ELSEVIER

Biophysical Chemistry 52 (1994) 191–218

Biophysical Chemistry

Effect of ethidium binding and superhelix density on the supercoiling free energy and torsion and bending constants of p30δ DNA

James B. Clendenning, Alexei N. Naimushin, Bryant S. Fujimoto, Doyle W. Stewart, J. Michael Schurr

Department of Chemistry, BG-10, University of Washington, Seattle, WA 98195, USA

Received 14 October 1993; accepted in revised form 3 March 1994

Abstract

Topoisomer distributions created by the action of topoisomerase I on p30δ DNA in the presence of various concentrations of ethidium are measured and analyzed using recently developed theory to obtain the twist energy parameter (E_T) that governs the free energy of supercoiling in each case. Competitive dialysis experiments to investigate the relative affinity of ethidium for linear and supercoiled DNAs at different binding ratios are assayed fluorometrically and the results are analyzed using related theory. The topoisomer distributions and fluorescence intensity ratios agree well with the theory, which is based on the assumption that the supercoiling free energy varies quadratically with the effective linking difference, regardless of ethidium binding or superhelix density. The topoisomer distribution experiments alone yield an average best-fit value, $E_T = 950 \pm 80$, independent of ethidium binding ratio from $r = 0$ to 0.082, while the combined topoisomer distribution and ethidium binding experiments yield an average best-fit value, $E_T = 1030 \pm 90$, which is essentially independent of ethidium binding ratio from $r = 0$ to 0.082 and superhelix density from $\sigma = 0$ to $(-)0.053$. One may conclude that the supercoiling free-energy-varies quadratically with effective linking difference over the entire range of observed ethidium binding ratios and superhelix densities. The independently measured torsion constant (α) of p30δ DNA is likewise essentially independent of superhelix density and ethidium binding ratio. The observed invariance of E_T and α implies that the bending constant κ_β is similarly invariant to superhelix density and ethidium binding ratio. The apparently ideal behavior displayed by p30δ DNA is not exhibited by pBR322 DNA, which is discussed in the following companion paper.

Keywords: Ethidium binding; Superhelix density; p30δ DNA

1. Introduction

The free-energy of supercoiling drives certain structural transitions in circular DNAs and also governs the interactions of supercoiled DNAs with unwinding ligands. In addition, it contains information about

SSDI 0301-4622(94)00038-L

the elastic constants for twisting and bending of DNA. Despite its fundamental importance, our knowledge of the free energy of supercoiling is rather limited, incomplete in some important respects, and contradictory in others. Supercoiling free energies have been determined by ligation methods, in which fluctuations in superhelix density are sampled by ligase or topoisomerase I (Topo I), and also by dye-binding methods. The action of Topo I on a circular DNA yields a distribution of topoisomers *around zero superhelix density* [1]. In the case of *long* DNAs with more than 2000 base-pairs (bp), the free energy of supercoiling obtained by analyzing such distributions varies quadratically with the linking difference (Δl). The length-independent proportionality constant, namely the twist energy parameter (E_T), that governs the free energy of supercoiling is usually found to be about $E_T = 1000 \pm 100$ by such measurements [2–4], although substantially (≈ 1.4-fold) larger values were reported by Shore and Baldwin [5,6]. (Our E_T corresponds to the NK/RT used by Wang and co-workers [3,4] and Shore and Baldwin [5,7], where N is the number of base-pairs). The enhanced affinity of negatively supercoiled DNAs for intercalated ethidium or chloroquine was used to determine E_T over a range of superhelix densities (σ, defined below) from native values ($\sigma = -0.05$ to -0.06) *at low binding ratios* to values nearer to $\sigma = 0$ *at higher binding ratios* [8–11]. More often than not, these dye-binding studies yielded $E_T \lesssim 500$, although higher values ($E_T = 890$, 1250) were also sometimes observed. This discrepancy between E_T values obtained from topoisomer distributions created near zero superhelix density in the absence of ethidium and most E_T values obtained from ethidium and chloroquine binding experiments, which are heavily weighted toward native superhelix density, raises the question of whether E_T might be affected directly by either ethidium binding or negative superhelix density, or both.

According to some Monte Carlo simulations, the free energy of supercoiling is expected to rise less rapidly than quadratically with the linking difference for superhelix densities beyond the range of *practically accessible* topoisomers in an equilibrium population, unless the effective diameter of the DNA (in the excluded volume sense) is 100 Å or more [12]. Taking account of transitions to underwound structures, such as separated strands, would be expected to further diminish the rise in supercoiling free energy with linking difference [13–15]. Indirect experimental evidence for transitions to underwound states (at low ionic strength) and a weaker than quadratic dependence of supercoiling free energy on linking difference in the intermediate and native ranges has been reported [16,17]. Thus, on the basis of such theoretical predictions and experimental precedents, one might expect a significant *decrease* in E_T upon increasing negative superhelix density from zero into the native range ($\sigma = -0.05$).

It has been argued that two previous ethidium binding studies [8,10] demonstrate the quadratic dependence of supercoiling free energy on linking difference from zero up to native superhelix density. However, those studies rely on the unproven assumption that E_T is unaffected by ethidium binding, which was never verified by independent experiments. Moreover, those studies provide essentially no information regarding either the quadratic variation of supercoiling free energy with linking difference or the value of E_T at low superhelix density, since the errors in estimates of E_T diverge at small superhelix density. Thus, they fail to establish the quadratic dependence over the full range of superhelix density.

In general, E_T depends upon the torsion constant α and the bending constant κ_β between base-pairs, as well as the interaction potential between non-adjacent subunits of the chain. Provided the interaction potential between non-adjacent subunits depends only on the subunit positions, but *not* on their torsion coordinates, the contributions of the torsion and bending coordinates to the partition function are separable except for the net twist, which is coupled to the net writhe by the topological constraint in Eq. (2) below [11]. Then, under conditions when the distribution of writhe in the unconstrained (nicked) circular DNA is gaussian, as suggested by all Monte Carlo simulations at low superhelix density and by some (but not all) Monte Carlo calculations at higher superhelix density [12,18–20], E_T is related to α and κ_β by

$$E_T = \left[(2\pi)^2/2k_BT\right]\left[\alpha\kappa_w/(\alpha+\kappa_w)\right], \tag{1}$$

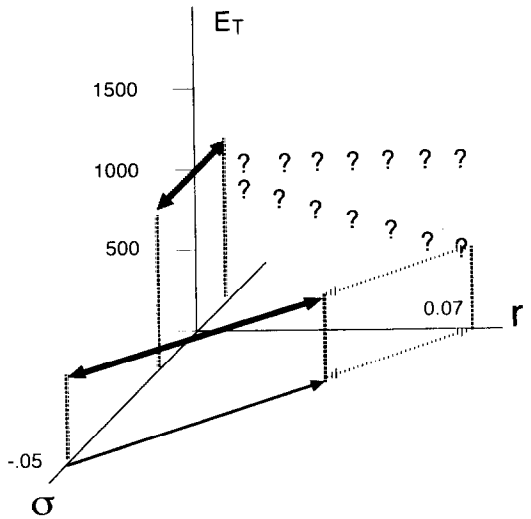


Fig. 1. Schematic illustration of consensus E_T values versus superhelix density (σ) and bound ethidium per base pair (r). Topoisomer distribution methods typically yield $E_T \approx 1000$ in the region near $\sigma = 0$ in the E_T–σ plane (stippled two headed arrow). Dye binding methods typically yield $E_T \approx 500$ along the diagonal (solid two headed arrow). E_T values determined by dye binding methods are heavily weighted toward the large σ end of that diagonal. Question marks indicate the region in the E_T–r plane that has not been previously examined.

where κ_w is the effective force constant for writhe of the unconstrained circle which is defined as Nk_BT divided by the variance in writhe of the unconstrained circle. κ_w depends upon both κ_β and the excluded volume interactions, which are primarily electrostatic. The long-range electrostatic potential and excluded volume interactions should be largely unaffected by modest levels of ethidium binding, due in part to the compensating release of 'condensed' Na^+ counterions in the sense described by Manning [21]. Under conditions where the excluded volume interactions remain constant, κ_w is very likely nearly proportional to κ_β [17], and should in any case be quite sensitive to changes in κ_β. Direct measurements of the torsion constants of *linear* DNAs in 0.1 M NaCl by fluorescence polarization anisotropy (FPA) show that ethidium binding up to 1 dye per 5 bp has no effect on α [22]. Recent force versus extension measurements on single *linear* DNA molecules in 0.1 M NaCl indicate that the persistence length $P = h\kappa_\beta/k_BT$ (where h is the rise per bp) is also unaffected by bound ethidium [23]. From these observations, one might expect that α, P, κ_β, κ_w and E_T should all be independent of ethidium binding, at least near zero superhelix density. However, FPA measurements on two *ethidium-relaxed native supercoiled* DNAs in 0.1 M NaCl indicate that α actually declines substantially with increasing ethidium binding to those DNAs [22]. This puzzling difference between linear and dye-relaxed supercoiled DNAs is also observed for chloroquine [11], and was attributed to metastable secondary structure prevailing in the dye-relaxed molecules [22]. In view of this complication, the above expectation that E_T will be independent of bound ethidium, at least near $\sigma = 0$, is by no means a foregone conclusion.

To better comprehend the domains of existing E_T measurements and possible origins of the discrepancy noted above, it is instructive to consider the diagram in Fig. 1, where consensus E_T values are plotted versus both superhelix density (σ) and ethidium binding ratio (r, bound ethidium per bp). The problem is to 'connect' the domain of the topoisomer distribution measurements near $\sigma = 0$ in the absence of dye with the domain of the dye-binding measurements, which extend across the σ–r plane on a diagonal line from ($\sigma = -0.05$, $r = 0$) toward ($\sigma = 0$, $r = 0.067$). In fact, the existing dye-binding measurements of E_T are weighted substantially toward the end of the line near native σ and rather small r. Although we are unable to extend the topoisomer distribution measurements to sample higher

superhelix densities, we can now extend such measurements in the direction of increasing bound ethidium. This is facilitated by use of a Topo I (from chicken erythrocytes) that does not require Mg^{2+}, so the reaction proceeds under conditions where the ethidium is practically all bound, and by the recent development of theory to analyze the resulting distribution of topoisomers. We can also extend the dye binding measurements farther toward higher binding ratios and closer to zero superhelix density, although uncertainties in the resulting E_T values diverge as one approaches the $\sigma = 0$ line. In this way, we can in a nearly continuous manner examine the effects of both ethidium binding and superhelix density on E_T. That is our primary objective. In addition, we measure the torsion constant α directly by FPA at different superhelix densities and ethidium binding ratios. Combining this information with the E_T results provides information about any variations of the bending constant κ_β with superhelix density or ethidium binding. That is our secondary objective.

2. Theory for topoisomer distribution experiments

2.1. Free energy difference between topoisomers in the absence of ethidium

Each topoisomer exhibits an integral number of turns l of one strand around the other. This linking number is partitioned between net twist (t) and net writhe (w) according to the topological constraint [24,25],

$$l = t + w. \tag{2}$$

In the absence of unwinding ligands, the linking difference is defined by $\Delta l = l - l_0$, where l_0 is the non-integral equilibrium (or intrinsic) twist of the unstrained (nicked) circular DNA with N base-pairs, and the superhelix density is defined by $\sigma = \Delta l / l_0$. It is found empirically that the free energy difference between topoisomers with linking numbers i and m *near zero superhelix density* is given by

$$\Delta F^\circ_{im} = k_B T (E_T/N)\left[(i - l_0)^2 - (m - l_0)^2\right], \tag{3}$$

wherein T is the absolute temperature, k_B is Boltzmann's constant, N is the number of base-pairs, and E_T is the twist energy parameter discussed above [2,3]. When the DNA is modelled by a chain of subunits connected by identical Hookean torsion and bending springs, and provided the distribution of writhe in the unconstrained (nicked) circular DNA is Gaussian, then Eqs. (1) and (3) also follow on theoretical grounds [6,11]. In the absence of ethidium, the ratios of topoisomer concentrations are governed simply by ΔF°_{im}, but in the presence of ethidium the situation is considerably more complex, as described in section 2.2.

2.2. Topoisomer distribution in the presence of ethidium

The binding of n_E ethidiums to a circular DNA reduces its intrinsic twist from l_0 to $l_0^{eff} = l_0 - n_E(\phi_E/360)$, where ϕ_E is the unwinding angle of ethidium, and alters its deformational strain accordingly. We define the effective linking difference in the presence of bound ethidium by $\Delta l_{eff} = l - l_0^{eff}$, and the corresponding superhelix density by $\sigma = \Delta l_{eff}/l_0^{eff}$. We assume that, even in the presence of ethidium, the distribution of topoisomers produced by the action of Topo I corresponds to an *equilibrium* distribution. When ethidium is present, the resulting distribution of topoisomers is centered around $\Delta l_{eff} = 0$, and is broadened by fluctuations in the amount of bound ethidium, as well as by the usual fluctuations in deformational strain energy. When the topoisomerization reaction is halted by removal of the enzyme, and the ethidium is then removed, the final distribution becomes centered around $\Delta l = -\langle n_E \rangle (\phi_E/360)$, where $\langle n_E \rangle$ is the average number of bound ethidiums prior to removal, but the width remains unchanged. It is necessary to account for this width, which is enhanced due to fluctuations

in the number of bound ethidiums (n_E) at the time of closure, in order to extract a value of E_T from the measured topoisomer distribution. A theory for the ratios of total concentrations of any two topoisomers created by the action of Topo I, or by ligation, in the presence of ethidium was developed recently [6], and is summarized in Appendix A. Each intercalated ethidium is assumed to add one more net torsion and bending spring, to increase the length by h_E, to decrease the intrinsic twist by ϕ_E degrees, and to exclude binding of another dye to its nearest neighbor intercalation sites. The additional net torsion and bending spring is assumed to be identical to those between base-pairs. This assumption is supported by the observed invariance of the torsion constant and persistence length of linear DNAs to intercalated ethidium [22,23]. Moreover, this assumption enabled a rather good fit of the remarkable stepped curve of average linking number (relative to the nearest integer to the intrinsic twist) vs. free ethidium concentration in the ligation medium for a 247 bp DNA [5,6]. For a long DNA, such as that studied here, the effects of moderate ethidium binding on the minimum bending strain energy required for circularization, on the contribution of the closure constraints, and on two additional factors discussed in Appendix A are entirely negligible. The increase in number of springs has only a very minor ($\sim 1\%$) effect on the supercoiling free energy, but is retained here for completeness. The ratio of total concentrations of two topoisomers is given by

$$\frac{c^{\mathrm{tot}}_{\mathrm{C}(i)}}{c^{\mathrm{tot}}_{\mathrm{C}(m)}} = \exp\left[-\left(\frac{E_T}{N}\right)\left[(i-l_0)^2-(m-l_0)^2\right]\right]\left(\frac{1+\mathrm{GPF}(i)}{1+\mathrm{GPF}(m)}\right), \tag{4}$$

wherein the grand partition function is

$$\mathrm{GPF}(i) = \sum_{n_E=1}^{n_{\max}} (c_E K_E)^{n_E} \exp\left[-\left(\frac{E_T}{N}\right)\left(\frac{\phi_E}{360}\right)^2\left(\frac{n_E^2 - 2n_E\left(n^*_{Ei}+n^{*2}_{Ei}/N\right)}{1+n_E/N}\right)\right] W(n_E), \tag{5}$$

c_E is the concentration of free ethidium, $K_E = \exp(-\Delta F^\circ_E/k_B T)$ is the intrinsic binding constant for ethidium,

$$n^*_{Ei} = -(i-l_0)(360/\phi_E) \tag{6}$$

is the number of bound ethidiums that relaxes Δl_{eff} to zero, and the degeneracy factor $W(n_E)$ is given by Eqs. (A 17)–(A 20), which apply for nearest-neighbor exclusion.

2.2.1. Center of the distribution

We define a function, NINT(X) ≡ nearest integer to X. In the absence of ethidium, the most probable topoisomer (C(m)) has linking number m = NINT (l_0), where $l_0 = N/\mathscr{P}$, and $\mathscr{P}$ is the average pitch of the helix. Previously measured values for long DNAs in vitro lie in the range $\mathscr{P} = 10.45 \pm 0.05$ bp/turn [2–6,7,26]. The center of the distribution resulting from the action of Topo I in the presence of ethidium is located near $l(n_E)_U = l_0 - \langle n_E\rangle_U\ (\phi_E/360)$, and the linking number of the most probable topoisomer is initially taken as m = NINT ($l(n_E)_U$). The quantity $\langle n_E\rangle_U$ is the average number of ethidiums bound to an unstrained (nicked circular or linear) plasmid under the Topo I reaction conditions, and is given by

$$\langle n_E\rangle_U = N c_E K_E (1-2r_U)^2/(1-r_U), \tag{7}$$

wherein $r_U = \langle n_E\rangle_U/N$. By considering a sufficiently large range of adjacent linking numbers, $i = m \pm 1$, $\pm 2, \ldots, \pm 10$, the precise value of m is rendered completely inconsequential.

2.2.2. Normalization

We need to express $c^{\mathrm{tot}}_{\mathrm{C}(m)}$ in Eq. (4) in terms of the total DNA concentration, c^{tot}_D. Combining Eqs. (3) and (4) with the conservation condition,

$$c^{\mathrm{tot}}_D = \sum_i c^{\mathrm{tot}}_{\mathrm{C}(i)} + c^{\mathrm{tot}}_{\mathrm{nc}}, \tag{8}$$

where c_{nc}^{tot} is the concentration of nicked circles, yields

$$c_{C(m)}^{tot} = \left(c_D^{tot} - c_{nc}^{tot}\right) \Bigg/ \left(\sum_i \exp[-\Delta F_{im}^{\circ}] \frac{1 + GPF(i)}{1 + GPF(m)} \right) \tag{9}$$

2.2.3. All bound assumption

Eq. (4) contains three unknowns, namely E_T, l_0, and the product $c_E K_E$. K_E is not known with certainty under conditions of the Topo I reaction, and c_E in the reaction mixture is not actually measured. However, the ethidium is practically *all bound* under our reaction conditions, so

$$c_E^{tot} = \sum_i c_{C(i)}^{tot}(\langle n_E(C(i))\rangle) + c_{nc}^{tot}(\langle n_E\rangle_U), \tag{10}$$

where $\langle n_E\rangle_U$ is given in Eq. (7), and $\langle n_E(C(i))\rangle$ is the average number of ethidiums bound per topoisomer with linking number i, which is calculated by direct averaging over the corresponding grand distribution function according to Eqs. (A 21) and (A 22). In other words, under the conditions of our experiment, the concentration of free ethidium contributes negligibly to the r.h.s. of Eq. (10). For a given choice of E_T and l_0, we can adjust $c_E K_E$ until Eq. (10) is satisfied.

2.2.4. Correction for differences in bound ethidium in the gel

The distribution of topoisomers is measured after removing the Topo I and ethidium and separating the topoisomers in an agarose gel. The gel is then equilibrated with a solution of much larger volume containing ethidium and photographed under UV light using negative-producing film. The negative is scanned using a microdensitometer and the density profile extracted. The area under any peak is proportional to the total amount of fluorescence due to bound ethidium, which in turn is a function of the total mass of DNA in that band and its superhelical strain. In fact, the effect of the different superhelical strains is practically negligible, but is taken into account for completeness. The ratio of the areas of two peaks in the densitometer tracing is related to the ratios of the number of bound ethidiums and amounts of DNA in the two bands by

$$\frac{A_i}{A_m} = \frac{\langle n_E(C(i))\rangle_{gel} c_{C(i)}^{tot}}{\langle n_E(C(m))\rangle_{gel} c_{C(m)}^{tot}}, \tag{11}$$

where $\langle n_E(C(j))\rangle_{gel}$ is the number of bound ethidiums per topoisomer j under staining conditions in the gel. Since extremely high accuracy is not required, we calculate $\langle n_E(C(j))\rangle_{gel}$ using the simple maximum term formula [11],

$$r_j = c_E K_E \exp\left[-2E_T(\phi_E/360)^2(r_j - r_j^*)\right](1 - 2r_j)^2/(1 - r_j) \tag{12}$$

wherein $r_j = \langle n_E(C(j))\rangle/N$, $r_j^* = n_{Ej}^*/N = -(j - l_0)(360/\phi_E)/N$, c_E is the known concentration of free ethidium in the stained gel. K_E is the intrinsic binding constant for ethidium in the gel, and $\phi_E = 26°$ [9,27,28]. Since the total concentration of ethidium in the stained gel vastly exceeds the DNA concentration at any point therein, we take c_E simply as the total concentration of ethidium in the stained gel, or equivalently in the equilibrated staining solution. In calculating $\langle n_E(C(i))\rangle_{gel}$, we assume the standard value $K_E = 3.5 \times 10^5$ M^{-1} [29]. The calculated ratios $\langle n_E(C(i))\rangle_{gel}/\langle n_E(C(m))\rangle_{gel}$ are very nearly 1.0 and rather insensitive to the choice of c_E or K_E. In any case, A_i/A_m is readily calculated for any theoretical value of the ratio $c_{C(i)}^{tot}/c_{C(m)}^{tot}$ via Eq. (11).

2.3. The fitting algorithm for topoisomer distributions

Given that $N = 4752$ and $\phi_E = 26°$ are known, there remain three unknowns in Eqs. (4) and (5), namely E_T, l_0, and the product $c_E K_E$. These are adjusted until the theoretical ratios of peak areas in

Eq. (11) optimally fit the experimentally measured ratios, in the sense that the sum of squared differences between experimental and theoretical ratios.

$$\mathrm{SSD} = \sum_i \left(\left(\frac{A_i}{A_m} \right)^{\mathrm{ex}} - \left(\frac{A_i}{A_m} \right)^{\mathrm{th}} \right)^2 \tag{13}$$

is a minimum, while at the same time the constraints on total DNA concentration (Eq. (9)) and total added ethidium (Eq. (10)) are maintained. The fitting algorithm performs two checks in regard to m. Specifically, it certifies that (1) the calculated total concentration $c^{\mathrm{tot}}_{C(m)}$ of the mth topoisomer is greatest, and (2) the calculated peak area ratio $A(m)/A(m) = 1.0$ is the largest. In all examples encountered in the present work, both checks were satisfied. In the event that another topoisomer m' should give a peak area ratio exceeding 1.0, one needs only to divide all peak areas by $A(m')/A(m)$, which is equivalent to taking m' in place of m. An outline of the fitting algorithm, which proceeds essentially by a sequential grid search method, are presented in Appendix B. The entire procedure for one ethidium concentration takes about two hours on a Convex C-1 minisupercomputer when the code is fully optimized using the Convex Fortran Vectorizing Compiler, and takes approximately 10 h on a Dec Station 3100. Described elsewhere [30] is a rapid and accurate approximate method for calculating GPF(i) that enables this algorithm to be performed on almost any small computer.

3. Theory for ethidium binding experiments

Equilibrium dialysis experiments are performed such that two *equal* volumes of solution are separated by a semi-permeable membrane. One volume contains linear or unstrained DNA at concentration $c^{\mathrm{tot}}_{\mathrm{U}}$, while the other contains native supercoiled circular DNA at concentration $c^{\mathrm{tot}}_{\mathrm{C}}$, which includes a small fraction of nicked circular DNA that is henceforth ignored. The native supercoiled DNA exists as a distribution of covalently closed circular molecules $C(i)$ that differ by 1 in their linking difference $\Delta l(i) = l_i - l_0$. The fraction of all topoisomers with linking number l_i is denoted by f_i. The most probable topoisomer is designated by $i = m$.

Sufficient ethidium bromide is added to one volume or the other to produce a total concentration $c^{\mathrm{tot}}_{\mathrm{E}}$ *averaged over both volumes*. The added ethidium bromide per base-pair ratio is then $\mathrm{EB/BP} = c^{\mathrm{tot}}_{\mathrm{E}} / [\frac{1}{2} N (c^{\mathrm{tot}}_{\mathrm{U}} + c^{\mathrm{tot}}_{\mathrm{C}})]$. The conservation condition for ethidium is

$$c^{\mathrm{tot}}_{\mathrm{E}} = \tfrac{1}{2} \left(c^{\mathrm{tot}}_{\mathrm{U}} N r_{\mathrm{U}} + c^{\mathrm{tot}}_{\mathrm{C}} N \sum_i f_i r_i \right) + c_{\mathrm{E}}, \tag{14}$$

where r_{U} and r_i are the ethidium binding ratios for the unstrained linear molecule and the ith topoisomer, respectively, and are defined following Eqs. (7) and (12) above. In these experiments c_{E} is practically negligible compared to the other terms in Eq. (14). We choose experimental conditions such that $c^{\mathrm{tot}}_{\mathrm{U}} = c^{\mathrm{tot}}_{\mathrm{C}}$. Then, according to Eq. (14) the average number of *added* ethidiums per base-pair becomes

$$\mathrm{EB/BP} = \tfrac{1}{2} \left(r_{\mathrm{U}} + \sum_i f_i r_i \right) + c_{\mathrm{E}} / \left(N c^{\mathrm{tot}}_{\mathrm{C}} \right). \tag{15}$$

The fluorescence intensity signal per *bound* ethidium is the same for the unstrained DNA as for supercoiled DNA and exceeds that of the *free* ethidium by a factor of 12 [11]. Hence, the free ethidium on either side makes an even smaller contribution to the total fluorescence than to the total ethidium

concentration, and is ignored in the fluorescence ratio. The theoretical fluorescence intensity ratio for equal aliquots of solution from either side of the membrane is then given by

$$\left(\frac{f_C}{f_U}\right)^{th} = \frac{c_C^{tot} N \sum_i f_i r_i}{c_U^{tot} N r_U} = \frac{\sum_i f_i r_i}{r_U}, \tag{16}$$

where the condition $c_C^{tot} = c_U^{tot}$ is employed in the final equality. The quantity r_i is given by

$$r_i = c_E K_E \exp\left[-E_T(\phi_E/360)^2\left(\frac{r_i^2 - r_i^{*2} - 2(r_i - r_i^*)}{(1+r_i)^2}\right)\right]\frac{(1-2r_i)^2}{(1-r_i)}, \tag{17}$$

wherein $r_i^* = -(\Delta l(i)/N)(360/\phi_E)$. Eq. (17) is obtained from the maximum term of the grand partition function in Eq. (A 16) in the manner described by Wu et al. [11]. It differs from Eq. (12) in that the increase in length, or number of bending and torsion springs, is taken into account, which yields slightly more accurate results. The quantity $r_U = \langle n_E \rangle_U / N$ is calculated from Eq. (7). The $f_i, i = 1, \ldots, p$, are presumed known, as described below. We adopt the standard value, $K_E = 3.5 \times 10^5\ M^{-1}$ [29], but it must be noted that the best-fit values of E_T, r_U, and r_j, $i = 1, \ldots, p$, are extremely insensitive to the choice of K_E, when the ethidium is practically all bound, as is the case here. This can be understood as follows. When the c_E terms on the right hand sides of Eqs. (14) and (15) are neglected, which is valid when they are sufficiently small compared to the remaining terms, then c_E occurs always jointly with K_E as the product $c_E K_E$ (cf. Eqs. (7) and (17)). Under such conditions any value of K_E is precisely compensated by c_E so as to yield the same final results for E_T, r_U, and r_j, $j = 1, \ldots, p$. There are in all $p + 3$ unknowns, namely r_i, $i = 1, \ldots, p$, r_U, c_E, and E_T. Relating these are $p + 3$ equations, namely Eqs. (7), (15), (16) (when the l.h.s. is equated to the experimental value), and Eq. (17) for $i = 1, \ldots, p$, so all of the unknowns, in particular E_T, can be obtained for *each* dialysis experiment at its particular value of EB/BP.

3.1. Topoisomer fractions

The topoisomer fractions are estimated according to

$$f_i = \exp\left[-(1000/2N)\Delta l(i)^2\right] / \sum_j \exp\left[-(1000/2N)\Delta l(j)^2\right]. \tag{18}$$

where i runs from 1 up to $p = 13$, and the $\Delta l(i)$ are chosen such that $\Delta l(7) = \Delta l(m)$ is the linking difference of the most probable topoisomer in the *observed* distribution. $\Delta l(m)$ is determined to the nearest integer by the band counting method described below. The average linking difference for this distribution is $\overline{\Delta l} = \sum_l f_i \Delta l(i) = \Delta l(m)$. Eq. (18) has a width twice that of a normal thermal distribution, close to those of the native supercoiled DNAs examined here. For such long DNAs, the final results are extraordinarily insensitive to the width of the distribution. Indeed, the same values of E_T (within 1 part in 10^5) are obtained for a distribution consisting of a single topoisomer with $\Delta l = \Delta l(7) = \overline{\Delta l}$, as for the entire distribution in Eq. (18).

3.2. The solution / fitting algorithm for ethidium binding experiments

For each experimental value of EB/BP we measured the ratio of fluorescence intensities $(f_C/f_U)^{ex}$. Then for each value of EB/BP, the values of the $p + 3$ unknowns, r_i, $i = 1, \ldots, p$, r_U, c_E, and E_T are determined by a grid search algorithm that minimizes the squared deviation,

$$SD = \left[(f_C/f_U)^{th} - (f_C/f_U)^{ex}\right]^2 \tag{19a}$$

to a value arbitrarily close to zero. Details of the algorithm are supplied in Appendix C. Known input parameters are $N = 4752$, $\phi_E = 26°$ and the f_i, which are calculated according to Eq. (18) using the measured value of $\Delta l(m) = \Delta l(7)$. $\Delta l(m)$ was -24 ± 1 turns for one preparation of p30δ DNA and -25 ± 1 turns for the other. This fitting algorithm does not require evaluation of the GPF(i), and consequently runs very rapidly.

We could also assume that a single value of E_T characterizes all (Q) of the data points. In this case, one has input values $(EB/BP)_q$ and $(f_C/f_U)_q^{ex}, q = 1, \ldots, Q$. The values of the $Q(p+2)+1$ unknowns, namely, $\{r_i^q, i = 1, \ldots, p, r_U^q$, and $(c_E)_q\}$, $q = 1, \ldots, Q$, and E_T are determined using essentially the same algorithm to minimize the sum of squared deviations.

$$\mathrm{SSD} = \sum_q \left[(f_C/f_U)_q^{th} - (f_C/f_U)_q^{ex} \right]^2, \tag{19b}$$

as described in Appendix C.

4. Materials and methods

4.1. DNA preparation

Five μg of plasmid p30δ were generously supplied by Professor Ernesto di Mauro, Department of Genetics and Molecular Biology, University of Rome. p30δ was constructed by cloning the approximately 700 bp BglII-to-SalI fragment from *S. cerevisiae*, which contains the left Ty1(D15)δ element described by Cameron et al. [31], into the region between the BamHI and SalI sites of pBR322 [32,33]. The resulting plasmid contains 4752 bp. p30δ was used to transform competent *E. coli* strain HB101 essentially as described by Miller [34]. The transformation, plating, growth, and harvesting of the cells, as well as cell lysis, DNA preparation, and plasmid isolation by banding in a CsCl density gradient followed essentially standard protocols described elsewhere [30]. Ethidium was removed by 8–10 extractions with buffered phenol, followed by dialysis alternately versus high salt (HSTE) (500 mM NaCl, 10 mM Tris, 1 mM Na_2 EDTA, pH 7.8 at 22°C) and normal salt (NSTE) (100 mM NaCl, 10 mM Tris, 1 mM Na_2 EDTA, pH 7.8 at 22°C) buffers at 4°C, terminating finally in the latter, which is the storage buffer. Dialysis was continued until ethidium fluorescence excited by a 351.1 nm Ar-ion laser beam was no longer visible. Samples prepared in this way and maintained at 4°C in the normal salt buffer typically contain $> 90\%$ covalently closed circles even after two years. Most of the present experiments were performed within 5 months of cell lysis. For the samples prepared in this way, $A_{260}/A_{280} \approx 2.0$.

4.2. Determination of the approximate linking difference of the most probable topoisomer

The approximate linking difference Δl_{app} of the most probable topoisomer (i.e. brightest band) in each sample in the *absence* of ethidium was determined by the band counting method [1]. Chloroquine was used instead of ethidium to shift l_0 in the gel, because it provides better resolution. Electrophoresis was performed, as described in Section 4.5 below, using TAE buffer (0.037 M Tris base, 0.019 M acetic acid, 1 mM EDTA, pH 8.5), which also enhances resolution. Δl_{app} is found simply by adding up the differences in Δl_{app} values between samples, beginning with the sample relaxed in the absence of ethidium. The standard deviations of such determinations are about one turn, due to our inability to always correctly identify the most probable topoisomer in each sample.

4.3. Linearization

Linearized samples required for competitive dialysis were prepared by reaction of the plasmids with EcoRI for 24 h under conditions recommended by the supplier (Bethesda Research Laboratories).

Table 1
Chloroquine concentration in the gel for topoisomer resolution

EB/BP	Δl_{app}	Chl (μg/ml)
0.0	0	0.72
0.0102	−4	1.36
0.0204	−6	0.0
0.0305	−10	0.36
0.0402	−15	0.36
0.0503	−19	0.72–1.36
0.0599	−23	2.0
0.0820	−30	4.2

Reaction progress was monitored by gel electrophoresis of aliquots withdrawn at intermediate times, and was judged complete in 10.5 h. The reaction mixture was extracted three times with buffered phenol and three times with diethyl ether saturated with water. The sample was then dialyzed alternately versus HSTE and NSTE buffers for four days, changing buffers every 24 h.

4.4. Preparation and reactions of topoisomerase I

Topo I was prepared from fresh chicken erythrocytes by modifications of the methods described by Hymer and Kuff [35] for obtaining the nuclei and by Bina-Stein et al. [36] for extracting the Topo I, as described elsewhere [30]. The Topo I extract is stable for up to 18 month when stored at 4°C, but loses activity relatively quickly upon dilution.

Topo I reactions were performed in 100 mM NaCl, 10 mM Tris-HCl, 1 mM Na_2 EDTA, 1.48 mM sodium phosphate plus various ethidium concentrations in total volumes of 1.608 mL. The total amount of DNA relaxed in each reaction was 100 μg. The *added* ethidium per bp ratios (EB/BP) for each reaction and the corresponding Δl_{app} values are listed in Table 1. Reactions were carried out at 37°C in 5 ml sterile culture tubes sealed with parafilm, covered with Al-foil, and placed in an H_2O bath with continuous shaking for 20 h. Reactions were halted by extracting three times with phenol equilibrated with 200 mM NaCl, 100 mM Tris, 0.2% w/v 8-hydroxyquinoline, pH 8.0, followed by three more extractions with diethylether saturated with H_2O. The samples were then placed in 10 mm Spectrapor dialysis tubing with a 12–14 kD molecular weight cutoff and dialyzed alternately versus HSTE and NSTE buffers every 24 h for 4 days at 4°C to remove ethidium and other contaminants. The samples were stored in NSTE buffer at 4°C. After the Topo I reaction, A_{260}/A_{280} ratios ranged from 1.90 to 1.96, with an average value of 1.95, and the A_{260} values ranged from 0.967 to 1.21.

4.5. Gel electrophoresis

Some 0.2 to 0.5 μg DNA were loaded into every other well (1 mm × 6 mm × 8 mm) in a 0.8% agarose gel containing TAE buffer plus chloroquine. The chloroquine concentrations used for each sample are given in Table 1. This gel was immersed in 2 L of the same buffer. Samples were electrophoresed for either 36 h at 50 V in a 28 cm gel or 42 h at 80 V in a 40 cm gel, which exhibited about the same initial and final currents. The buffer was recirculated at 20 mL/min. The concentrations of chloroquine in the gel and the approximate linking numbers of the most probable topoisomer in each case are indicated in Table 1. After electrophoresis, chloroquine was removed by replacing the buffer with 2 L of distilled H_2O, circulating for 1.5 h, then replacing the distilled H_2O and circulating again for 1.5 h. After decanting the distilled water, the gel was immersed in 2 L of 0.5 μg/ml ethidium bromide in the dark, covered with plastic wrap and left to equilibrate for at least 12 h in the dark.

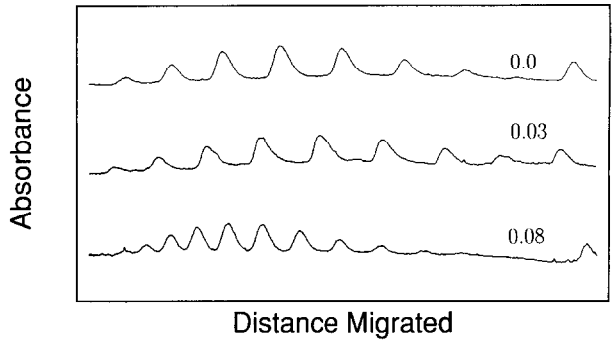


Fig. 2. Typical densitometer tracings of gel lanes containing p30δ relaxed by topoisomerase 1 from chicken erythrocytes in the presence of 0.0 (upper), 0.0305 (middle), or 0.0820 (lower) added ethidiums per base-pair (EB/BP). Topoisomers were resolved in 0.8% agarose gels, stained with ethidium, photographed, and the negatives scanned, as described in the text. Each lane contains 0.3 μg DNA.

4.6. Photography

The stained gel was aligned on a Fotodyne 3-4500 transilluminator and photographed using Kodak type 665 negative film and a Kodak Wratten 22A UV filter. The film response was shown experimentally to be proportional to the amount of DNA in each band for samples containing 0.2, 0.3, 0.4, and 0.5 μg per lane. Within experimental error, identical area ratios were obtained in each case.

4.7. Densitometry

Negatives were scanned using a scanning densitometer (Hoeffer Scientific Instruments, model G5300) connected to an IBM XT computer. Some 2000 data points, each averaging 10 readings per point, were collected for each scan. Three such scans were performed parallel to the direction of migration, one in the center and one nearer each edge. After transfer to the computer, peaks were integrated from valley to valley using the program Spectra Calc (Galactic Industries Corp.). Typical scans are shown in Fig. 2, and the ratios of peak areas obtained therefrom are listed in Table 2.

Table 2
Average relative areas for each reaction

$\Delta l(i) - \Delta l(m)$	EB/BP					
	0.0	0.0102	0.0204	0.0308	0.0599	0.0820
4		0.070			0.055	0.240
3	0.176	0.275	0.141	0.282	0.247	0.413
2	0.463	0.625	0.386	0.472	0.544	0.749
1	0.807	0.989	0.737	0.851	0.776	0.984
0	1.00	1.00	1.00	1.00	1.00	1.00
−1	0.866	0.742	0.983	0.967	0.870	0.766
−2	0.466	0.397	0.719	0.639	0.712	0.469
−3	0.220	0.127	0.349	0.324	0.208	0.215
−5			0.100	0.116	0.112	

4.8. Competitive dialysis

Competitive dialysis experiments were performed using an 8-cell vessel (Hoeffer Scientific Instruments). The two sample chambers in each cell were separated by a 12–14 kD molecular weight cutoff membrane (Hoeffer Scientific Instruments). The cell cleaning and sample loading protocols are described elsewhere [30]. Ethidium concentration in the stock solution was determined from the measured absorbance using $A_{460} = 4220$ cm^2/mol [37]. Equal volumes (1 mL) of solutions containing either 50 μg/mL supercoiled DNA or 50 μg/mL linear DNA plus the prescribed concentration of ethidium were placed in opposing chambers of each cell. Each sample port was sealed with parafilm, and the entire vessel was sealed with alternating layers of parafilm and Saran wrap to prevent entry or loss of water, and finally wrapped in Al-foil. The vessel was placed together with either a pre-equilibrated metal block and crumpled Al-foil or pre-equilibrated sand in a 3 L beaker, which was then immersed in a water bath maintained at 37°C. The sample was equilibrated for 48 h by shaking in a direction perpendicular to the membrane. Control experiments, in which supercoiled p30δ was dialyzed versus ethidium plus buffer, and in which supercoiled p30δ plus ethidium was dialyzed vs. buffer, showed that equilibration is complete in less than 36 h, regardless of which side initially contains the ethidium.

4.9. Fluorescence measurements

Fluorescence intensities of each sample were measured at 37°C by using a Perkin-Elmer 650-10s spectrofluorimeter. Excitation was at 575 nm, emission at 645 nm, and both slit widths were 5 nm. Within each experiment, three measurements were made on each sample using three different photomultiplier gain and sensitivity settings that were previously shown to give linear response in control experiments. In all cases, the three measurements yielded identical fluorescence intensity ratios for the supercoiled and linearized samples from a given chamber. The measured fluorescence intensity ratios are presented in Table 3.

4.10. Torsion constant measurements

Torsion constants were measured by time-resolved fluorescence polarization anisotropy (FPA) of intercalated ethidium. The experimental methods and data analysis protocols were as described previ-

Table 3
Fluorescence intensity ratios for p30δ

EB/BP	f_C/f_U [a]	f_C/f_U [b]	r [c]
0.0042		1.96	0.0056
0.0084	1.80		0.0105
0.0125		1.72	0.0154
0.0168	1.55		0.0198
0.0175		1.67	0.0213
0.0250	1.49		0.0290
0.0250		1.53	0.0293
0.0320		1.39	0.0362
0.0420	1.24		0.0448
0.0420		1.27	0.0453
0.0500	1.18		0.0521
0.0500		1.18	0.0521
0.0590	1.13		0.0601

[a] Experimental fluorescence intensity ratios for p30δ preparation 1 with $\Delta l = -24$ turns.
[b] Experimental fluorescence intensity ratios for p30δ preparation 2 with $\Delta l = -25$ turns.
[c] Binding ratio for the supercoiled species predicted by the fitting routine.

Table 4
Torsion constants of topol reaction products [a]

EB/BP	$\alpha \times 10^{12}$ (dyne cm)
0.0	4.20
0.0102	4.17
0.0308	4.47
0.0406	4.39
0.0500	4.48
0.0599	4.57
ave.	4.38 ± 0.15

[a] Reported values are obtained by assuming a negligible contribution of bending to the FPA, hence they are proportional lower bounds.

ously [22,38]. In all cases, the samples were measured in the NSTE buffer at 20°C. It is assumed here that the tumbling correlation functions are 1.0, which corresponds to infinite bending rigidity, so the present α values are proportional lower bounds. If our current best guess for the dynamic persistence length, $p_d = 1500$ Å, were adopted, then all of the present α values would be multiplied by the factor 1.35 [38,39]. That is not done here, since we are interested primarily in *relative* changes in α. The values reported here are proportional to a despite being lower bounds.

Measurements were performed on the series of p30δ samples that had been relaxed by Topo I in the presence of various amounts of ethidium bromide, but after removal of that ethidium bromide followed by subsequent re-addition to give finally 1 ethidium per 200 bp. The results are presented in Table 4. Measurements were also performed on two samples of native supercoiled p30δ from the dialysis experiments, specifically those with EB/BP = 0.0168 ($r = 0.0198$) and EB/BP = 0.0500 ($r = 0.0521$). These results are presented in Table 5. For these latter samples, depolarization by excitation transfer makes a significant contribution to the FPA signal. A theoretical procedure for taking account of excitation transfer was developed and tested on linear DNAs by [22]. From those results we estimate that the ratio of corrected (for excitation transfer) to uncorrected torsion constants is 1.16 for $r = 0.0198$ and 1.50 for $r = 0.0521$. In the present work, the uncorrected torsion constant in each case is simply scaled by the appropriate factor to obtain the corrected value, which is also listed in Table 5.

5. Results and discussion

5.1. Results of the topoisomer distribution experiments

The experimental and best-fit peak area ratios for topoisomer distributions created in the presence of 0, 0.03, and 0.08 EB/BP are compared in Fig. 3. In these and all other cases examined, the fit is excellent. The optimum values of E_T and pitch $P = N/l_0$ obtained from such fits are listed along with

Table 5
Torsion constants of partially relaxed (by ethidium) native topoisomer distributions

EB/BP	α uncorrected (dyne cm)	α corrected [a] (dyne cm)	r [b]
0.0168	4.02×10^{-12}	4.65×10^{-12}	0.0198
0.0500	3.09×10^{-12}	4.64×10^{-12}	0.0521

[a] The torsion constant is corrected for excitation transfer.
[b] Binding ratio for the supercoiled supercoiled species predicted by the fitting routine.

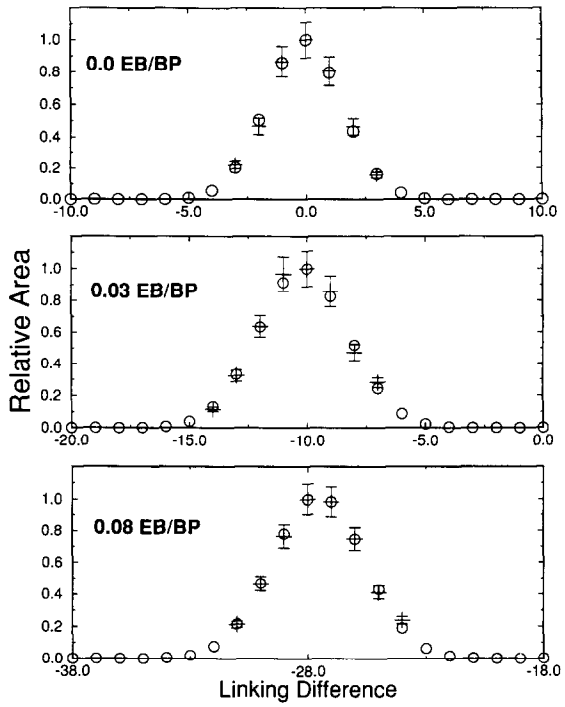


Fig. 3. Relative peak areas in the densitometer scan versus shifted linking difference $\Delta l_i - \Delta\, l_m - \mathrm{NINT}(\langle n_E \rangle_U \phi_E / 360)$. Crosses are experimental data obtained by relaxing p30δ in the presence of 0.00 (upper), 0.0305 (middle), or 0.0820 (lower) added EB/BP. Each data set is from the average of three scans similar to the corresponding scan shown in Fig. 2. Error bars are uncertainties calculated using Eqs. (D.1)–(D.4). Circles are relative areas calculated via Eq. (11) using the optimum parameters $E_T = 900$, $P = 10.446$ bp/turn (upper); $E_T = 820$, $P = 10.442$ bp/turn (middle); $E_T = 1100$, $P = 10.443$ bp/turn (lower), along with $N = 4752$ bp and $\phi_E = 26°$. Optimum parameters were determined using the algorithm, described in Appendix B.

the corresponding SSD in Table 6. Relative errors in E_T are estimated in Appendix D, and found to be 9.6%. E_T values are plotted versus added EB/BP (EB $\approx r$) in the topoisomerization medium in Fig. 4. All data points lie within 1.5 standard deviations of the best-fit horizontal line given by $E_T = 950 \pm 80$. Within experimental error, for topoisomers with superhelix density near zero, E_T is unaffected by ethidium binding up to a binding ratio $r = 0.082$, which corresponds to 1 dye per 12 bp. Evidently, ethidium binding to p30δ DNA *at low superhelix density* causes no structural changes that affect the energetics of supercoiling.

Table 6
Optimum values of E_T and P

EB/BP	E_T	P (bp/turn)	SSD
0.0	900	10.446	2.715×10^{-3}
0.0102	950	10.449	2.780×10^{-3}
0.0204	950	10.457	1.903×10^{-3}
0.0305	820	10.442	7.611×10^{-3}
0.0599	970	10.458	1.374×10^{-2}
0.0820	1100	10.443	2.956×10^{-3}
ave.	950 ± 80	10.449 ± 0.006	

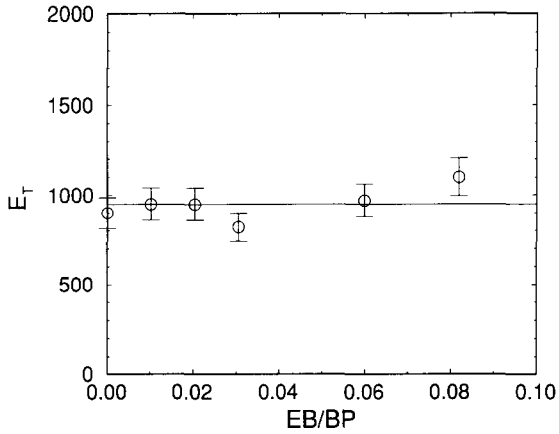
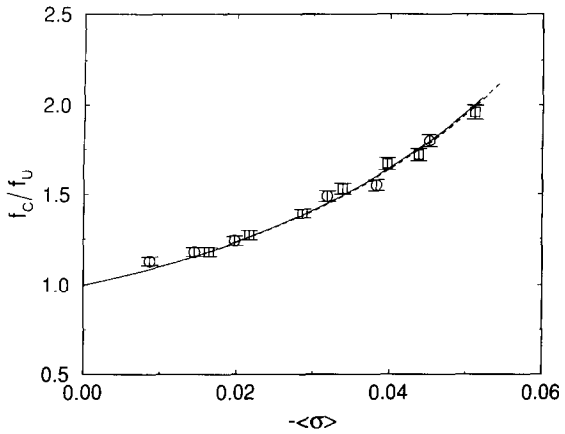


Fig. 4. Optimum E_T versus added ethidium per base-pair in the topoisomerase 1 reaction medium (EB/BP). Circles are optimum E_T values for the samples listed in Table 6. Error bars reflect the estimated 9.6% uncertainty calculated for the sample with 0.0820 EB/BP via Eq. (D.5). Solid line is the average $E_T = 950$.

The best-fit values of the intrinsic pitch all lie remarkably close to the average value, $\bar{P} = 10.449 \pm 0.006$, as indicated in Table 6. Evidently, ethidium binding to p30δ causes no structural changes that affect the intrinsic pitch. A plot of the much less precise Δl_{app} values in Table 1 versus EB/BP (not shown) yields a straight line with a slope that would correspond to an unwinding angle $\phi_E = 28° \pm 2°$, if all the ethidium were bound [30]. In view of the experimental errors and the small systematic error due to the all-bound assumption, this value agrees satisfactorily with the consensus value, $\phi_E = 26°$ that was employed in the data analysis.

5.2. *Width of the topoisomer distribution*

The width of the topoisomer distribution is due to two factors: thermal fluctuations in linking number about the effective intrinsic twist l_0^{eff} at fixed n_E, and fluctuations in l_0^{eff} due to fluctuations in n_E at the time of enzymatic closure. The question arises whether the total variance (σ_{tot}^2) of the linking number distribution is approximately the sum of the variance (σ_T^2) of the linking number about l_o^{eff} at fixed n_E and the variance (σ_E^2) of l_0^{eff} due to fluctuations in n_E. If so, how is it apportioned among those? These questions are addressed for the 0.08 EB/BP sample as follows. The total variance of the distribution is estimated by fitting the observed area ratios A_i/A_m under the (erroneous) assumption that no dye is present, so $c_E = 0$, which gives an effective value of $E_T(E_T^{eff})$ that simply characterizes the Gaussian width of the entire distribution. We obtain $E_T^{eff} = 620$, which corresponds to a total variance $\sigma_{tot}^2 = N/(2E_T^{eff}) = 3.83$. The variance of the linking number about l_0^{eff} at fixed n_E is just $\sigma_T^2 = N/(2E_T)$, wherein $E_T = 1100$ is the actual best-fit value for this sample. Hence, $\sigma_T^2 = 2.16$. The variance in the distribution of $l_0^{eff} = l_0 - n_E(\phi_E/360)$ is obtained from the variance in the number of ethidiums bound to the *nicked circular* DNA at the time of closure, $\sigma_E^2 = (\langle n_E^2\rangle_U - \langle n_E\rangle_U^2)(\phi_E/360)^2$. The quantity $\langle n_E\rangle_U$ is calculated using Eqs. (A.21) and (A.16) with $E_T = 0$, and $\langle n_E^2\rangle_U$ is calculated using the analogous equation to (A.21) with n_E^2 in place of n_E together with (A.16), both with $E_T = 0$. For both calculations, $c_E K_E$ is taken to be the optimum value that was obtained by fitting the area ratios to obtain E_T. We find $\sigma_E^2 = 1.54$. The sum of the individual variances, $\sigma_T^2 + \sigma_E^2 = 3.70$, is nearly equal to the total variance, $\sigma_{tot}^2 = 3.83$. Thus, for the 0.08 EB/BP sample, we find that $(\sigma_E^2/\sigma_{tot}^2)(100) = 40\%$ of the total variance (of the linking number distribution) is due to fluctuations in bound ethidium. More importantly, we may conclude that in long DNAs, such as that considered here, fluctuations in n_E are very nearly indepen-



dent of fluctuations in linking number about l_0^{eff} at fixed n_E, provided the superhelix density is near zero. This conclusion contrasts sharply with that reached for short DNAs containing 237 and 247 bp, in which dye binding is strongly perturbed by the lowest possible levels of supercoiling [6].

5.3. Results of ethidium binding experiments

The fluorescence intensity ratios (f_C/f_U) for sample 1 of p30δ in Table 2 are collectively fitted to determine the common optimum value of E_T and the corresponding values of r_{iq}, $i = 1, \ldots, p$, and r_U for each experiment. The same is done for sample 2. The measured fluorescence intensity ratios for both samples 1 and 2 of p30δ are plotted versus the average superhelix density $\langle\sigma\rangle$ in Fig. 5. The average superhelix density is defined as: $\langle\sigma\rangle = [\langle l - l_0\rangle + N\langle r\rangle(\phi_E/360)]/[l_0 - N\langle r\rangle(\phi_E/360)]$, where the averages are taken over the distribution of topoisomers. The theoretical curves for samples 1 and 2 (which nearly coincide) are also plotted in Fig. 5. Both sets of data lie acceptably close to their theoretical curves, except for the datum at the lowest superhelix density, corresponding to $r = 0.063$, which deviates by about two standard deviations. We may conclude that, up to at least $r = 0.052$ and possibly still higher r values, a single common value of E_T fits each entire data set rather well. Moreover, the optimum E_T values for samples 1 and 2 are 1030 and 1040, respectively, which are practically identical. Clearly, p30δ behaves both reproducibly and in a manner that is consistent with the theory over the experimental range of r values.

It is noteworthy that samples 1 and 2 follow essentially the same curve, despite the fact that their Δl_{app} values differ (i.e. -24 versus -25 turns). At relatively low binding levels, the f_C/f_U ratios are expected to exhibit nearly universal behavior as a function of $\langle\sigma\rangle$ for the following reason. The average superhelix density of each topoisomer is given by

$$\begin{aligned}\sigma(i) &= \{l_i - [l_0 - Nr_i(\phi_E/360)]\}/[l_0 - Nr_i(\phi_E/360)] \\ &= N(\phi_E/360)(r_i - r_i^*)/[l_0 - N(\phi_E/360)r_i],\end{aligned} \tag{20}$$

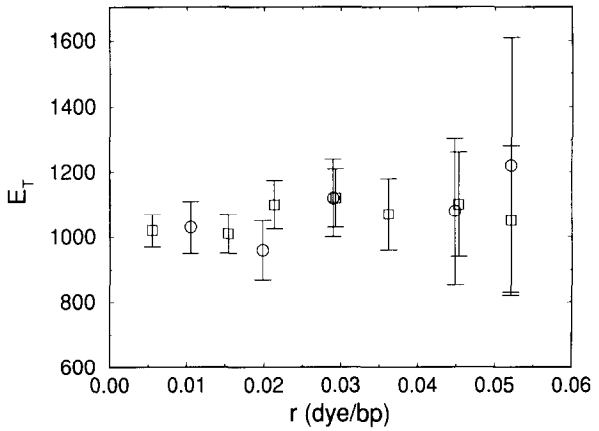


Fig. 6. Optimum E_T versus bound ethidiums per base-pair (r) in competitive dialysis experiments. Circles denote preparation 1 with $\Delta l = -24$ turns; squares denote preparation 2 with $\Delta l = -25$ turns. E_T and r values are shown for each experimental condition. Optimum E_T values for preparations 1 and 2 were determined using the algorithm described in Appendix C along with the values $\phi_E = 26°$, $K_E = 3.5 \times 10^5$ M^{-1}, $N = 4752$ bp and $P = 10.45$ bp/turn to minimize SD in Eq. (19a).

which is essentially proportional to $r_i - r_i^*$ at low binding levels. Under conditions when the ethidium is practically all bound and the binding ratios are small, r_i is a universal function of both $r_i - r_i^*$, or $\sigma(i)$, and $c_E K_E$ (c.f. Eq. (17)). r_U is the same universal function of $c_E K_E$ (c.f. Eq. (7)). According to Eq. (16), the universal dependence on $c_E K_E$ cancels out of the fluorescence intensity ratio, which then depends only on the $\sigma(i)$ of the various topoisomers. Since we find that the binding behavior is essentially independent of the width of the topoisomer distribution, it is expected that the fluorescence intensity ratio should depend essentially only on the average (or median) superhelix density, $\langle\sigma\rangle$, as observed.

5.4. Combined results of topoisomer distribution and ethidium binding experiments

The rather good agreement of the observed topoisomer distributions (Fig. 3) and fluorescence intensity ratios (Fig. 5) with their respective best-fit theoretical predictions argues strongly that the supercoiling free energy of p30δ varies quadratically with the linking difference over the range of ethidium binding ratios from $r = 0$ to 0.083 and over the range of superhelix densities from $\sigma = 0$ to -0.053. Affirmation of this conclusion requires that the same optimum value of E_T be obtained under every set of conditions examined.

Each fluorescence intensity ratio is analyzed individually to determine both the value of E_T and the r value of the supercoiled DNA that apply for a particular added EB/BP ratio. The resulting E_T values are plotted versus r in Fig. 6 and versus $\langle\sigma\rangle$ in Fig. 7. The latter also contains the E_T values derived from topoisomer distribution experiments. The standard deviations of E_T values obtained from dye-binding experiments diverge as r approaches r^*, so values obtained for $r \geqslant 0.063$ (or $-\langle\sigma\rangle \leqslant 0.006$) are unreliable. Nevertheless, the agreement in Fig. 7 between E_T values obtained via topoisomer distribution experiments on one hand and dye-binding experiments on the other is remarkably good. For p30δ DNA, E_T is evidently unaffected by ethidium binding *at any superhelix* density examined. Fig. 7 shows that E_T for p30δ DNA is independent of the average superhelix density from about $\langle\sigma\rangle = 0$ to -0.053, which is essentially the entire range from relaxed to native values. This remarkable invariance of E_T to changes in ethidium binding ratio and superhelix density strongly affirms the quadratic dependence of supercoiling free energy on linking difference that was assumed in the theory. For p30δ, then, we find no evidence of any decrease in E_T with increasing negative superhelix density. This result, which may be

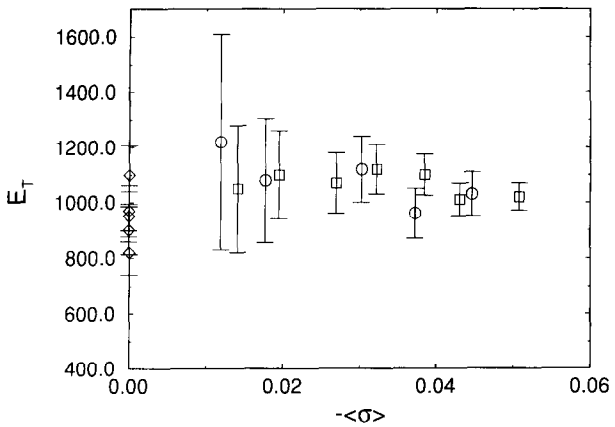


Fig. 7. Optimum E_T versus negative average superhelix density ($-\langle\sigma\rangle$). Circles denote results of competitive dialysis experiments on preparation 1 with $\Delta l = -24$ turns and squares denote results for preparation 2 with $\Delta l = -25$ turns. Diamonds denote topoisomer distribution results. E_T and $-\langle\sigma\rangle$ values are shown for each experimental condition. Optimum E_T values for preparations 1 and 2 were determined using the algorithm described in Appendix C along with the values $\phi_E = 26°$, $K_E = 3.5 \times 10^5$ M^{-1}, $N = 4752$ bp and $P = 10.45$ bp/turn to minimize SD in Eq. (19a). Optimum E_T values for the topoisomerization reaction products were determined using the algorithm described in Appendix B along with the values $\phi_E = 26°$ and $N = 4752$ bp to minimize SSD in Eq. (13).

peculiar to p30δ DNA, contradicts both theoretical predictions [12–15] and some experimental evidence [16,17] to the contrary. Further discussion of this point is deferred to the following companion paper, after the rather different results obtained for pBR322 are presented. The present results for p30δ are summarized schematically in Fig. 8. The essential point is that E_T for p30δ DNA remains constant at about 1030 ± 90 independent of ethidium binding ratio (r) and superhelix density (σ). There is no significant discrepancy in this case between the results obtained from topoisomer distribution and dye-binding experiments.

In view of Eq. (1), the invariance of E_T implies that the torsion constant α and bending constant κ_β are either similarly invariant, or change in mutually compensating ways so as to hold E_T constant. The measured torsion constants α are plotted versus $\langle\sigma\rangle$ in Fig. 9. For the two measurements in the presence of ethidium, the α values in Fig. 9 are corrected for excitation transfer. These results indicate that α, too, is practically independent of ethidium binding ratio from $r = 0.020$ to 0.052 and superhelix density from $\sigma = 0$ to -0.05. In this regard native supercoiled p30δ differs considerably from native *supercoiled* pBR322 and pJMSII DNAs, whose torsion constants decline appreciably with increasing r. Other significant differences between pBR322 and p30δ DNAs are described in the following companion paper. According to Eq. (1), the observed invariance of both E_T and α to r and σ implies that the writhing constant k_w and hence also the bending constant κ_β are likewise invariant to superhelix density over the range from $\sigma = 0$ to -0.05 and to ethidium binding ratio from $r = 0.020$ to 0.052. The implied invariance of κ_β to ethidium binding confirms the conclusion of Smith et al. [23]. The implied invariance to superhelix density is apparently without precedent.

6. Conclusion

The present results indicate that p30δ DNA exhibits practically ideal behavior in the sense that its supercoiling free energy varies quadratically with the effective linking difference over a wide range of

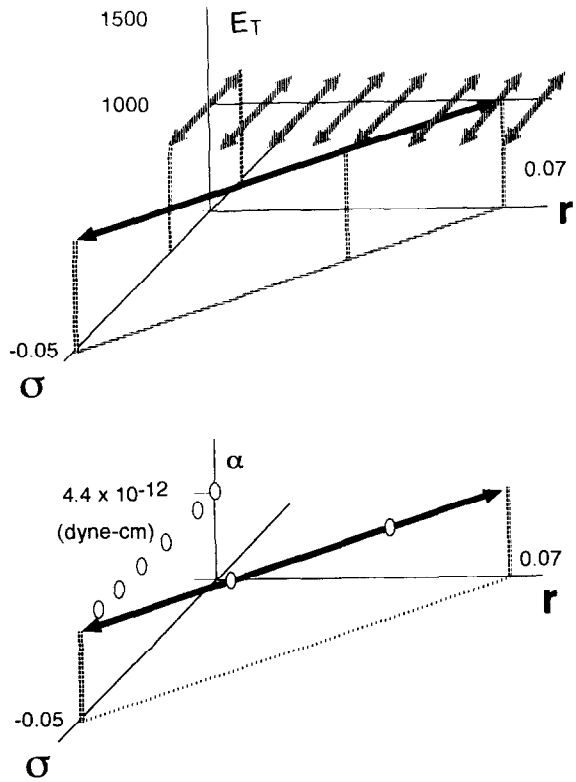


Fig. 8. Schematic illustration of results for p30δ. The upper figure displays E_T values versus superhelix density (σ) and bound ethidium per base pair (r). The topoisomer distribution method yields $E_T \approx 1000$ in the region near $\sigma = 0$ in the $E_T - \sigma$ plane as well as slightly in, in front of and behind the $E_T - r$ plane from $r = 0$ up to $r = 0.082$ (lined two headed arrows). The ethidium-binding method yields $E_T \approx 1000$ along the diagonal (solid two headed arrow). E_T values determined by the ethidium-binding method are weighted toward the higher $-\sigma$ values along that diagonal. The lower figure displays the torsion constant (α) versus superhelix density (σ) and bound ethidium per base pair (r).

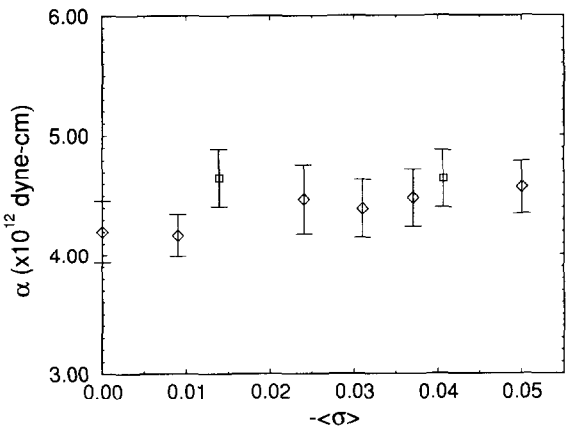


Fig. 9. Torsion constant (α) versus negative average superhelix density ($-\langle\sigma\rangle$). Diamonds denote torsion constants of p30δ samples with different $-\langle\sigma\rangle$ prepared by relaxation in the presence of different amounts of ethidium. (see also Tables 1 and 4). Squares denote torsion constants of native p30δ ($\Delta l = -24$ turns) with $r = 0.0198$ and $r = 0.0523$ bound ethidiums per base-pair (see also Table 5).

ethidium binding ratios and superhelix densities, and the values of its twist energy parameter E_T, torsion constant α, and bending constant κ_β are invariant to superhelix density from $\sigma = 0$ to -0.053 and to ethidium binding ratios from $r = 0$ up to 0.05 or higher. Neither ethidium binding nor supercoiling affect the secondary structure in such a way as to significantly alter the fundamental elastic moduli of the DNA. This conclusion does *not* extend to pBR322 DNA, which is discussed in the following companion paper.

Appendix A. Theory

Reactions and free energies

A rigorous development of the theory is presented elsewhere [6], and only a brief summary is presented here. The system is assumed to be sufficiently dilute that all solute species obey Henry's Law. The free energies of all DNAs are calculated relative to a relaxed circular species (R) containing a nick at a particular location. The closed circular topoisomer with minimum strain free energy has linking number m, and is denoted by C(m). The reaction

$$R \rightleftharpoons C(m) \tag{A.1}$$

has standard free energy change ΔF°_{om}, which represents the free energy required to align and seal the nick. The reaction to form any other topoisomer,

$$R \rightleftharpoons C(i) \tag{A.2}$$

has standard free energy change

$$\Delta F^\circ_{0i} = \Delta F^\circ_{0m} + \Delta F^\circ_{im}, \tag{A.3}$$

where ΔF°_{im} is the difference in free energy between topoisomers C(i) and C(m), and is given by

$$\Delta F^\circ_{im} = \mu^\circ_{C(i)} - \mu^\circ_{C(m)} = -k_B T \ln\left[\frac{Q(1_{C(i)}, N_S, V, T)}{Q(1_{C(m)}, N_S, V, T)}\right] \tag{A.4}$$

wherein the $\mu_{C(j)}{}^\circ$ are chemical potentials in the 1 molecule/cm^3 standard state, and $Q(1_{C(i)}, N_s, V,T)$ is the canonical partition function for one molecule of C(i) and $N_s = V/v_s$ solvent molecules in a volume $V = 1$ cm^3. The partial molecular volume of solvent is $\bar{v}_s$.

If each molecule, whether solute or solvent, consists of rigid-body subunits connected to each other by sufficiently weak restoring forces that their rigid body translations and rotations obey *classical* statistical mechanics, then all kinetic energy factors cancel completely out of the partition function ratio. That is, all factors depending on masses and moments of inertia no longer appear. In that case, one has finally

$$\Delta F^\circ_{im} = -k_B T\left[\frac{Z(1_{C(i)}, N_S, V, T)}{Z(1_{C(m)}, N_S, V, T)}\right] \tag{A.5a}$$

where the configuration integral,

$$Z(1_j, N_S, V, T) = \int \cdots \int \prod_{l=1}^{v_j} (d\boldsymbol{r}_l d\Omega_l) \prod_{p=1}^{N_S} \prod_{n=1}^{v_S} (d\boldsymbol{r}_n^p \, d\Omega_n^p) \, e^{-\mathscr{U}/k_B T} \tag{A.5b}$$

is taken over all positions ($\boldsymbol{r}$) and orientations (Ω) of the v_j rigid subunits of the molecule j and the v_s rigid subunits of each of the N_s solvent molecules, and $\mathscr{U} = \mathscr{U}(\{1_j, N_s\})$ is the potential energy function of the positions and orientations of those same subunits.

We assume that: (i) the twisting and bending coordinates separate completely from the remaining coordinates, and (ii) fluctuations in *net* twist are coupled by the topological constraint to fluctuations in *net* writhe, which are assumed to exhibit a Gaussian distribution in an equilibrium population of nicked circles, as suggested by all Monte Carlo simulations at low superhelix density and some (but not all) Monte Carlo simulations at higher superhelix density [12,10–20]. With these assumptions, it is found that

$$\Delta F^{\circ}_{im} = k_B T(E_T/N)\left[(i-l_0)^2-(m-l_0)^2\right] - k_B T\, ln\left[\frac{Z_R(1_{C(i)}, N_S, V, T)}{Z_R(1_{C(m)}, N_S, V, T)}\right], \tag{A.6}$$

where for long DNAs, $E_T = [(2\pi)^2/2k_BT][\alpha\kappa_w/(\alpha+\kappa_w)]$ [6,10], and Z_R $(1_{C(i)}, N_s, V, T)$ denotes the configuration integral over the remaining coordinates. We assume that the latter is independent of linking difference, so $Z_R(1_{C(i)}, N_s, V, T) = Z_R(1_{C(m)}, N_s, V, T)$, and finally

$$\Delta F^{\circ}_{im} = k_B T(E_T/N)\left[(i-l_0)^2-(m-l_0)^2\right] \tag{A.7}$$

Thus, ΔF_{im} is just the difference in superhelical strain free energy between C(i) and C(m).

Each topoisomer may bind n_E ethidiums (E) to form a specific complex.

$$C(m) + n_E E \rightleftharpoons C_J(m, n_E), \tag{A.8}$$

$$C(i) + n_E E \rightleftharpoons C_J(i, n_E), \tag{A.9}$$

where J denotes the specific arrangement of bound ethidiums among its possible intercalation sites. It is assumed that the binding forces are sufficiently weak that the relative coordinates between the DNA and the intercalated ethidium obey classical statistical mechanics. In this case, no new rigid subunits are formed, and the total number of rigid subunits in the complex is the same as in the free dye plus DNA. Each intercalated ethidium is assumed to add one more net torsion and bending spring, to increase the length by h_E, to decrease the intrinsic twist by ϕ_E degrees, and to exclude binding of another dye to its nearest-neighbor intercalation sites. The additional torsion and bending spring is assumed to be identical to those between base-pairs. Evidence for the validity of this assumption comes from the observed invariance of the torsion constant and persistence length to intercalated ethidium [22,23], as discussed in Section 1 and the success of this assumption in fitting curves of average linking number (relative to the nearest integer to the intrinsic twist) versus free ethidium concentration in the ligation medium in the case of small ($N \lesssim 250$ bp) circular DNAs [5,6]. With these assumptions, the standard free energy change for reaction (A.9) is given by [6]

$$\begin{aligned}\Delta F^{\circ}_{C(i,n_E)} = {} & n_E \Delta F^{\circ}_E + k_B T\left(\frac{E_T}{N}\right)\left(\frac{\phi_E}{360}\right)^2\left(\frac{n_E^2 - 2n_E\left(n^*_{Ei} + n^{*2}_{Ei}/N\right)}{(1+n_E/N)}\right) \\ & + \kappa_\beta(2\pi)^2 n_E/[2N(N+n_E)] - k_B T\,\ln\left\{\left[(N+n_E)/N\right]^{-3/2}\right\} \\ & - k_B T\,\ln\left(\frac{\left\{\left[\alpha+\kappa_w(\kappa_\beta, N+n_E)\right]^{-1/2} f(\kappa_\beta, N+n_E, T)\right\}}{\left\{\left[\alpha+\kappa_w(\kappa_\beta, N)\right]^{-1/2} f(\kappa_\beta, N, T)\right\}}\right),\end{aligned} \tag{A.10}$$

wherein $n^*_{Ei} \equiv -(i-l_0)\,(360/\phi_E)$. When positive, n^*_{Ei} is the number of bound ethidiums required to completely relax the effective linking difference to zero, at which point the superhelical strain vanishes. ΔF_E° is the intrinsic free-energy change for binding one ethidium to a single isolated site in a relaxed circular DNA. The second term in Eq. (A.10) is the change in the minimum strain free energy of supercoiling due to: (i) unwinding the DNA by n_E ϕ_E, and (ii) the increase in number of twisting and bending springs due to the bound ethidium. It is mentionable that the ratio of the total number of mean persistence lengths after dye-binding to that before dye binding is $(N+n_E)/N$, provided that the new springs have identical bending torque constants, despite the fact that the rise per ethidium is not

identical to the rise per bp (i.e. $h_E = 2.7$ Å $\neq h = 3.4$ Å [40]). The third term in Eq. (A.10) with κ_β represents the decrease in minimum bending energy of circularization due to the addition of springs as ethidium binds. The fourth term arises from closure constraints on the twisting and bending coordinates in the configuration integral for the complex. For long DNAs, the correct form of this term is not completely known, due to the intractable nature of the configuration integral over the bending coordinates. Here we assume that its form is identical to that for small circles [6]. The fifth (last) term arises in part from the normalization factor of the assumed Gaussian distribution function for writhe and in part from the factor $f(\kappa_\beta, N, T)$ by which the contribution of bending fluctuations of a long circular DNA differs from that of small circular DNA evaluated at the same N. The factor $f(\kappa_\beta, N, T)$ is a weak function of N in the sense that N does not occur in the exponent. The contributions of the third and fourth terms were numerically found to be negligibly small for a 247 bp DNA, and should make an even smaller relative contribution for longer DNAs. The fifth term is similarly expected to be negligibly small. Hence, the last three terms in Eq. (A.10) are neglected in this work.

Topoisomer distribution

We seek the probability of topoisomer C(i) relative to that of topoisomer C(m) in an equilibrium reaction mixture, namely $c^{tot}_{C(i)}/c^{tot}_{C(m)}$, where $c^{tot}_{C(i)}$ is the total concentration of topoisomer C(i). The equilibrium constant expression for reaction (A.2) is

$$c_{C(i)}/c_R = \exp[-(\Delta F^\circ_{0m} + \Delta F^\circ_{im})/k_B T] \tag{A.11}$$

while that for reaction (A.9) is

$$c_{C_J(i,n_E)}/c_{C(i)} = c_E^{n_E} \exp(-\Delta F^\circ_{C(i,n_E)}/k_B T), \tag{A.12}$$

where c_l denotes the concentration of species l. Summing Eq. (A.12) over all configurations J and all numbers of bound ethidiums from 1 to n_{max} yields the *grand partition function*, GPF(i), for the binding of ethidium to a covalently closed circular DNA with linking number i,

$$\frac{c_{C(i,n_E \geq 1)}}{c_{C(i)}} = \sum_{n_E=1}^{n_{max}} c_E^{n_E} \exp(-\Delta F^\circ_{C(i,n_E)}/k_B T) W(n_E) = \mathrm{GPF}(i), \tag{A.13}$$

wherein $c_{C(i,\ n_E \geq 1)}$ is the concentration of all complexes of topoisomers C(i) with $n_E \geq 1$ bound ethidiums and $W(n_E)$ is the number of distinct configurations that satisfy nearest-neighbor exclusion when n_E ethidiums are bound. The total concentration of topoisomer C(i) is the concentration of all complexes of C(i) plus the concentration of all uncomplexed C(i),

$$\begin{aligned} c^{tot}_{C(i)} &= c_{C(i)} + c_{C(i,n_E \geq 1)} = c_{C(i)}[1 + \mathrm{GPF}(i)] \\ &= c_R \exp[-\Delta F^\circ_{0m}/k_B T] \exp[-\Delta F^\circ_{im}/k_B T][1 + \mathrm{GPF}(i)], \end{aligned} \tag{A.14}$$

wherein the last line was obtained by using Eq. (A.11). Thus, the ratio of total concentrations of topoisomers C(i) and C(m) is finally given by

$$\frac{c^{tot}_{C(i)}}{c^{tot}_{C(m)}} = \exp\left[-\left(\frac{E_T}{N}\right)\left[(i-i_0)^2 - (m-l_0)^2\right]\right]\left(\frac{1 + \mathrm{GPF}(i)}{1 + \mathrm{GPF}(m)}\right). \tag{A.15}$$

where

$$\mathrm{GPF}(i) = \sum_{n_E=1}^{n_{max}} (c_E K_E)^{n_E} \exp\left[-\left(\frac{E_T}{N}\right)\left(\frac{\phi_E}{360}\right)^2 \left(\frac{n_E^2 - 2n_E(n^*_{Ei} + n^{*2}_{Ei}/N)}{1 + n_E/N}\right)\right] W(n_E), \tag{A.16}$$

and $K_E = \exp[-\Delta F_E^\circ/k_B T]$ is the intrinsic binding constant for ethidium.

The exact combinatorial expression for $W(n_E)$ is given by Reiter and Epstein [41],

$$W(n_E) = \sum_{c=c_{min}}^{c=c_{max}} \frac{f!}{c!(f-c)!}\mu\frac{(n_E-1)!}{(n_E-c)!(c-1)!}, \tag{A.17}$$

where c is the island cluster index. An island cluster is a group of contiguously bound ligands with at least one free site on its right. The sum runs from the minimum to the maximum number of clusters. $f = N - kn_E$ is the number of free sites remaining after the binding of n_E ligands. k is the number of (contiguous) sites excluded by one bound ethidium. Here $k = 2$, the sites being the intercalation site and its nearest-neighbor (on the right) binding site. The maximum (c_{max}) and minimum (c_{min}) numbers of clusters are,

$$c_{max} = \begin{cases} n_E, & \text{for } f \geqslant n_E \\ f, & \text{for } f < n_E \end{cases}, \tag{A.18}$$

$$c_{min} = \begin{cases} 0, & \text{for } n_E = 0, \\ 1, & \text{for } n_E > 0 \text{ and } f \neq 0, \end{cases} \tag{A.19}$$

and μ is given by

$$\mu = \frac{N}{f}, \quad \text{for } f > 0. \tag{A.20}$$

We note that Eq. (A.17) is undefined when $f = 0$, that is when the DNA becomes saturated with ligand. This is because 'island clusters' are no longer defined. Reiter and Epstein attempt to correct this by setting $\mu = N/N_E$ and $c_{min} = c_{max} = 0$ for $f = 0$. However, this causes the second combinatorial factor to become undefined, so the situation is not improved by those additional definitions, hence we do not include them above. As a consequence of this we must sometimes choose n_{max} in Eq. (A.16) to be slightly different than the actual value. In the case of p30δ DNA, $N = 4752$ and $k = 2$, hence Eq. (A.17) fails when $n_E = 2376$. This is the last of the approximately 1.88×10^6 terms in the partition function. If we choose $n_{max} = 2375$, then Eq. (A.17) holds, but we lose the last term in the partition function. This choice has no significant effect at the binding levels considered here, because the neglected state has effectively zero probability. In fact, at the highest binding levels treated in this work, deletion of the last 1.4×10^5 states (i.e. set $n_{max} = 2000$) has absolutely no effect on any properties calculated using Eqs. (A.15) and (A.16). The missing state can also be included, if necessary. For $f = 0$, one finds by inspection that $W(n_E) = 1$. Eq. (A.16) gives exactly the same numerical values of GPF(i) as do matrix methods for the case when $\Delta F_{im}^\circ = 0$, or when ΔF_{im}° varies linearly with i and m.

The average number of bound ethidiums in *complexes* of each topoisomer are calculated along with GPF(i) according to

$$\langle n_E(C(i, n_E \geqslant 1))\rangle = \frac{1}{GPF(i)} \sum_{n_E=1}^{n_{max}} n_E(c_E K_E)^{n_E} \otimes \exp\left(-\frac{E_T}{N}\left(\frac{\phi_E}{360}\right)^2 \frac{\left[n_E^2 - 2n_E(n_{Ei}^* + n_{Ei}^{*2}/N)\right]}{1 + n_E/N} \right) W(n_E) \tag{A.21}$$

and finally the average number of bound ethidiums, whether or not in complexes, is calculated according to

$$\langle n_E(C(i))\rangle = \langle n_E(C(i, n_E \geqslant 1))\rangle GPF(i)/[1 + GPF(i)]. \tag{A.22}$$

For the dye concentrations considered here, $GPF(i) \ggg 1.0$, so $\langle n_E(C(i))\rangle = \langle n_E(C(i, n_E \geqslant 1))\rangle$.

Appendix B. Fitting algorithm for topoisomer distributions

The fitting algorithm outlined below begins with the selection of input trial values of the unknowns $c_E K_E$, l_0, and E_T. For convenience, we fix $K_E = 3.5 \times 10^5$ M (which could well be incorrect for our DNA) and adjust c_E. This is permissible since c_E and K_E occur effectively only as the product $c_E K_E$ under conditions when the ethidium is practically all bound, as is the case here. Reasonable initial choices are $c_E = 0.01 c_E^{tot}$, $l_0 = N/10.45$, and $E_T = 3000$. The steps in the algorithm are as follows.

(1) Select trial values of c_E, l_0, and E_T.
   (a) Calculate $\langle n_E \rangle_U$ by Eq. (7).
   (b) Construct a set of trial linking differences $\{\Delta l_i\}$ by setting $\Delta l_m = \text{NINT}\,(l_0 - \langle n_E \rangle_U\,(\phi_E/360)) - l_0$, and taking the remaining Δl_i as $\Delta l_m \pm 1, 2, \ldots, 10$.
   (c) Calculate n^*_{Ei} by Eq. (6) for all topoisomers.
   (d) Calculate GPF(i) by Eq. (5), $\langle n_E(C(i), n_E) \geqslant 1 \rangle$ by Eq. (A.21), and $\langle n_E(C(i)) \rangle$ by Eq. (A.22) for all topoisomers.
   (e) Calculate $c^{tot}_{C(m)}$ by Eq. (9).
   (f) Calculate $c^{tot}_{C(i)}$ by Eq. (4) for each topoisomer $i \neq m$.
   (g) Calculate the theoretical bound ethidium concentration on the r.h.s. of Eq. (10) and compare with the added c_E^{tot} on the l.h.s.
   (h) If the difference between the added and theoretical ethidium concentrations exceeds 1% of the added, then the current trial c_E value is adopted as a lower bound. If this difference is less than -1% of the added ethidium concentration, then the current trial value of c_E is adopted as an upper bound. The next trial value of c_E is taken as one half the sum of the upper and lower bounds.
   (i) Proceed to (1a) above with the new trial c_E and repeat steps (1a) to (1i) until the absolute value of the difference between the added and theoretical ethidium concentrations is less than or equal to 1% of the added. The free ethidium concentration c_E for which this occurs is saved.

(2) The set of relative concentrations $\{c^{tot}_{C(i)}/c^{tot}_{C(m)}\}$ is checked to verify that the largest value is 1.0. If it exceeds 1.0, each member of the set $\{c^{tot}_{C(i)}\}$ is divided by the largest concentration and the position of m is updated. This did not happen during the calculations reported here.

(3) The theoretical area ratios are computed from the theoretical topoisomer ratios by Eq. (11). This correction has a negligible effect on the final results. In this part of the calculation it is assumed that $K_E = 3.5 \times 10^5$, as reported for linear DNAs under similar conditions by MacGregor and Clegg [29].

(4) The theoretical distribution of area ratios is checked to determine whether any ratio is greater than 1.0. If so, then all members of the set $\{A_i/A_m\}$ are divided by the largest value and the position of m is updated. This did not happen during the calculations reported here.

(5) SSD is calculated according to Eq. (13).

(6) A new value of E_T is chosen and the entire procedure is repeated beginning with step (1a). The last value of c_E, which yielded convergence for the previous value of E_T, is adopted as the first trial value. Convergence at the new and all subsequent values of E_T occurs in a single pass. The physical reason for this fortunate circumstance is that in the relaxed limit a population of long topoisomers binds very nearly the same average number of ethidiums as the corresponding linear or nicked circular DNA, hence the average number of bound ethidiums is practically independent of E_T. This circumstance does *not* hold for short circular DNAs.

(7) Step (6) is repeated for a range of E_T values until the value that minimizes SSD is found.

(8) The value of E_T that minimizes SSD is used along with its associated value of c_E and a new trial value of l_0 beginning again with step (1a). Again convergence is achieved in one pass for each value of l_0. The SSD versus l_0 surface is degenerate with successive minima separated by one turn. The accepted value of the pitch is 10.45 ± 0.02 bp/turn. Three adjacent minima are found that bracket this value and the closest minimum is taken as the 'correct' one. This has no effect on the optimum values of E_T and $c_E K_E$.

(9) A final round of iterations is performed to adjust E_T to minimize SSD for the optimal value of l_0. This results in no more than a 5% change in E_T.

Appendix C. Solution / fitting algorithm for ethidium binding experiments

One begins with values of the added EB/BP and $(f_C/f_U)^{ex}$ for the experiment under consideration. In addition, one must know the topoisomer fractions f_i, which can be calculated from the independently measured $\Delta l(m) = \Delta l(7)$ by Eq. (18). One could also represent the distribution by a single topoisomer with $\Delta l(m)$, which has a negligibly small effect on the final results, as noted in the text. We take $K_E = 3.5 \times 10^5\ M^{-1}$, which has almost no bearing on the final results, since c_E and K_E occur effectively only as the product $c_E K_E$, when the ethidium is nearly all bound, as it is here. One requires initial trial values of E_T and c_E. A reasonable initial choice is $c_E = 0.01\ c_E^{tot}$. The steps in the algorithm to determine E_T for any single experiment are as follows.

(1) Select a new trial value of E_T.
  (a) Input the experimental value of EB/BP and $(f_C/f_U)^{ex}$.
  (b) Choose a trial c_E.
  (c) Solve Eq. (7) in the form $r_U = c_E K_E\,[1 - 2r_U^2/(1 - r_U)]$. This is done by adjusting r_U until the absolute difference between the l.h.s. and r.h.s. is less than 10^{-9}.
  (d) Solve Eq. (17) for each r_i by adjusting r_i until the absolute difference between the l.h.s. and r.h.s. is less than 10^{-9}.
  (e) Calculate the theoretical bound ethidium per bp on the r.h.s. of Eq. (15) and compare with the added EB/BP on the l.h.s.
  (f) If the r.h.s. of Eq. (15) is greater than 1.00001 times EB/BP, then the current trial value of c_E is adopted as an upper bound. Conversely, if the r.h.s. is smaller than 0.99999 times EB/BP, then the current trial value of c_E is adopted as a lower bound. In either case, the next trial value of c_E is taken as one half the sum of upper and lower bounds, and we return to (1c).
  (g) If the absolute difference between the r.h.s. of Eq. (15) and EB/BP is less than 0.01 times EB/BP, then we have found c_E, so proceed to step (2).

(2) Calculate $(f_C/f_U)^{th}$ by Eq. (16).

(3) Calculate SD according to Eq. (19a).

(4) Return to step (1) above, and repeat. Continue until the minimum SD is located.

(5) Check that SD is close to zero, which it must be for a solution for an individual data point.

The same algorithm with minor modifications is employed to estimate the optimum E_T that fits the entire data set. The modifications are as follows. The experimental values $(EB/BP)_q$ and $(f_C/f_U)_q^{ex}$ now bear the index $q = 1, \ldots, Q$. The index q is set to 0 in step (1) and is incremented by 1 at the start of step (1a). The theoretical values r_i^q, $i = 1, \ldots, p$, r_U^q, $(f_C/f_U)_q^{th}$ and SD_q also bear an index q. A new step (3a) inserted after step (3) is as follows:

(3a) Add SD_q to the accumulating SSD (cf. Eq. (19b)). If $q < Q$ return to step (1a), but if $q = Q$, pass on to step (4)

In step (4), SD is replaced by SSD. Step (5) is omitted.

Appendix D. Error analysis

Uncertainties in E_T values from topoisomer distributions

Errors in E_T arise primarily from uncertainties in the peak area ratios. Errors in the latter are estimated in the following way. The raw data consist of three densitometer scans down the 'lane' of a

particular distribution. The main errors arise from a systematic variation in DNA concentration across every band, which may be as much as 20%, misalignment of the scan in the sense that it does not sample the same relative position in every band, and random fluctuations of DNA concentration about the average systematic distribution within each band. We consider two different ways in which the data for two topoisomers, i and m, can be combined to yield the peak area ratio. One could take the ratio A_i/A_m of areas in each particular scan, and average those over the three scans to obtain $\langle A_i/A_m \rangle$, which has variance

$$\sigma_{1i}^2 = \tfrac{1}{3} \sum_{s=1}^{3} \left[\left(\frac{A_i}{A_m} \right)_s - \left\langle \frac{A_i}{A_m} \right\rangle \right]^2, \tag{D.1}$$

where the subscript s denotes scans. This variance is unaffected by any systematic variation in DNA concentration across *every* band, provided the scans are exactly aligned, but is sensitive to scan misalignment. One could also average the results of the three scans to obtain the $\langle A_i \rangle = \frac{1}{3}\Sigma_s (A_i)_s$, then take the ratio $\langle A_i \rangle / \langle A_m \rangle$, which has variance

$$\sigma_{2i}^2 = \frac{1}{\langle A_m \rangle^2} \sigma_i^2 + \frac{\langle A_i \rangle^2}{\langle A_m \rangle^4} \sigma_m^2. \tag{D.2}$$

This variance is comparatively insensitive to scan misalignment, but rather sensitive to systematic variation in DNA concentration across every band. In practice, we find that $\langle A_i/A_m \rangle \approx \langle A_i \rangle / \langle A_m \rangle$ in every case, but $\sigma_{1i}^2 < \sigma_{2i}^2$. Thus, for the area ratios we actually take the quantity

$$Q_i = \tfrac{1}{2}(\langle A_i/A_m \rangle + \langle A_i \rangle / \langle A_m \rangle). \tag{D.3}$$

which has the approximate variance

$$\sigma_{Qi}^2 = \tfrac{1}{4}(\sigma_{1i} + \sigma_{2i})^2. \tag{D.4}$$

In (D.4) it is assumed that the covariance between $\langle A_i/A_m \rangle$ and $\langle A_i \rangle / \langle A_m \rangle$ is approximately $\sigma_{1i}\, \sigma_{2i}$.

The errors in E_T arising from those (σ_{Qi}) in peak area ratios are estimated by evaluating E_T $(Q_1, \ldots, Q_{i-1}, Q_i + \sigma_{Qi}, Q_{i+1}, \ldots)$ for the circumstance when Q_i is increased by σ_{Qi}, repeating this exercise for all bands i, and setting

$$\sigma_{E_T}^2 = \sum_i \left[E_T^{\text{opt}} - E_T(Q_1, \ldots, Q_i + \sigma_{Qi}, \ldots) \right]^2 \tag{D.5}$$

where $E_T^{\text{opt}} = E_T\,(Q_1, \ldots, Q_{i-1}, Q_i, Q_{i+1}, \ldots)$. This computation is extremely time-consuming, since it requires minimization of SSD for every $Q_i + \sigma_{Qi}$. Thus, we performed this full error analysis only for the sample relaxed in 0.08 EB/BP. The result, $\sigma_{E_T}/E_T = 9.6\%$ is assumed to apply equally to all samples and is used to generate the error bars in Fig. 4.

Uncertainties in E_T values from individual competitive dialysis experiments

The error in E_T determined from a single competitive dialysis experiment arises primarily from errors in the fluorescence intensity ratio, in the concentrations of DNA and ethidium, and in Δl. Experimental standard deviations were determined to be 2% for f_C/f_U, 1% for the concentrations of linear and supercoiled DNAs, 2% for ethidium concentration (serial dilution), and 1 turn for Δl. Under the assumption that each of these variables is distributed as a Gaussian function with the experimental standard deviation around its measured value for a particular experiment, 1000 replicas of each experiment were simulated by generating random numbers with the appropriate Gaussian distributions.

For each such replica (j) the best-fit value of E_{T_j} was determined. From the 1000 best-fit values of E_{T_j} for a given 'experiment', the mean value

$$\langle E_T \rangle = \sum_{j=1}^{1000} E_{T_j} \tag{D.6}$$

and standard deviation

$$\sigma_{E_T} = \left[\frac{1}{1000} \sum_{j=1}^{1000} \left(E_{T_j} - \langle E_T \rangle \right)^2 \right]^{1/2} \tag{D.7}$$

were calculated. The resulting σ_{E_T} values range from 50 for EB/BP = 0.0042 up to 230 for EB/BP = 0.0500, and diverge, as EB/BP (actually r) approaches r^*. The σ_{E_T} values appear as the error bars in Figs. 6 and 7.

Acknowledgements

This work was supported in part by a Grant (P01-GM32681) from the National Institutes of Health.

References

[1] W. Keller, Proc. Natl. Acad. Sci. USA 72 (1975) 4876–1880.
[2] D.E. Pulleyblank, M. Shure, D. Tang, J. Vinograd and H.P. Vosberg, Proc. Natl. Acad. Sci. USA 72 (1975) 4280–4284.
[3] R.E. Depew and J.C. Wang, Proc. Natl. Acad. Sci. USA 72 (1975) 4275–4279.
[4] D. Horowitz and J.C. Wang, J. Mol. Biol. 173 (1984) 75–91.
[5] D. Shore and R.L. Baldwin, J. Mol. Biol. 170 (1983) 957–1007.
[6] J.B. Clendenning and J.M. Schurr, Biopolymers 34 (1994) 849–868.
[7] D. Shore and R.L. Baldwin, J. Mol. Biol. 170 (1983) 957–981.
[8] W. Bauer and J. Vinograd, J. Mol. Biol. 47 (1970) 419–435.
[9] W. Bauer, Ann. Rev. Biophys. Bioeng. 1 (1978) 287–313.
[10] T.S. Hsieh and J.C. Wang, Biochemistry 14 (1975) 527–535.
[11] P.-G. Wu, L. Song, J.B. Clendenning, B.S. Fujimoto, A.S. Benight and J.M. Schurr, Biochemistry 27 (1988) 8128–8144.
[12] K.V. Klenin, A.V. Vologodskii, V.V. Anshelevich, V. Yu. Klishko, A.M. Dykhne and M.D. Frank-Kamenetskii, J. Mol. Biol. 217 (1991) 413–419.
[13] C.J. Benham, J. Chem. Phys. 92 (1990) 6294–6305.
[14] C.J. Benham, J. Mol. Biol. 225 (1992) 835–847.
[15] C.J. Benham, Cell Biophys. 10 (1987) 193–204.
[16] S.A. Kozyavkin, A.I. Slesarev, S.R. Malkhosyan and I.G. Panyutin, Eur. J. Biochem. 191 (1990) 105–113.
[17] D. Kowalski, D.A. Natale and M.J. Eddy, Proc. Natl. Acad. Sci. USA 85 (1988) 9464–9468.
[18] A.V. Vologodskii, V.V. Anshelevich, A.V. Lukashin and M.D. Frank-Kamenetskii, Nature 280 (1979) 294–298.
[19] M. Le Bret, Biopolymers 19 (1980) 619–637.
[20] Y. Chen, J. Chem. Phys. 75 (1981) 2447–2453.
[21] G.S. Manning, Biophys. Chem. 7 (1977) 95–102.
[22] P.-G. Wu, B.S. Fujimoto, L. Song and J.M. Schurr, Biophys. Chem. 41 (1991) 217–236.
[23] S.B. Smith, L. Finzi, and C. Bustamante, Science 258 (1992) 1122–1126.
[24] F.B. Fuller, Proc. Natl. Acad. Sci. USA 68 (1971) 815–819.
[25] J.H. White, Am. J. Math. 91 (1969) 693–728.
[26] W.H. Taylor and P.J. Hagerman, J. Mol. Biol. 212 (1990) 363–376.
[27] J.C. Wang, J. Mol. Biol. 89 (1974) 783–801.
[28] D.E. Pulleyblank and A.R. Morgan, J. Mol. Biol. 91 (1975) 1–13.
[29] R.B. MacGregor and R.M. Clegg, Biopolymers 26 (1987) 2103–2106.

[30] J.B. Clendenning, Ph.D. Thesis, University of Washington (1993).
[31] J.R. Cameron, E.Y. Loh and R.W. Davis, Cell 16 (1979) 739–751.
[32] E. di Mauro, M. Caserta, R. Negri and F. Carnevalli, J. Biol. Chem. 260 (1985) 152–159.
[33] R. Negri, F. Della Seta, E. di Mauro and G. Camilloni, Topological Evidence for Allosteric Transitions in DNA Secondary Structure, preprint.
[34] J.H. Miller, in: Guide to molecular cloning techniques, eds. S.L. Berger and A.R. Kimmel (Academic Press, New York, 1987) pp. 145–170.
[35] W.C. Hymer and E.L. Kuff, J. Histochem. Cytochem. 12 (1964) 359–363.
[36] M. Bina-Stein, T. Vogel, S.D. Singer, and M.F. Singer, J. Biol. Chem. 251 (1976) 7363–7366.
[37] J.C. Wang, J. Mol. Biol. 43 (1969) 263–272.
[38] J.M. Schurr, B.S. Fujimoto, P. Wu and L. Song, in: Topics in fluorescence spectroscopy, Vol. 3. Biochemistry Applications, ed. J.R. Lakowicz (Plenum Press, New York, 1992) pp. 137–229.
[39] B.S. Fujimoto and J.M. Schurr, Nature 344 (1990) 175–178.
[40] M. Hogan, N. Dattagupta and D. Crothers, Biochemistry 18 (1979) 280–288.
[41] J. Reiter and I.B. Epstein, J. Phys. Chem. 91 (1989) 4813–4820.