[11] DNA-Cellulose Chromatography

By BRUCE ALBERTS and GLENN HERRICK[1]

Many proteins that function in association with intracellular DNA recognize purified DNA as a substrate and bind tightly to it at physiological ionic strengths *in vitro*. When DNA-free, crude extracts are passed through a column consisting of DNA adsorbed onto an inert cellulose matrix ("DNA-cellulose"), these proteins are specifically retained. After an extensive wash to remove unbound proteins, the DNA-binding proteins are recovered in active form by elution with buffers containing either an elevated salt concentration, competing nucleic acids, or specific biological inducers. With an appropriate choice of conditions, most adventitious binding of non-DNA-related proteins to DNA can be avoided.[2-4] This type of procedure, independently derived in this laboratory[2] and by Dr. R. Litman,[5] has been successfully used in the purification of a bacterial RNA polymerase[2,6,7] and DNA polymerase,[2,5] and for the isolation of nucleases specific for ultraviolet-irradiated DNA.[8] Another type of application lies in the screening of crude extracts for new types of DNA-binding proteins, as witnessed by the discovery and subsequent isolation of gene 32-protein of bacteriophage T4, a new type of DNA-binding protein with an essential role in the replication and recombination of DNA.[9,10]

[1] The work presented in this article was supported by grants from the United States Public Health Service (GM-14927) and the American Cancer Society (E-510).

[2] B. M. Alberts, F. J. Amodio, M. Jenkins, E. D. Gutmann, and F. J. Ferris, *Cold Spring Harbor Symp. Quant. Biol.* **33**, 289 (1968).

[3] B. M. Alberts, *Fed. Proc., Fed. Amer. Soc. Exp. Biol.* **29**, 1154 (1970).

[4] The majority of the *E. coli* proteins which bind to DNA-cellulose in the presence of 0.15 *M* NaCl appear to have a DNA-related function, since they are absent in DNA-less mini-cells (unpublished experiments). Likewise, the percentage of the total proteins synthesized after T4 bacteriophage infection of *E. coli* which binds to DNA is 5- to 8-fold greater than in uninfected cells, consistent with the much larger fraction of genes in the T4 system coding for DNA functions.[2]

[5] R. M. Litman, *J. Biol. Chem.* **243**, 6222 (1968).

[6] R. R. Burgess, A. A. Travers, J. J. Dunn, and E. K. F. Bautz, *Nature (London)* **221**, 43 (1969).

[7] E. K. F. Bautz and J. J. Dunn, *in* "Procedures in Nucleic Acid Research" (G. L. Cantoni and D. R. Davies, eds.), Vol. 2, in press. Harper, New York.

[8] J. C. Kaplan, S. R. Kushner, and L. Grossman, *Proc. Nat. Acad. Sci. U.S.* **63**, 144 (1969).

[9] J. Tomizawa, N. Anraku, and Y. Iwama, *J. Mol. Biol.* **21**, 247 (1966).

[10] B. M. Alberts and L. M. Frey, *Nature (London)* **227**, 1313 (1970).

General Preparative Procedures

Cellulose

In our experiments with *Escherichia coli,* T4 bacteriophage-infected *E. coli,* and mammalian cell extracts, control DNA-free cellulose columns usually bind only from 0.1 to 0.2% of the applied protein as elutable counts. This low background requires, in addition to special treatment of the extract (see below), a highly purified grade of cellulose. For our work, we have used Munktell 410 cellulose (available in 500-g lots from Bio-Rad, Richmond, California). This cellulose is first washed several times with boiling ethanol to remove pyridine (remaining from extractions carried out by the manufacturer), and then quickly washed by suspension and filtering at room temperature successively in 0.1 *M* NaOH, 1 m*M* Na_3, EDTA, and 10 m*M* HCl solutions. After washing with H_2O to neutrality, the cellulose is lyophilized to dryness and stored. As a variation on this procedure, we have reduced the ethanol-extracted cellulose at 60° for 1 hour in 50 m*M* NaOH plus 2 g/liter $NaBH_4$[11] prior to extensive aqueous washes as described above. For our purposes, this latter treatment has proved to be unnecessary. Unwashed Whatman CF-11 cellulose has been used successfully for making DNA-cellulose for large-scale purification of *E. coli* RNA polymerase[7]; however, the quantity of elutable proteins binding nonspecifically to this cellulose is not known.

DNA-Cellulose

To make DNA-cellulose, a solution of DNA, at 1–3 mg/ml, 10 m*M* Tris·HCl, pH 7.4,[12] plus 1 m*M* Na_3 EDTA (Tris·EDTA), is transferred to a glass beaker. Clean dry cellulose is added with occasional stirring until the paste thickens (~1 g cellulose per 3 ml). The lumpy mixture is spread out in an evaporating dish and left at room temperature covered with gauze until dry. After grinding to a powder with a glass rod, any remaining water is removed by overnight lyophilization. The thoroughly dry powder is then suspended in about 20 volumes of Tris·EDTA and left at 4° for a day. After two quick washes to remove free DNA, the DNA-cellulose is stored as a frozen slurry in Tris·EDTA plus 0.15 *M* NaCl. About one-half of the original DNA is normally bound to the cellulose (500–1500 μg per packed milliliter), as determined by DNA released from an aliquot heated at 100° for 20 minutes in Tris·EDTA

[11] J. Porath and S. Hjertén, *in* "Methods of Biochemical Analysis" (David Glick, ed.), Vol. 9, p. 193. Wiley (Interscience), New York, 1962.

[12] All pH values reported for Tris buffers are those measured at 20°.

buffer. No DNA is bound to the cellulose without drying. Moreover, if the air-drying step is omitted and a damp paste is lyophilized instead, the fraction of DNA bound to the cellulose decreases by about half. It appears that intimate contact between DNA and cellulose is required for adsorption, and this is absent when local regions of DNA solution dry from the frozen state.

In preparing native DNA-cellulose, DNA of molecular weight about 10×10^6 daltons is probably optimal, since in this case the viscosity is low enough to permit quite concentrated solutions of DNA to be dried with cellulose. DNA-cellulose with a high DNA content may also be obtained using dilute DNA solutions, but repeated cycles of drying are then necessary.

To make single-stranded mammalian DNA-celluloses, a 2 mg/ml solution of DNA is divided into 10-ml aliquots in test tubes and denatured for 15 minutes at 100° in 10 m*M* K_2HPO_4, 1 m*M* Na_3 EDTA. After rapid quenching, the pooled aliquots are made 20 m*M* in Tris·HCl buffer, pH 7.4, mixed with cellulose, and processed as previously described. Because of the facility with which bacteriophage DNA's renature during drying, single-stranded bacteriophage DNA-celluloses must be made by an alternative procedure. The concentrated DNA is denatured with 0.5 *M* NaOH at 100° for 10 minutes, cooled to 0°, mixed with excess dry cellulose, and immediately lyophilized. The dried DNA-cellulose is neutralized by addition of an appropriate quantity of 1 *M* KH_2PO_4 and then rehydrated and washed in Tris·EDTA.[13]

When adsorption-temperature profiles are measured on DNA which has been adsorbed to cellulose and then removed by violent agitation (sonication or shear in 1 m*M* phosphate buffer, pH 7.4), native and denatured DNA's are found to remain double and single-stranded, respectively, throughout the drying process.

The complex formed by drying DNA with cellulose is slowly reversible in aqueous suspension, and at equilibrium very little DNA remains adsorbed. However, the rate of desorption at low temperature is insignificant compared to the duration of a typical chromatographic experiment. This is shown in Table I, where the time required for loss of one-half of adsorbed native calf thymus DNA from DNA-cellulose is listed as a function of temperature. Note that the half-time of adsorbed DNA at 37° is about 400 hours, whereas column fractionations require less than 12 hours at 4°. The stability of this native DNA-cellulose at 37° was relatively insensitive to changes in NaCl concentration (10 m*M* to 5 *M*) and pH in the range pH 5 to pH 9.

[13] Unpublished results of P. Christner, this laboratory, Princeton University (1969).

TABLE I
STABILITY OF NATIVE DNA-CELLULOSE[a]

Temperature (°C)	Time (hours) for loss of one-half adsorbed DNA	
	1 m*M* Na_3 EDTA, 10 m*M* sodium phosphate, pH 7.4	0.20 *M* NaCl, 1 m*M* Na_3 EDTA, 10 m*M* sodium phosphate, pH 7.4
37	400	400
50	110	120
70	11	20
90	<0.08	0.3

[a] Calf thymus DNA (Worthington Biochemicals) was adsorbed to washed, Munktell 410 cellulose. The resulting DNA-cellulose (560 μg/packed ml) was washed and resuspended in the buffer indicated. After various times of incubation, the DNA released into the supernatant of a low speed spin was determined by absorbance measurements. Estimated initial rates are shown. The first-order rate constant for DNA loss decreases with increasing incubation times, indicating that individual DNA molecules are bound to the cellulose with varying affinities [*Cold Spring Harbor Symp. Quant. Biol.* **33,** 289 (1968).]

Similar complexes between high molecular weight denatured DNA and cellulose are quite stable at NaCl concentrations above 50 m*M*, while R_{17} RNA[13a] and sonicated native DNA complexes with cellulose break down more rapidly. Drying tRNA with cellulose does not produce any detectable complex. It appears that a weak adsorption normally occurs at several sites on a single nucleic acid molecule, and that the cumulative effect of many such interactions holds long nucleic acid molecules to the cellulose more firmly than shorter molecules that have fewer sites of interaction.

In order for nucleic acid-cellulose complexes to be useful in the fractionation of proteins, the adsorbed nucleic acid must be accessible for protein binding. Since a brief pancreatic DNase treatment suffices to remove all adsorbed DNA from DNA-cellulose, nucleic acids adsorbed to cellulose remain freely accessible to proteins in solution. DNase treatment thus provides a convenient alternative to boiling for determination of the DNA content of DNA-cellulose.

[13a] Columns of bacteriophage R_{17} RNA-cellulose (300 μg/ml) have been used successfully for the fractionation of RNA-binding proteins from a RNase I deficient strain of *E. coli*. In this case, 10 m*M* $MgCl_2$ was added to all column buffers to help stabilize the RNA-cellulose matrix [N. Sigal, unpublished experiments, this laboratory, Princeton University (1970)].

Columns

For most analytical work, enough DNA-cellulose suspension for about one packed milliliter is diluted to 20 ml with 50 m*M* NaCl-containing buffer and left to settle, at room temperature without deaeration, into a 5-mm diameter column. Packed columns are transferred to 4° and, after rinsing for several hours, are loaded and run at 2 ml/hour, which is close to their gravity flow rate. DNA-cellulose columns do not appear to be damaged by running dry under gravity, and, if necessary, they can be re-used repeatedly. For preparative work, column beds of 100 ml and larger have been employed.

DNA-Free Crude Extracts

In initial experiments with DNA-cellulose columns, we investigated the chromatographic behavior of DNA and RNA polymerase enzymatic activities from crude extracts of *E. coli*.[2] When clarified cell-sonicates were chromatographed, a major portion of both DNA and RNA polymerase activities washed through the DNA-cellulose column, bound to the endogenous DNA of the extract. It appears, therefore, that DNA-free crude extracts are necessary for efficient use of this method. We currently prepare such extracts from both bacterial and mammalian cells by one of two alternative procedures: (1) pancreatic DNase treatment followed by extensive dialysis against an EDTA-containing buffer to remove divalent cations and inactivate DNase (which would otherwise digest DNA-cellulose), or (2) precipitation of the DNA with polyethylene glycol (PEG) and high salt[14] (to dissociate bound proteins), followed by dialysis of the protein-rich supernatant back to lower ionic strengths. The PEG procedure should be used if chromatography is carried out in the presence of Mg^{2+}. Both the pancreatic DNase and the PEG pass through a DNA-cellulose column; they are therefore readily removed from the fractions of interest.[14a]

DNase-Treated Extracts. To 5×10^{10} *E. coli* cells, or 10^7 mammalian tissue culture cells, 2 ml of extract buffer is added (10 m*M* $MgCl_2$, 2 m*M* $CaCl_2$, 1 m*M* Na_3 EDTA, 1 m*M* β-mercaptoethanol, and 20 m*M* Tris·HCl, pH 8.1 or 7.4. [For mammalian cell extracts, where the protein concentration is low, 500 μg/ml bovine serum albumin (BSA, Calbiochem, A

[14] K. R. Yamamoto, B. M. Alberts, R. Benzinger, L. Lawthorne, and G. Treiber, *Virology* **40**, 734 (1970).

[14a] An alternative procedure for nucleic acid removal which should allow chromatography in the presence of Mg^{2+} is DNA (and RNA) digestion with micrococcal nuclease; since this enzyme requires Ca^{2+}, it can be readily inactivated by use of the specific chelating agent ethyleneglycol-*bis*-(aminoethyl ether) tetraacetic acid (EGTA) [S. Spiegelman, A. Burny, M. R. Das, J. Keydar, J. Schlom, M. Travnicek, and K. Watson, *Nature (London)* **227**, 1029 (1970)].

grade) and 10% glycerol are also added.] After the cells are broken by sonication, bacterial extracts are digested for either 60 minutes at 10° or 15 minutes at 25° with 20 μg/ml of pancreatic DNase I (Worthington Biochemical, code D); with [^{3}H]thymidine-labeled DNA of *E. coli* and T4-infected *E. coli* cells, it was found that 96–99% of these DNA's are rendered acid soluble under these conditions. For mammalian extracts, 100 μg/ml of DNase I is used. Addition to radioactively labeled mammalian extracts of 0.4 ml of an unlabeled extract of T4 bacteriophage-infected *E. coli* facilitates complete (>98% acid soluble) pancreatic DNase digestion of mammalian DNA; without these added nucleases, 4% of the DNA remains acid insoluble.[14b]

It is sometimes desirable to carry out pancreatic DNase digestions at salt concentrations high enough to dissociate endogenous DNA + protein complexes. We therefore determined the effect of NaCl concentration on the activity of pancreatic DNase under our buffer conditions. The results are presented in Fig. 1. Although a marked inhibition of activity is seen at elevated NaCl concentrations,[15] we have found >98% digestion of endogenous mammalian DNA in the presence of 0.60 *M* NaCl after incubation for 20 minutes at 20° with 100 μg/ml of DNase. By increasing the DNase concentration and/or the time and temperature of incubation to compensate for the loss of DNase activity, conditions which should be suitable for complete digestion at higher salt concentrations can be derived from Fig. 1.

After digestion with DNase, extracts are clarified by centrifugation for 1 hour at 38,000 rpm in the Spinco No. 40 rotor and dialyzed against three 5-hour changes of a buffer containing 50 m*M* NaCl, 1 m*M* β-mercaptoethanol, 1 m*M* Na_3 EDTA, and 20 m*M* Tris·HCl, pH 8.1 or 7.4. Just prior to loading onto DNA-cellulose, the dialyzed extracts are recentrifuged to remove a light precipitate and made 10% (w/w) in glycerol. These last two steps are important for low blank values on DNA-free cellulose columns.

PEG-Treated Extracts. Bacterial or mammalian cells are broken by sonication, as previously described, in an extract buffer consisting of 1.7 *M* NaCl, 1 m*M* Na_3 EDTA, 1 m*M* β-mercaptoethanol, and 20 m*M* Tris·HCl, pH 8.1 or 7.4. After a low speed centrifugation to remove debris, 30% PEG 6000 is added to a final concentration of 10% (w/w). After 30

[14b] We have recently found a histone-cleaving activity in the T_4 extract. Thus, the enhancement of mammalian DNA digestion by the T_4 extract may be due not to "added nucleases," but to exposing DNA otherwise masked by histones. Because of histone proteolysis, as well as the danger of a low level of general proteolysis, we now prefer either the high-salt DNase I or the high-salt PEG method for mammalian DNA removal. The histone-cleaving activity is not inhibited by phenylmethyl sulfonylfluoride.

[15] M. Kunitz, *J. Gen. Physiol.* **33**, 363 (1950).

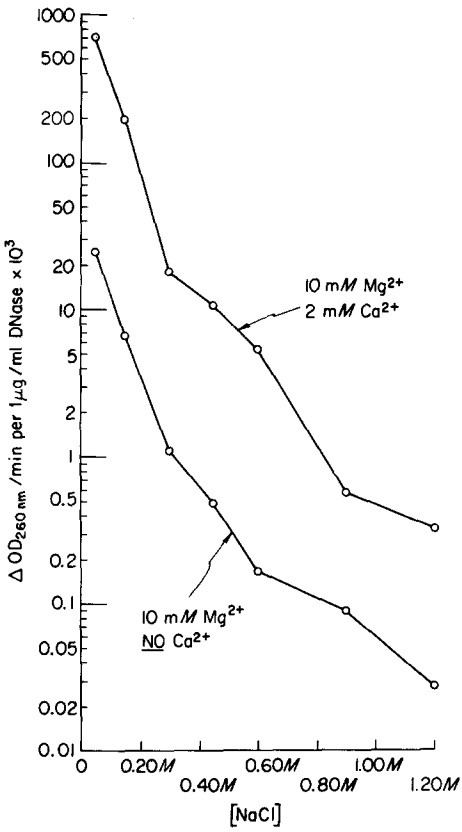

FIG. 1. Effect of NaCl and Ca^{2+} on the activity of pancreatic DNase I. The activities were measured by the rate of increase of DNA absorbance as described by M. Kunitz [*J. Gen. Physiol.* **33**, 363 (1950)]. In order to obtain measurable rates of hydrolysis, DNase concentrations from 0.1 to 25 μg/ml were used; the rates are normalized to a concentration of 1 μg/ml. The digestion mixture contained 50 μg/ml of native calf thymus DNA (Worthington), 10 m*M* $MgCl_2$, 1 m*M* Na_3 EDTA, 10% glycerol, 500 μg/ml bovine serum albumin, 20 m*M* Tris·HCl, pH 7.4 and, when noted, 2 m*M* $CaCl_2$. The striking stimulation of DNase by Ca^{2+} seen here (up to 30-fold) is apparently not found under all reaction conditions [J. Shack and B. S. Bynum, *J. Biol. Chem.* **239**, 3843 (1964)]. In another experiment (not shown) the fraction of DNA hydrolyzed per unit time was independent of the initial DNA concentration (between 50 and 200 μg/ml); thus, the exact concentration of the extract is not crucial in the DNase procedure described in the text.

minutes at 0°, the extract is clarified by a 15 minute centrifugation at 10,000 *g*. The supernatant is dialyzed and prepared for DNA-cellulose chromatography as described for DNase-treated extracts.

If desired, different conditions can be used for the PEG precipitation of DNA: the effect of varying the PEG and NaCl concentrations on the

separation of DNA and proteins is shown in Fig. 2. With 1.7 *M* NaCl and 10% PEG, 98% of the DNA, but only 12% of the soluble protein, precipitates; at lower salt concentrations decreased separation between DNA and protein is observed.

Chromatography

Standard Procedure

Prior to use, DNA-cellulose columns are rinsed for several hours at 4° with a standard buffer containing 50 m*M* NaCl, 1 m*M* Na_3 EDTA, 1 m*M* β-mercaptoethanol, 10% (w/w) glycerol, 100 μg/ml BSA, and

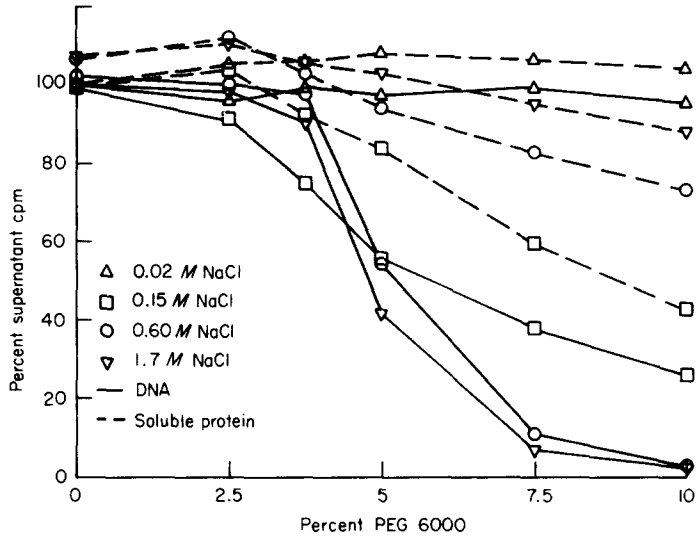

FIG. 2. Precipitation of DNA as a function of polyethylene glycol and salt concentrations. *Escherichia coli* B were infected with T4 bacteriophage and [^{3}H] thymidine and [^{14}C]leucine were used to label T4 DNA and early proteins, respectively. An extract was prepared as described for PEG precipitation, except that the concentration of NaCl was 50 m*M* and the extract was clarified by centrifugation for 1 hour at 38,000 rpm in a Spinco No. 40 rotor. To 100 μl of supernatant, 200 μl of 1.5 × concentrated PEG-NaCl solution was added. After 15 minutes at 0°, a visible precipitate was pelleted at 10,000 *g* for 15 minutes. Samples of the supernatants were applied to GF/A glass filters (Whatman) and rinsed with cold 5% trichloroacetic acid, followed by cold ethanol and warm diethylether rinses [D. Pettijohn, *Eur. J. Biochem.* 3, 25 (1967)]. The samples were counted by standard scintillation techniques, and the data were corrected for overlaps. Treatment of a 10-fold dilution of the extract at 0.60 *M* NaCl (not shown) gave less severe loss of protein (12% vs. 35% at 10% PEG) and nearly equivalent precipitation of DNA. High molecular weight DNA's precipitate at lower PEG concentrations than the sonicated DNA shown here; a cleaner separation of DNA and protein should therefore be obtainable after gentle lysis procedures.

20 m*M* Tris·HCl (pH 8.1 or 7.4) in quartz-distilled water.[15a] This buffer is also used for all subsequent rinse and elution steps. After up to 5 ml of DNA-free extract are loaded onto a 1-ml column, unbound proteins are removed with a rinse of about 5 ml (5 column bed-volumes). Bound proteins can then be eluted by increasing the salt concentration in the standard buffers; either step or gradient elution can be used. All operations are carried out at 4° at a flow rate of 2 ml/hour, and fractions are collected every 15 minutes.

For routine chromatographic analyses, salt elution steps of 2–3 column bed-volumes of buffer each have been used. In order to illustrate the type of results obtained, an elution profile for [^{3}H]leucine-labeled mammalian proteins chromatographed on calf thymus DNA-cellulose is shown in Fig. 3. A mixture of native and denatured DNA's was used to ensure retention of as many kinds of proteins as possible. [The vast majority of detectable *E. coli*, T4 bacteriophage and mouse cell DNA-binding proteins bind to DNA irrespective of its source; therefore, for screening experiments we have often used calf thymus DNA, which is commercially available (Worthington Biochemicals)]. In Fig. 3, 6 elution steps of 0.10 *M*, 0.15 *M*, 0.25 *M*, 0.60 *M*, 1.0 *M*, and 2.0 *M* NaCl were employed on a 4-ml column. It can be seen that a discrete peak of proteins appears at each salt concentration. Analysis by sodium dodecyl sulfate (SDS)-containing polyacrylamide gel electrophoresis reveals that there are at least 4 or 5 distinct protein bands in a peak, and that, by and large, different proteins are eluted at each salt concentration.[16] Thus, as observed earlier for bacterial systems,[2,3] salt elution results in a significant fractionation of the many mammalian proteins binding to DNA-cellulose.

Depending on the extract chromatographed, 0.2–0.8% of the radioactive protein applied remains on the DNA-cellulose as counts nonelutable with salt after a 2.0 *M* NaCl rinse. Any proteins remaining on the DNA at

[15a] In order to study the effect of pH as a variable in DNA-cellulose chromatography, two parts of a doubly-labeled, radioactive extract of mouse and bacteriophage T_4-infected cells were loaded onto identical native DNA-cellulose columns at pH 7.2 and 9.6, respectively (Tris-phosphate, pH determined at 4°). After the columns were rinsed to background, they were partially eluted by changing the pH at constant NaCl concentration. From the column loaded at pH 7.2, 26% of the mouse-cell and 22% of the T_4-induced proteins originally bound were eluted by a rinse with pH 9.6 buffer. From the column loaded at pH 9.6, however, only 6.6% of the mouse-cell and 3.7% of the T_4-induced proteins originally bound were eluted by a rinse with pH 7.2 buffer. Polyacrylamide gel electrophoresis patterns of the proteins remaining on the two columns after these pH changes were quite similar, while the gel patterns of the proteins eluted by the pH shifts were very dissimilar. Thus, it appears that, while most of the DNA-binding proteins bind over a wide pH range, the binding of a minority of species is pH-sensitive.

[16] Unpublished data of G. Herrick (1970).

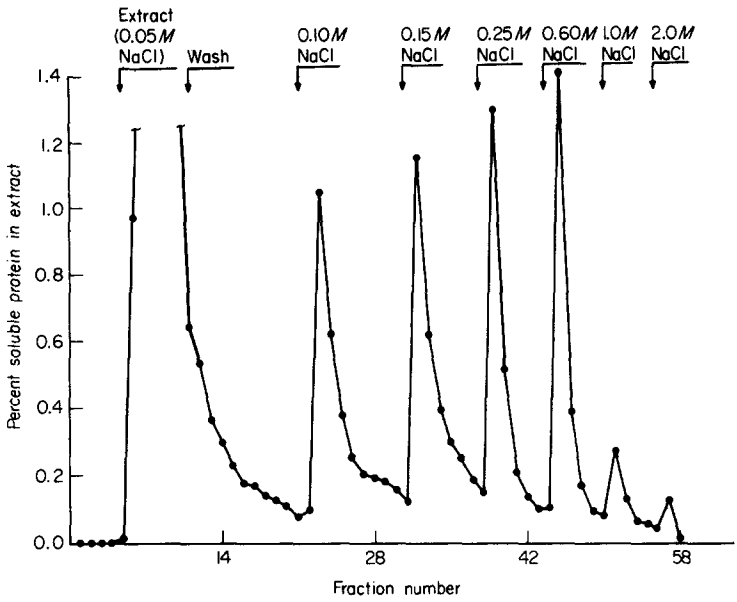

FIG. 3. DNA-cellulose chromatography of mouse fibroblast (3T3) proteins. A 2-ml extract, representing 10^7 fibroblasts (labeled in tissue culture with [^{3}H]leucine for 30 hours at 37°) was prepared for chromatography by the DNase method at pH 7.4 (see p. 202), and applied to a column consisting of 4 ml packed volume of calf thymus DNA cellulose (1.7 mg each of native and denatured DNA). An overlaying, DNA-free cellulose pad of 4 ml packed volume was removed just after the extract had passed through the column. Six NaCl elution steps were then performed as indicated.

this point should be recoverable by washing the column with 10 m*M* $MgCl_2$, 20 m*M* Tris·HCl, pH 7.4, followed by an incubation and rinse with 20 μg of pancreatic DNase per milliliter in this buffer to remove the DNA and its bound proteins from the cellulose. Most of the residual proteins are not removed by this step, however, and seem to be firmly bound to the cellulose.[17]

When DNA-cellulose chromatography is carried out for the first time on a particular type of extract, several important controls should be run:

1. A DNA-free cellulose column should be used in parallel with the

[17] The fact that the vast majority of DNA-binding proteins thus far observed may be reversibly eluted from a DNA column at sufficiently high ionic strengths implies that electrostatic attractions play a crucial role in their interaction with DNA. While many DNA-binding proteins carry a net negative charge in the physiological pH range, they apparently share a positively charged groove designed to fit the DNA phosphate-sugar chain.

DNA-cellulose column to measure spurious binding. (High blank values will be obtained unless particulate matter has been removed completely from the extract.) When low blank values are crucial, as for the detection of minor protein components, it is helpful to cover the DNA-cellulose with a column of DNA-free cellulose, which is removed with a Pasteur pipette after the extract has been loaded.

2. The DNA content of the DNA-cellulose should be compared before and after the chromatographic fractionation. In our experience, providing that divalent cations are efficiently removed from DNase-treated extracts, more than 95% of the DNA is left on DNA-cellulose after each run.

3. To determine whether the quantity of DNA-cellulose used and the contact time are sufficient for a given extract, the breakthrough from an initial DNA-column should be rechromatographed on a second column, and the DNA-binding fractions obtained assayed for the proteins of interest.

Variations

Detection of DNA-Associated Proteins with Low DNA Affinity. A very substantial DNA affinity is required for a DNA-associated protein to be recognized as such by these procedures. Any protein with a dissociation constant for DNA (K) greater than some maximum value will be poorly retarded by DNA-cellulose and be lost in the 50 mM NaCl rinse. The classical theory of partition chromatography of Martin and Synge[18] enables a simple relationship to be derived between R, the relative movement of protein and solvent through a DNA-cellulose column, and

$$K = \frac{(\text{free protein})(\text{free DNA-sites})}{(\text{DNA-bound protein})} = K_{\text{eff}}\,(\text{free DNA-sites}).$$

Note that an "effective" equilibrium constant, K_{eff}, has been introduced; K_{eff} is equivalent to the partition coefficient for the protein in a "theoretical plate" consisting of an imaginary DNA-containing stationary phase immobilized on the column and a mobile solvent phase of equal volume [$K_{\text{eff}} = \alpha(A_S/A_L)$ in the notation of Martin and Synge[18]]. Following the published derivation, one obtains

$$\frac{1}{R} = \frac{A_L}{A}(K_{\text{eff}} + 1)$$

where A_L/A = actual fraction of total column volume occupied by the aqueous phase; and R = (movement in the column of the position of

[18] A. J. P. Martin and R. L. M. Synge, *Biochem. J.* **35**, 1358 (1941).

maximum protein concentration)/(simultaneous movement of the surface of the developing fluid in the empty part of the tube above column). From this result, it follows that the position of maximum protein concentration will just leave the bottom of the column when the number of column liquid-volumes which have followed the protein onto the top of the column is equal to $1 + K_{eff}$ [alternatively written as $1 + K/$(free DNA sites)]. Application of this equation to a chromatogram, such as the one in Fig. 3, assuming one DNA binding site per 10 base pairs, reveals that those proteins remaining on the column after the 50 mM NaCl rinse must have a $K < 10^{-5}$ M (binding energy > 7 kcal/mole). Many proteins have an affinity only for selected regions of the DNA, and for them the estimate of one binding site per 10 base pairs is much too high. These proteins must of course bind even more tightly in order to be detected.

It is possible that proteins having low DNA affinities under physiological salt conditions would bind strongly enough to DNA to be detected if DNA-cellulose columns were loaded and rinsed as in Fig. 3, but without NaCl present. However, many non-DNA-related, positively charged proteins would also bind to the DNA at such low ionic strengths and reduce the value of the fractionation.[19] Another possible approach would be to rinse a small volume of extract in 50 mM NaCl buffer through a long thin column of DNA-cellulose. The protein fractions which trail behind the main breakthrough peak from such a column would represent proteins with a weak DNA affinity. The potential resolution of this method could be enhanced by increasing the concentration of DNA on the cellulose, for example, by using multiple cycles of drying during DNA-cellulose preparation.

The Use of Cofactors for Binding and Elution. Some important DNA-binding proteins may represent too small a fraction of the total cell protein to be noticeable against the background of major proteins binding to DNA-cellulose. This is certainly the case with regulatory proteins, such as the lactose and bacteriophage lambda repressors.[20,21] Detection of these proteins in DNA-cellulose eluates demands either specific assays for them or special purification methods such as selective DNA-binding and/or elution from the DNA with the appropriate biological inducer. Likewise other proteins which function on DNA require cofactors (e.g., ATP, Mg^{2+}, and *S*-adenosylmethionine[22]) in order to bind to DNA. Such proteins will appear in DNA-cellulose eluates only if the appropriate co-

[19] B. H. J. Hofstee, *Biochim. Biophys. Acta* **91**, 340 (1964).
[20] W. Gilbert and B. Müller-Hill, *Proc. Nat. Acad. Sci. U.S.* **58**, 2415 (1966).
[21] M. Ptashne, *Nature (London)* **214**, 232 (1967).
[22] R. Yuan and M. Meselson, *Proc. Nat. Acad. Sci. U.S.* **65**, 357 (1970).

factors are added to the extract, and they might be specifically elutable by cofactor removal.

The Use of Solubilizing Detergents. Neither 1% Triton X-100, 1% Brij 58, nor 1% Tween 80 changes the DNA-binding of bacterial proteins when added to crude soluble extracts, and to all rinse and elution buffers.[23] The DNA-cellulose remains intact throughout this procedure. These nonionic detergents are therefore useful solubilizing agents for particulate protein fractions, which can then be screened for possible DNA-binding proteins that are not normally soluble. Prior to analysis by electrophoresis on SDS-containing polyacrylamide gels, Triton X-100 is removed from DNA-cellulose eluates by a single extraction at 20° with an equal volume of isoamyl alcohol.

The Use of Competing Nucleic Acids. Proteins exhibiting selective DNA affinity can be isolated and identified by the use of competing nucleic acids in DNA-cellulose chromatography. Since free nucleic acids in solution are not retained by DNA-cellulose, proteins can be partitioned between one type of nucleic acid in solution (mobile phase) and DNA immobilized on the cellulose (stationary phase). With an appropriate choice of conditions, the proteins remaining bound to the DNA-cellulose will be those preferring the DNA on the column. These are then recovered by salt elution.[2]

Such experiments have been carried out by two different procedures: the competing nucleic acid can be added to the extract and rinse buffers, or it can be used for the first elution step after the usual 50 m*M* NaCl rinse. An advantage of the latter method is that DNA-binding proteins of both specific and nonspecific classes are recovered in purified form from the same DNA-cellulose column. However, it should be noted that a protein with a very low dissociation constant will exchange slowly in either type of competition experiment. Thus, regardless of its relative affinity for the two nucleic acids used, it will be eliminated with the nonspecific DNA-binding proteins if the competing nucleic acid is added to the extract, but remain bound with the specific DNA-binding proteins if the competing nucleic acid is instead used for the first elution.

Using RNA as the competing nucleic acid, one can separate a class of proteins with a high specificity for DNA from other proteins which bind to either RNA or DNA[24] (see, for example, Table II and Fig. 5, below). Other possibilities include competition between native and de-

[23] Unpublished results of B. M. Alberts and F. Amodio (1968).

[24] An RNA affinity under these conditions does not preclude a DNA function *in vivo*, since, for example, *E. coli* RNA polymerase can bind to both RNA and DNA *in vitro* [C. Yegian and M. W. Konrad, *J. Mol. Biol.* **16**, 94 (1966)].

natured DNA's and competition between two DNA's from different organisms.[25]

As an alternative method for isolating DNA-binding proteins having selective DNA affinities, two separate DNA-cellulose columns can be run in series. In this case, the initial column is designed to screen out DNA-binding proteins lacking specificity for the type of DNA used for the second column. This simple approach suffers from the fact that, unlike the partition method, it requires nearly absolute DNA-binding specificity on the part of the proteins to be isolated.

Applications

In this section, some general examples of the use of DNA-cellulose chromatography will be presented to illustrate the type of results obtainable.

Preparative-Scale Purification of DNA-Binding Proteins

DNA-cellulose chromatography should serve as a mild and effective first step toward the complete purification of many DNA-binding proteins. For example, preparative-scale procedures have been employed for purification of T4 bacteriophage gene 32-protein,[10] *E. coli* RNA polymerase,[7] and *Micrococcus luteus* DNA polymerase[5] on single-stranded, double-stranded, and nicked-irradiated double-stranded DNA-celluloses, respectively. In the case of T4 gene 32-protein, 200 ml of dialyzed DNase-treated crude extract from 50 g of T4-infected cells are applied at 100 ml/hour to a 20-ml denatured calf thymus DNA-cellulose column. Elution between 0.60 *M* and 2.0 *M* NaCl, followed by DEAE-cellulose chromatography, gives 8 mg of homogeneous protein in 70% overall yield.[10]

Identification of New DNA-Binding Proteins with Specific Biological Roles

DNA-cellulose chromatography enables selective purification of proteins which function on DNA, without requiring a specific functional assay. Thus, DNA-binding proteins whose functions are unknown can be recognized and catalogued according to some convenient physical

[25] The regulatory proteins that control DNA expression are expected to require a special nucleotide sequence for complex formation with DNA *in vivo* and might be expected to recognize the DNA containing this sequence *in vitro*. Selective DNA-binding has in fact been observed with both the lactose and phage lambda repressors,[20,21] and DNA-cellulose chromatography has recently enabled detection of DNA-binding proteins with apparent sequence specificity from both rat [L. J. Kleinsmith, J. Heidema, and A. Carrol, *Nature* (*London*) **226**, 1025 (1970)] and frog [M. Crippa, *Nature* (*London*) **227**, 1138 (1970)].

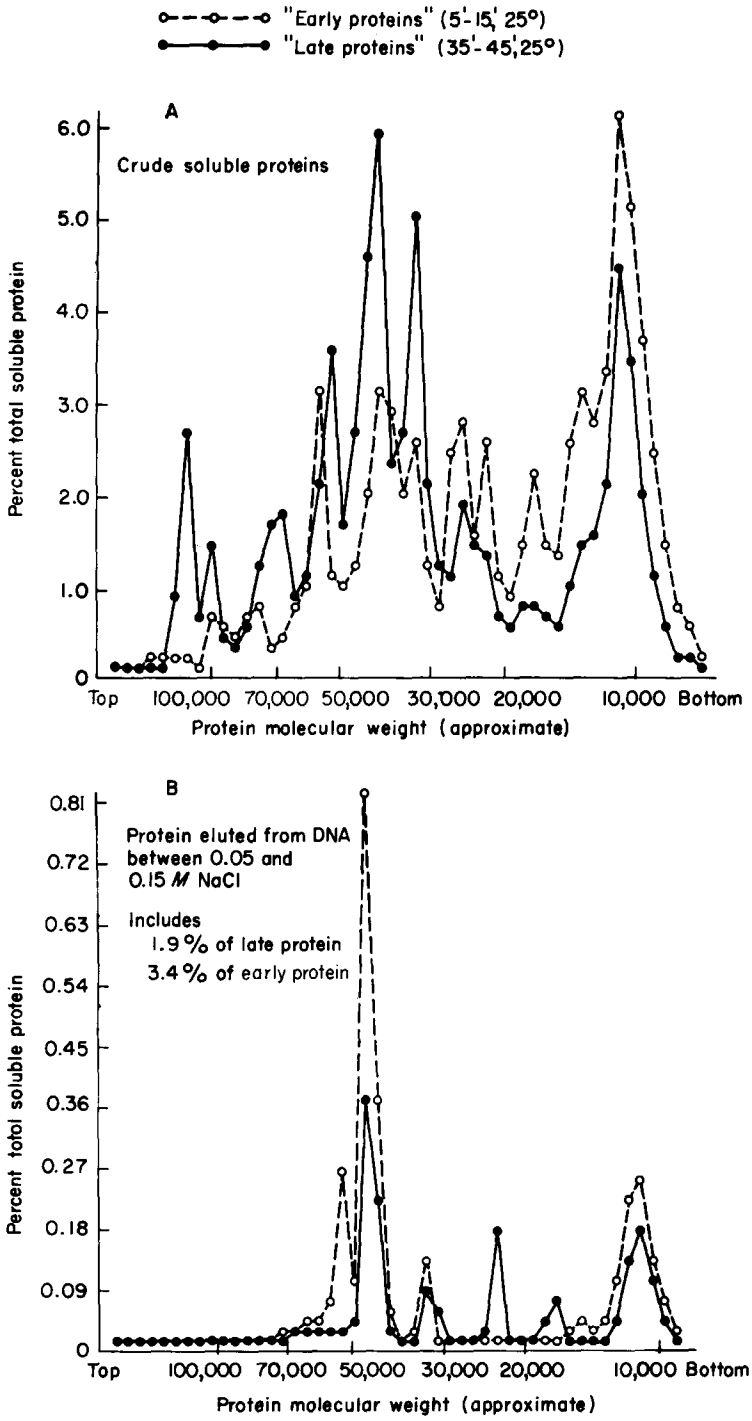
"Early proteins" (5'-15', 25°)
"Late proteins" (35'-45',25°)
A
Crude soluble proteins
Percent total soluble protein
6.0
5.0
4.0
3.0
2.0
1.0
0
Top
100,000
70,000
50,000
30,000
20,000
10,000
Bottom
Protein molecular weight (approximate)
B
Protein eluted from DNA between 0.05 and 0.15 M NaCl
Includes
1.9 % of late protein
3.4 % of early protein
Percent total soluble protein
0.81
0.72
0.63
0.54
0.45
0.36
0.27
0.18
0.09
0
Top
100,000
70,000
50,000
30,000
20,000
10,000
Bottom
Protein molecular weight (approximate)

FIG. 4. A and B.

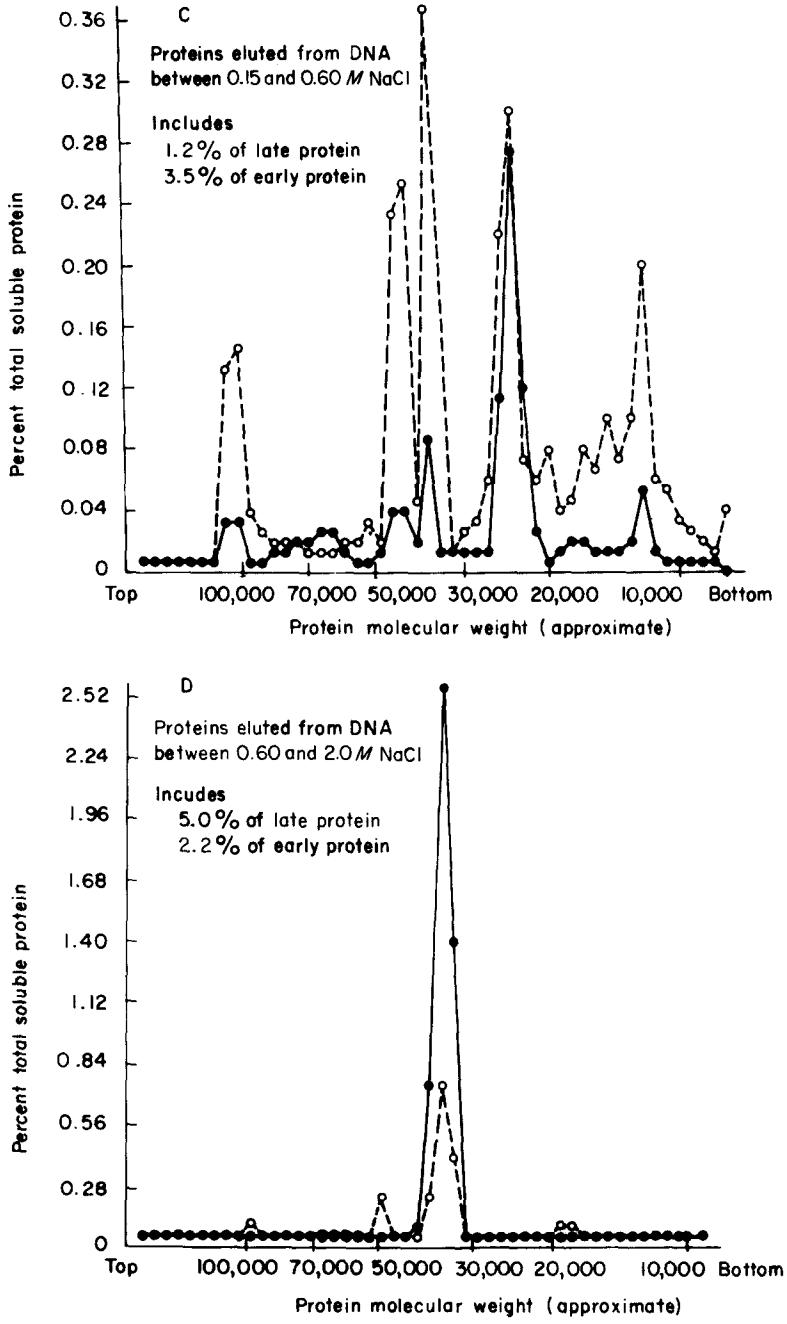
C
Proteins eluted from DNA
between 0.15 and 0.60 M NaCl
Includes
1.2% of late protein
3.5% of early protein
Percent total soluble protein
0.36
0.32
0.28
0.24
0.20
0.16
0.12
0.08
0.04
0
Top
100,000 70,000 50,000 30,000 20,000 10,000
Bottom
Protein molecular weight (approximate)
D
Proteins eluted from DNA
between 0.60 and 2.0 M NaCl
Incudes
5.0% of late protein
2.2% of early protein
Percent total soluble protein
2.52
2.24
1.96
1.68
1.40
1.12
0.84
0.56
0.28
0
Top
100,000 70,000 50,000 30,000 20,000 10,000
Bottom
Protein molecular weight (approximate)

FIG. 4. C and D.

property (e.g., electrophoretic mobility on polyacrylamide gels). If a mutant organism with a defect in DNA function has an alteration in one of its proteins which binds to DNA, the corresponding mutant and normal proteins can be purified and a detailed study made of their interactions with DNA. A combination of genetic and biochemical data concerning this protein is then used to unscramble the molecular events which underlie its biological role.

This type of analysis has been carried out with T4 bacteriophage-infected *E. coli*.[2,3] In order to construct a catalog of the total T4 DNA-binding proteins, proteins synthesized early after infection were ^{3}H-labeled, and those synthesized late were ^{14}C-labeled. The mixed extract was subjected to DNA-cellulose chromatography. Electrophoretic analysis of the separate DNA-binding fractions eluted with NaCl are given in Fig. 4. About 20 distinct protein bands can be seen among the DNA-binding proteins. By examination of a variety of amber and temperature-sensitive T4 mutants, the major protein eluting with 2.0 *M* NaCl (Fig. 4D) has been identified as the product of T4 gene 32 while the 110,000 mol. wt. component in Fig. 4C appears to be the product of T4 gene 43 (T4 DNA polymerase).[26] We are currently screening available mutants in order to identify as many of the remaining DNA-binding proteins as possible as the products of known bacteriophage genes.

[26] A. DeWaard, A. V. Paul, and I. R. Lehman, *Proc. Nat. Acad. Sci. US.* **54**, 1241 (1965).

FIG. 4. SDS-containing polyacrylamide gel electrophoresis of double-labeled T4 bacteriophage-induced "early" and "late" DNA-binding proteins. In this experiment, *Escherichia coli* cells infected with T4 bacteriophage at 25° were labeled from 5 to 20 minutes post-infection with [^{3}H]leucine ("early proteins") and from 35 minutes to 45 minutes post-infection with [^{14}C]leucine ("late proteins"). Extracts were prepared by the DNase procedure, and chromatographed at pH 8.1 on an equal mixture of native T4 and denatured calf thymus DNA-cellulose. Samples from the DNA-binding peaks obtained were concentrated to 0.2 ml by vacuum dialysis, made 1% in SDS and 0.14 *M* in β-mercaptoethanol, heated at 100° for 2 minutes, and, electrophoresed on gels (10% acrylamide, 0.3% ethylene diacrylate in 0.1% SDS, 9 cm × 0.5 cm) at 3 mA/gel for 16 hours [A. L. Shapiro, E. Viñuela, and J. V. Maizel, *Biochem. Biophys. Res. Commun.* **28**, 815 (1967)]. Two-millimeter gel slices were depolymerized by shaking for 12 hours at 32° in 1-dram vials containing 0.5 ml of 0.50 *M* NaOH and were counted after addition of 4 ml of Bray's scintillation fluid [G. A. Bray, *Anal. Biochem.* **1**, 279 (1960)] containing 20 g/liter of Cab-o-sil (Grade M-5, Cabot Corp., Edison, New Jersey). Data were corrected for channel overlaps. The approximate molecular weight scale was constructed from the migrations of proteins of known molecular weights (BSA, pepsin, trypsin, and lysozyme) run on a parallel gel. From B. M. Alberts [*Fed. Proc., Fed. Amer. Soc. Exp. Biol.* **29**, 1154 (1967)].

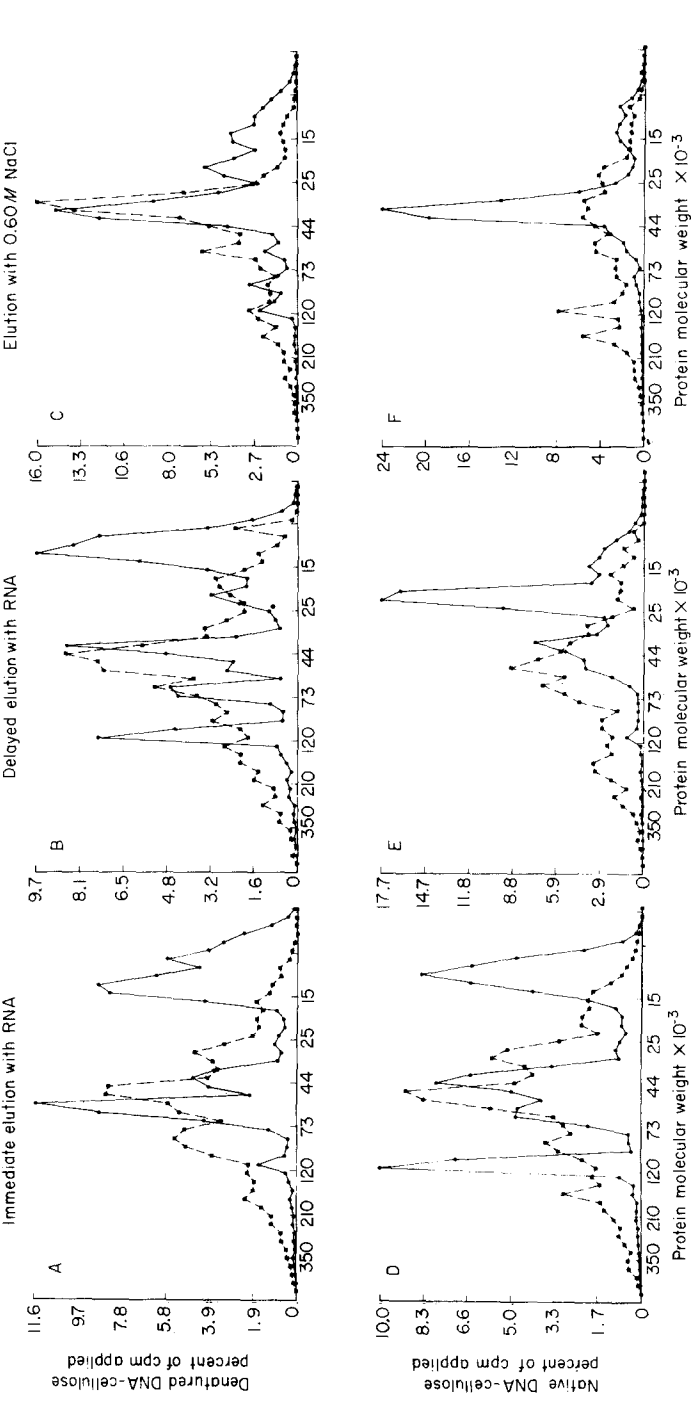

FIG. 5. SDS-containing polyacrylamide gel electrophoresis of **T4** bacteriophage and mouse fibroblast DNA-binding proteins. Samples from the experiment described in Table II were analyzed as in Fig. 4, except that electrophoresis was through gels of larger pore size (5% acrylamide, 0.135% ethylene diacrylate) for 12 hours at 2 mA per gel. Solid lines are ^{14}C-labeled T4 proteins; dashed lines indicate ^{3}H-labeled mouse proteins.

Classification of DNA-Binding Proteins by Selective Nucleic Acid Affinity

Because of a lack of both mutants and specific functional assays, the pattern of DNA-binding proteins in mammalian cells seems hopelessly complex. Traditionally, attempts have been made to reduce the complexity by grouping the proteins present in mammalian nucleoprotein according to solubility or isoelectric properties. DNA-binding proteins might alternatively be grouped on the basis of their relative nucleic acid affinities. Such a classification system can be readily constructed through the use of competing nucleic acids in DNA-cellulose chromatography (see p. 210).

As the first step in devising a classification system based on selective nucleic acid affinities, T4 bacteriophage and mouse fibroblast DNA-binding proteins have been bound to DNA-cellulose columns and fractionated

TABLE II

SOLUBLE T4 BACTERIOPHAGE AND MOUSE FIBROBLAST PROTEINS IN VARIOUS FRACTIONS ELUTED FROM DNA-CELLULOSE[a]

	Percent of soluble protein applied to columns					
	Denatured DNA-cellulose[b]		Native DNA-cellulose		DNA-free cellulose	
Fraction	T4	Mouse	T4	Mouse	T4	Mouse
Immediate elution with RNA[c]	7.0	3.1	8.0	2.6	0.06	0.03
Delayed elution with RNA[d]	6.8	2.0	4.5	0.90	0.38	0.06
Elution with 0.15 *M* NaCl	0.77	0.50	0.56	0.18	0.01	0.01
Elution with 0.60 *M* NaCl	3.7	1.4	8.1	0.68	0.06	0.05
Elution with 2.0 *M* NaCl	3.7	0.58	0.23	0.21	0.03	0.07
Uneluted	0.72	0.82	0.88	0.77	0.27	0.35
Eluted, total	22[e]	7.6	21[e]	4.6	0.54	0.22
Eluted, by NaCl	8.2	2.5	8.9	1.1	0.10	0.13

[a] The extract was prepared from 5×10^9 *Escherichia coli* infected with T4 bacteriophage (multiplicity of infection = 5, labeled with [^{14}C]leucine 5–20 minutes post-infection at 25°) mixed with 10^7 mouse fibroblasts (labeled with [^{3}H]leucine for 30 hours at 37°) in 4 ml of extract buffer, pH 7.4. The DNase extract procedure was employed.

[b] The columns were each 1 ml packed volume, and contained 1.1 mg denatured, 0.5 mg native, and no calf thymus DNA, respectively; each received 1/4 of the total extract. Chromatography was at pH 7.4 (Tris·HCl).

[c] Eluted in the first 1.5 column volumes of a 500 μg/ml tRNA rinse (DNA-free, unfractionated yeast tRNA) in 50 m*M* NaCl buffer.

[d] Eluted by the remainder of the tRNA rinse required to bring the counts to background (8 column-volumes).

[e] These numbers are greater than observed at pH 8.1 (see footnote 15a).

with an RNA rinse. After all proteins elutable with RNA were removed, the remaining proteins were collected with stepwise salt elutions. Native DNA-cellulose, denatured DNA-cellulose, and DNA-free cellulose columns were run simultaneously. The results are summarized in Table II. Polyacrylamide gel electrophoretic patterns of "immediate RNA," "delayed RNA," and 0.60 *M* NaCl eluting fractions from both native and denatured DNA-cellulose columns are displayed in Fig. 5. It can be seen that, while a substantial fraction of bacteriophage and mouse DNA-binding proteins are elutable with RNA, many different "DNA-specific" proteins are also present. Note that there are differences in the proteins binding to native and denatured DNA-cellulose.

Surprisingly, the number of different "DNA-specific" proteins detected in mouse cell extracts is not much greater than the number synthesized by T4 bacteriophage. It would appear, therefore, that a meaningful analysis of mammalian DNA-binding proteins can be developed using this technique.