Proc. Nat. Acad. Sci. USA
Vol. 68, No. 6, pp. 1222–1225, June 1971

Purification and DNA-Binding Properties of the Catabolite Gene Activator Protein

(phosphocellulose/DEAE-cellulose/*lac* operon DNA/cyclic AMP)

A. D. RIGGS, G. REINESS, AND G. ZUBAY

Department of Biology, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, California 91010; and Department of Biological Sciences, Columbia University, New York, N.Y. 10036

Communicated by Norman Davidson, March 29, 1971

ABSTRACT **A protein required for the activation of the *lac* operon has been extensively purified and partly characterized. This protein, called CGA protein (catabolite gene activator protein, sometimes named CAP), is a dimer with subunits of 22,000 daltons. Purified CGA protein has a substantial affinity for DNA; this affinity is greatly strengthened by cAMP and strongly inhibited by cGMP. Other studies have shown that these cyclic nucleotides compete for a binding site on CGA protein. The opposing effects of the two cyclic compounds in DNA–CGA protein binding show a parallel behavior to their effects on the expression of the *lac* operon. Thus cAMP, in addition to CGA protein, is required for expression of the *lac* operon, whereas cGMP inhibits the expression. The obvious inference is that CGA protein activates the *lac* operon by binding to the DNA under the influence of cAMP. Thus, CGA protein seems to be a new type of regulatory protein: a DNA-binding activator.**

The phenomenon of catabolite repression in *Escherichia coli* (1) has led to the delineation of a positive control system sensitive to the intracellular and extracellular level of adenosine cyclic 3′:5′-monophosphate (cAMP) (2). Recently, a protein called the catabolite gene activator protein (CGA protein) has been shown (3) to stimulate the DNA-directed synthesis of the enzymes of the *lac* operon, a catabolite-sensitive operon. The CGA protein is effective only in the presence of cAMP. Mutants defective in CGA protein can express *in vivo* neither *lac* nor other catabolite-sensitive genes (4), which suggests a common mechanism of activation by CGA protein. Studies with altered promoters of the *lac* operon show that certain properties of an intact promoter are needed for effective action of CGA protein (5). Since the action of CGA protein is intimately involved with the promoter part of the operon, it seems likely that CGA protein is involved in the initiation of messenger synthesis from the *lac* operon, perhaps by forming a part of the initiation complex. Previous attempts to detect interaction between CGA protein and RNA polymerase have been unsuccessful. Here, we report the complete purification of CGA protein and demonstrate it to be a DNA-binding protein. This result suggests that binding of CGA protein to promoter is required for expression of the *lac* operon.

Abbreviation: SDS, sodium dodecyl sulfate; CGA protein, catabolite gene activator protein. This protein has frequently been referred to as CAP, a less logical choice since it implies activation of the *catabolite*. The word "protein" is necessary in the abbreviation to eliminate possible confusion with a trinucleotide (cytidine-guanosine-adenosine). In general, this journal prefers 3-letter to Greek-letter designations of this kind of compound.

METHODS

Purification of CGA protein

The procedure reported here is a significant modification of an already reported procedure (3). About 200 g of frozen *E. coli* strain 514 is homogenized in 700 ml of buffer I (0.01 M Tris-acetate (pH 8.2)–0.01 M $Mg(OAc)_2$–0.06 M KCl–6 mM mercaptoethanol) and centrifuged for 30 min at 16,000 × *g*. The sediment containing the bacteria is homogenized and recentrifuged. The final sediment is resuspended in 260 ml of buffer I. The suspension of cells is lysed in an Aminco pressure cell at pressures between 4000 and 8000 psi. The lysate is centrifuged for 30 min at 30,000 × *g* in a small Sorvall rotor. The resulting supernatant is centrifuged for 4 hr at 30,000 rpm in a Spinco no. 30 rotor. The resulting supernatant is dialyzed for 16 hr against buffer II [0.01 M K_2HPO_4–CH_3COOH (pH 7.0)–6 mM mercaptoethanol]. This solution is passed over a 2.5 × 15 cm phosphocellulose column (Whatman P11, medium fibrous powder, 7.4 meq/g), previously equilibrated with buffer II. After the column is rinsed progressively with 200 ml of buffer II and 100 ml of buffer II + 0.4 M KCl, the active fraction, which constitutes about 1% of the protein put on the column, is eluted in buffer II + 0.50 M KCl. The active fraction, detected by ultraviolet absorption, is pooled and dialyzed against buffer III [0.01 M KH_2PO_4–KOH (pH 7.7)–6 mM mercaptoethanol] overnight. This solution is passed over a 1.4 × 13 cm DEAE-cellulose column previously equilibrated with buffer III. The active fraction, which constitutes about 10% of the protein put on this column, passes through the column with no holdback. The remainder of the protein is retained by the column. Total yield of protein is 300–1000 μg in about 20 ml.

Acrylamide gel electrophoresis

The CGA protein solution was concentrated about 5-fold prior to electrophoresis by placing a dialysis sac containing the CGA protein solution in a beaker of dry G-200 Sephadex for 16 hr. 10% polyacrylamide gels containing 0.1% sodium dodecyl sulfate (SDS) were prepared by the procedure of Weber and Osborn (6) and prerun without samples for 2 hr at 3 mA per gel in Weber and Osborn's running buffer. CGA protein preparations, obtained as above, were layered on the gels in 100 μl of sample buffer [containing 0.1% SDS–0.01

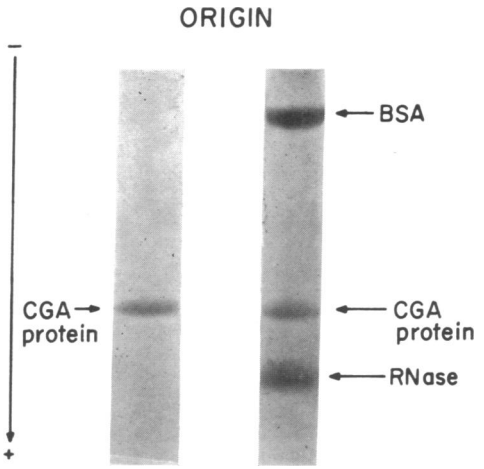

FIG. 1. Acrylamide gel electrophoresis of CGA protein. 10% polyacrylamide gels containing 0.1% SDS were prepared according to Weber and Osborn (6). The gel at the *left* contains only a CGA protein preparation. To the gel at the *right* were added, in addition to CGA protein, the marker proteins bovine serum albumin (BSA; 5 μg) and RNase (2.5 μg).

M sodium phosphate (pH 7.2)–10% (v/v) glycerol–0.002% bromophenol blue–0.14 M β-mercaptoethanol] that had been incubated at 65°C for 30 min. The gels were run for 2.5–3 hr at 8 mA per gel, after which time the tracking dye had migrated about 50 mm. Gels were removed from the tubes and stained for 4 hr in 0.25% Coomassie Brilliant Blue in methanol–water–acetic acid 5:5:1. Gels were then soaked overnight in 7.5% acetic acid–5% methanol and destained electrophoretically in the same solution. Mobilities of protein bands were determined according to Weber and Osborn and the molecular weight of CGA protein was estimated.

DNA-binding methods

The membrane filter technique developed for *lac* repressor (7) was used with only minor changes. A typical experiment will be described here. Any variations in this basic procedure are indicated in the figure legends. The appropriate volume of CGA protein solution was mixed with 0.05 μg of λh80*dlac* [^{32}P]DNA in a total volume of 1.3 ml that contained buffer IV [10 mM KCl–3 mM $Mg(OAc)_2$–0.1 mM EDTA–0.1 mM dithiothreitol–50 μg/ml of bovine serum albumin–5% dimethyl sulfoxide–10 mM Tris·HCl (pH 7.4 at 24°C)]. After incubation at room temperature (about 24°C) for 30 min, a time more than adequate to reach equilibrium, 0.4-ml samples were filtered in duplicate through 25-mm nitrocellulose membrane filters (Schleicher and Schuell, B-6). The filtering rate was such that the sample passed through in about 15 sec. The filters were washed two times with 0.4 ml of buffer IV without either bovine serum albumin or dithiothreitol. The data points shown here represent the average of two filters. The filters had been treated with 0.5 M KOH for 30 min at room temperature to help reduce DNA binding in the absence of CGA protein (11). For most experiments, this background was less than 5% of the total counts filtered (see Fig. 2).

[^{32}P]DNA from λh80*dlac* was prepared as was described (7). DNA concentrations were measured spectrophotometrically at 260 nm; an extinction coefficient of 0.02 $cm^2/\mu g$ was assumed. CGA protein was prepared as described above;

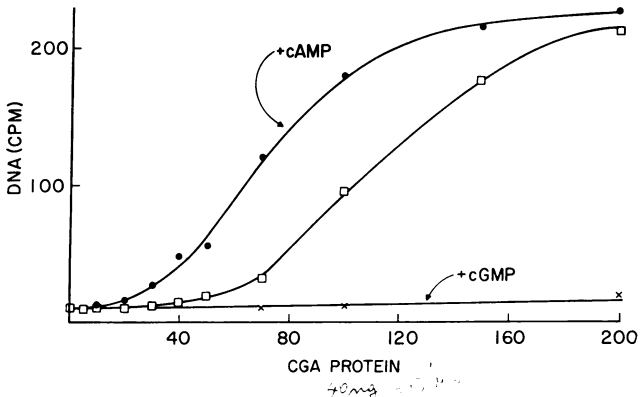

FIG. 2. Bound DNA as a function of the amount of CGA protein. A fixed amount of λh80*dlac* [^{32}P]DNA (0.05 μg) was mixed with the indicated volume of CGA protein solution in a final volume of 1.3 ml, containing buffer IV. The abscissa units are μl of a CGA protein solution that contained about 0.4 μg/ml of CGA protein. Each filter received 0.4 ml, containing a total of 260 cpm. (□) Buffer alone; (●) 3.7×10^{-4} M cAMP; (×) 3.7×10^{-4} M cGMP.

the concentrations of CGA protein given here are based on total protein estimated from the UV absorption at 280 nm; an extinction coefficient of 0.5 cm^2/mg was assumed. Low protein concentrations and the presence of dithiothreitol make other methods difficult.

RESULTS AND DISCUSSION

CGA protein purification and properties

Only two major purification steps, phosphocellulose and DEAE-cellulose chromatography, are needed to give apparently pure CGA protein. This fortunate result was the first hint that CGA protein interacts with DNA, since many other DNA-binding proteins, among them the *lac* repressor (8), bind to phosphocellulose, whereas more than 90% of *E. coli* proteins do not. As shown in Fig. 1, electrophoresis of the purified CGA protein in SDS–acrylamide gels yields only a single band, corresponding to a polypeptide of about 22,000 daltons. In the absence of SDS, the protein has a molecular weight of about 45,000 (3); thus, it seems to be a dimer. CGA protein binds cAMP with a bimolecular formation constant, K_f, of about 0.6×10^5 liters/mol, as measured by equilibrium dialysis. cGMP inhibits the binding of cAMP to CGA protein and also antagonizes its stimulation of an *in vitro* system that synthesizes β-galactosidase (3).

During purification, about 90% of the activity that stimulates the *in vitro* synthesis of β-galactosidase is lost. It is not possible to measure cAMP binding throughout the preparation, because binding cannot be detected in crude extracts. However, the cAMP binding of partially purified or highly purified extracts indicates that most of the CGA protein molecules bind cAMP. The "activation" activity of CGA protein thus seems to be more labile than its cAMP-binding activity, an important fact in light of results to be presented below.

CGA protein is a DNA-binding protein

A membrane-filter technique was used to study DNA–CGA protein complex formation. This method consists of mixing unlabeled protein with radioactively labeled DNA and passing the solution through a nitrocellulose filter. Little of

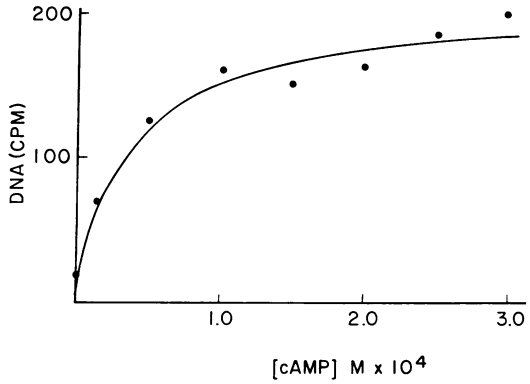

FIG. 3. Bound DNA as a function of cAMP concentration. CGA protein and λh80*dlac* [^{32}P]DNA were present at 0.01 and 0.05 μg, respectively. Only the concentration of cAMP in buffer IV was varied. Each filter received 0.4 ml, containing a total of 706 cpm.

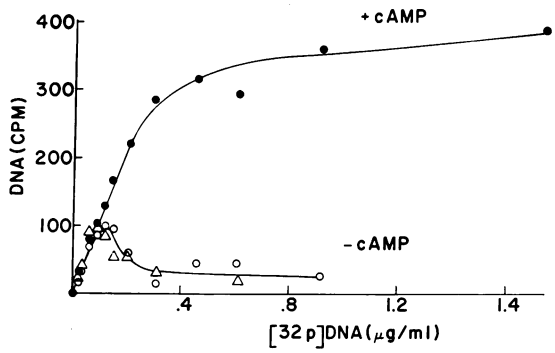

FIG. 4. Binding curves with CGA protein fixed and DNA varied. In these experiments, the CGA protein concentration was fixed at 0.04 μg/1.3 ml and the λh80*dlac* [^{32}P]DNA concentration was varied as indicated. In one set of experiments, (●), cAMP was present at a concentration of 1.6×10^{-4} M; in the others, (○) and (△), no cAMP was present. Each filter received 0.4 ml, containing [^{32}P]DNA at the concentration indicated. For experiments (●) and (○), the DNA used had a specific activity of 3.4×10^3 cpm/μg. For experiment (△), the DNA had a specific activity of 1.0×10^5 cpm/μg, but to facilitate comparison, the data have been normalized to the same specific activity as for the other curves. For each DNA concentration, the DNA retained in the absence of CGA protein was determined and subtracted. Only the binding due to CGA protein is shown.

the free, native DNA binds to the filter, but DNA bound to protein adheres strongly to the filter. The amount of DNA bound to protein is measured by the retention of radioactivity on the filter. This simple and sensitive technique has proven to be very effective in studies of interaction between DNA and the *lac* repressor (7, 9).

The binding of λh80*dlac* DNA as a function of CGA protein concentration is shown in Fig. 2. Binding occurs in buffer alone, but it is clearly stimulated by cAMP. In the early portion of the curve, when DNA is in excess of CGA protein (see below), five to ten times more DNA is bound in the presence of 3.7×10^{-4} M cAMP than in its absence. cGMP, on the other hand, eliminates DNA binding. The 5′-mononucleotides, AMP and GMP, have no significant effect on the binding curves, even at higher concentrations (6×10^{-4} M).

The dependence of DNA binding on cAMP concentrations was studied at a low weight ratio of CGA protein to DNA, where a marked enhancement of binding is seen. These results are shown in Fig. 3. Half-maximum binding of DNA to protein occurs at a cAMP concentration of about 2×10^{-5} M. The dependence of binding on cAMP concentration is quantitatively very similar to the dependence shown by the DNA-directed cell-free system for the synthesis of β-galactosidase of the *lac* operon (10), and is also in agreement with equilibrium dialysis data (3).

cAMP not only increases the amount of DNA–CGA protein complex, but it produces a complex that can be qualitatively distinguished from that formed in its absence. The relevant data are presented in Fig. 4, in which the binding is studied as a function of DNA concentration at a fixed concentration of CGA protein. The curves with and without cAMP are indistinguishable at very low DNA concentrations. As the DNA concentration is increased, however, a marked difference begins to appear between the two curves. Thus, when cAMP is present, the amount of filter-bound DNA increases to a plateau. In the absence of cAMP, a much lower maximum amount of filter-bound DNA is obtained, and there is a sharp decrease in DNA bound as the DNA concentration is further increased. The curve with cAMP present is consistent with the notion that one molecule of CGA protein binding to DNA (or several molecules binding cooperatively) is sufficient to make it adhere to the filter. In contrast, when no cAMP is present, the binding curve can be explained most simply by assuming that the binding of more than one molecule of CGA protein is required to make a molecule of DNA adhere to the filter. As the amount of DNA is increased, so that there is only one or no molecule of CGA protein bound to each DNA molecule, the DNA no longer adheres to the filter. Certainly the complex formed in the absence of cAMP is different from that formed when cAMP is present.

Since the curves in Fig. 2 are distinctly sigmoidal, it is worth pointing out that at least two interpretations are consistent with our present data: (*a*) there is adjacent cooperative binding, or (*b*) CGA protein reversibly dissociates into subunits, but only oligomeric CGA protein binds to DNA. Additional work will be required to distinguish between these two possibilities.

Only a small fraction of CGA protein is involved in the formation of the DNA complex

The data shown in Fig. 4 can be used to estimate the concentration of active CGA protein. That a plateau is reached means that essentially all molecules capable of binding DNA have done so. Yet, in the presence of cAMP, 0.4 μg/ml of DNA is sufficient (extrapolating from the early linear portion of the curve) to saturate 0.031 μg/ml of CGA protein. If we assume that one CGA protein per DNA molecule is sufficient to cause retention, 0.4 μg/ml of DNA of molecular weight 30×10^6 corresponds to only 6×10^{-4} μg/ml of a protein of molecular weight 45,000. Thus, only about 2% of the CGA protein is active in binding DNA. If there is cooperative binding of several molecules, then a higher percentage is active. Nevertheless, it seems reasonably certain that most molecules cannot bind to DNA. This result is consistent with our other observations that a considerable inactivation of CGA protein occurs during purification.

Lack of expected specificity

Mutants defective in CGA protein are viable; it is only certain catabolite-sensitive genes that can not be expressed (4). Thus, our working hypothesis was, and remains, that *in vivo* CGA protein binds specifically with catabolite-sensitive promoters. However, under the conditions of our binding experiments, we have not been able to demonstrate such specificity. It was hoped that the only catabolite-sensitive promoter in λh80*dlac* would be that of the *lac* operator. Thus, CGA protein might not have bound to DNA from λh80, but in fact we have observed no difference between the binding of [^{32}P]DNA from λh80*dlac* and [^{33}P]DNA from λh80. Unlabeled λh80 DNA also competes well with [^{32}P]DNA from λh80*dlac* for CGA protein. In such competition experiments, no significant differences were seen between the DNAs from λh80*dlac*, λh80, salmon sperm, *Clostridium perfringens*, *Micrococcus luteus*, and poly(dA-dT). *E. coli* rRNA and tRNA, however, compete much less than DNA, so there is specificity for DNA. Under the conditions of low ionic strength of the assay used here, the *lac* repressor also binds to DNA other than the *lac* operator, though some specificity for λh80*dlac* is still observed (11). The *lac* repressor shows greater specificity, though weaker binding, at higher ionic strengths, so a most pressing problem for the future will be to study DNA–CGA protein complex formation at higher ionic strengths.

However, since we have been unable to demonstrate specificity, one should keep in mind that specificity of action need not be based solely on specificity of binding. Perhaps only catabolite-sensitive promoters *require* CGA protein binding.

Conclusion: a new type of regulatory protein

The opposing effects of cAMP and cGMP on DNA–CGA protein complex formation parallels the effects of these cyclic compounds on the activation of *in vitro* synthesis of β-galactosidase from λh80*dlac* DNA. For the latter system, cAMP is a required activator and cGMP inhibits the activation by cAMP (10). Other studies have shown that cAMP and cGMP compete for binding sites on CGA protein. The activation and DNA-binding properties of CGA protein are also similar, in that both seem to be preferentially labile during purification when compared to cAMP-binding activity. Another regulatory protein, the *lac* repressor, has often shown a similar preferential loss of DNA-binding activity during purification (7). It thus seems almost certain that the biochemically significant action of cAMP is to stimulate CGA protein complex formation with DNA, and that this leads to both increased mRNA (2) and enzyme induction (3). It now appears that there are three distinct types of regulatory proteins: Type I includes the DNA-binding repressors such as *lac*, phage λ, and phage 434 (7, 12, 14). Type II is the σ factor, which interacts with RNA polymerase and activates transcription (13). Type III is CGA protein, an activator that interacts with DNA.

NOTE ADDED IN PROOF

Recent preparations of CGA protein, made with care and and assayed promptly, give linear rather than sigmoidal DNA binding curves and show an almost total dependence on cAMP. However, we are still unable to demonstrate specificity for the *lac* promoter.

We are indebted to Tetteh Blankson and to Joan Roberts for excellent technical assistance. This work was supported by grants from the National Institutes of Health (5-R01-GM-16648-03), the American Cancer Society (E-545), and the National Science Foundation (B01 8733-000) to G. Zubay, and by a grant from the National Institutes of Health (1-R01-HD-04420-01) to A. D. Riggs. G. R. thanks the National Institutes of Health for a predoctoral fellowship.

1. For reviews of the subject see Magasanik, B., in *Cold Spring Harbor Symp. Quant. Biol.*, **26**, 249 (1961); Magasanik, B., in *The Lac Operon*, ed. D. Zipser and J. R. Beckwith, Cold Spring Harbor, N.Y., Laboratory of Quantitative Biology (1970).
2. Pastan, I., and R. Perlman, *Science*, **24**, 339 (1970).
3. Zubay, G., D. Schwartz, and J. Beckwith, *Proc. Nat. Acad. Sci. USA*, **66**, 104 (1970).
4. Schwartz, D., and J. Beckwith, in *The Lac Operon*, ed. D. Zipser and J. R. Beckwith, Cold Spring Harbor, N.Y., Laboratory of Quantitative Biology (1970).
5. Silverstone, A. E., R. R. Arditti, and B. Magasanik, *Proc. Nat. Acad. Sci. USA*, **66**, 773 (1970).
6. Weber, K., and M. Osborn, *J. Biol. Chem.*, **244**, 4406 (1969).
7. Riggs, A. D., H. Suzuki, and S. Bourgeois, *J. Mol. Biol.*, **48**, 67 (1970).
8. Riggs, A. D., and S. Bourgeois, *J. Mol. Biol.*, **34**, 361 (1968).
9. Riggs, A. D., R. F. Newby, and S. Bourgeois, *J. Mol. Biol.*, **51**, 303 (1970).
10. Zubay, G., D. A. Chambers, and L. C. Cheong, in *The Lac Operon*, ed. D. Zipser and J. R. Beckwith, Cold Spring Harbor, N.Y., Laboratory of Quantitative Biology (1970).
11. Lin, S., and A. D. Riggs, *Nature*, in press.
12. Pirrotta, V., and M. Ptashne, *Nature*, **222**, 541 (1969).
13. Burgess, R. R., A. A. Travers, J. J. Dun, and E. K. F. Bautz, *Nature*, **221**, 43 (1969).
14. Gilbert, W., and B. Müller-Hill, *Proc. Nat. Acad. Sci. USA*, **58**, 2415 (1967).