The cAMP binding site is located entirely within the amino-terminal domain and lies in a deep pocket formed between a β-roll structure and a long α-helix. Each cAMP molecule is completely buried within the protein and interacts with amino acid side chains from both subunits of the dimer (14, 15).

We show here that there are homologous sequences between CAP and R_{II} and suggest that the cAMP-binding domain is a conserved structure between these two cAMP-dependent proteins. This homology allows us to locate the cAMP-binding sites in R_{II} by analogy with the CAP structure and deduce some of the amino acids interacting with the cAMP molecule.

METHODS

The amino acid sequence of R_{II} (1) was compared with the sequence of CAP (16, 17) to search for the cAMP binding sites. The sequences were compared by using computer programs run on a VAX 11/750; one written by Peter Brick includes the scoring tables of McLachlan (18) and the other, written by I.T.W., calculates statistics for the numbers of identical and homologous amino acids in all possible alignments of the two protein sequences (19).

RESULTS

The best alignments of the two homologous domains of R_{II} with the CAP sequence are shown in Fig. 1 with a double entry of glutamate-265. The homology is most apparent over the region of CAP, residues 30–89, that includes β-strands 3–7, part of β-strand 2, and the loop between β-strands 6 and 7. The secondary structure of one subunit of CAP is shown in Fig. 2. The alignment with the first homologous region of R_{II} includes two deletions of a single amino acid residue each: one in β-strand 6 and one in the loop between β-strands 6 and 7. There are 18 identical and 9 homologous amino acids in this stretch of 60 residues, giving a total homology of 45% (Table 1). The second homologous region of R_{II} contains one deletion in the loop between β-strands 6 and 7 and an extra seven residues (Ser-Lys-Thr-Lys-Val-Asn-Lys) in the surface loop between β-strands 4 and 5. This insertion in a surface loop is consistent with the observation (1) that trypsin cleaves lysyl bonds in this portion of native R_{II}. There are 23 identical and 5 homologous amino acids or 46.7%

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: R_{II}, regulatory subunit of cAMP-dependent protein kinase type II; CAP, *Escherichia coli* catabolite gene activator protein.

‡ To whom correspondence should be addressed.

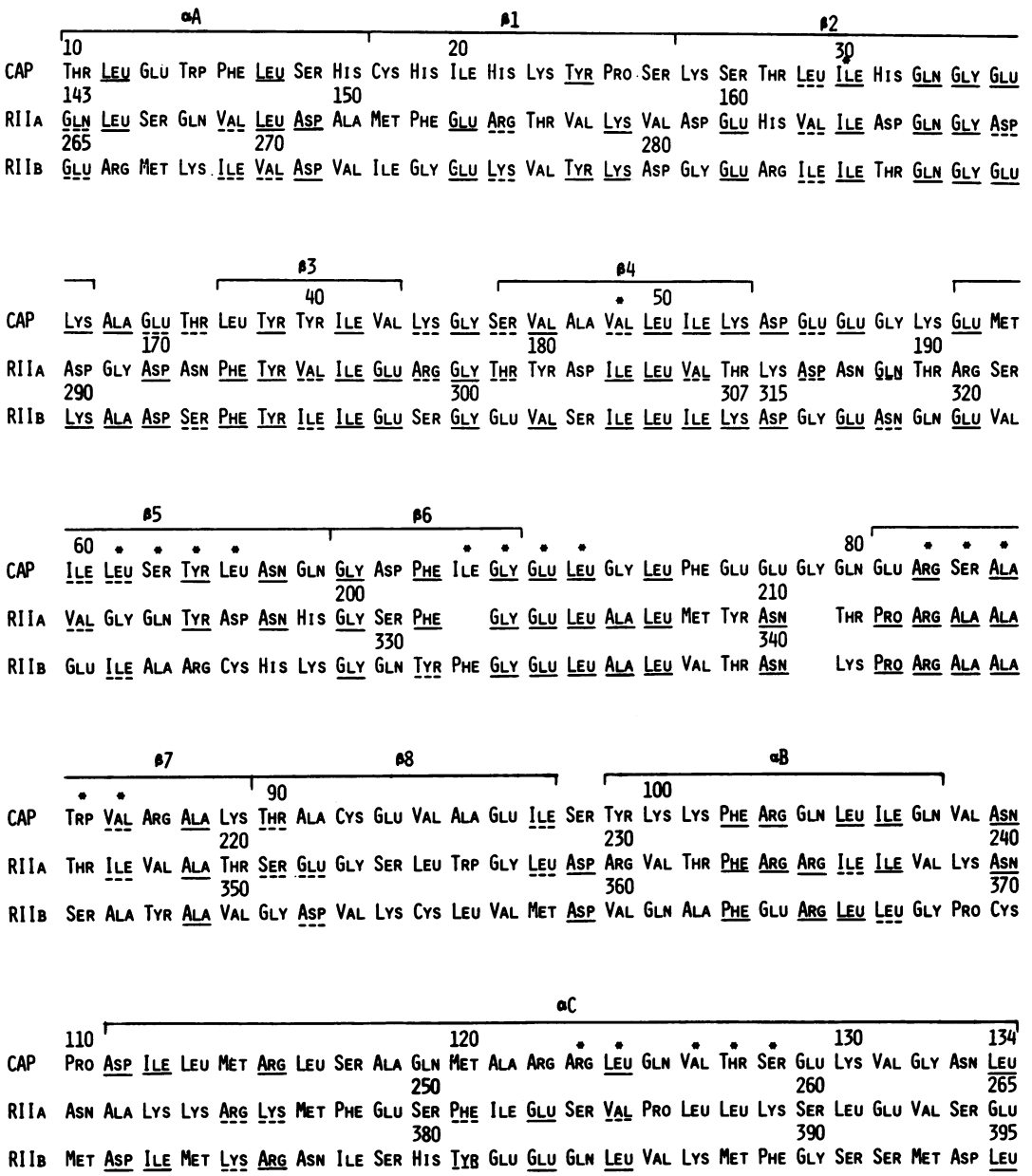

FIG. 1. Alignment of two adjacent regions of the R_{II} sequence with the CAP sequence. $R_{II}A$ and $R_{II}B$ are homologous regions of R_{II} sequence. Solid underlining indicates amino acids that are identical to one or more residues in the other sequences; dashed underlining indicates closely similar amino acids. The positions of the α-helices and β-strands of the CAP molecule are indicated. *, Amino acids that are close to cAMP in CAP.

total homology as defined in Table 1. This alignment in Fig. 1 also illustrates the extensive homology between the two regions of R_{II}. Note that all but one of the four deletions and insertions occur in the region of loops in the CAP structure and thus would be easily accommodated in the β-roll structure.

The alignment also shows some evidence of homology over a larger region, from CAP α-helix A through the eight β-strands to α-helix B. There may also be an alignment of R_{II} with helix C of the CAP, but the level of sequence homology is very low. The larger region, CAP residues 10–106, has 39.2% homology with residues 143–237 of R_{II} and 35.1% homology with the R_{II} region residues 265–367 in the carboxyl-terminal domain. Note also that the mean percentage homology for all random alignments of CAP and R_{II}, without deletions, is 14.2% over 60 residues with a SD of 4.6%, so that the above alignments are very significant.

There is no significant homology of CAP with the amino-terminal 90 residues of R_{II} that are not expected to be involved in binding cAMP. Similarly, any homology of the carboxyl-terminal domain of CAP with R_{II} is insignificant; this domain of CAP is involved in specific binding to DNA.

DISCUSSION

The amino acid sequence of R_{II} contains two extensive regions of homology with the sequence of CAP. The amino-terminal large domain of CAP that includes an eight-stranded β-roll structure is homologous with the two regions of R_{II} sequence that are very similar to each other (1). These R_{II} sequences are located in the carboxyl-terminal two-thirds of the molecule that is known to bind two molecules of cAMP (4).

The homology between CAP and R_{II} is greatest over the region from β-strand 2 to β-strand 7. Also the internal homology between the two regions of R_{II} is greater over this stretch: 46.7% for the CAP range of 30–89, as compared with 38.1% for the 97 residues that align with CAP 10–106 (from helix A

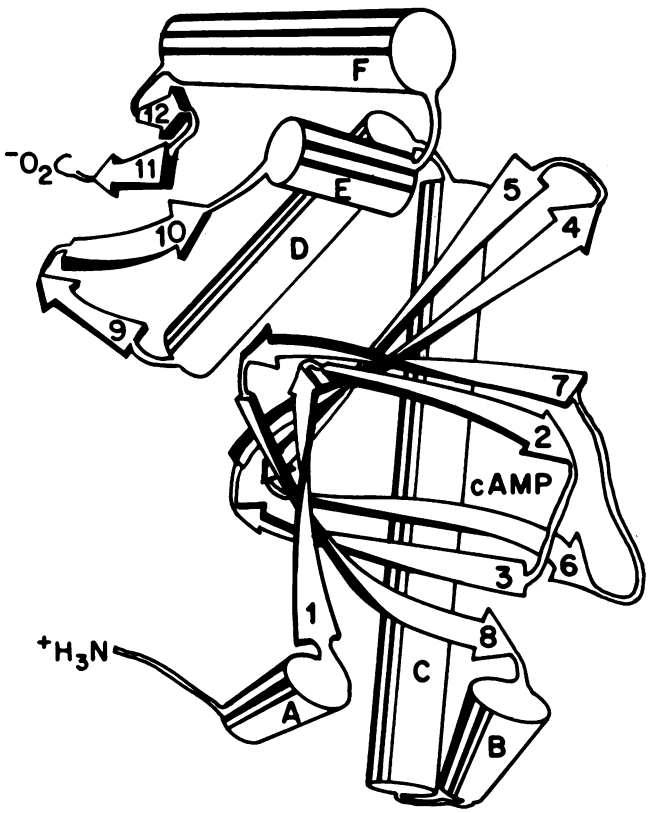

FIG. 2. Schematic drawing illustrating the secondary structure and cAMP-binding site of one subunit of CAP. The α-helices are indicated by cylinders and the β-strands are indicated by arrows. (Reproduced from ref. 15.)

through the β-roll to helix B). The cAMP molecule in the CAP structure is close to the amino acids on the β-strands marked in Fig. 1. Amino acid side chains from β-strands 2, 4, 5, 6, and 7 form a pocket around the cAMP molecule as illustrated in Figs. 2 and 3.

Thus, we expect that the protein kinase regulatory subunit, R_{II}, contains two structural domains in the carboxyl-terminal 250 residues that are extremely similar to the β-roll structure of CAP. As in the case of CAP, each of the β-rolls of R_{II} presumably contains within it one binding site for cAMP. By comparing the sequence of R_{II} with the homologous structure of CAP, it is possible to identify some of the interactions between cAMP and R_{II} that may be expected.

Table 1. Number of identical and homologous amino acids for the best alignment of CAP and R_{II}

Sequences aligned	Identical amino acids, no.	Homologous amino acids, no.	% homology*
	β-Strand 2 to β-strand 7		
CAP 30–89 vs. R_{II} 163–220	18	9	45.0
CAP 30–89 vs. R_{II} 285–350	23	5	46.7
	α-Helix A to α-helix B		
CAP 10–106 vs. R_{II} 143–237	23	15	39.2
CAP 10–106 vs. R_{II} 265–367	26	8	35.1
R_{II} 163–220 vs. R_{II} 285–350	24	4	46.7
R_{II} 143–237 vs. R_{II} 265–367	31	6	38.1

* For both identical and homologous amino acid residues within the alignment of Fig. 1.

The cAMP molecule in the CAP structure is buried within a deep pocket formed by residues from the β-roll and two long α-helices, one from each subunit of the dimer (15). The phosphate group of cAMP interacts with the arginine and serine side chains from β-strand 7, and the adenine ring interacts with the threonine and serine side chains on α-helix C (Fig. 4).

Some of the interactions between R_{II} and the ribose phosphate moiety of cAMP can be identified because the portion of the cAMP binding site in CAP that interacts with that part of cAMP is clearly homologous to R_{II}. Arginines-213 and -343 of R_{II} are homologous to and presumably play the same role as arginine-82 of CAP, whose side chain binds the cyclic phosphate. The hydroxyl of serine-83 in CAP also interacts with this phosphate, but the corresponding residues in R_{II} are both alanine. The other residues on β-strands 4, 5, 6, and 7, indicated in Fig. 1 as interacting with cAMP, primarily bind the ribose phosphate. Consistent with this identification of the cAMP-binding site is the observation that the reaction of one cysteine residue with dinitrothiobenzoic acid is blocked by bound cAMP (20). This is probably cysteine-326, which lies in a region homologous to β-strand 5 of CAP and forms part of the cAMP-binding pocket. Neither predicted cAMP-binding site in R_{II} is close to the single tryptophan residue (no. 226), which agrees with fluorescence experiments (21).

Most of the specific interactions between CAP and the ad-

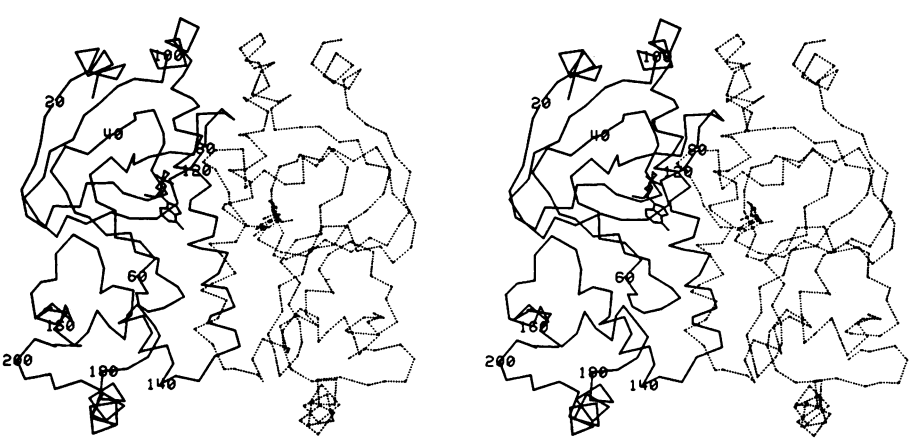

FIG. 3. Stereo view of the α-carbon backbone of the CAP dimer and the two bound molecules of cAMP. (Reproduced from ref. 15.)

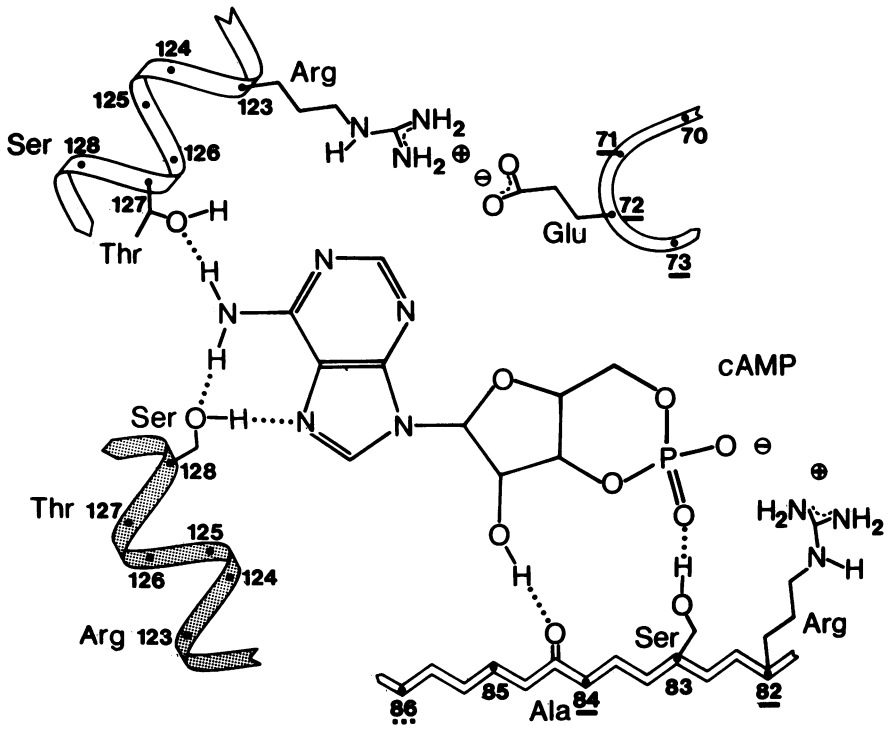

FIG. 4. Schematic drawing of the probable interactions of cAMP in one CAP binding site. The shaded helix is in the opposite subunit of the CAP dimer. The cAMP-binding sites in R_{II} are expected to be similar to those in CAP, except in the region of the CAP helices. Underlined numbers indicate amino acid residues of CAP that are similar to R_{II} residues for the alignment in Fig. 1; solid underlining indicates amino acids that are identical in CAP, $R_{II}A$, and $R_{II}B$ while dashed underlining indicates similar amino acids. (Modified from ref. 15.)

enine of cAMP are interactions with the two C α-helices of the dimer. Unfortunately, unambiguous identification of the regions in R_{II} that are homologous to α-helix C is not obvious from the sequence comparison. Detailed model building will be required (at least). Evidence that R_{II} may contain a region analogous to α-helix C of CAP is provided by the fact that 8-azidoadenosine 3',5'-cyclic monophosphate can be crosslinked through the C-8 position to tyrosine-381 (22). This residue is in the region of R_{II} that we would expect to be homologous to α-helix C of CAP. In CAP, the C-8 position of the adenine is indeed near to the C α-helices.

The homology between CAP and R_{II} that is described here is sufficiently high that it should be possible to construct a model of the cAMP-binding domains of R_{II} assuming that the identical and similar residues have the same structure in the two proteins and that many of the other residues will have closely similar structures. It should then be possible to account for the binding of various analogs of cAMP to R_{II}.

EVOLUTION OF cAMP SENSITIVITY

In both prokaryotes and eukaryotes, changes in the levels of cAMP act as a second messenger to signal changes in glucose concentration. When glucose levels drop, adenylate cyclase is stimulated and cAMP levels rise. In bacteria, a drop in glucose levels results in switching on of the appropriate catabolic genes—e.g., the arabinose, lactose or galactose operons—depending on what sugars are in the medium. In contrast, a drop in sugar levels in higher eukaryotes results in a breakdown of glycogen. These different responses to the same problem are required in the two cases because bacteria do not rely on glycogen and cows are not swimming in solutions of arabinose.

It appears that a common ancestral precursor protein capable of binding cAMP has evolved into the appropriate receptor or transducer of cAMP levels in both bacteria and mammals. In prokaryotes, the cAMP-binding domain is attached to a DNA-binding domain that shares structural and sequence homologies with the DNA-binding portion of bacterial and viral repressors (14, 19, 23). In mammals, the cAMP-binding domain is part of a subunit that regulates, in a cAMP-dependent manner, the activity of protein kinase, the first enzyme in a cascade that results in activation of the phosphorylase-catalyzed breakdown of glycogen. Thus, the modular "design" of proteins allows homologous cAMP-binding domains to switch on transcription of catabolic genes in bacteria and activate protein kinase in cows when glucose levels fall and cAMP levels rise.

We thank Drs. E. G. Krebs, K. A. Walsh, and S. B. Smith for their interest and helpful discussions. The research was supported by National Science Foundation Grant PCM-81-10880 to T.A.S. and in part by National Institutes of Health Grant GM 15731 to K. A. Walsh.

1. Takio, K., Smith, S. B., Krebs, E. G., Walsh, K. A. & Titani, K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2544–2548.
2. Krebs, E. G. & Beavo, J. A. (1979) *Annu. Rev. Biochem.* **48**, 923–959.
3. Beavo, J. A., Bechtel, P. J. & Krebs, E. G. (1975) *Adv. Cyclic Nucleic Res.* **5**, 241–257.
4. Corbin, J. D., Sugden, P. H., West, L., Flockhart, D. A., Lincoln, T. M. & McCarthy, D. (1978) *J. Biol. Chem.* **253**, 3997–4003.
5. Builder, S. E., Beavo, J. A. & Krebs, E. G. (1980) *J. Biol. Chem.* **255**, 2350–2354.
6. Takio, K., Walsh, K. A., Neurath, H., Smith, S. B., Krebs, E. G. & Titani, K. (1980) *FEBS Lett.* **114**, 83–88.
7. Potter, R. L. & Taylor, S. S. (1979) *J. Biol. Chem.* **254**, 2413–2418.
8. McKay, D. B. & Steitz, T. A. (1981) *Nature (London)* **290**, 744–749.
9. Zubay, G., Schwartz, D. & Beckwith, J. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 104–110.
10. Epstein, W., Rothman-Denes, L. B. & Hesse, J. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2300–2304.

11. deCrombrugghe, B., Chen, B., Anderson, W., Nessley, P., Gottesman, M., Pastan, I. & Perlman, R. (1971) *Nature (London) New Biol.* **231,** 139–142.
12. Musso, R. E., DiLauro, R., Adhya, S. & deCrombrugghe, B. (1977) *Cell* **12,** 847–854.
13. Ogden, S., Haggerty, D., Stoner, C. M., Kolodrubetz, D. & Schlief, R. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 3346–3350.
14. Steitz, T. A., McKay, D. B. & Weber, I. T. (1982) in *Nucleic Acid Research: Future Development*, ed. Watanabe, I. (Academic, Japan), in press.
15. McKay, D. B., Weber, I. T. & Steitz, T. A. (1982) *J. Biol. Chem.* **257,** 9518–9524.
16. Aiba, H., Fujimoto, S. & Ozaki, N. (1982) *Nucleic Acids Res.* **10,** 1345–1361.
17. Cossart, P. & Gicquel-Sanzey, B. (1982) *Nucleic Acids Res.* **10,** 1363–1378.
18. McLachlan, A. D. (1971) *J. Mol. Biol.* **61,** 409–424.
19. Weber, I. T., McKay, D. B. & Steitz, T. A. (1982) *Nucleic Acids Res.* **10,** 5085–5102.
20. Armstrong, R. N. & Kaiser, E. T. (1978) *Biochemistry* **17,** 2840–2845.
21. LaPorte, D. C., Builder, S. E. & Storm, D. R. (1980) *J. Biol. Chem.* **255,** 2343–2349.
22. Kerlavage, A. R. & Taylor, S. S. (1980) *J. Biol. Chem.* **255,** 8483–8488.
23. Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F. & Matthews, B. W. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 3097–3100.