Purification and Properties of D Protein: A Transcription Factor of *Escherichia coli*

(RNA polymerase/bacteriophage λ)

SUDHAMOY GHOSH AND HARRISON ECHOLS

Department of Molecular Biology, University of California, Berkeley, Calif. 94720

Communicated by Howard K. Schachman, October 5, 1972

ABSTRACT D factor is a protein that stimulates DNA transcription *in vitro* by RNA polymerase. For DNA from bacteriophage λ , we present evidence that D protein interacts with DNA to increase the frequency and improve the specificity of initiation by RNA polymerase.

Recent evidence indicates that purified RNA polymerase from *Escherichia coli* rarely (if ever) initiates RNA chains *in vitro* with the same fidelity achieved *in vivo* (1). DNA of phage λ is particularly useful for the study of this problem because of the availability of highly specific assays for "proper" transcription. *In vivo*, *E. coli* RNA polymerase probably initiates transcription of λ DNA at only two promoter sites—termed immediate-early—in the absence of phage-specified proteins; the resultant RNA synthesis from the early gene region of λ DNA is blocked by λ cI protein (" λ repressor") (2-4). Thus, properly initiated transcription *in vitro* should satisfy two criteria: the RNA should represent the early genes, and RNA synthesis should be repressed by cI protein.

As judged by these criteria, the fidelity of transcription of λ DNA is poor under typical conditions of RNA synthesis *in vitro* (5, 6). As a consequence, the demonstration of effective repression *in vitro* with purified components has required one of two approaches: a sedimentation analysis of the RNA product to separate properly initiated chains from the improperly initiated "background" (5), or the selective use of an inhibitor of initiation, rifampicin, to eliminate many improperly initiated chains (6). Early experiments on repression of RNA synthesis in a crude preparation indicated the presence of a host factor that could improve repression by cI protein (7); this result suggested that there might exist an *E. coli* transcription factor able to improve the specificity of initiation on λ DNA (6). This report describes the purification and properties of an *E. coli* protein that may be such a specificity factor. This protein—termed D factor—binds tightly to DNA, stimulates RNA synthesis from several templates, and improves the fidelity of transcription from λ DNA.

MATERIALS AND METHODS

Source of Proteins and DNA. Preparation of λ and λ b2 DNA was described (6). *E. coli* DNA was the gift of William Haseltine (8); P2 DNA was the gift of Kathy Barrett (9); T7 DNA was the gift of Anthony Cashmore (10). Calf-thymus DNA and "chicken blood" DNA were commercial products obtained from Worthington and Calbiochem,

respectively. Purifications of λ cI protein and RNA polymerase have been described (6); cI protein was purified through the phosphocellulose step (6).

Synthesis and Characterization of RNA. RNA was synthesized and characterized essentially as described (6). The standard reaction mixture (0.25 ml) contained: 0.4 mM each of ATP, UTP, and GTP, 0.04 mM [3 H]CTP, 125 mM KCl, 0.1 mM EDTA, 0.2 mM dithiothreitol, 20 mM Tris·HCl (pH 7.9), 1 mM potassium phosphate, 0.5 μ g of RNA polymerase, 2 μ g of DNA, 5 μ g of bovine-serum albumin, and 1 mM Mg(OAc)₂. Specific activity of [3 H]CTP was 10,000 cpm/nmol. The optimal stimulatory activity of D protein was obtained at 1-3 mM Mg⁺⁺ and 100-125 mM KCl, although stimulation was also observed at higher Mg⁺⁺ concentrations. Incubation was usually done at 26-27° for 15 min. RNA synthesis was generally measured by insolubility of the RNA product in trichloroacetic acid (11). The RNA product was also characterized by hybridization to λ DNA or to *hyb80imm434* DNA. Since *hyb80imm434* DNA is homologous to λ DNA mainly for the early gene region, RNA hybridizable to *hyb80imm434* DNA is defined operationally as early gene RNA (6).

Purification of D Protein. *E. coli* cells were grown in a Biogen and harvested (6). A polylysogen of strain CR302 (6) was generally used; however, the presence of λ prophages was not necessary for active D factor. Purification of D factor was guided by the DNA-binding capacity of the protein. The assay for DNA-binding was a membrane filter assay (12), done essentially as described (13). The standard binding mixture (0.25 ml) contained: 0.5 μ g of 32 P-labeled λ DNA, 10 mM Tris·HCl (pH 7.4), 14 mM 2-mercaptoethanol, 0.1 mM EDTA, 20 mM KCl, and 60 μ g of chicken-blood DNA.

For preparation of D protein, 80 g of frozen cells were suspended in 200 ml of "lysis buffer" [10 mM Tris·HCl (pH 8.0)-0.2 M KCl-2 mM EDTA-10 mM 2-mercaptoethanol-5% glycerol]; 90 mg of egg-white lysozyme was added; and the mixture was incubated at 10-15° for 30 min. All subsequent operations were done between 0 and 5°. To the lysed cell suspension, 7 ml of 1 M magnesium acetate and 12 ml of 25% Triton X-100 were added, and the mixture was centrifuged for 2.5 hr in a Spinco model L 30 rotor at 29,000 rpm. The supernatant fraction was collected, and the residue was resuspended in "lysis buffer" (minus lysozyme) and centrifuged again as above. The two supernatant frac-

tions were combined (270 ml) and subjected to a polyethylene glycol-Dextran phase separation procedure. NaCl (64 g) was added and dissolved, and a mixture of polyethylene glycol (18 g) and Dextran (12 g) was added and stirred gently until the solid material was dissolved. After centrifugation at $12,000 \times g$ for 10 min to separate the phases, the lower Dextran layer (112 ml) was diluted to 1.5 liters with 10 mM potassium phosphate (pH 7.0) and adjusted to pH 4.2 with acetic acid. After 3 hr, the precipitate was collected by centrifugation, suspended in 200 ml of 50 mM potassium phosphate (pH 7.5)–0.1 mM EDTA–0.7 mM 2-mercaptoethanol, adjusted to pH 7.5 with 1 N NaOH, and centrifuged at 10,000 rpm for 30 min. The supernatant fraction was made 70% saturated with solid $(\text{NH}_4)_2\text{SO}_4$; the precipitate was collected by centrifugation and dissolved in 100 ml of the phosphate-EDTA-mercaptoethanol buffer. The $(\text{NH}_4)_2\text{SO}_4$ precipitation and suspension were repeated once more. The resultant $(\text{NH}_4)_2\text{SO}_4$ fraction was centrifuged at $100,000 \times g$ for 1 hr; the supernatant fraction was dialyzed against "PC buffer" [10 mM potassium phosphate (pH 7.0)–0.1 mM EDTA–5 mM 2-mercaptoethanol–10 mM KCl–5% glycerol] and subjected to chromatography on a phosphocellulose column (50 ml volume). The column was washed with 150 ml of PC buffer with 0.1 M KCl, and D protein was eluted with the same buffer except that the KCl concentration was 0.6 M. Fractions showing DNA-binding activity were pooled (110 ml), protein was precipitated by the addition of 50 g of $(\text{NH}_4)_2\text{SO}_4$, and the precipitate was dissolved in 2 ml of "BP buffer" [10 mM potassium phosphate (pH 7.0)–14 mM 2-mercaptoethanol–0.2 mM EDTA–0.4 M KCl–5% glycerol] for further purification by Biogel P 150 column chromatography (See Fig. 1). For some experiments, the Biogel fraction was subjected to additional phosphocellulose chromatography. The protein was adsorbed to a 10-ml column in PC buffer with 20 mM KCl; the column was washed with 150 ml of PC buffer with 0.1 M KCl, and D protein was eluted by a linear gradient from 0.1 M–1 M KCl. D protein eluted at about 0.6 M KCl.

RESULTS

General properties of D protein

D factor has two properties of particular interest for the study of transcription: the capacity to bind tightly to DNA and the capacity to stimulate RNA synthesis. Our fractionation

TABLE 1. Effect of template on stimulation of RNA synthesis by D factor

DNA template	D protein ($\mu\text{g/ml}$)	Relative stimulation
λ	8	2.2
<i>E. coli</i>	15	2.5
P2	12	2.1
T7	4	1.5
Calf thymus	4	2.0

Concentrations of DNA and of RNA polymerase were 8 $\mu\text{g/ml}$ and 2 $\mu\text{g/ml}$, respectively, except for T7 DNA where the concentrations were 4 $\mu\text{g/ml}$ and 1.2 $\mu\text{g/ml}$, respectively. The relative stimulation is the ratio of RNA synthesis in the presence of D protein to that in the absence of D protein. Concentrations of D protein are those that produced maximal stimulation under the conditions used.

FIG. 1. Elution profile of D protein from a Biogel P150 column. The D-factor fraction from prior phosphocellulose chromatography (2 ml) was applied to a Biogel P150 column (36 cm $\times$ 2.5 cm) and fractionated with BP buffer (see Methods). The protein concentration of each fraction (3.6 ml/fraction) was estimated from $A_{280\text{nm}}$ values ($\circ$ — $\circ$). The DNA-binding activity was determined with 1- μl aliquots from each fraction ($\square$ — $\square$). Binding of 0.5 μg of λ [^{32}P]DNA (100% binding) corresponds to 2.0 on the ordinate. Stimulation of RNA synthesis was measured with 15- μl aliquots from each fraction ($\bullet$ — $\bullet$). The RNA polymerase activity in the absence of D factor corresponds to 1 on the ordinate.

studies indicate strongly that these two activities are associated with a single protein, which we term D protein because of the DNA-binding property. Fig. 1 shows the association of DNA-binding and stimulation of transcription after separation by size with a Biogel P-150 column. The peak fractions in this experiment revealed only a single major protein species, as judged by polyacrylamide gel electrophoresis in sodium dodecyl sulfate (14). Further fractionation of the protein from the Biogel column by gradient elution from a phosphocellulose column and by glycerol gradient sedimentation produced no separation of the two activities. In one purification, a substantial amount of an additional protein was found in the Biogel fraction; this protein was removed by phosphocellulose chromatography without any change in the activities of the D factor.

D factor is a small, heat-stable protein. By glycerol gradient sedimentation, we estimated a molecular weight of 12,000 (15). Both activities of D protein are retained after 5 min at 100° . Although we have studied mainly λ DNA, D protein can also bind to other DNA and stimulate RNA synthesis. Table 1 summarizes transcription results with DNA from several different sources. Although D protein stimulated transcription from all templates studied, the amount of D protein required for maximal stimulation varied with the DNA source.

Interaction of D protein with DNA

The capacity of D protein to stimulate RNA synthesis might result from an interaction with DNA or with RNA polymerase. The DNA-binding activity of D protein suggests a primary interaction with DNA. This concept is supported by two additional findings: (i) the critical stoichiometric feature for stimulation is the ratio of D protein to DNA, and (ii) the melting temperature of λ DNA is increased in the presence of D protein.

FIG. 2. Effect of D protein on RNA synthesis as a function of the ratio of DNA to polymerase. Incubation mixtures were as described in *Methods* except for the indicated variation in amounts of λ DNA (per 0.25-ml reaction mixture). RNA synthesis was measured by insolubility of the product in trichloroacetic acid. $\circ$ — $\circ$, RNA synthesis without D protein; $\blacksquare$ — $\blacksquare$, 1 μ g of D protein; $\bullet$ — $\bullet$, 2 μ g of D protein; $\blacktriangle$ — $\blacktriangle$, 6 μ g of D protein.

Fig. 2 shows the relationship between the stimulatory activity of D protein and the DNA/protein ratio. For 2 μ g of DNA, maximum RNA synthesis is observed with 2 μ g of D protein; for 6 μ g of DNA, maximum RNA synthesis is observed with 6 μ g of D protein. For lower protein/DNA ratios, stimulation is less; for high protein/DNA ratios, D protein inhibits RNA synthesis. We conclude that D factor probably affects transcription through an interaction with the DNA template. From the stoichiometry and the estimated molecular weight of D protein, an approximate ratio of 3000 D protein molecules per 1 λ DNA molecule is optimal for RNA synthesis.

Some DNA-binding proteins affect the thermal denaturation of DNA. For example, T4 gene-32 protein lowers the melting temperature (16) and histone raises it (17). Fig. 3 shows the effect of D protein on the thermal denaturation of λ DNA. The shift in melting temperature indicates a strong interaction between D protein and λ DNA. Since the con-

FIG. 3. Thermal denaturation profiles of λ DNA in the presence and absence of D protein. Relative hyperchromicity at 260 nm is shown as a function of temperature. Solutions were dialyzed in 10 mM Tris-HCl (pH 7.4)–1 mM EDTA overnight before the thermal denaturation experiment. 20 μ g of λ DNA per ml was used. $\circ$ — $\circ$, plus 12.5 μ g of D protein per ml; $\bullet$ — $\bullet$, no D protein.

TABLE 2. Effect of D factor on specificity of transcription—early gene RNA

Presence of D protein	cpm RNA hybridized to		
	λ DNA	hyb80-imm434 DNA	% RNA from early genes
—	280	155	55
+	560	465	83

hyb80imm434 DNA is complementary to λ RNA mainly for the early gene region. % RNA from early genes = $100 \times (\text{RNA hybridized to hyb80imm434 DNA})/(\text{RNA hybridized to } \lambda \text{ DNA})$. For this experiment, the concentrations of RNA polymerase, D factor, and λ b2 DNA were 4 μ g/ml, 6.4 μ g/ml, and 8 μ g/ml, respectively, and the specific activity of [3 H]CTP was 8×10^6 cpm/nmol.

formation of D protein at the high temperature is not likely to be the same as at the transcription temperature, it is difficult to interpret this result further.

Effect of D protein on initiation specificity of RNA synthesis

The capacity of D protein to stimulate RNA synthesis is compatible with three general possibilities: D protein might increase the initial start frequency, the propagation rate, or the release of terminated enzymes for restarts. Experiments with rifampicin, an inhibitor of initiation, indicate that D protein increases the initial start frequency. D protein stimulates RNA synthesis if it is added before rifampicin and nucleoside triphosphates (Fig. 4). In contrast, D protein exerts no stimulatory effect if it is added 1 min after the RNA synthesis reaction is started in the presence of rifampicin. Rifampicin should prevent restarts, but should allow D protein to affect propagation at any time; thus, the failure of

FIG. 4. Effect of D factor on initiation of RNA synthesis. The incubation mixture was as described in *Methods*, with the following modifications: 0.1 M KCl, 2 mM magnesium acetate, 12 μ g/ml of λ DNA. The reaction mixtures without nucleoside triphosphates were first incubated at 17° for 10 min to provide synchronous initiation, and then the reaction was started by the addition of triphosphates and rifampicin. The concentrations of D protein and rifampicin were 8 μ g/ml and 20 μ g/ml, respectively. $\circ$ — $\circ$, RNA synthesis without D protein; $\bullet$ — $\bullet$, RNA synthesis when D protein was added during the first incubation without nucleoside triphosphates; $\blacktriangle$ — $\blacktriangle$, RNA synthesis when D protein was added 1 min after the reaction was started by the addition of nucleoside triphosphates.

D protein to act after addition of rifampicin indicates that D protein increases the frequency of primary initiation events.

An increase in initiation frequency might result from more starts at the same sites used in the absence of D protein or from an effect of D protein on the specificity of initiation. Two observations on the effect of D protein support the latter possibility: (i) there is an increase in the fraction of RNA from the early gene region of λ DNA, and (ii) there is an increase in the fraction of RNA that can be repressed by *cI* protein.

Table 2 gives data on the region of DNA transcribed in the presence and absence of D protein. The fraction of RNA from the early genes is estimated by hybridization of the RNA product to DNA from phage *hyb80imm434*, which is homologous to λ DNA mainly for the early gene region (6). Essentially all of the stimulation in RNA synthesis by D protein corresponds to early gene RNA.

Table 3 gives data on the repressibility of RNA synthesized from λ DNA in the presence and absence of D protein; this experiment provides a more stringent criterion for initiation specificity because *cI* protein acts at or near the initiation stage of RNA synthesis (5, 6). The amount of RNA synthesis that escapes repression increases only slightly in the presence of D protein, whereas total RNA synthesis increases 2.2-fold. Data are included for a *λimm434* DNA template to show the specificity of the repression effect; although mainly λ DNA, *λimm434* DNA is resistant to the repression activity of λ *cI* protein because it lacks the specific binding sites for *cI* protein (5, 6, 18).

From the results presented in Fig. 4 and Tables 2 and 3, we conclude that D protein increases the frequency of initiation by RNA polymerase at the immediate-early promoter sites of λ DNA.

DISCUSSION

Possible mechanisms for D factor activity

D factor probably interacts with λ DNA to increase the frequency and improve the specificity of initiation by RNA polymerase. We can say nothing about the mechanism for this activity and will only comment briefly on the possibilities. The general possibilities for the mechanism of activity of D protein (as for any transcription factor) can be summarized by two questions: Does D protein activate "proper" sites or inactivate "improper" ones? Does the action of D protein lead to more binding by RNA polymerase at sites at which initiation invariably occurs, or does D protein increase the probability that a bound polymerase will initiate an RNA chain? The evident interaction of D protein with DNA is compatible with the possibility that D protein simply covers up sites at which improper initiation occurs. Alternatively, D protein might induce a structural alteration in DNA that favors the proper initiation sites—through a facilitation either of binding or of initiation by RNA polymerase.

Comparison of D factor with other factors that stimulate RNA synthesis

D protein resembles, in many respects, transcription factor H (19). Both proteins have low molecular weight, are thermostable, and stimulate RNA synthesis. However, H protein does not appear to exhibit the close association with DNA that characterizes D protein in both crude extracts and the purified state. D protein differs markedly from M factor

TABLE 3. Effect of D factor on specificity of transcription—repressibility

DNA template	Presence of D protein	cpm in RNA		% Repression of RNA
		—rep	+rep	
λ	—	1030	570	45
λ	+	2340	810	66
<i>λimm434</i>	—	1230	1300	None
<i>λimm434</i>	+	2850	3130	None

% Repression of RNA = $100 \times [\text{the difference between RNA in the absence (—rep) and presence (+rep) of } cI \text{ protein}] / (\text{RNA in the absence of } cI \text{ protein})$. For this experiment, the concentrations of RNA polymerase, D factor, λ b2 DNA, and *cI* protein were 4 $\mu\text{g/ml}$, 6.4 $\mu\text{g/ml}$, 8 $\mu\text{g/ml}$, and 8 $\mu\text{g/ml}$, respectively. Specific activity of [^3H]CTP was 4000 cpm/nmol.

(20), which has a much higher molecular weight, is thermostable, and probably acts through association with RNA polymerase (Ramakrishnan, T. & Echols, H., submitted to *J. Mol. Biol.*). D protein probably also differs from other recently described transcription factors (1, 21–24). D protein bears a superficial resemblance to the host factors active in the reaction of Q β RNA replicase (25).

Possible relevance of D factor to *in vivo* transcription

In the absence of compelling genetic evidence, one can question the relevance of any *in vitro* experiments to *in vivo* processes. There are two general arguments that are consistent with a role for D factor *in vivo*. First, the evident failure of RNA polymerase to perform *in vitro* as it does *in vivo* indicates that additional factors exist for the transcription process. The catabolite gene activator (CGA) protein is an example of such a factor (26, 27). Second, the strong interaction of D protein with DNA *in vitro* indicates that an association with DNA *in vivo* is plausible. D protein is associated with the nucleic acid (Dextran) phase in a phase partition and appears to have a strong positive charge, eluting from phosphocellulose at very high salt concentration (0.6 M). There is a superficial resemblance between D protein and histone in their low molecular weight, heat stability, capacity for strong DNA-binding, and effect on thermal denaturation of DNA (17).

On the negative side, nonprotein factors can influence the specificity of transcription of λ DNA *in vitro*. The superhelical density of circular λ DNA can produce large changes in the magnitude and specificity of RNA synthesis (Botchan, P. B., Wang, J. C. & Echols, H., in preparation). Quite accurate transcription can be achieved by selective inhibition of RNA synthesis by rifampicin (6), or by a combination of an extremely brief period of RNA synthesis and rifampicin (28).

Thus, although we find intriguing the possibility of a histone-like bacterial protein with a role in specificity of transcription, we cannot provide any compelling evidence for this idea.

We thank Peter and Michael Botchan for advice and Kathy Barrett, Anthony Cashmore, and William Haseltine for the gift of DNA samples. S. G. is on leave from the Bose Institute, Calcutta 9, India. This research was supported in part by GM 17078 from the Institute of General Medical Sciences.

1. Losick, R. (1972) *Annu. Rev. Biochem.* **41**, 409-446.
2. Echols, H. (1971) *Annu. Rev. Biochem.* **40**, 827-854.
3. Echols, H. (1972) *Annu. Rev. Genet.* **6**, in press.
4. Ptashne, M. (1971) in *The Bacteriophage λ* , ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 221-238.
5. Steinberg, R., & Ptashne, M. (1971) *Nature New Biol.* **230**, 76-80.
6. Wu, A. M., Ghosh, S., Echols, H. & Spiegelman, W. G. (1972) *J. Mol. Biol.* **67**, 407-421.
7. Echols, H., Pilarski, L. M. & Cheng, P. Y. (1968) *Proc. Nat. Acad. Sci. USA* **59**, 1016-1023.
8. Haseltine, W. A. (1972) *Nature* **235**, 329-333.
9. Barrett, K. J., Gibbs, W. & Calendar, R. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 2986-2990.
10. Chamberlin, M. J., McGrath, J. & Waskell, L. (1970) *Nature* **228**, 227-231.
11. Wu, A. M., Ghosh, S. & Echols, H. (1972) *J. Mol. Biol.* **67**, 423-432.
12. Riggs, A. D., Bourgeois, S., Newby, R. F. & Cohn, M. (1969) *J. Mol. Biol.* **34**, 365-368.
13. Wu, A. M., Ghosh, S., Willard, M., Davison, J. & Echols, H. (1971) in *The Bacteriophage λ* , ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 589-598.
14. Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406-4412.
15. Martin, R. G. & Ames, B. N. (1961) *J. Biol. Chem.* **236**, 1372-1379.
16. Alberts, B. & Frey, L. (1970) *Nature* **227**, 1313-1318.
17. Shih, T. Y. & Bonner, J. (1970) *J. Mol. Biol.* **48**, 469-487.
18. Ptashne, M. (1967) *Nature* **214**, 232-234.
19. Jacquet, M., Cukier-Kahn, R., Pla, J. & Gros, F. (1971) *Biochem. Biophys. Res. Commun.* **45**, 1597-1607.
20. Davison, J., Pilarski, L. M. & Echols, H. (1969) *Proc. Nat. Acad. Sci. USA* **63**, 168-174.
21. Travers, A. A., Kame, R. I. & Schleif, R. F. (1970) *Nature* **228**, 748-751.
22. Mahadik, S. P. & Srinivasan, P. R. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1898-1901.
23. Leavitt, J. C., Moldave, K. & Nakada, D. (1972) *J. Mol. Biol.* **70**, 15-40.
24. Murooka, Y. & Lazzarini, R. A. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 2336-2340.
25. Kuo, C.-Y. & August, J. T. (1972) *Nature New Biol.* **237**, 105-108.
26. Chen, B., de Crombrughe, B., Anderson, W. B., Gottesman, M. E., Pastan, I. & Perlman, R. L. (1971) *Nature New Biol.* **233**, 67-70.
27. Eron, L. & Block, R. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1828-1832.
28. Blattner, F. R. & Dahlberg, J. E. (1972) *Nature New Biol.* **237**, 227-232.