Quantification of DNA–protein interaction by UV crosslinking

Gyongyi Molnar, Nuala O’Leary, Arthur B. Pardee and David W. Bradley

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

Received March 13, 1995; Revised and Accepted July 4, 1995

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 $2 \times 10^{-9}$ 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 $2.4 \times 10^{-9}$ and $1.9 \times 10^{-9}$ 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% CO2 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 MycoTest 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,
1.5 mM MgCl₂, 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 MgCl₂, 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 [³²P]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/μg = 1.37 × 10¹³ c.p.m/M correction factor. Using this calculation, the activities of the probes prepared in the different experiments were between 1.5 × 10⁵ and 6 × 10⁵ c.p.m/μg.

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 background, the lower and upper part of the gel were separated, 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 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 \times 10^{18} \text{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:

$$d + p \leftrightarrow dp$$  

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

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

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:

$$i + p \leftrightarrow ip$$  

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):

$$p + dp + ip = p_t$$  

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

$$[dpc] = e[dp]$$  

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

$$\frac{[dpc]}{[d]} = \frac{e_p}{K_d} \left( \frac{K_i}{[i]} \right) \left( \frac{K_i}{[i] + K_i} \right) - \frac{1}{K_d} \left( \frac{K_i}{[i] + K_i} \right) [dp]$$  

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:

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

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

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

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

$$m = -\frac{1}{K_d} + \frac{i}{K_d K_i}$$  

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:

$$b = \frac{e_p}{K_d} (\frac{K_i}{[i] + K_i})$$  

Rearranging terms:

$$b = (\frac{e_p}{K_d}) ([i] / K_i) + 1)^{-1}$$  

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

$$b = \frac{e_p}{K_d} - e_p \frac{1}{K_d K_i} [i]$$
Table 1. Equilibrium binding data

<table>
<thead>
<tr>
<th>Crosslinked MT3 (c.p.m.)</th>
<th>DPC80</th>
<th>DPC107</th>
<th>Noncrosslinked MT3 (c.p.m.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>179</td>
<td>168</td>
<td>173</td>
<td>188</td>
</tr>
<tr>
<td>225</td>
<td>343</td>
<td>345</td>
<td>277</td>
</tr>
<tr>
<td>297</td>
<td>335</td>
<td>546</td>
<td>344</td>
</tr>
<tr>
<td>408</td>
<td>517</td>
<td>763</td>
<td>454</td>
</tr>
<tr>
<td>430</td>
<td>725</td>
<td>836</td>
<td>512</td>
</tr>
<tr>
<td>468</td>
<td>509</td>
<td>656</td>
<td>619</td>
</tr>
<tr>
<td>780</td>
<td>1000</td>
<td>1656</td>
<td>1092</td>
</tr>
<tr>
<td>1113</td>
<td>1385</td>
<td>1897</td>
<td>1641</td>
</tr>
<tr>
<td>1103</td>
<td>1128</td>
<td>2195</td>
<td>1434</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>A. DPC80</th>
<th>Crosslinked/Free (10^{-3})</th>
<th>Crosslinked (10^{-12} M)</th>
<th>B. DPC107</th>
<th>Crosslinked/Free (10^{-3})</th>
<th>Crosslinked (10^{-12} M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>3.39</td>
<td>8.10</td>
<td>12.4</td>
<td>2.07</td>
<td>1.94</td>
<td>2.00</td>
</tr>
<tr>
<td>3.17</td>
<td>4.51</td>
<td>9.08</td>
<td>2.60</td>
<td>3.97</td>
<td>3.99</td>
</tr>
<tr>
<td>2.94</td>
<td>3.26</td>
<td>9.88</td>
<td>3.43</td>
<td>3.87</td>
<td>6.31</td>
</tr>
<tr>
<td>2.17</td>
<td>2.60</td>
<td>7.92</td>
<td>4.72</td>
<td>5.98</td>
<td>8.82</td>
</tr>
<tr>
<td>1.80</td>
<td>2.89</td>
<td>5.18</td>
<td>4.97</td>
<td>8.38</td>
<td>9.67</td>
</tr>
<tr>
<td>1.56</td>
<td>1.45</td>
<td>3.16</td>
<td>5.41</td>
<td>5.88</td>
<td>7.58</td>
</tr>
<tr>
<td>1.06</td>
<td>1.50</td>
<td>3.62</td>
<td>9.02</td>
<td>11.6</td>
<td>19.1</td>
</tr>
<tr>
<td>1.06</td>
<td>1.30</td>
<td>2.32</td>
<td>12.9</td>
<td>16.0</td>
<td>21.9</td>
</tr>
<tr>
<td>0.92</td>
<td>1.08</td>
<td>2.71</td>
<td>12.8</td>
<td>13.0</td>
<td>25.4</td>
</tr>
<tr>
<td>A, 0.16 μg/μl; B, 0.08 μg/μl; C, 0.04 μg/μl salmon sperm DNA.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 $2.35 \times 10^{-9}$ and $1.89 \times 10^{-9}$ M for DPC80 and DPC107, respectively.

The value of $K_d$ did not change significantly when higher order series terms ($i/k_i$) of equation 10 were added to equation 11. Even with these higher terms, the y-intercept is still $-1/K_i$. 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 ($2.35 \times 10^{-5}$ and $1.59 \times 10^{-5}$ 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 ($p_o$) 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 DPC107 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.
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 \( c \) 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 \( 1.6 \times 10^{-5} \) of the nuclear extract protein, which is \( \approx 4000 \) 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 \( 1.4 \times 10^{-10} \) moles. This value corresponds to \( 1.5 \times 10^{-10} \) g bound protein, which is equivalent to \( 2 \times 10^{-5} \) g crude nuclear extract. Therefore the DPC80 protein is \( 7.5 \times 10^{-6} \) 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 I 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.

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 \( 2.83 \times 10^{18} \) 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 \( 0^\circ \)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 32P 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 (dj) concentrations. The dissociation constant was within 5% to calculation with non-crosslinked DNA (dnc). 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 Keq 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 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 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 transcriptional regulation of thymidine kinase (manuscript in preparation).
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 $2.35 \times 10^{-9}$ and $1.89 \times 10^{-9}$ 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: $K_d = 1.93 \times 10^{-9}$ M for purified Egr-1 and $K_d = 2.20 \times 10^{-9}$ 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^-15 l (i.e., 10 × 10 × 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**

This work was supported by National Institute of Health grant #GM24571 and Tobacco Council research grant # 3415 (to A.B.P.), and Richard B. Carter fellowship (to D.W.B.). The authors thank Drs Peng Liang and Ulla Hansen for their comments and suggestions. D.W.B. is grateful to both his undergraduate advisor, Dr Alan F. Horwitz for teaching instrumentation skills and to his graduate dean, Dr John Collier for the opportunity to apply these skills.

**REFERENCES**