Quantification of DNA–protein interaction by UV crosslinking

Gyongyi Molnar, Nuala O'Leary, Arthur B. Pardee and David W. Bradley[1],*

Dana-Farber Cancer Institute, Division of Cell Growth and Regulation, 44 Binney Street, Boston, MA 02115, USA and [1]BioMolecular Assays, Inc., 25 Olympia Avenue, Unit S, Woburn, MA 01801-6307, USA

ABSTRACT

Measurement of the affinity of a protein for a promoter sequence is critical when assessing its potential to regulate transcription. Here we report that the DNA protein crosslinking (DPC) assay can be used to measure affinity, amount and molecular weight of DNA binding proteins to specific and non-specific DNA sequences. By applying a theoretical analysis to evaluate the binding data, it was shown that the affinity constants of two proteins (named DPC80 and DPC107) to the MT3 region of the mouse thymidine kinase promoter were 2×10^{-9} M, which is 10^4 times higher than to non-specific DNA. Similar affinity constants were found when the purified proteins corresponding to DPC80 and DPC107 instead of nuclear extracts were used to assess the reliability of the DPC assay. A value for crosslinking efficiency was determined as 0.07, however, it is not needed for computation of the DNA–protein affinity, but with it the abundance of a binding protein can be estimated. In summary, the DPC assay is useful for quantifying DNA binding proteins and thereby judging their influence on transcription.

INTRODUCTION

To study the protein–DNA interaction at the promoter region is especially important for understanding the regulation of gene expression. Despite this, with most methods only the quality of protein–DNA partnership can be examined. In the traditional assays, such as the electromobility shift assay (1), Southwestern blot (2) or the DNA affinity precipitation method (3–11), application of the non-specific competitor gives information about the specificity of the binding, but does not describe the relative affinity and the abundance of the binding proteins.

The UV crosslinking between protein and DNA has been a well known technique to study their interaction (12). This method is used in a wide variety of applications to find DNA regulatory proteins or study their features (13–18) as well as to identify RNA binding proteins (19,20). Irradiation with UV creates covalent bindings between the reactive groups of DNA and amino acids such as thymine and cysteine, serine, methionine, lysine, arginine, histidine, tryptophan, phenylalanine or tyrosine (4,21). The DNA–protein crosslinking assay (DPC) is based on this covalent linkage (3) which makes the complex resistant to denaturing agents, and therefore it can be measured on denaturing gels. Here

we report that the mathematical analysis of the characteristics of complex formation makes the determination of qualitative and quantitative relations possible.

In order to quantitate the protein–DNA interactions, the mouse thymidine kinase system was chosen as a model. DNA footprint analysis showed that the protein binding site of the thymidine kinase promoter extends from −174 to +159 bp relative to the translation start site. In this region three sites can be distinguished called MT1, MT2 and MT3. The binding pattern of these binding sites were partially similar in gel shift assays, indicating partial identity with each other. One of these binding complexes named Yi was able to bind to the promoter in a cell cycle dependent manner (22,23). To find and analyze these binding proteins, two different approaches were applied. By molecular cloning we found two different proteins that bind to MT3 (24). The other method which is described in this article takes advantage of the photocrosslink formation between MT3 DNA and the associated proteins for quantitative and qualitative analysis.

Here we report that UV crosslinking in crude BPA31 nuclear extracts detects binding of two different proteins called DPC80 and DPC107 to the MT3 region of the mouse thymidine kinase promoter. With Scatchard analysis the affinity constants of these complexes appeared to be 2.4×10^{-9} and 1.9×10^{-9} M, respectively. Since we could measure similar K_d values with purified proteins corresponding to DPC80 and DPC107 instead of nuclear extracts, we suggest that the DPC analysis can be applied to crude extracts and does not require the application of purified proteins. Our data show that UV crosslinking is a useful method to calculate not only the affinity constant, but also the abundance of the protein, and these values together with estimation of the molecular weight contribute to a more precise description of a DNA binding protein.

MATERIALS AND METHODS

Cell culture

Benzo(a)pyrene transformed Balb/c 3T3 cells (BPA31) (25,26) were incubated at 37°C with 10% CO_2 in Dulbecco's modified Eagle's medium supplemented with 10% bovine calf serum, 4 mM glutamine, 10 mM HEPES buffer (pH 7.3) and penicillin–streptomycin. The culture was mycoplasma free, as determined by MycoTect kit (Gibco).

Preparation of nuclear extract

Nuclear extract was prepared as described earlier (27). Briefly, the cells were homogenized in Buffer I (10 mM HEPES pH 7.9,

---

* To whom correspondence should be addressed

1.5 mM MgCl$_2$, 10 mM KCl, 0.3 M sucrose, 0.1 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF), and after centrifugation the pellet was extracted in Buffer II (20 mM HEPES pH 7.9, 25% glycerol, 0.42 M NaCl, 1.5 mM MgCl$_2$, 0.2 mM EDTA, 0.1 mM EGTA, 0.5 mM DTT, 0.5 mM PMSF). Finally, the clear supernatant was dialyzed against Buffer III (20 mM HEPES pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF).

Preparation of probes

The double stranded MT3:

```
TCG ACA CCC ACG GAC TCT CGG TGC T
    GT GGG TGC CTG AGA GCC ACG AAG CT
```

element of the mouse thymidine kinase promoter (22) was end-labeled by Klenow fragment and then separated from unincorporated [^{32}P]dCTP using a G-25 Sephadex spin column. The specific activity of the probe was calculated from the activity of a small aliquot and the amount of labeled DNA. This value was modified by 1 c.p.m./ng = 1.37 × 10^{13} c.p.m./M correction factor. Using this calculation, the activities of the probes prepared in the different experiments were between 1.5 × 10^5 and 6 × 10^5 c.p.m./ng.

DNA–protein crosslinking (DPC) assay

Nuclear extract containing 20 µg protein was mixed with end-labeled MT3 oligonucleotide in the presence or absence of salmon sperm DNA or poly[d(I–C)] as non-specific competitor. The reaction buffer contained 5 mM Tris pH 7.5, 50 mM NaCl, 5 mM EDTA, 5 mM DTT, 25% glycerol, and the final volume of the reaction mixture was 12.5 µl. The binding assay was carried out in a 'V' bottom 96 well microtiter plate placed on ice for 20 min, and the samples were irradiated with 1.5 J of ultraviolet light (Stratalinker, Stratagene) at the optimal wavelength (260 nm) for crosslinking DNA to proteins (28). Samples were then mixed with 12.5 µl Laemmli buffer, heated at 95°C for 5 min, and finally electrophoresed on a 10% SDS–polyacrylamide gel. To prevent high background, the lower and upper part of the gel were separated, and both were stained by Coomassie Brilliant Blue, dried and exposed to X-ray film. For quantitation, developed films were aligned with dried gels and bands corresponding to DPC80, DPC107 and free DNA were located. Portions of the gel containing the radioactive bands were excised and then counted in liquid scintillation cocktail.

Mobility-shift DNA binding assay and *in situ* crosslinking

The mobility-shift DNA binding assay was performed as described earlier (22) with some slight modifications. MT3 probe (5 × 10^{-9} M) was incubated with nuclear extract containing 40 µg protein in the presence of 1.2 µg poly(d[I–C]). The reaction mixtures were loaded onto a 4% non-denaturing agarose gel, using low melting point agarose. After electrophoresis the Yi band was determined by autoradiography and cut out for *in situ* crosslinking studies. For this purpose, the gel slices were irradiated with 3 J of 260 nm UV light, then boiled in SDS-loading dye and resolved in a 10% polyacrylamide gel.

RESULTS

Reduction of non-specific binding

In the initial DNA–protein crosslinking experiments we used the MT3 region of the murine thymidine kinase promoter as a probe
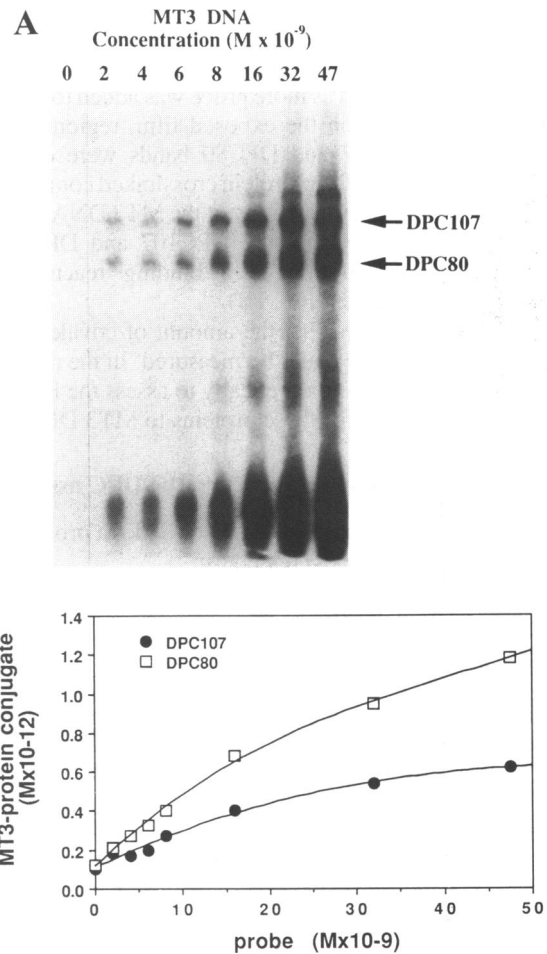
Figure 1. (**A**) Increasing probe (MT3) activity leads to increasing DPC80 and DPC107 activity on SDS–PAGE. BPA31 nuclear extract (20 µg) was combined and crosslinked with varying concentrations of ^{32}P-labeled MT3 oligonucleotide in the presence of 1 µg poly (d[I–C]). (**B**) The activity of conjugated probes plotted against probe concentration shows a hyperbolic curve. The DPC80 and DPC107 activities were measured (see A), expressed as conjugated MT3 concentration and plotted as a function of total MT3 activities.

and crude nuclear extract from BPA31 cells as the source of binding proteins. We detected several bands that were decreased by non-specific competitors. Finally two bands persisted despite an extremely high 2500- or 500-fold excess mass of salmon sperm DNA or poly[d(I–C)], respectively. These bands were called DPC107 and DPC80 from the estimated molecular weights shown on denaturing gel after DNA–Protein Cross-linking (DPC; Fig. 1A). The actual molecular weights of these protein components were calculated by subtracting the molecular weight of MT3 as 93 and 66 kDa. These proteins appeared to be sequence-specific, high-affinity binding molecules, and so they were chosen for further quantitative study.

Generation of a hyperbolic binding curve

After non-specific binding was reduced, the potential for the DPC assay to quantitate DNA–protein interaction was assessed. First, to determine whether the DNA–protein conjugates can be measured, it was necessary to generate a simple dosage curve by varying the DNA concentration.

The DPC assay was carried out on a constant amount of nuclear protein extract from exponentially growing BPA31 cells and increasing amounts of MT3 DNA probe. The DPC107 and DPC80 signals increased as more probe was added to the reaction mix (Fig. 1A). Based on the exposed film, regions of the gel containing the DPC107 and DPC80 bands were excised and counted. The moles of DNA–protein crosslinked conjugates were calculated from the specific activity of the MT3 DNA probe (7.85 × 10^{18} c.p.m./M). A plot of the DPC107 and DPC80 moles against probe concentration in a binding reaction yielded hyperbolic shaped curves (Fig. 1B).

These curves indicated that the amount of covalently-linked, DNA–protein conjugate could be measured. In the next step the DPC assay was analyzed theoretically to assess the feasibility of measuring the affinities of these proteins to MT3 DNA.

Application of equilibrium theory to the DPC assay

When MT3 DNA (d) binds to the 93 or 66 kDa proteins (p) the reaction equation can be expressed as:

$$d + p \leftrightarrow dp \qquad 1$$

For the sake of simplicity, (p) represents a single DNA-binding protein, although in crude nuclear extracts, 93 (DPC107) and/or 66 kDa (DPC80) molecules could actually be part of a protein complex. At equilibrium the dissociation constant (K_d) is

$$K_d = \frac{[d][p]}{[dp]} \qquad 2$$

In order to reduce non-specific binding, salmon sperm DNA or poly[d(I–C)] was included in the binding reaction. These non-specific competitors were expected to act as inhibitors of the MT3-protein binding reaction and so a competitive inhibitor (i) was implicated in the theoretical analysis by adding the following equations:

$$i + p \leftrightarrow ip \qquad 3$$

$$K_i = \frac{[i][p]}{[ip]} \qquad 4$$

where K_i is the dissociation constant for binding between the non-specific competitor (i) and protein (p). As will become evident below, the efficiency of crosslinking between the non-specific competitor and protein is not required for determination of K_d or K_i.

By rearranging equations 2 and 4 and using the fact that the total protein (p$_t$) is equal to the free protein (p) plus that bound to the DNA (dp) and non-specific competitor (ip):

$$p + dp + ip = p_t \qquad 5$$

one can derive an equation similar to that derived by Scatchard (29), which describes a linear relationship between the bound DNA and the ratio of the bound to free DNA:

$$\frac{[dp]}{[d]} = \frac{p_t}{K_d}(\frac{K_i}{[i] + K_i}) - \frac{1}{K_d}(\frac{K_i}{[i] + K_i})\,[dp] \qquad 6$$

After UV irradiation a fraction of the binding proteins and DNA are covalently linked. The amount of covalently linked DNA–protein conjugate (dpc) is proportional to the amount of bound DNA–protein complex (dp), and so

$$[dpc] = e[dp] \qquad 7$$

where (e) is defined as the efficiency of crosslinking.

For the DNA–protein crosslinking assay, the amount of covalently linked conjugate (dpc), not the spontaneously associated complex (dp) is measured. By substituting equation 7 into equation 6, one obtains

$$\frac{[dpc]}{[d]} = \frac{ep_t}{K_d}(\frac{K_i}{[i] + K_i}) - \frac{1}{K_d}(\frac{K_i}{[i] + K_i})\,[dpc] \qquad 8$$

Plots of [dpc]/[d] versus [dpc], allow one to calculate the binding protein abundance (if e is known) and the DNA–protein affinities. Note that the efficiency of crosslinking (e) remains only in the y-intercept term of equation 8.

Theoretical determination of DNA–protein affinities without the efficiency of crosslinking

First the specific and non-specific DNA affinities will be derived. The slope (m) of equation 8 contains the dissociation constants and can be written as:

$$m = - K_d^{-1}(\frac{[i]}{K_i} + 1)^{-1} \qquad 9$$

Equation 9 can be expanded as a series (30) if [i] > K_i

$$m = - K_d^{-1}[1 - ([i]/K_i) + ([i]/K_i)^2 - ([i]/K_i)^3 + ...] \qquad 10$$

As the non-specific inhibitor concentration becomes lower, the first two terms of this series will dominate (i/K < 1/4), and the Scatchard slope (m) can be approximated by:

$$m = - \frac{1}{K_d} + \frac{[i]}{K_d K_i} \qquad 11$$

Equation 11 predicts that the Scatchard slope (m) will be proportional to the inhibitor concentration (i) with a y-intercept of $-1/K_d$ and an x-intercept of K_i.

Equations 8 and 11 allow one to use DNA–protein crosslinking data to calculate the apparent affinity between a binding protein and DNA without knowing the efficiency of crosslinking. By empirically varying the non-specific competitor concentration (i) in a series of binding reactions, one can construct a set of Scatchard plots. The slopes of these Scatchard plots can be graphed against the non-specific competitor concentration (i) as prescribed by equation 11, then K_i and K_d can be determined from the x- and y-intercepts.

Theoretical determination of abundance

The abundance of the binding protein can be derived similarly from the y-intercept (b) of equation 8:

$$b = \frac{ep_t}{K_d}(\frac{K_i}{[i] + K_i}) \qquad 12$$

Rearranging terms:

$$b = (ep_t / K_d)([i / K_i] + 1)^{-1} \qquad 13$$

After expanding in a power series (30) and selecting the first two terms:

$$b = \frac{ep_t}{K_d} - \frac{ep_t}{K_d K_i}[i] \qquad 14$$

Equation **14** can be used to determine the total amount of binding protein (p_t), if the efficiency of crosslinking (e) is known. First, the non-specific competitor concentration must be varied in a series of binding reactions. The y-intercepts of the resulting Scatchard plots can be graphed against the non-specific competitor concentration. After extrapolation to [i] = 0, the y-intercept value (b) becomes:

$$b = \frac{ep_t}{K_d} \qquad 15$$

and then one can calculate the total amount of binding protein (p_t) with known (b), (e), and K_d values.

Before applying the theoretical analysis to the experimental data, an assumption must be made that the free DNA and non-crosslinked DNA concentrations were similar. Before cross-linking, the total DNA (d_t) can be expressed as follows:

$$d_t = dp + d \qquad 16$$

Again, (d) is the free DNA and (dp) is the bound DNA. After crosslinking and electrophoresis, the expression for the total DNA (d_t) must be adjusted to

$$d_t = dpc + d_{nc} \qquad 17$$

where (d_{nc}) is the non-crosslinked DNA which runs to the bottom of the gel and (dpc) is the covalently linked DNA–protein

conjugate (i.e. the DPC80 or DPC107 bands). If (dpc) and (dp) are small compared to (d_{nc}) and (d), then

$$d_t \cong d \cong d_{nc} \qquad 18$$

and the non-crosslinked DNA can be used as an estimate of the free DNA. The validity of this assumption is discussed later.

Calculation of DNA–protein affinities using the adjusted Scatchard equation

In order to test the theoretical analysis with DNA–protein crosslinking data, the MT3 oligonucleotide concentration was varied in a binding reaction, while the amount of the nuclear extract was kept at constant level. Three dosage curves were created using three different concentrations of salmon sperm DNA as non-specific competitor. After crosslinking and electro-phoresis, the bands corresponding to DPC80, DPC107 and non-crosslinked DNA were excised and measured by scintillation counting (Table 1).

To calculate the radioactive concentration of crosslinked and free DNA, the specific activity of the DNA (6.92×10^{18} c.p.m./M) and the reaction volume (1.25×10^{-5} l) were used. Then the crosslinked/free ratio was obtained from the molarities of these fractions (Table 2).
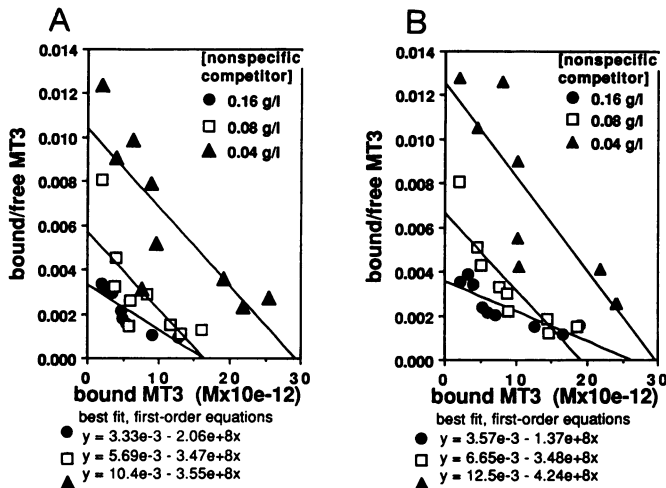
Table 1. Equilibrium binding data

| Crosslinked MT3 (c.p.m.) | | | | | | Noncrosslinked MT3 (c.p.m.) | | |
| DPC80 | | | DPC107 | | | | | |
A	B	C	A	B	C	A	B	C
179	168	173	188	167	179	52 807	20 739	14 002
225	343	345	277	389	400	70 961	76 047	38 018
297	335	546	344	440	698	101 184	102 916	55 272
408	517	763	454	657	870	188 125	198 628	96 338
430	725	836	512	756	887	238 771	250 750	161 520
468	509	656	619	772	883	300 300	351 050	207 828
780	1000	1656	1092	1233	1896	734 033	666 533	457 600
1113	1385	1897	1641	1602	2089	1 048 850	106 905	817 300
1103	1128	2195	1434	1260	2075	1 201 100	104 595	811 300

20 µg BPA31 nuclear extract was UV crosslinked with different concentration of MT3 probe, in the presence of A, 0.16 µg/µl; B, 0.08 µg/µl; C, 0.04 µg/µl salmon sperm DNA. The crosslinked and free MT3 activities are shown in c.p.m.

Table 2. Equilibrium binding: calculations of crosslinked/free ratio and crosslinked MT3 molarity

A. DPC80						B. DPC107					
Crosslinked/Free			Crosslinked			Crosslinked/Free			Crosslinked		
(10^{-3})			(10^{-12} M)			(10^{-3})			(10^{-12} M)		
A	B	C	A	B	C	A	B	C	A	B	C
3.39	8.10	12.4	2.07	1.94	2.00	3.56	8.05	12.8	2.17	1.93	2.07
3.17	4.51	9.08	2.60	3.97	3.99	3.90	5.12	10.5	3.20	4.50	4.62
2.94	3.26	9.88	3.43	3.87	6.31	3.40	4.28	12.6	3.98	5.09	8.07
2.17	2.60	7.92	4.72	5.98	8.82	2.41	3.31	9.03	5.25	7.60	10.1
1.80	2.89	5.18	4.97	8.38	9.67	2.14	3.02	5.49	5.92	8.74	10.2
1.56	1.45	3.16	5.41	5.88	7.58	2.06	2.20	4.25	7.16	8.93	10.3
1.06	1.50	3.62	9.02	11.6	19.1	1.49	1.85	4.14	12.6	14.3	21.9
1.06	1.30	2.32	12.9	16.0	21.9	1.57	1.50	2.56	19.0	18.5	24.2
0.92	1.08	2.71	12.8	13.0	25.4	1.19	1.21	2.56	16.6	14.6	24.0

A, 0.16 µg/µl; B, 0.08 µg/µl; C, 0.04 µg/µl salmon sperm DNA.

A

best fit, first-order equations
● y = 3.33e-3 - 2.06e+8x
□ y = 5.69e-3 - 3.47e+8x
▲ y = 10.4e-3 - 3.55e+8x

B

best fit, first-order equations
● y = 3.57e-3 - 1.37e+8x
□ y = 6.65e-3 - 3.48e+8x
▲ y = 12.5e-3 - 4.24e+8x

Figure 2. Scatchard plot of equilibrium binding data of DPC80 (A) and DPC107 (B). BPA31 nuclear extract (20 μg) was UV crosslinked to various concentration of radioactive MT3 probe in the presence of 0.16 g/l (●), 0.08 g/l (Π) or 0.04 g/l (△) salmon sperm DNA. The crosslinked and free MT3 oligonucleotides were measured (Table 1), the molar equivalents calculated (Table 2, A and B) and the crosslinked/free ratios plotted against crosslinked molarity, as described by Scatchard.

best fit equations
□ DPC80   y = - 4.26x10e+8 + 1.32x10e+9 X R^2=0.874
● DPC107   y = - 5.30x10e+8 + 2.43x10e+9 X R^2=0.981

Figure 3. Derivation of dissociation constant (K_d) from Scatchard plots. The slopes (m) from the DPC80 (Π) and DPC107 (●) Scatchard plots (Fig. 3) were graphed against the concentration of salmon sperm DNA. The best fit, first order equations are shown.

Scatchard plots for the MT3 dosage experiments were generated by plotting the crosslinked/free against the MT3 crosslinked to DPC80 (Fig. 2A) or DPC107 (Fig. 2B) as prescribed by equation 8. First order equations fitting these data were computed (Cricket Graph, Macintosh) and best fit lines were drawn. As predicted, the slope (m) was negative and decreased in absolute value as the non-specific competitor concentration (i) was increased.
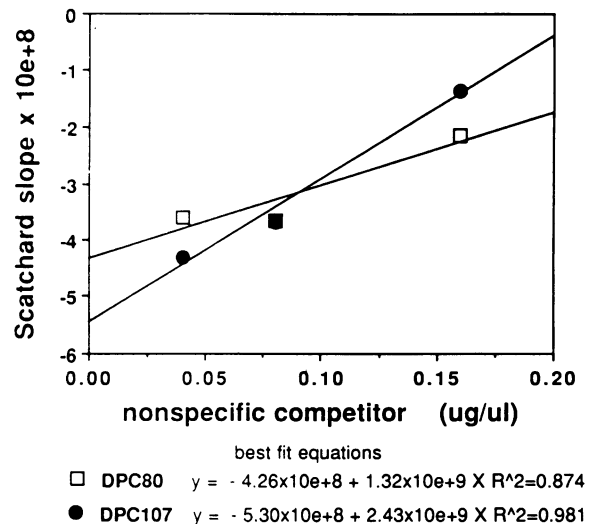
In order to calculate the dissociation constants, the Scatchard slopes for DCP80 and DPC107 shown on Figure 2 were plotted against the non-specific competitor concentration (Fig. 3). As shown in equation 11, the y-intercepts of these lines are the negative inverse of the dissociation constants with MT3: they were calculated to be 2.35×10^{-9} and 1.89×10^{-9} M for DPC80 and DPC107, respectively.

The value of K_d did not change significantly when higher order series terms ($[i]/K_i)^n$ of equation 10 were added to equation 11. Even with these higher terms, the y-intercept is still $-1/K_d$. Thus the assumption that higher order terms could be ignored in equation 11 appears to be valid.

To find the dissociation constant for non-specific DNA (K_i), after setting m = 0, the x intercept of equation 11 is equal to K_i. Correspondingly, the x-intercepts (i.e. K_i) for DPC80 and DPC107 from Figure 3 were 0.323 and 0.218 g/l, respectively. Converting to molarity after dividing K_i (g/l) by the molecular weight of MT3 (MW:13 727), the dissociation constants of DPC80 and DPC107 with salmon sperm DNA (2.35×10^{-5} and 1.59×10^{-5} M) were 10^4-fold greater than the dissociation constant relevant to MT3.

Estimation of DPC80 crosslinking efficiency

The y-intercept of equation 8 contains three known terms; K_i, K_d and (i), and two unknown terms, the total protein (p_t) and the crosslinking efficiency (e). The crosslinking efficiency is needed to calculate the abundance of the binding protein.

In order to find (e), the DPC protein binding in the Yi complex had to be identified. Nuclear extracts prepared from exponentially growing BPA31 cells were mixed with the MT3 probe and electrophoresed through a non-denaturing low-melting-point (LMP) agarose gel. Then the binding band (Yi) was excised, irradiated and electrophoresed through a denaturing polyacrylamide gel (Fig. 4, lane 1–3). To locate the DPC107 and DPC80 positions, extracts from exponentially growing cells were crosslinked to MT3 and loaded on the adjacent lanes (Fig. 4, lane 4). The *in situ* crosslinked band comigrated with DPC80. Therefore DPC80 is almost certainly the DNA binding protein of the Yi complex.

As a control to ensure that the LMP agarose did not alter the observed mobility of the crosslinked proteins, 50 μl of 1% LMP agarose was mixed with crosslinked samples before electrophoresis. The bands in these samples had the same mobilities as crosslinked samples alone (Fig. 4, lane 5).

To estimate the crosslinking efficiency between MT3 and the DPC80 protein, a mobility shift assay was used to measure the total amount of MT3 probe bound to the protein and a DPC assay was used to measure the UV conjugated amount of MT3. In this experiment the reaction volume was doubled, keeping the DNA and protein concentrations as they were in the DPC assay. Then half of the reaction mixture was UV irradiated and analysed on a denaturing gel, while the second half was examined on an electromobility shift assay. The Yi band activity gave 3617 c.p.m., the DPC80 conjugate had 255 c.p.m., thus the crosslinking efficiency (e) was calculated as 0.07. This value falls within the typical range of crosslinking efficiencies for DNA binding proteins (21).

The dissociation of binding proteins from DNA during non-denaturing electrophoresis is not significant, because of a 'caging effect' of the surrounding polyacrylamide (11), and so a shifted band should represent the total amount of DNA binding protein, if DNA is in excess and the protein–protein interaction is negligible.
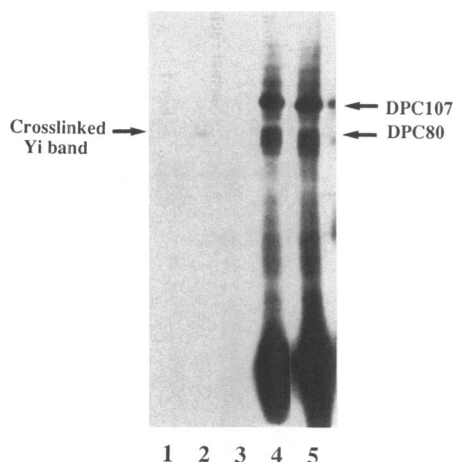
Figure 4. *In situ* crosslinking of mobility-shift Yi band resulted in a DPC80 band. Section of Yi band from non-denaturing electromobility gel was UV irradiated, then separated again on 10% denaturing gel (lanes 1–3). Exponentially growing BPA31 nuclear extracts were crosslinked to end-labeled MT3 probe (lanes 4, 5) to identify migration distance of DPC80 and DPC107 bands. LMP agarose was mixed with sample before electrophoresis (lane 5).

$y = 1.1580e\text{-}2 - 5.4714e\text{-}2x \quad R^2 = 0.921$

Figure 5. Derivation of DPC80 abundance from y-intercepts of Scatchard plots. The y-intercepts from DPC80 Scatchard plots of three binding experiments were plotted against their corresponding salmon sperm DNA concentrations. The best fit first-order polynomials were drawn.

The crosslinking efficiency of DPC107 was not measurable. DPC107 was not found in the UV irradiated Yi band, probably because either its binding to the MT3 site is independent of the Yi complex or the amount of the protein in the complex is too low to detect by *in situ* crosslinking.

Estimation of DPC80 abundance

One step remains before the abundance of DPC80 can be estimated, now that a value for (e) has been obtained. In accordance with equation **14**, the y-intercepts of the DPC80 Scatchard plots Figure 2A were plotted against the non-specific competitor concentrations (Fig. 5).

After extrapolation to [i] = 0, the y-intercept value was 0.0116. Using equation **15** the total amount of DPC80 protein was found to be 1.6×10^{-5} of the nuclear extract protein, which is ~4000 molecules per cell. Thus, as one would expect for a potential transcriptional regulator, this protein is a very small fraction of the total nuclear protein mass.

Then the abundance of DPC80 was verified by calculation from the dosage data of Figure 1. The MT3 saturates DPC80 at roughly 1.4×10^{-16} moles. This value corresponds to 1.5×10^{-10} g bound protein, which is equivalent to 2×10^{-5} g crude nuclear extract. Therefore the DPC80 protein is 7.5×10^{-6} of the total nuclear protein mass, which is only 2-fold lower than that calculated from y-intercept and K_d. The subtotal saturation of DPC80 in Figure 1 partly accounts for this difference.

Variation of protein concentration in a binding reaction

An assumption made in this analysis was that the proteins did not interact with one another. The protein (p) of the mass action equation **1** could actually represent part of a preformed complex. If binding proteins were in rapid equilibrium, a theoretical analysis under these conditions would be more complex than the simple derivations shown here.
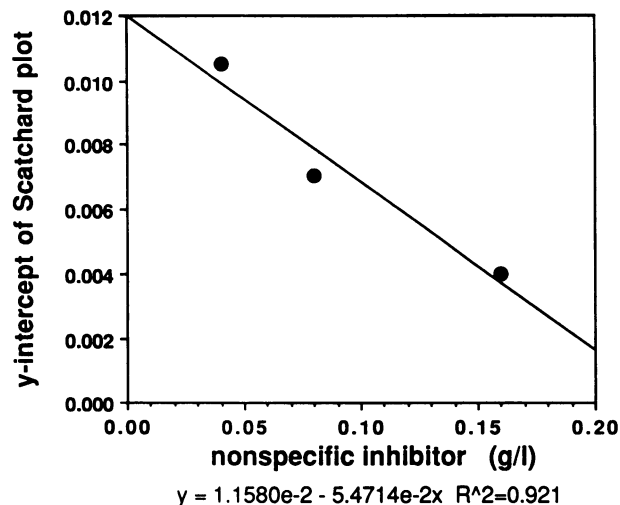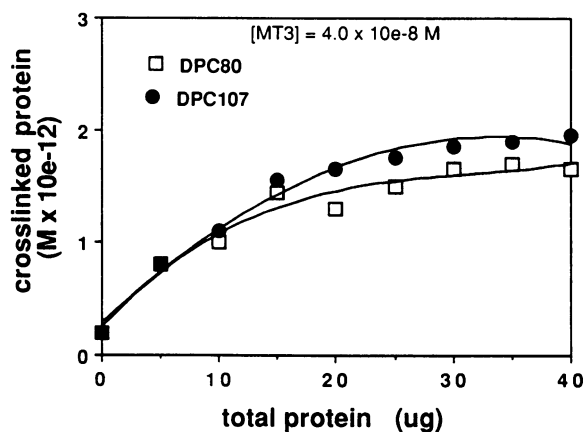
Figure 6. Variation of the protein concentration demonstrated saturable dose-dependence curve. Various amount of nuclear extracts (0–40 µg) were mixed with constant, 4×10^{-8} M labeled MT3 oligonucleotide before UV crosslinking. The crosslinked DNA activities are plotted against protein concentrations.

To test this assumption, different nuclear extract concentrations were applied, with constant MT3 probe concentration. After UV irradiation the DPC80 and DPC107 activities were measured and mole amounts were calculated from the specific activity (2.83×10^{18} c.p.m./mole), generating hyperbolic shaped curves (Fig. 6). If proteins were interacting weakly with one another under this binding condition, the protein dosage curves would be sigmoidal, and therefore the derivation seems valid for this system.

Optimization of conditions and additional controls

In further experiments the optimal binding temperature and incubation time were determined. At 0°C, the time of incubation up to 60 min before crosslinking did not significantly affect the

formation of DPC80 or DPC107. In these experiments the energy and time of irradiation were 1.5 J and 10 min, respectively. As the temperature was increased to 21 or 37°C, the intensity of both DPC80 and DPC107 bands decreased, and additionally the non-specific binding increased (data not shown). Moreover, freezing and thawing samples on dry ice and returning to 37°C once or four times over 60 min also had no effect on binding ability.

Other standard controls for UV crosslinking (21) were performed. No crosslinked bands were observed when BSA was used in a binding reaction nor when nuclear proteins were left out. Furthermore, no bands were seen from non-irradiated reaction mixtures, demonstrating that UV irradiation caused the bands to appear rather than degradation of the probe and subsequent reincorporation of the ^{32}P into a high molecular weight protein (data not shown).

As a control to ensure that small differences in UV energy for individual samples were not altering the levels of crosslinked DNA–protein conjugate, the amount of crosslinking energy was varied between 1.0 and 2.0 J. Levels of DNA–protein conjugate did not change appreciably in this range. Thus a dose of 1.5 J was used for all assays unless otherwise specified. These results were consistent with others (28,31), that showed the amount of crosslinked product increases and then saturates as the irradiation time is increased.

Validity of the assumption that non-crosslinked DNA equals free DNA

Levels of non-crosslinked DNA were used in place of free DNA in equation (18) when Scatchard plots were produced. The non-crosslinked DNA migrates to the bottom of the denaturing gel, and unless the efficiency of crosslinking is 100%, some of this non-crosslinked DNA was actually bound to DPC80 but became separated during electrophoresis. Therefore the measured non-crosslinked DNA is slightly larger than the actual free DNA.

Since the crosslinking efficiency of DPC80 is known, it is possible to calculate the free DNA. Affinity was recalculated using the free DNA (d) or the total DNA (d_t) concentrations. The dissociation constant was within 5% to calculation with non-crosslinked DNA (d_{nc}). Therefore the assumption that values of total DNA, free DNA and non-crosslinked DNA were close (equation **18**) appears to be reasonable.

Test of the UV crosslinking method using purified protein

In order to further evaluate the reliability of this newly developed method, we need to compare the K_d values presented above with those measured with purified proteins in a simple well-defined protein–DNA interaction system. To this end we used two different MT3 binding proteins that were found by expression screening of a mouse 3T3 cDNA library using the MT3 oligonucleotide as a probe. One of these proteins is the already known early response gene Egr-1 which we reported to regulate the MT3 site (24). The other MT3 binding protein called FC18 is a ~100 kDa nuclear protein that may also contribute to the transcriptional regulation of thymidine kinase (manuscript in preparation). We have already cloned these proteins and proved that they specifically bind the MT3 site and affect the transcrip-
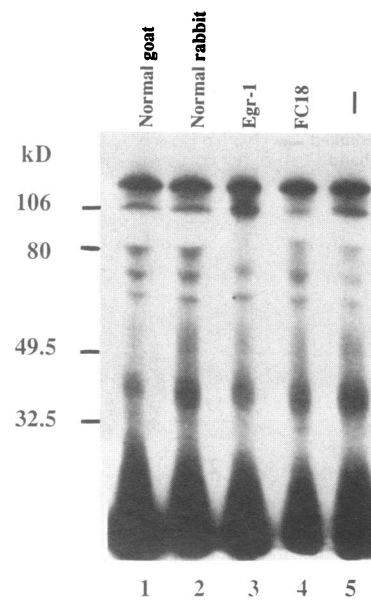
Figure 7. Antibodies against Egr-1 and FC18 decrease the DPC80 and DPC107 bands, respectively. Nuclear extracts (20 µg) from exponentially growing A31 cells were incubated with pre-immune goat (lane 1) and rabbit (lane 2) sera, with antibody against Egr-1 (lane 3) or FC18 (lane 4), or without antibody (lane 5). Then the reaction mixtures were substituted with 1.2 × 10^{-9} M of ^{32}P-labeled MT3 probe and crosslinked by UV irradiation.

tion of TK; we assume that these two proteins are identical to those detected as DPC80 and DPC107.

The assumptions that the DPC80 protein is Egr-1 and that DPC107 represents FC18 were supported by UV crosslink experiments, combined with the addition of antibody against them (Fig. 7). Nuclear extracts from exponentially growing A31 cells were treated with antibody against Egr-1 (lane 3) or FC18 (lane 4) or pre-immune sera (lane 1 and 2), and then MT3 probe was added, and the samples were irradiated with UV. The antibody against Egr-1 inactivated the MT3 binding site of the protein and abolished the DPC80 band. The antibody against FC18 strongly decreased the intensity of the DPC107 band. In contrast, the pre-immune sera (lane 1, 2) did not have any effect on the appearance of the bands compared to lane 5 where neither pre-immune nor immune sera were added to the same cell extract.
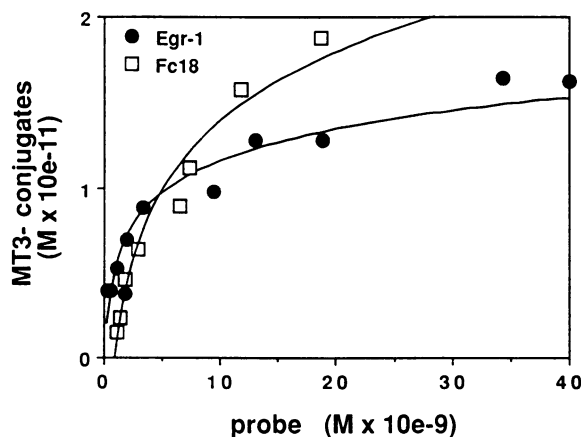
For UV crosslink experiments we used the isolated full-length Egr-1 as a fusion protein bound to glutathione-S-transferase (25). One ng of purified protein was used to create a dose–response curve between 0.23 × 10^{-9} M and 43 × 10^{-9} M probe concentration without adding inhibitor (Fig. 8, lane 1–8 and Fig. 9). Applying the above equations the calculated K_{dEgr-1} from the y-intercept was 1.78 × 10^{-9} M (Fig. 10) which is very close to that found for DPC80 (2.4 × 10^{-9} M) in nuclear extract.

The FC18 protein was isolated from exponentially growing A31 cell extract using antibody purification (32). We added 2 ng purified protein and the ^{32}P labeled MT3 probe in a concentration range from 1.07 × 10^{-9} to 18.6 × 10^{-9} M and applied UV crosslinking (Fig. 9). The Scatchard equation resulted in a K_{dFC18} of 2.20 × 10^{-9} M (Fig. 10) which is identical to that of DPC107 measured in crude nuclear extract.

The molecular weights of both Egr-1/MT3 and FC18/MT3 conjugates in crosslinking condition were found similar to that of
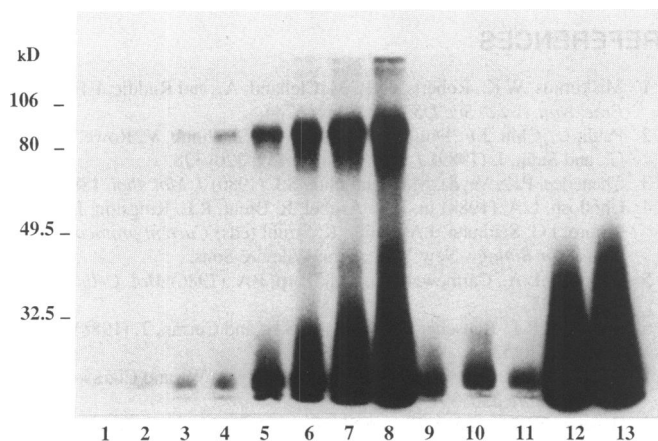
Figure 8. Bacterially expressed and purified Egr-1 protein binds to the MT3 probe in a dose-dependent manner in a UV crosslinking experiment. Increasing concentrations of MT3 probe (0.23 × 10⁻⁹ – 43.6 × 10⁻⁹ M; see Fig. 9) and 1 ng Egr-1 protein were combined in lanes 1–8. The ~96 kDa band was UV irradiation dependent (lane 9 has the same conditions as lane 5 but without UV). Lane 10 represents equivalent MT3 probe as in lane 6 but without protein. In lanes 12 and 13 the addition of 1 μg albumin did not result in formation of MT3 binding complex; the probe activities are the same as in lanes 7 and 8, respectively.

Figure 10. Scatchard plots of Egr-1 and FC18 protein were computer analysed, and the free/bound MT3 probe ratio was plotted against the bound concentration of MT3 oligonucleotide.
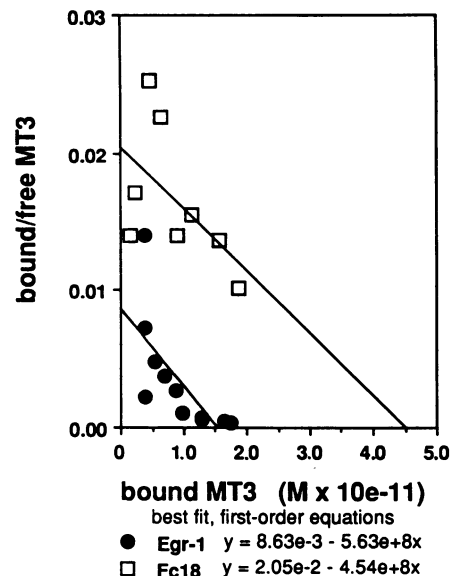

Figure 9. The binding activity of two purified MT3 binding proteins: Egr-1 and FC18 show a hyperbolic dose-response curve after photocrosslink. The protein–DNA conjugates from Figure 8 and an experiment with FC18 were graphed against the total concentration of MT3 probe.

DPC80 and DPC107, respectively (Figs 7 and 8). The crosslinking efficiency was determined from the data of Figure 9. When the MT3 bound fraction of the total proteins (6.25 × 10⁻¹⁰ M Egr-1) was calculated using equations 2 and 5, the total and bound probe concentrations were 34 × 10⁻⁹ and 1.84 × 10⁻¹¹ M, respectively. Assuming one to one DNA–protein stoichiometry and excess of the probe compared to the protein, the crosslinking efficiency was found e = 0.04. This value is similar to that of the crude extract (0.07).

DISCUSSION

DNA–protein crosslinking is employed here to obtain quantitative information about DNA-binding proteins that are responsible for

regulation during growth or differentiation (33). We report that the MT3 element of the mouse thymidine kinase promoter binds two different proteins (DPC80 and DPC107) in BPA31 nuclear extract using UV crosslinking. We increased the specificity of binding by adding a non-specific competitor such as salmon sperm DNA or poly[d(I–C)] to the system.

When one parameter is kept constant (e.g. protein), while the other one is varied (DNA probe), generation of a dose–response curve between the specifically interacting molecules gives a hyperbola. For mathematical analysis of these curves we applied the Scatchard equation which is suitable to determine the affinity constant between two molecules and the number of interacting molecules under equilibrium conditions (29). Thus, the calculated K_d values of DPC80 and DPC107 were 2.35 × 10⁻⁹ and 1.89 × 10⁻⁹ M, respectively.

In UV crosslink experiments combined with antibody treatment we demonstrated that these two proteins are identical with Egr-1 and FC18. Screening a λgt11 expression library with MT3 oligonucleotide, we had found that both proteins bind MT3 (24). We purified and applied them in UV crosslink experiments to measure their K_d values. In this simplified system the K_d values measured in nuclear extracts were fairly well reproduced: K_d = 1.93 × 10⁻⁹ M for purified Egr-1 and K_d = 2.20 × 10⁻⁹ M for purified FC18. Taken together, these data support the hypothesis that the method is sufficient to determine the K_d values in a crude extract.

The molecular weight of DNA binding protein complexes also can be estimated by the DPC method. The DPC80 band includes the 14 kDa MT3 oligonucleotide and a 66 kDa protein component, while the DPC107 represents a 93 kDa protein after subtraction of the molecular weight of DNA partner. The molecular weights of the pure Egr-1 and FC18 protein on SDS–PAGE gel were found to be 80 and 100 kDa, respectively. The underestimation of the molecular weights appears to be the result of using different techniques. In the DPC assay the

molecular weight of the proteins were determined from the conjugate mass, while the molecular weight of Egr-1 and FC18 proteins were measured in a pure system.

Thirdly, the amount of crosslinking protein also can be estimated if the crosslinking efficiency is known or is assumed. In this way the abundance of DPC80 was measured as 7.5×10^{-6} of the total nuclear mass.

Based on these results, the UV crosslink method has three advantages to quantitate protein–DNA interaction. One is to determine dissociation constants between protein and DNA. Secondly, the UV crosslink experiment makes possible a direct measure of the molecular weight of interacting proteins, since the denaturing gel separates the crosslinked DNA binding protein from other members of the complex. In contrast, the mobility shift assay, which occasionally is also applied for this DNA binding purpose (5,34,35), does not determine the molecular weight of the DNA binding protein, because the complex identified by a non-denaturing gel may contain several other proteins that do not bind directly to DNA. Moreover, the DPC assay is also able to detect proteins that cannot be seen by gel shift assay, as was found with DPC107 protein whose MT3 binding ability was detected only by the DPC assay. Finally, the DPC assay is also suitable to estimate the amount of the binding protein from nuclear extracts, or even the number of molecules present in one cell can be calculated.

Biological relevance of equilibrium data

The above calculated affinity values are within a range expected for biological regulators. Typically, the nuclear extraction procedure yields ~2.5×10^{-11} g of protein per cell. Without taking into account the proteins lost during purification, there are 6×10^{-21} M of DPC80 or ~4000 DPC80 molecules per cell. Assuming that the volume of a cell is 10^{-12} l (i.e. $10 \times 10 \times 10$ μm) and the DPC80 binding sites had the same affinity as MT3, 70 % of the binding sites would be filled. In contrast only 3 in every 10 000 binding sites would be filled non-specifically. The repetitive elements and highly ordered arrangements of the eukaryotic genome decrease the availability of non-specific binding sites, and so these affinities are within a biologically meaningful range (36–41). Regulation by cooperativity, protein modification or abundance could further increase the ratio of bound to free sites.

In conclusion, the DPC assay was useful for finding and quantifying binding proteins to the TK promoter. We suggest that it could be useful for other transcriptional systems as well.

ACKNOWLEDGEMENTS

REFERENCES

1 Miskimins, W.K., Roberts, M.P., McClelland, A., and Ruddle, F.H. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6741–6744.
2 Pauli, U., Chiu, J.F., Ditullio, P., Kroeger, P., Shalhoub, V., Rowe, T., Stein, G., and Stein, J. (1989) *J. Cell. Physiol.* **139**, 320–328.
3 Chatterjee, P.K., Vayda, M.E., and Flint, S.J. (1986) *J. Mol. Biol.* **188**, 23–37.
4 Chodosh, L.A. (1988) In: F.M. Asubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (eds) *Current protocols in Molecular Biology*. New York : John Wiley & Sons.
5 Chodosh, L.A., Carthew, R.W., and Sharp, P.A. (1986) *Mol. Cell. Biol.* **6**, 4723–4733.
6 Franza, B.R.J., Rauscher, F.J., Josephs, S.F., and Curran, T. (1988) *Science*, **239**, 1150–1153.
7 Franza, B.R.J., Josephs, S.F. Gilman, M.Z., Ryan, W., and Clarkson, B. (1987) *Nature* **330**, 391–395.
8 Herman, T.M., Lefever, E., and Shimkus, M. (1986) *Anal. Biochem.* **156**, 48–55.
9 Hilton, M.D., and Whiteley, H.R. (1985) *J. Biol. Chem.* **260**, 8121–8127.
10 Kasher, M.S., Pintel D., and Ward D.C. (1986) *Mol. Cell Biol.*, **6**, 3117–3127.
11 Shimkus, M., Levy. J., and Herman, T., (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2593–2597.
12 Laemmli, U. (1970) *Nature*, **227**, 680–685.
13 Distel, R.J., Ro, H.S., Rosen, B.S., Groves, DL., and Spiegelman, B.M. (1987) *Cell*, **49**, 835–844.
14 Praseuth, D., Perrouault, L., Le Doan, T., Chassignol, M., Thuong, N., and Helene, C. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1349–1353.
15 Miller, T., and Kaiser, E.T. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5429–5433.
16 Allen, T.D., Wick, K.L., and Matthews, K.S. (1991) *J. Biol. Chem.* **266**, 6113–6119.
17 Wick, K.L., and Matthews K.S. (1991) *J. Biol. Chem.* **266**, 6106–6112.
18 Telfer, J., and Rudd C.E. (1991) *Science*, **254**, 439–441.
19 Marciniak, R.A., Garcia-Blanco, M.A., and Sharp, P.A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 3624–3628.
20 Kelly, S., Sheng, N., and Dennis, D. (1990) *J. Biol. Chem.* **265**, 7787–7792.
21 Merrill, B.M., Williams, K.R., Chase, J.W., and Konigsberg, W.H. (1984) *J. Biol. Chem.* **259**, 10850–10856.
22 Dou, Q.-P., Fridovich-Keil, J.L., and Pardee, A.B. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1157–1161.
23 Bradley, D.W., Dou, Q.-P., Fridovich-Keil, J.L., and Pardee A.B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9310–9314.
24 Molnar, G., Crozat, A., and Pardee A.B. (1994) *Mol. Cell. Biol.* **14**, 5242–5248.
25 Holley, R., Baldwin, J., Kieman, J., and Messmer, T. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3229–3232.
26 Oshiro, Y., and Di Paolo, J.A. (1974) *J. Cell. Physiol.* **83**, 193–202.
27 Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
28 Hockensmith, J.W., Kubasek, W.L., Vorachek, W.R., and von Hippel, P.H. (1986) *J. Biol. Chem.* **261**, 3512–3518.
29 Scatchard, G. (1949) *Annals N. Y. Acad. Sci.* **51**, 660–672.
30 Boas, M. L. (1966) *Mathematical Methods in the Physical Science*. John Wiley & Sons, Inc. New York. 21.
31 Chatterjee, P.K., Yang, U.-C., and Flint, S.J. (1986) *Nucleic Acids Res.* **14**, 2721–2735.
32 Harlow, E., and Lane, D. (1988) *Antibodies a Laboratory Manual*. Cold Spring Harbor Laboratory Chapter 13.
33 Bradley, D.W. (1991) *Cell Cycle Regulation of Binding Activities to the Murine Thymidine Kinase Promoter in Normal and Transformed Cells*. Harvard University doctoral thesis, UMI Press, Ann Arbor, Michigan. 97–119.
34 Fried, M.G., and Crothers, D.M. (1981) *Nucleic Acids Res.* **9**, 6505–6525.
35 Fried, M.G., and Crothers, M.D. (1984) *J. Mol. Biol.* **172**,, 263–282
36 Britten, R.J., and Davidson, E.H. (1969) *Science* **165**, 349–357.
37 Britten, R.J., and Davidson, E.H. (1971) *Q. Rev. Biol.* **46**, 111–133.
38 Callan, H.G. (1967) *J. Cell. Sci.* **2**, 1–7.
39 Crick, F. (1971) *Nature* **234**, 25–27.
40 Davidson, E.H. (1976) *Gene Activity in Early Development*. Academic Press. New York. 26.
41 Davidson E.H., and Britten, R.J. (1973) *Q. Rev. Biol.* **48**, 565–613.