A Structural Transition in Duplex DNA Induced by Ethylene Glycol

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · OCTOBER 2008

Impact Factor: 3.3 · DOI: 10.1021/jp802139a · Source: PubMed

3 AUTHORS, INCLUDING:

J. Michael Schurr
University of Washington Seattle
178 PUBLICATIONS **4,100** CITATIONS

SEE PROFILE

Published in final edited form as:
J Phys Chem B. 2008 October 23; 112(42): 13367–13380. doi:10.1021/jp802139a.

A Structural Transition in Duplex DNA Induced by Ethylene Glycol

Greg P. Brewood, **Theresa Aliwarga**, and **J. Michael Schurr**
Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98195-1700

Abstract

The twist energy parameter (E_T) that governs the supercoiling free energy, and the linking difference (Δl) are measured for p30 δ DNA in solutions containing 0 to 40 w/v% ethylene glycol (EG). A plot of E_T vs. $-\ln a_w$, where a_w is the water activity, displays the full (reverse) sigmoidal profile of a discrete structural transition. A general theory for the effect of added osmolyte on a cooperative structural transition between two duplex states, 1□ 2, is formulated in terms of parameters applicable to individual base-pairs subunits. The resulting fraction of base-pairs in the 2-state (f_2^0), is incorporated into expressions for the effective torsion and bending elastic constants, the effective twist energy parameter (E_T^{eff}), and the change in intrinsic twist (δl_0). Fitting the expression for E_T^{eff} to the measured E_T -values yields reasonably unambiguous estimates of $E_{T_1}$and E_{T_2}, the midpoint value $(\ln a_w)_{1/2}$, and midpoint slope $(\partial E_T/\partial \ln a_w)_{1/2}$, but does *not* yield unambiguous estimates of the equilibrium constant (K_0), the difference in DNA-water preferential interaction coefficient ($\Delta\Gamma$), or the inverse cooperativity parameter, J. Fitting a non-cooperative model (assumed J=1.0) to the data yields, $K_0 = 0.067$, and $\Delta\Gamma = -30.0$ per base-pair (bp). Essentially equivalent fits are provided by models with a wide range of correlated J, $\Delta\Gamma$, and K_0 values. Other results favor $\Delta\Gamma$ in the range -1.0 to 0, which then requires $K_0 \geq 0.914$, and a cooperativity parameter, $1/J \geq 30.0$ bp. The measured δl_0 and circular dichroism (CD) at 272 nm are found to be compatible with curves predicted using the same f_2^0-values that best-fit the E_T -data. At least 7 to 10 % of the base-pairs are inferred to exist in the 2-state in 0.1 M NaCl in the complete absence of added osmolyte. Compared with the 1-state, the 2-state has a ~2.0- to 2.1-fold greater torsion elastic constant, a ~0.70-fold smaller bending elastic constant, a ~0.91-fold smaller E_T -value, a ~0.2 % lower intrinsic twist, a somewhat lower CD near both 272 and 245 nm, and less water and/or more EG in its neighborhood. However, the *relative* change in preferential interaction coefficient associated with the transition is likely rather slight.

Introduction

DNA secondary structures are typically classified into three main families, specifically the right-handed A and B families and the left-handed Z-family. Numerous conformational substates within the right-handed families have also been detected by various methods, although specific structural details are lacking in most cases.[1] Average properties that reflect changes in secondary structure respond to practically every conceivable perturbation, including the presence of small neutral osmolytes.[1–25] Such changes appear to arise from shifts of population (base-pairs) among different conformational substates. Conformation-sensitive properties include the elastic constants, α and κ_β, for torsion and bending, respectively, the twist energy parameter, E_T, that governs the supercoiling free energy, the intrinsic twist, l_0 (turns), the intrinsic binding constant for ethidium, and the circular dichroism spectrum.[1] All of these properties can be measured for samples as small as 20 μg of DNA.

schurr@chem.washington.edu, Telephone: 206 543 6681, FAX: 206 585 8665.

At 37 °C, increasing concentrations of the neutral osmolytes, ethylene glycol (EG) and acetamide, caused initially sigmoidal decreases in the respective curves of E_T vs. −ln a_w, where a_w is the water activity.[17] These data, which previously extended only to ~20 w/v% osmolyte, [17] could be satisfactorily modeled by a two-state transition of independent subunits (or domains) of unknown size. However, the terminal lower plateau of the transition could *not* be located from those data, and neither the value of $(-\ln a_w)_{1/2}$ nor the slope, $(\partial E_T/\partial (-\ln a_w))_{1/2}$, at the midpoint could be reliably assessed. Moreover, quantitative information about properties of the terminal state, which prevails in ~40 w/v% EG, could not be obtained.

A principal goal of the present study is to measure a complete curve of E_T vs. −ln a_w from 0 to 40 w/v% EG for a single preparation of p30δ DNA. Also obtained is a corresponding curve of δl_0 vs. −ln a_w, where δl_0 is the change in intrinsic twist from the value prevailing in 0 w/v % EG. Circular dichroism (CD) spectra in 0, 10, 20, 30, and 40 w/v% EG were also recorded. Particular objectives are: (i) to develop a general model of a *hydration-coupled* two-state transition, which expresses the various thermodynamic quantities on a per base-pair basis and takes into account cooperativity of the transition; (ii) to fit and compare different versions of this model to the E_T vs. −ln a_w data to extract relevant model parameters; (iii) to identify the values of −ln a_w and $\partial E_T/\partial$ (−ln a_w) at the midpoint of the transition; (iv) to estimate what fraction of the sequence exhibits the terminal state *in the complete absence of added osmolyte*; (v) to estimate the torsion and bending elastic constants, α and κ_β, of the terminal state that predominates in ~40 w/v% EG from the measured values of α determined previously for 0 and 20 w/v% EG[17,25] and the value of E_T for 40 w/v% EG; (vi) to ascertain δl_0 for that same terminal conformation; (vii) to show that a model, which gives a good fit of the E_T vs. −ln a_w data, also gives an acceptable fit of the somewhat noisier δl_0 vs. −ln a_w data with no further adjustment of parameters, when a suitable estimate is adopted for the difference in intrinsic twist between the two states involved; (viii) to show that the CD data are consistent with the same model; and (ix) to learn something about the cooperative domain size at the midpoint of the transition and the change, ΔΓ, in the preferential interaction coefficient *per base-pair* for water. Because our data lack sufficient precision, not all of these objectives are reached unambiguously. Nevertheless, significant new information emerges.

Theory

An osmolyte-induced structural transition

We consider a duplex DNA that contains N subunits, each of which consists of a base-pair (bp) plus its associated backbone atoms. Each subunit is assumed to exist in either of two conformational states, 1 or 2. The instantaneous *conformational* configuration of the entire DNA is specified by the N consecutive state labels, k = 1 or 2, of its N subunits, for example,

…111122222221111111111222221111111222 2…

Such a specific configuration is denoted by the index I, and a DNA with the Ith configuration is denoted by D_I. In purely aqueous buffer (0 w/v% neutral osmolyte), the standard state chemical potential of D_I is $\mu^0_{D_I}(c^\square_w)$, where $c^\square_w$ denotes the concentration (molecules per unit volume) of water in that buffer. Adding a small neutral osmolyte (os) to concentration c_{os} generally decreases both c_w and the water activity, a_w. It also typically alters the standard state molecular chemical potential of D_I to $\mu^0_{D_I}(c_w, c_{os})$. This standard state chemical potential is that part of the total chemical potential, $\mu_{DI}(c_w,c_{os})$, that is independent of the concentration of D_I. In its hypothetical 1.0 M standard state, each D_I molecule experiences only the environment of the infinitely dilute solution, where $c_{D_I} = c^\infty_{D_I} \to 0$. Under such conditions, the solution surrounding the DNA has only two components, w and os, and at constant T and P it has only

a single *independent* chemical potential, which we take to be, $\mu_w(c_w, c_{os})=\mu_w^0(c_w^{\square})+kT\ln a_w$. To first order, the variation of $\mu_{D_I}^0(c_w, c_{os})$ with ln a_w is given by

$$\begin{aligned}\mu_{D_I}^0(c_w, c_{os}) &= \mu_{D_I}^0(c_w^{\square}) \\ &\quad - \Gamma_{D_I|w}\cdot kT\ln a_w\end{aligned} \tag{1}$$

where k is Boltzmann's constant, T is absolute temperature, and

$$\begin{aligned}\Gamma_{D_I|w} &\equiv -\left(\partial\mu_{D_I}^0/\partial\mu_w\right)_{T,P,c_{D_I}^{\infty}} \\ &= -(1/kT)\left(\partial\mu_{D_I}^0/\partial\ln a_w\right)_{T,P,c_{D_I}^{\infty}}\end{aligned} \tag{2}$$

is a *preferential interaction coefficien*t (PIC) of D_I with water. In equation (1) it is *assumed* that $\Gamma_{D_I|w}$ is approximately constant over the range of ln a_w encountered in the present experiments, and that ln a_w for the purely aqueous buffer is approximately zero. $\Gamma_{D_I|w}$ can be expressed in molecular terms as, [26,27]

$$\Gamma_{D_I|w}=c_w\left(\int_0^{\infty}d\mathbf{r}\left(g_{D_I|w}(r)-g_{D_I|os}(r)\right)\right) \tag{3}$$

where $g_{D_I|w}(r)$ and $g_{D_I|os}(r)$ are the D_I-w and D_I-os pair correlation functions, respectively, r is the distance from the (arbitrarily chosen) center of the macromolecule to the (arbitrarily chosen) centers of neighboring *w* or *os* molecules, and d**r** is a volume element.

We invoke the heuristic assumption that the integrand in equation (3) arises primarily from two additive contributions, namely (i) the different volumes that are totally excluded by D_I to water and osmolyte centers, and (ii) osmolyte-water exchange (or displacement) reactions at osmolyte accessible sites outside those regions, but presumably still near the surface of D_I. These assumptions yield,[27]

$$\Gamma_{D_I|w}=X+S \tag{4}$$

where $X=(V_{os}^{ex}-V_w^{ex})/\overline{V}_w$ is the difference between the volumes excluded by D_I to osmolyte centers, V_{os}^{ex}, and water centers, V_w^{ex}, in units of the partial molecular volume of water, $\overline{V}_w$. S is the osmolyte-water exchange contribution, and is given by

$$S=\sum_{i=1}^{L}(\nu-(c_w/c_{os})K_i a_{os}(a_w)^{-\nu}/(1+K_i a_{os}(a_w)^{-\nu})) \tag{5}$$

where the i-sum runs over a lattice of L sites of equal volume, $\bar{V}_{os}$, that completely fill the space accessible to osmolyte centers, $v \equiv \bar{V}_{os}/\bar{V}_w$, and K_i is the equilibrium constant for the exchange reaction, wherein a single osmolyte is removed from a site in the distant solution, and placed in the ith site near D_I, while v water molecules are transported in the reverse direction. Finally, a_{os} is the osmolyte activity. X varies with the area and rugosity of the DNA surface and the osmolyte size, which define the region that is *inaccessible* to osmolyte centers due to hard-core repulsion, whereas S depends upon the net affinity of osmolyte for its *accessible* sites outside the region of total exclusion, and could in principle be either positive (net exclusion) or negative (net attraction). Some of the sites accessible to osmolyte centers permit surface contact between the osmolyte and DNA. An important point is that X and S in equation (4) arise from excess or deficit concentrations of water *near the DNA surface* relative to its bulk average concentration. Because both contributions, X and S, in equation (4) arise from a region close to the surface of D_I, it is reasonable to assume that $\Gamma_{DI|w}$ is a sum of contributions originating from the surfaces of the individual *subunits*,

$$\Gamma_{D_I|w} = \sum_{k=1}^{N} \Gamma_{kj|w} \tag{6}$$

where $\Gamma_{kj|w}$ is the contribution of the kth subunit, which is in state j (j=1 or 2) in the Ith conformational configuration, and is the PIC per bp.

It is also assumed that $\mu^0_{D_I}(c_w^{\square})$ can be parsed into contributions from base-pairs in the 1-state and those in the 2-state, plus an additional free energy, F_J, associated with 12 or 21 junctions,

$$\mu^0_{D_I}(c_w^{\square}) = \sum_{k=1}^{N} \mu^0_{kj}(c_w^{\square}) + n_J F_J \tag{7}$$

where $\mu^0_{kj}(c_w^{\square})$ is the contribution of the kth base-pair in state j, and n_J is the total number of 12 plus 21 junctions in the Ith conformational configuration. Incorporating equations (6) and (7) into (1) yields finally,

$$\mu^0_{D_I}(c_w, c_{os}) = \sum_{k=1}^{N} \left(\mu^0_{kj}(c_w) - \Gamma_{kj|w} \cdot kT\ln(a_w)\right) + n_J F_J \tag{8}$$

For simplicity, it is now assumed that the difference, $\mu^0_{k2}(c_w^{\square}) - \mu^0_{k1}(c_w^{\square})$, takes the same value, $\Delta\mu^0(c_w^{\square})$, regardless of whether the kth subunit is a GC or an AT base-pair. Likewise, the difference, $\Gamma_{k2|w} - \Gamma_{k1|w}$, is assumed to take the same value, $\Delta\Gamma$, for both GC and AT base-pairs. These assumptions are approximations of unknown accuracy, and may need to be revised after data pertaining to the composition dependence of $\Delta\mu^0(c_w^{\square})$ and $\Delta\Gamma$ become available.

The configuration, D_1, wherein every base-pair is in state 1, is adopted as a reference state. The concentration of configuration D_I relative to that of D_1 is found to be:

$$c_{D_I}/c_{D_I} = \exp\left[-\left(\mu^0_{D_I}(c_w, oc_s) - \mu^0_{D_I}(c_w, c_{os})\right)\right] = (B)^{n_2} J^{n_J} \tag{9}$$

wherein

$$B = K_0 (a_w)^{\Delta\Gamma} \tag{10}$$

The quantity, $K_0 = \exp\left[-\Delta\mu^0\left(c^0_w\right)\right]$ is the intrinsic equilibrium constant for the 1□ 2 transition of a single base-pair in the purely aqueous buffer with no osmolyte, $\Delta\Gamma$ is the difference in the PIC *per base-pair* between the 2 and 1 states, and $J \equiv \exp[-F_J/kT]$ is the inverse cooperativity parameter, which is assumed to be insensitive to a_w.

By summing over all distinct configurations, $I(n_2)$, that exhibit exactly n_2 subunits in state 2, we obtain the total concentration of DNAs with n_2 subunits in state 2,

$$c_{n_2} = \sum_{I(n_2)} c_{D_I} = c_{D_I} \cdot B^{n_2} \cdot \sum_{I(n_2)} J^{n_J} \tag{11}$$

where n_J denotes the number of junctions in the I(n_2)-th configuration. A further summation over n_2 yields the total DNA concentration, $c_{tot} = \sum_{n_2 \geq 0} c_{n_2} = c_{D_1} \chi$, where

$$\chi \equiv \sum_{n_2 \geq 0} B^{n_2} \sum_{I(n_2)} J^{n_J} \tag{12}$$

is the conformational grand partition function.

The average fraction of subunits in state 2 is given by

$$f^0_2 \equiv (1/Nc_{tot}) \sum_{n_2 \geq 0} n_2 c_{n_2} = (1/N) B \partial \ln \chi / \partial B \tag{13}$$

and the average number of 12 plus 21 junctions per DNA molecule is given by

$$\langle n_J \rangle = J \partial \ln \chi / \partial J \tag{14}$$

The grand partition function, χ, in equation (12) is evaluated simply by summing over all possible 2^N distinct states of the N subunits, which can be effected by using a transfer matrix, **M**. For a circular DNA,

$$\chi = Tr\left(\mathbf{M}^N\right) = \lambda_+{}^N + \lambda_-{}^N \cong \lambda_+{}^N \tag{15}$$

where

$$\mathbf{M} = \begin{pmatrix} 1 & JB \\ J & B \end{pmatrix} \tag{16}$$

and

$$\lambda_\pm = (1/2)\left(B+1 \pm \left((B-1)^2 + 4BJ^2\right)^{1/2}\right) \tag{17}$$

are its eigenvalues. The last equality in equation (15) applies in the limit of large N.

Use of equations (17) and (15) in (13) yields the fraction of subunits in the 2-state,

$$\begin{aligned} f_2^0 &= (1/2)\left(1 + \frac{B-1}{\left[(B-1)^2 + 4BJ^2\right]^{1/2}}\right) \\ &= (1/2)\left(1 + \frac{K_0(a_w)^{\Delta\Gamma} - 1}{\left[\left(K_0(a_w)^{\Delta\Gamma} - 1\right)^2 + 4K_0(a_w)^{\Delta\Gamma} J^2\right]^{1/2}}\right) \end{aligned} \tag{18}$$

When J=1, equation (18) simplifies to the non-cooperative expression employed previously.[17] From equation (18) one obtains the slope,

$$\begin{aligned} \partial f_2^0 / \partial \ln B &= B(B+1)J^2 \\ &\quad / \left[(B-1)^2 + 4BJ^2\right]^{3/2} \end{aligned} \tag{19}$$

The midpoint of the transition, $f_2^0 = 1/2$, occurs when B=1.0, or equivalently when

$$(\ln a_w)_{1/2} = -(\ln K_0)/\Delta\Gamma \tag{20a}$$

At the midpoint the slope becomes $\left(\partial f_2^0 / \partial \ln B\right)_{1/2} = 1/4J$. Using $d \ln B = \Delta\Gamma d \ln a_w$, this becomes

$$\left(\partial f_2^0 / \partial \ln a_w\right)_{1/2} = \Delta\Gamma / 4J \tag{20b}$$

Evidently, the midpoint slope in equation (20b) is equally sensitive to $\Delta\Gamma$ and $1/J$. Experimental measurements of $(\ln a_w)_{1/2}$ and $\left(\partial f_2^0 / \partial \ln a_w\right)_{1/2}$ provide only two numbers, which are

insufficient to determine K_0, $\Delta\Gamma$, and $1/J$. Estimation of all three parameters must rely on data in the wings of the transition. Unfortunately, our data lack sufficient precision to provide separate estimates of $\Delta\Gamma$ and $1/J$. It can be shown from equations (10) and (19) that the curve of f_2^0 vs. ln a_w is symmetric for all choices of K_0, $\Delta\Gamma$, and J, in the sense that $\left(\partial f_2^0/\partial \ln a_w\right)_{1/2+\Delta}$ $=\left(\partial f_2^0/\partial \ln a_w\right)_{1/2-\Delta}$, where Δ is any difference between f_2^0 and the midpoint value of 1/2 (unpublished analysis).

Use of equations (17) and (15) in (14) gives the average number of 12 plus 21 junctions,

$$\langle n_J\rangle = N2BJ^2 \Big/ \left(\lambda_+\left[(B-1)^2+4BJ^2\right]^{1/2}\right) \tag{21}$$

At the midpoint, when B = 1.0, equations (17) and (21) give,

$$\langle n_J\rangle_{1/2} = NJ/(1+J) \tag{22}$$

and the average size of domains of consecutive 1-states or 2-states at the midpoint is given in base-pairs by,

$$d = N/\langle n_J\rangle = 1+1/J \tag{23}$$

Clearly, the domain size is determined entirely by $1/J$ with no contribution from $\Delta\Gamma$, as expected.

Equations (6)–(23) generalize linear lattice models of two-state structural transitions to incorporate the effects of a small cosolute (the osmolyte), which relative to water may be either net attracted to, or excluded from, the neighborhood of a subunit of the macromolecule in either of its two possible conformations. This theory could readily be extended to admit an effect of osmolyte on the junction free energy, F_J.

Free energy of supercoiling

For a given topoisomer, the number of turns, l, of one strand around the other is an integral topological invariant called the *linking number*. The state of deformation of a supercoiled DNA is characterized by its linking difference, $\Delta l = l - l_0$, where l_0 is the generally non-integral intrinsic twist. The linking number is partitioned between twist (t) and writhe (w) according to, $l = t + w$,[28,29] so

$$\Delta l = t - l_0 + w \tag{24}$$

Both experiments[17, 30–34] and simulations[24,35,36] pertaining to long DNAs ($N \geq 2000$ bp) in ~0.1 M monovalent salt indicate that the free energy change to vary the linking difference from $\Delta m = m - l_0$ to Δl is given by

$$\Delta G_{sc}=kT(E_T/N)\left(\Delta l^2-\Delta m^2\right) \tag{25}$$

wherein E_T is independent of Δl and Δm, especially over the thermally populated range of topoisomers. Equation (25) also accurately represents the present data for p30 δ in all concentrations of EG from 0 to 40 w/v%.

It was previously proposed that, for DNAs in ~0.1 M monovalent salt, E_T could be approximated by,[4]

$$E_t^{eff}=\left((2\pi)^2/2kT\right)\alpha^{eff}B_w\kappa_\beta^{eff} / \left(\alpha^{eff}+B_w\kappa_\beta^{eff}\right) \tag{26}$$

where α^{eff} and κ_β^{eff} are the effective elastic constants for torsion and bending of the springs that are imagined to exist between base-pairs, and B_w is a formerly unknown constant. Recent simulations of a 4910 bp model DNA under our standard topoisomerization conditions validated equation (26) *within the simulation errors*, and established that $B_w \cong 0.594$.[37]

Variation of the effective elastic constants with EG concentration

We imagine that the kth subunit is connected to the (k+1)th subunit by Hookean torsion and bending springs, whose minimum energy positions and torque constants depend upon the conformational state of the kth subunit. In regard to twisting springs, the minimum energy twists are, respectively, φ_1^0 and φ_2^0, and the corresponding torque constants for twisting are α_1 and α_2. Similarly, in regard to bending springs, the minimum energy bends are, respectively, β_1^0 and β_2^0, and the corresponding torque constants for bending are $\kappa_{\beta 1}$ and $\kappa_{\beta 2}$.

When the numbers of springs of types 1 and 2 are fixed at their unperturbed values, n_1^0 and n_2^0, the effective elastic constants are found to be

$$1/\alpha^{eff}=\left(1-f_2^0\right)/\alpha_1+f_2^0/\alpha_2 \tag{27}$$

and

$$1/\kappa_\beta^{eff}=\left(1-f_2^0\right)/\kappa_{\beta_1}+f_2^0/\kappa_{\beta_2} \tag{28}$$

where $1-f_2^0$ and $f_2^0=n_2/N$ are the unperturbed fractions of subunits (base-pairs) prevailing in the absence of any deformational strain. However, in reality the 1□ 2 equilibrium generally shifts in such a way as to reduce the deformational strain energy, so f_2 is not exactly equal to its unperturbed value, f_2^0, in equations (18), (27), and (28). The relevant theory to incorporate torsional strain into the present model of the conformational equilibrium is presented in the Appendix. The final result for $f_2 = n_2/N$ is given by (A20) with F (n_2) given by (A19) and λ given by (A17). The quantity, ∂U (n_2, Δt) ∂n_2, is the slope of the torsional strain energy for a given fixed net twist, $\Delta t = t - l_0$, with respect to n_2 (the number of subunits in state 2), and is

given explicitly in terms of n_2, α_1, α_2, $\Delta\varphi^0 \equiv \varphi_2^0 - \varphi_1^0$, and Δt in equation (A21). In principle, equation (A20) can be solved by iteration to find f_2 for any fixed choice of α_1, α_2, $\Delta\varphi^0$, and Δt. It is shown via equations (A22) and (A23) that in the present case of extremely small $\Delta\varphi^0$ and small Δt, corresponding to $\Delta l = +2$ turns, the value of n_2 must lie very close to n_2^0, so that the 1□ 2 equilibrium is practically unshifted by the net twist. The similarly small net bending strain in a topoisomer with $\Delta l = +2$ is also believed to exert no significant effect on the 1□ 2 equilibrium. Consequently, equations (27) and (28) should apply to very high accuracy in the present study.

Substituting equations (27) and (28) into the inverse of equation (26) and rearranging somewhat yields,

$$1/E_T^{eff} = \left(1 - f_2^0\right)/E_{T_1} + f_2^0/E_{T_2} \tag{29}$$

where $E_{T_j} = \left((2\pi)^2/2kT\right)\left(\alpha_j B_w \kappa_j / \left(\alpha_j + B_w\kappa_j\right)\right)$ is the expected value of E_T, when the entire DNA is in state j, j=1 or 2. Curves of E_T^{eff} vs. ln a_w can be calculated by using equations (18) and (29) for arbitrary choices of E_{T1}, E_{T2}, K_0, J, and $\Delta\Gamma$. Fitting these equations to the measured values of E_T vs. −ln a_w *in principle* allows determination of all five of these parameters. However, the present data lack sufficient precision to determine both J and $\Delta\Gamma$, which can be distinguished only in the wings of the transition, as noted earlier.

At the midpoint of the transition,

$$\left(1/E_T^{eff}\right)_{1/2} = (1/2)\left(1/E_{T_1} + 1/E_{T2}\right) \tag{30}$$

and

$$\begin{aligned}\left(\partial E_T^{eff}/\partial \ln a_w\right)_{1/2} &= \left(\partial E_t^{eff}/\partial f_2^0\right)_{1/2} \cdot \left(\partial f_2^0/\partial \ln a_w\right)_{1/2} \\ &= \left(0.5/E_{T_1} + 0.5/E_{T_2}\right)^{-2} \cdot \left(1/E_{T_1} - 1/E_{T_2}\right) \cdot \Delta\Gamma/4J\end{aligned} \tag{31}$$

Fortunately, the best-fit values of E_{T_1} and E_{T_2} and also the best-fit values of $(\ln a_w)_{1/2}$ and $\left(\partial E_T^{eff}/\partial \ln a_w\right)_{1/2}$ are rather insensitve to the large, but highly correlated, uncertainties in K_0, $\Delta\Gamma$, and $1/J$, as shown below.

Materials and Methods

DNA Preparation

The p30δ plasmid was described previously.[17] However, recent direct sequencing indicates unambiguously a length, N=4932 bp, rather than the 4752 bp inferred from restriction digests by the group that constructed the plasmid.[38] Isolation and purification of the p30δ were carried out as described previously.[17] The native p30δ was dialyzed into a storage buffer (50 mM KCl, 50mM Tris, 0.1 mM EDTA, pH 7.5) and kept at 4 °C. The A260/A280 was ~1.90. Analysis by gel electrophoresis indicated a supercoiled:nicked ratio greater than 3:1.

Circular dichroism spectra

A portion of the p30δ sample was linearized by Eco RI. Completion of the circular → linear conversion was confirmed by gel electrophoresis in 0.8% agarose. The enzyme was extracted from the linearized DNA sample several times with buffered phenol, and the sample was then dialyzed into the CD buffer (100 mM NaCl, 10 mM Tris, 1mM EDTA, and 2.5 w/v% glycerol, plus various w/v% EG). The water activities of these buffers differ negligibly from those of the Topo I buffers with the same w/v% EG, which are described subsequently. CD spectra of these samples, corrected for any solvent or instrumental contribution, were measured at 37 ° C, as described elsewhere.[41] The absorbance of these CD samples at 258 nm was 0.204.

Topoisomerization reactions

Supercoiled p30δ was relaxed at 37 °C for 6 hours by calf-thymus topoisomerase I (from Invitrogen) in the presence of 0 to 40 w/v% EG. This topoisomerase I remains active up to 43 w/v% EG, but is rendered inactive by 47 w/v% EG. The topoisomerase I was added in three separate aliquots at 0, 2.0, and 4.0 hours to ensure complete topisomerization of the sample. The total final concentration of topoisomerase I in the sample is 0.1 units/μL. The final concentrations of all other components of the Topo I buffer were: 50 mM KCl, 50mM Tris, 0.1 mM EDTA, 10 mM $MgCl_2$, 30 μg/mL bovine serum albumin, 0.5 mM dithiothreitol, 2.5 w/v% glycerol, 25 μg/mL DNA, and various w/v% EG. Reactions were halted by extracting the enzyme once with a mixture of phenol, 0.1 w/v% 8-hydroxyquinoline, 48 v/v% chloroform, and 2 v/v% isoamyl alcohol. The phenol/8-hydroxyquinoline was preequilibrated with a buffer containing 0.2 mM NaCl and 0.1 mM Tris at pH 8.0. Every topoisomerization reaction was replicated at least four times. The ~60 μL volume of each replicate reaction yielded three 20 μL aliquots of reaction products, each of which was run in a separate lane in gel electrophoresis. The topoisomerization reactions performed for each particular condition thus yielded a total of at least 12 lanes for gel analysis. The large number of gel lanes was required to achieve acceptable statistical accuracy of the results.

Gel electrophoresis

Gel electrophoresis was performed in 0.8 w/v% agarose, as described previously.[17] Sufficient (0.08 μg/mL) chloroquine was added to reduce the effective intrinsic twist, and thereby shift the linking differences of all visible topoisomers into the slightly positive range. At the conclusion of the run, the gel was soaked twice in distilled water (2L, 1 hr) to remove the chloroquine. The gel was then stained by soaking in 1 L of 0.5 μg/mL ethidium bromide for 45 minutes in the dark, and finally soaked twice in distilled water to reduce the background level of ethidium. Under the prevailing conditions, ethidium uptake and fluorescence are proportional to the local concentration of the topoisomers in each band.[11]

Photoimaging and quantitation

The gel was illuminated by 488 nm light, and a digital image of the ethidium fluorescence (λ_{max} = 640 nm) was recorded by a Molecular Dynamics Fluoro Imager-SI Optical Scanner. Plots of fluorescence intensity vs. distance migrated (in pixels) are shown in Figure 1 for p30δ samples that were relaxed in 0 and 40 w/v% EG. The integrated fluorescence intensity of each band was determined with the aid of the ImageQuant 5.1 program from Molecular Dynamics. The baseline and limits of integration for each peak were set manually. The ratio of integrated intensities of any two bands in the same lane is taken to be equal to the ratio of their equilibrium concentrations under the conditions of their particular topoisomerization reaction.[11] Reliable data were typically obtained for 6–8 topoisomers in each lane.

Analysis of the topoisomer distribution data

The linking number of the most populous topoisomer is the nearest integer to the intrinsic twist, l_0, and is denoted by NINT(l_0). The excess linking number, l_{ex}, of a given topoisomer with linking number, l, is defined by,

$$l_{ex}=l - \mathrm{NINT}(l_0) \tag{32}$$

The bands in each gel lane are indexed according to their l_{ex} values (…−3, −2, −1,0,+1,+2,+3, …). The ratio, $c_{l_{ex}}/c_0$, of the equilibrium concentration of a given topoisomer with l_{ex} to that of the most populous topoisomer is given by,

$$c_{l_{ex}}/c_0=\exp\left[-(E_T/N)\left((l_{ex}+\Delta l_0)^2-(\Delta l_0)^2\right)\right] \tag{33}$$

where $\Delta l_0 \equiv \mathrm{NINT}\,(l_0) - l_0$ is the linking difference of the most populous topoisomer. Equation (33) is fitted to the measured ratios of integrated intensities, $I_{l_{ex}}/I_0 = c_{l_{ex}}/c_0$, by an iterative nonlinear least-squares protocol with E_T and Δl_0 as adjustable parameters.[17] Optimum fits for the two sets of data in Figure 1 are shown in Figure 2. Such good fits typify every data set analyzed. After fitting the $c_{l_{ex}}/c_0$ data for each of the 10–15 satisfactory gel lanes for each reaction condition, the resulting E_T and Δl_0 values were averaged and reported along with the standard deviations of their means in Table 1.

Obtaining δl_0 as a function of EG concentration

Both l_0 and Δl_0 may vary with EG concentration (c_{EG} (w/v%)), and are now written as l_0 (c_{EG}) and Δl_0 (c_{EG}) to indicate that. In the case of 0 w/v% EG, c_{EG} = 0, and one has l_0 (0) and Δl_0 (0). The difference between experimental linking differences,

$$\begin{aligned}\Delta\Delta l_0 \equiv\ & \Delta l_0(c_{EG}) \\ & - \Delta l_0(0) \\ =\ & \mathrm{NINT}(l_0(c_{EG})) \\ & - l_0(c_{EG}) \\ & - (\mathrm{NINT}(l_0(0)) - l_0(0))\end{aligned} \tag{34}$$

reflects a change in l_0 (c_{EG}) with increasing c_{EG}. The change in intrinsic twist upon adding osmolyte is defined by[17]

$$\begin{aligned}\delta l_0 \equiv\ & l_0(c_{EG}) \\ & - l_0(0) \\ =\ & \mathrm{NINT}(l_0(c_{EG})) \\ & - \mathrm{NINT}(l_0(0)) - \Delta\Delta l_0\end{aligned} \tag{35}$$

The variation in NINT (l_0 (c_{EG})) with EG concentration can be ascertained directly by running topoisomerization reaction products for the conditions c_{EG} and 0 in the same gel, as shown in Figure 1. Although NINT (l_0 (40)) − NINT (l_0 (0))= 0 for the two lanes in Figure 1, only slight differences in integrated intensities could lead to values of −1 or −2. However, a shift of NINT

(l_0 (40)) − NINT (l_0 (0)) to −1 would be accompanied by a change of −1.0 in the corresponding value of $\Delta\Delta l_0$, and a shift to −2 would be accompanied by a change of −2 in the corresponding value of $\Delta\Delta l_0$, so the value reckoned for δl_0 would remain unaltered in any case. The mean values of the measured δl_0 for each condition are listed in Table 1. Standard error propagation is used to calculated the standard deviations, $\sigma_{\Delta\Delta l_0}$ and $\sigma_{\delta l_0}$, of the mean $\Delta\Delta l_0$ and δl_0 values from the standard deviations, $\sigma_{\Delta l_0}$, of the mean Δl_0 values for each condition. In calculating the mean Δl_0 and standard deviation $\sigma_{\Delta l_0}$ for all data sets *at a given condition*, the most populous topoisomer must be chosen to yield the same value of NINT (l_0 (c_{EG})) for each of those data sets. The $\sigma_{\delta l_0}$ values are all rather small, as indicated in Figure 4 by the error bars at ± $\sigma_{\delta l_0}$, which in most cases are separated by less than the symbol diameter.

Water activities

Thermodynamic activities (a_w) of water in EG-containing solutions were estimated as follows. For EG concentrations below 15 w/v%, the a_w were derived from vapor osmometry measurements of the topoisomerization (Topo I) buffers, as reported previously.[17] The vapor pressure osmometry method apparently becomes less reliable for EG concentration ≥ 15 w/v %, so a_w -values for those solutions were reckoned in the following way. First, ln a_w was determined from published freezing point (T_f) data on EG-water solutions[39] according to the following equation,

$$-RT\ln a_w = -\int_{273}^{T_f} dT\,\Delta\overline{H}_{fus}(273)/T - \int_{273}^{T_f} dT\left(C_P^{liq} - C_P^{ice}\right) + \int_{273}^{T_f} dT\left(C_P^{liq} - C_P^{ice}\right)273/T \quad (36)$$

The constants, R=1.9872 cal/(mol-K), C_p^{liq}=18.18 cal/(mol − K), C_p^{ice}=8.64 cal/(mol − K), and $\Delta\bar{H}_{fus}$ (273)= 1435.6 cal/mol, were taken from the Handbook of Chemistry and Physics.[40] The −ln a_w values obtained from equation (36) apply strictly to EG-water solutions at 273 K, but are assumed to apply also at 310 K. The −ln a_w vs. w/v% EG data were accurately fitted by the relation, f(x)=0.0035 + 0.0018 x + 0.0001 x^2, where f ≡ −ln a_w and x ≡ w/v%EG. This relation is expected to apply for EG-water solutions. However, the buffers used in topoisomerization experiments contained also 2.5 w/v% glycerol in addition to other buffer components. Consequently, the −ln a_w values of those buffers exceed those of the corresponding EG-water mixtures. Thus, *for Topo I buffers*, the best-fit curve was shifted upward so as to practically coincide with the vapor pressure osmometry measurements at 0, 5, and 10 w/v% EG. The resulting curve, $f_{topo}(x)$=0.011 + 0.0018 x + 0.0001 x^2, was used to estimate the −ln a_w values for all Topo I buffers, which are presented in Table 1. Because the concentration of small-ions plus small neutral buffer components, including the 2.5 w/v% glycerol, are practically the same in the CD buffer as in the Topo I buffer, the same water activities are expected to apply to those solutions as well.

Results and Discussion

The best-fit values of E_T are plotted vs. −ln a_w in Figure 3, where the reverse sigmoidal shape is plainly apparent. It is desired to fit equation (29), with f_2^0 given by equation (18), to these data. There are five adjustable parameters, namely E_{T_1}, E_{T_2}, K_0, ΔΓ, and J. From the shape of

the data "curve", one might presume that it could provide reasonable estimates of four quantities, namely the upper and lower plateau values, E_{T_1} and E_{T_2}, the midpoint value, $(-\ln a_w)_{1/2}$, and the slope at the midpoint, $\left(\partial E_T^{eff}/\partial \ln a_w\right)_{1/2}$. This is indeed the case, as shown below.

Fit of a non-cooperative model

Initially, it was assumed that $J = 1.0$, corresponding to a *non-cooperative* model, and a four-dimensional grid search was undertaken to find optimum values of E_{T_1}, E_{T_2}, $\Delta\Gamma$, and K_0. This yielded a fairly robust value, $E_{T_1} = 979$, and a plausible range of E_{T_2} -values from 880 to 894. Trial values of E_{T_2} within this range were selected, the optimum values of K_0 and $\Delta\Gamma$ were determined for each selection, and the corresponding value of $\chi^2 \equiv \sum_{j=1}^{10}\left(E_{T_j}^{\exp} - E_{T_j}^{th}\right)^2/\sigma_j^2$ was computed in each case. The minimum value, $\chi_0^2 = 12.91$ was found for $E_{T_1} = 979$, $E_{T_2} = 887$, $K_0 = 0.067$, and $\Delta\Gamma = -30.0$. These best-fit parameters are collected in Table 2. The probabililty, $F(\chi^2 > \chi_0^2) = 0.250$, indicates an acceptable fit of the theory to the experimental data. This is confirmed by visual comparison of the best-fit curve with the data in Figure 3.

The standard deviation of E_{T_2} was estimated as that deviation from the optimum value, 887, for which $\exp\left[-\chi^2/2\right]/\exp\left[-\chi_0^2/2\right] = \exp[-1/2] = 0.607$, where χ^2 applies for the deviated value of E_{T_2}. This protocol yields $\sigma_{E_{T_2}} \cong 5.0$. It may be inferred that there is no *statistically significant* difference among E_{T_2} -values in the range 884 to 890. The variances, $\left\langle(\delta\Delta\Gamma)^2\right\rangle$ and $\left\langle(\delta K_0)^2\right\rangle$, and normalized covariance, $\tau \equiv \langle(\delta\Delta\Gamma)(\delta K_0)\rangle / \left(\left\langle(\delta K_0)^2\right\rangle^{1/2}\left\langle(\delta\Delta\Gamma)^2\right\rangle^{1/2}\right)$, were estimated from elements of the curvature tensor of the χ^2 surface (as a function of $\delta\Delta\Gamma$ and δK_0) at its minimum, as described elsewhere.[41] The estimated relative error in K_0 is, $\left\langle(\delta K_0)^2\right\rangle^{1/2}/K_0 \cong 0.23$, and that in $\Delta\Gamma$ is, $\left\langle(\delta\Delta\Gamma)^2\right\rangle^{1/2}/\Delta\Gamma \cong 0.08$. However, these values cannot be simply interpreted, because the normalized covariance, $\tau \cong -0.8$, is large and negative, which implies that deviations of K_0 from its most probable value are strongly anti-correlated with deviations in $\Delta\Gamma$ from its most probable value, as expected.[41] Perhaps more relevant is the fact that the implied value of $(-\ln a_w)_{1/2}$ increases with decreasing E_{T_2} from 0.083 (for $E_{T_2} = 892$) to 0.090 (for $E_{T_2} = 887$) to 0.097 (for $E_{T_2} = 882$). Given the estimated uncertainty, $\sigma_{E_{T_2}} \cong 5.0$, in E_{T_2}, the associated relative uncertainty in $(-\ln a_w)_{1/2}$ is then $\pm 0.007/0.090 = \pm 0.08$.

Predictions of corresponding cooperative models

Different models with the same values of E_{T_1} and E_{T_2}, $(-\ln a_w)_{1/2}$, and $\left(\partial E_T^{eff}/\partial \ln a_w\right)_{1/2}$ give almost equally good fits to our data, and are henceforth referred to as *corresponding* models. As can be seen from equations (20a,b) and (31), it is possible to vary J subject to the constraint of constant midpoint slope by introducing a compensating change in $\Delta\Gamma$, and in turn to vary $\Delta\Gamma$ subject to the constraint of constant $(\ln a_w)_{1/2}$ by introducing a compensating change in K_0. Thus, there is a continuum of corrersponding models that all give nearly equivalent fits to our data. If the data in the wings of the transition were sufficiently precise, it would be possible to narrow considerably the allowed ranges of J, $\Delta\Gamma$, and K_0, but that is not the case here.

In all events, the *negative* midpoint slope, $\left(\partial f_2^0/\partial \ln a_w\right)_{1/2} = \Delta\Gamma/4J = -30.0$, indicates that $\Delta\Gamma$ must be negative, since J is inherently positive. Thus, there is evidently less water and/or more EG in the neighborhood of a subunit in state 2 than in the neighborhood of a subunit in state

1. In view of equation (4), either the exchange contribution (S) or the excluded volume contribution (X), or both, must be smaller in state 2 than in state 1. X is reduced primarily by decreasing the (hard-core) surface area of the DNA, and S is decreased either by increasing the equilibrium constants for osmolyte-solvent exchange at one or more sites, or by reducing the number of exchange sites, which is equivalent to a reduction in DNA surface area. Thus, DNA in state 2 has either a smaller hard-core surface area or a higher affinity of one or more exchange sites for EG, or both.

An absolute value, $|\Delta\Gamma| \equiv |\Gamma_{2|w} - \Gamma_{1|w} = 30.0$ *per bp*, seems much too large for an osmolyte like EG, which is net excluded (relative to water) from the DNA surface.[42] Studies of the effect of EG on the melting temperature indicated that $\Delta\Gamma \cong -4.0$ for that transition between very different structures with surfaces of rather different size and atomic composition.[43] Because the present transition appears to involve only rather modest changes in structure, with doubtless much smaller changes in surface size and atomic composition, $|\Delta\Gamma|$ should be considerably less than 4.0. Studies of the effect of sucrose and other polyols on the B□ Z transition of poly(dG-dm^{5}C) were combined with an independent estimate of J to infer that $\Delta\Gamma_{suc} = -2.5$, when the osmolyte is *sucrose*.[44] The value of $(\ln a_w)_{1/2}$ for EG exceeds that for sucrose by the ratio 8.5/3.9.[44] In view of equation (20a), this suggests that for EG, $\Delta\Gamma = \Delta\Gamma_{suc}/(8.5/3.9) = -1.15$. Because that B□ Z transition involves a much more extensive structural rearrangement than the present transition, it is most probable that $\Delta\Gamma$ for EG and the present transition lies in the range, $-1.0 \leq \Delta\Gamma \leq 0.0$.

For illustrative purposes, we now assume that $\Delta\Gamma = -1.0$ for the present transition, in which case the value, $1/J = 30.0$, is required to match the midpoint slope of the best-fit non-cooperative curve of E_T^{eff} vs. $-\ln a_w$. A value, $1/J \geq 1.0$, implies that the transition is cooperative. In addition, $-\ln K_0$ must decrease by 30-fold in order to attain the same midpoint value, $(\ln a_w)_{1/2} = -\ln K_0/\Delta\Gamma = -0.09010$, as that of the non-cooperative curve. This change yields, $K_0^c = 0.914$, for the cooperative model, which is denoted by the superscript c. The predicted curve of E_T^{eff} vs. $-\ln a_w$ for the cooperative model with $\Delta\Gamma = -1.0$, $1/J = 30.0$, and $K_0^c = 0.914$ is displayed in Figure 3, and the relevant parameters are listed in Table 2. The apparent fit to the data is practically as good as that for the original non-cooperative curve, despite the fact that it is merely a "prediction" that preserves the same values of E_{T1}, E_{T2}, $(-\ln a_w)_{1/2}$, and $\left(\partial E_t^{eff}/\partial \ln a_w\right)_{1/2}$, rather than a true fit.

If $\Delta\Gamma = -1.0$, as assumed here, then $1/J = 30.0$, and the average domain size at the midpoint is $d = 31$ base-pairs. There is no information to preclude the possibility that $|\Delta\Gamma| \ll 1.0$, in which case $1/J$ and d would exceed 30 base-pairs by many fold.

Fraction of base-pairs in the 2-state

The fractions of base-pairs in the 2-state that are predicted by equation (18) for the best-fit non-cooperative model ($J = 1.0$, $\Delta\Gamma = -30.0$, $K_0 = 0.067$) and the corresponding cooperative model ($\Delta\Gamma = -1.0$, $J = 1/30.0$, $K_0^c = 0.914$) are compared in Figure 3. These curves are practically identical, except in the wings, and they also converge to identical limiting values sufficiently far in the wings.

We can estimate the fractions, f_2^0, of base-pairs existing in the 2-state under standard 0.1 M NaCl conditions, where $a_w = 0.996$ (i.e. in the complete absence of ethylene glycol and glycerol) in the following way. When this value of a_w is substituted into equation (18), it yields, $f_2^0 = 0.070$ for the best-fit non-cooperative model, and $f_2^0 = 0.105$ for the corresponding cooperative model with $\Delta\Gamma = -1.0$. Thus, the best-fit non-cooperative model predicts that ~7%

of the base-pairs exist in the 2-state under standard 0.1 M NaCl conditions, and this estimate rises with increasing assumed cooperativity of the transition to ~ 10% of the base-pairs when J = 1/30.0!

Variation of the intrinsic twist with −ln a_w

The experimental values of δl_0 are plotted vs. −ln a_w in Figure 4. The effect of increasing EG from 0 to 40 w/v% is to reduce the intrinsic twist by ~0.84 ± 0.1 turn out of ~472 total turns. Thus, the relative change in intrinsic twist (~0.18%) is very slight.

It was shown previously that, if l_{01} and l_{02} are the intrinsic twists of states 1 and 2, respectively, then[17]

$$\delta l_0 = -\Delta_a + (l_{02} - l_{01}) f_2^0 \tag{37}$$

where f_2^0 is given by equation (18), and $\Delta_a \equiv (l_{02} - l_{01}) f_{20}^0$, where f_{20}^0 is the value of f_2^0, when −ln a_w = 0.011, which corresponds to the Topo I buffer with 0 w/v% EG. With this value of Δ_a, equation (37) yields δl_0= 0, when −ln a_w = 0.011, as required.

Our next goal is to obtain a reasonable estimate of $l_{02} - l_{01}$, and with that to show that the measured δl_0 vs. −ln a_w data are consistent with model parameters already determined by fitting equations (29) and (18) to the E_T vs. −ln a_w data. A theoretically predicted curve of δl_0 vs. −ln a_w is generated in the following way. First, f_2^0 is calculated for arbitrary values of −ln a_w according to equation (18) for the non-cooperative model with its best-fit parameters (J=1.0, $\Delta\Gamma = -30.0$, K_0 =0.067). Then $l_{02} - l_{01}$ is approximated simply by the measured value of δl_0 at 40 w/v% EG, namely $l_{02} - l_{01} = -0.836$. Use of these results in equation (37) yields the solid curve in Figure 4. This "predicted" curve accounts fairly well for the observed trend of the data, despite the absence of adjustable parameters. The predicted $|\delta l_0|$ in 40 w/v% EG is ~ 10% smaller than the measured value, due to the fact that ~10 % of the total drop, $|l_{02} - l_{01}|$, has already "occurred" by the first point at −ln a_w = 0.011 (i.e. 0 w/v% EG), where δl_0 is defined to vanish. Hence, the good agreement in the tail of the curve is due in part to the fact that the measured $|\delta l_0|$ at 40 w/v% EG, which was used to estimate $|l_{02} - l_{01}|$, fortuitously exceeds the mean of the last three plateau $|\delta l_0|$-values (at 35, 37.5, and 40 w/v% EG) by just the right amount. Any method of selecting δl_0 for 35, 37.5, and 40 w/v% EG that yields a reasonable fit of the terminal plateau, together with f_2^0 values computed using the already determined best-fit parameters, will predict a δl_0 vs. −ln a_w curve in fairly good agreement with the observed trend of the data. When f_2^0 values are calculated for the corresponding cooperative model ($\Delta\Gamma = -1.0$, J = 1/30.0, K_0^c=0.914) and inserted in equation (27), the dashed curve in Figure 4 is obtained. Again, the predicted curve follows the trend of the data rather well, although an increase in the assumed value of $|l_{02} - l_{01}|$ would improve the agreement with experiment. These results demonstrate that the δl_0 vs. −ln a_w data are consistent with the same model that fits the E_T vs. −ln a_w data.

Deviations of the δl_0 data from any smooth curve are quite large compared to the statistical errors of the individual values for each concentration of EG. This suggests the presence of some unidentified and apparently random sample-to-sample error that greatly exceeds the statistical errors from fluctuations within the data set for any given concentration of EG. In this regard, the E_T vs. −ln a_w data appear at first glance to be different, with much less indication of excess variance beyond that due to statistical fluctuations within a given data set. However, in some cases the full data set for a given reaction condition consists of multiple data sets, each comprising results for ~12 gel lanes. In such cases, the variance of the mean E_T-values of the

multiple data sets significantly exceeds the variances of the E_T-values for any single data set (~12 gel lanes). This implies a significant, but apparently random, reproducibility error from one 12-lane data set to another, even though the prevailing conditions are nominally identical. Conceivably, both the excess variance of δl_0 and the reproducibility error in E_T arise from some ultra-slowly relaxing metastable secondary structure(s) with different intrinsic twist(s) and E_T -values that prevail in circular DNAs subsequent to relaxation of native superhelical stress. Our laboratory has previously reported many examples of such metastable secondary structure appearing subsequent to changes in superhelical stress caused by linearization, binding of intercalating dyes, or binding of single-strand binding protein from E. coli.[2–8] The existence of metastable secondary structure in p30δ DNA immediately after relaxation of superhelical stress and acceleration of its equilibration by topiosomerase I[41] will be described in a forthcoming paper.

Distinctive aspects of the CD spectra

Each raw CD spectrum was smoothed by standard Fourier filtering techniques, and both the raw and smoothed spectra were reported elsewhere.[41] This filtering/smoothing yields good visual fits to the raw data at the maximum near 272 nm and the minimum near 245 nm, but not at wavelengths much longer than 272 nm or much shorter than 245 nm. The CD *decreases* with increasing w/v%EG at both 272 and 245 nm, as indicated in Table 3, and no isosbestic point was observed at any wavelength.

At all wavelengths ≥ 240 nm, the spectra for 10, 20, and 30 w/v% EG lie closer to each other than to the curves for 0 and 40 w/v% EG, which is contrary to expectation for a simple sigmoidal transition.

The failure of the present CD spectra to approach plateau values at low and high w/v% EG differs qualitatively from the behavior observed for previously studied structural transitions.[44–50] One may ask whether and how it might be possible for such behavior to be compatible with a structural transition. One possibility is that the CD at 272 nm consists of an ordinary sigmoidal transition superposed on a linearly descending contribution. Unfortunately, the present CD data lack sufficient accuracy for quantitative analysis, except near 272 nm, where the separation between the spectra is the greatest. We consider the possibility that the CD values at 272 nm *relative* to that in 0 w/v% EG could be expressible as, $S(y) = -ay + bf_1^0 + cf_2^0$, where y ≡ w/v% EG/100, $-a$ is the slope of the linearly descending contribution, and b and c are the additional contributions per unit fraction of base-pairs in the 1- and 2-states that are superposed on the linear contribution. The $S(y) \equiv CD_{272}(y)/CD_{272}(0)$ are tabulated in Table 3. The above expression can be simplified to give,

$$S(y) = -ay + b + df_2^0(y) \tag{38}$$

where $d \equiv c - b$ is the amplitude of the sigmoidal transition. The $f_2^0(y)$ are reckoned from equation (18) using the best-fit non-cooperative model (J = 1.0, $\Delta\Gamma = -30.0$, $K_0 = 0.067$) obtained for the E_T vs. $-\ln a_w$ data. The optimum parameters, a=1.7479, b=0.9639, d=0.2983 were obtained by a least-squares fit of equation (38) to the experimental $S(y)$ values. The best-fit theoretical values, $S^{th}(y)$, were computed from equation (38) using the optimum values of a, b and d, and are also presented in Table 2. In every case, the agreement with experiment is within ~1.5 % or less, which is comparable to the relative errors in the CD measurements. This indicates that equation (38) with f_2^0-values reckoned from equation (18) for the best-fit non-cooperative model gives a satisfactory account of the CD data. In that sense, the CD_{272} data are consistent with the independently determined non-cooperative model.

The origin of the $-ay$ term in equation (39) is unknown. In principle, it could arise from changes in coordinates to which the CD is sensitive, but the twisting and bending elastic constants and intrinsic twist are not. Perhaps such coordinates pertain to water molecules or ions in contact with the bases, whose displacement could significantly alter the excited and/or ground states of the bases, and thereby affect their CD spectra.

Comparison with previous results from this laboratory

The present results differ quantitatively from those of an earlier study[17] in two regards: (i) the present E_T -values for 0 and 10 w/v% EG lie *below* the corresponding previous values by ~3–4 %; and (ii) the present E_T -value for ~ 20 w/v% EG lies *above* the corresponding previous value by ~ 4 %. These deviations significantly exceed the estimated statistical errors, which are indicated by the error bars in Figure 3. Consequently, the midpoint slope, $(\partial E_T/\partial \ln a_w)_{1/2}$, in the present study is substantially lower than that obtained previously.

The origins of the differences in results between the two studies are unknown. It is noteworthy that the compositions of the present and previous Topo I buffers differ in certain regards, due primarily to a difference in composition of the supplier's buffer in which the Topo I is shipped. Specifically, the present buffer contained more bovine serum albumin (30 vs. 19 μg/mL, less glycerol (2.5 vs. 4.7 w/v%), and less Topo I activity (0.1 vs. 0.65 units/μL).

Torsion and bending elastic constants of the terminal state in 40 w/v% EG

Due to experimental difficulties, torsion elastic constants could not be measured directly in the Topo I buffer, or even in TE buffer (100 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, pH 8.0) with 40 w/v% EG, but could be measured in TE buffer with 0 and 20 w/v% EG. Moreover, these measurements were made at 40 °C, instead of 37 °C. The resulting α^{eff} -values are listed in Table 4. All of the other α^{eff} -values in Table 4 are estimated by correcting these two values. First, these α^{eff} are multiplied by 1.034 to correct from 40 to 37 °C. This factor was obtained from the slope, $\partial\alpha^{eff}/\partial T$, defined by previous measurements of α^{eff} vs. T in TE buffer with 0 w/v% EG.[14] The Topo I buffer contains 9.4 mM $MgCl_2$ in ~71.4 mM monovalent "salt". It was found previously that adding 5.5 mM $MgCl_2$ to 30 mM monovalent salt inceased α^{eff} by 1.23-fold. The Topo I buffer has ~1.7-fold more Mg^{2+}, but also 2.4-fold more monovalent cations, which act to reduce the concentration of the Mg^{2+} near the DNA. In the absence of further information, the α^{eff} -values in TE buffer at 37 °C are simply multiplied by 1.23 to estimate the values in the corresponding Topo I buffer, which are also listed in Table 3. This procedure likely underestimates the actual α^{eff} -values in the Topo I buffers. Implicit in such a correction are three assumptions: (i) Mg^{2+} affects the torsional rigidities of the 1- and 2-states similarly; (ii) Mg^{2+} does not alter the two-state equilibrium, apart from its (negligibly small) effect on the water activity; and (iii) the effect of Mg^{2+} is not significantly modulated by the presence of 20 w/v% EG.

Incorporating the α^{eff} and f_2^0 values appropriate for Topo I buffer containing first 0 and then 20 w/v% EG into equation (27) yields two simultaneous equations that can be solved for α_1 and α_2. The prevailing fractions, f_2^0, of base-pairs in the 2-state *in the Topo I buffer* at 37 °C can be estimated via equation (18) using the parameters for either the best-fit non-cooperative model or the corresponding cooperative model with $\Delta\Gamma = -1.0$, which are listed in Table 2.

The resulting values of f_2^0, α_1, and α_2 are listed under their respective categories in Table 4. *In Topo I buffer*, we find $\alpha_2/\alpha_1 = 1.97$ for the best-fit non-cooperative model, and $\alpha_2/\alpha_1 = 2.11$ for the corresponding cooperative model with $\Delta\Gamma = -1.0$.

It is also desired to estimate α_1 and α_2 in the TE buffer at 37 °C for purposes of comparison with previously measured torsion elastic constants. If the 1□ 2 equilibrium is not affected by

the Mg^{2+} in the Topo I buffer, as assumed here, but only by the water activity, then the same values of the parameters (J, $\Delta\Gamma$ and K_0) should also apply *in the TE buffer* at 37 °C. Hence, the f_2^0-values at 37 °C *in the TE buffer* are reckoned from the known values of $-\ln a_w$ via equation (18), and then α_1, and α_2 are reckoned as before. The resulting values of f_2^0, α_1, and α_2 for the best-fit non-cooperative and corresponding cooperative models *in TE buffer* are also listed in Table 4. The estimated torsion elastic constant of the 2-state in TE buffer, namely α_2 = (10.70–11.40) $\times 10^{-12}$ erg, lies in the range, (9.0–12.0)$\times 10^{-12}$ erg, that has been observed for sufficiently small circular DNAs (N ≤ 250 bp) and for DNAs under tensile stress in similar buffers.[24]

The bending elastic constants in the Topo I buffer are estimated by rearranging equation (26) to yield (for j=1 or 2),

$$\kappa_{\beta_j} = (1/B_w)(2kT)E_{T_j}\alpha_j \Big/ \left((2\pi)^2\alpha_j - 2kTE_{T_j}\right) \tag{40}$$

We consider the best-fit non-cooperative model first. Inserting E_{T_1} =979, α_1=6.08×10^{-12} erg, and B_w =0.594 yields κ_{β_1} =5.49×10^{-12} erg, and inserting E_{T_2} =887, α_2 =11.99×10^{-12} erg, and B_w =0.594 yields $\kappa_{\beta 2}$ =3.86×10^{-12} erg. The ratio, $\kappa_{\beta 2}/\kappa_{\beta 1}$ = 0.70, is independent of the value of B_w, provided that B_w is truly constant. For the corresponding cooperative model with $\Delta\Gamma$ = −1.0, similar calculations are performed using α_1=5.95×10^{-12} and α_2 =12.54×10^{-12} erg to obtain κ_{β_1} =5.56×10^{-12} erg, κ_{β_2} =3.83×10^{-12} erg, and $\kappa_{\beta_2}/\kappa_{\beta_1}$ =0.69. The *relative* decrease in bending elastic constant is virtually the same for either model.

Whenever intrinsic bends contribute negligibly to the mean squared curvature, the effective persistence length can be related to the effective bending elastic constant according to, $P^{eff} = h\kappa_\beta/kT$, where h=3.4 ×$10^{-8}$ cm is the rise per bp, which is *assumed* to be practically the same for the 2 -state as for the 1-state. Evidence that intrinsic bends in a native DNA contribute negligibly to its P^{eff} has been presented,[51] and that finding is *assumed* to apply here to both the 1- and 2-states. The estimated persistence lengths are then P_1= 436 Å and P_2 =307 Å for the non-cooperative model *in the Topo I buffer* at 37 °C.

The transition involves a large fraction of the total sequence

Upon undergoing the transition, the change in intrinsic twist is extremely slight (~0.84 turn) and the changes in E_T and the CD spectrum are relatively modest. These features would be consistent with a scenario, wherein only a small fraction of the total sequence undergoes the transition with increasing EG in the Topo I buffer. However, such a model cannot be reconciled with the observed substantial (2.0- to 2.1-fold) *increase* in torsion elastic constant from α_1 to α_2. An estimate of the absolute minimum fraction of the sequence undergoing the transition (f_2^{tr}) can be obtained by assuming that in ~ 40 w/v% EG a fraction ($1 - f_2^{tr}$) has α_1 =6.08×10^{-12} erg, and the fraction f_2^{tr} has $\alpha_2 = \infty$, so the effective torsion elastic constant is given by $1/\alpha_{eff} = 1/11.99 \times 10^{-12} = \left(1 - f_2^{tr}\right)/\alpha_1 + f_2^{tr}/\infty$, which gives f_2^{tr}=0.493. *At least* 49 % of the sequence *must* undergo the transition! Moreover, if α_2 takes any value between ∞ and 11.99×10^{-12} erg (Table III), then f_2^{tr} must take an appropriate value between 0.493 and 1.0. If α_2 = 12.0×10^{-12} erg, corresponding to the largest observed values of the torsion elastic constant, which have been measured for both small circular DNAs with N ≤ 250 bp and large linear DNAs under tension,[24] then the fraction, f_2^{tr}=1.00, of the entire sequence must undergo the transition. Thus the absolute *minimum* fraction undergoing the transition *in any case* is

substantial, and with a reasonable estimated upper limit for α_2, the corresponding lower limit fraction undergoing the transition is 1.00. In the absence of any information to the contrary, it is assumed in this work that the entire sequence undergoes the transition.

Comparison with prior studies of DNA hydration

Structural transitions induced in dilute calf-thymus DNA at 27 °C by increasing concentrations of several salts were studied extensively by CD spectroscopy.[47,50] CD spectra of the NaCl, KCl and less concentrated NH_4Cl solutions could in each case be represented as linear combinations of two limiting spectra, which were qualitatively assigned as normal B (low to moderate salt) and normal C (high salt) spectra.[47] Although these limiting B- and C-spectra varied somewhat from one salt to another, the C-spectra all exhibited modest *negative* values in the 260–278 nm 'band', instead of the large *positive* values exhibited by the B-form, but at 245 nm there was little or no significant difference between the B- and C-spectra. The CD spectra in LiCl and CsCl could not be simply represented in terms of two components, and it was suggested that a third A-type component was also involved.[50] However, the CD spectrum inferred for the A-form is practically identical to that of the later discovered left-handed Z-helix, rather than that eventually associated with A-helix.[49] Subsequent topological studies of PM2 DNA in 3.0 M CsCl, 6.2 M LiCl, and 5.4 M NH_4Cl indicated that the decrease in CD at 275 nm was a universal function of the *increase*, $\Delta\varphi^0$, in the intrinsic succession angle above that in 0.05 M NaCl, regardless of which salt was used to induce the structural transition.[48] Moreover, the net change in $\Delta\varphi^0$ to go from the B-spectrum in 0.05 M NaCl to the C-spectra in 3.0 M CsCl, 6.2 M LiCl, or 5.4 M NH_4Cl was rather small, in the range +0.7 to +0.8 degree/bp, corresponding to a change of −0.20 to −0.23 bp/turn in the helix repeat. For this reason, the B- and C-conformations in solution are now commonly referred to as, respectively, the B(10.4) and B(10.2) conformations.

The disappearance of B-DNA was correlated with a decrease in the net preferential hydration, which occurs with increasing salt concentration and depends almost exclusively on the water activity, independent of the particular salt used.[50] Sedimentation studies indicated that increasing salt concentration induced a more compact structure, due either to greater flexibility or intrinsic curvature or to weaker interduplex repulsions.[50]

One may inquire whether the present state 2 induced by 40 w/v% EG is related in any way to the B(10.2) and/or Z- or A-type structures induced by high concentrations of various salts. Although both EG and salts reduce a_w, their effects on the CD spectrum differ significantly. Salts reduce the CD primarily at 260–278 nm, but not at 245 nm, whereas EG lowers the CD significantly in both regions, as indicated in Table 3. Also, for the same water activity, salts lower the CD at 272–275 nm considerably more than does EG. For example, a_w is approximately 0.78 in both 40 w/v% EG and 5.70 molal (m) NaCl, yet the CD in 40 w/v% EG is reduced by ~0.59-fold, whereas that in 5.70 m NaCl is reduced by ~0.42-fold.[47] Interestingly, the sigmoidal part of the change in CD_{272} associated with the 1 □ 2 transition is *positive*, although it is overwhelmed by the negatively sloped linear contribution, upon which it is superposed. The change in intrinsic twist induced by 40 w/v% EG, namely $\Delta\varphi^0 = (0.84)(360)/4932 = -0.061$ degree/bp, is *extremely* small and *negative*, in contrast to the small (but much larger) *positive* changes, $\Delta\varphi^0 = +0.07$ to +0.08 degree/bp, induced by 3.0 M CsCl, 6.2 M LiCl, and 5.4 M NH_4Cl.[48] Finally, the torsion elastic constants of native and modified sequences from pBR322 *declined* by 0.60- to 0.73-fold between 0.1 M and 5.25 M NaCl.[9] In contrast, addition of 40 w/v% EG to 0.1 M NaCl TE buffer causes a ~10.70/5.16 = 2.1-fold *increase* in torsion elastic constant. These findings indicate that the present 2-state is a distinct subconformation within the B-family with its own characteristic properties and is not obviously related to any previously reported structure induced by high concentrations of NaCl, KCl, LiCl, CsCl, or NH_4Cl. The different states induced by high salt concentration on one hand and by

40 w/v% EG on the other both exhibit smaller apparent persistence lengths than that prevailing in 0.1 M NaCl with 0 w/v% EG, but the origins of such behavior are presently unknown, and could in principle differ between the two dehydration conditions.

The dielectric constant takes the values 72.7, 67.8, and 61.7 in, respectively, 0, 20 and 40 w/v % EG. Previous studies of the effects of ethylene glycol on the melting transition[43] and of sucrose on the B □ Z transition[44] suggested that the effect of such polyols to lower the dielectric constant did not contribute significantly, compared to simple dehydration, to stabilize the final states. This circumstance is assumed to prevail in the present work.

Conclusion

Increasing the EG concentration from 0 to 40 w/v% induces a sigmoidal transition of duplex DNA to an alternative duplex state. This transition is satisfactorily modeled by a two-state, 1 □ 2, transition. Compared to the 1-state, the 2-state exhibits a ~2.0- to 2.1-fold larger torsion elastic constant, a ~0.70-fold smaller bending elastic constant, a 0.2 % lower intrinsic twist, and a somewhat lower CD near both 272 and 245 nm. Even in the absence of EG, a small, but significant, fraction (7 to 10 %) of the base-pairs is predicted to exist in the 2-state. The change in preferential interaction coefficient, $\Delta\Gamma = \Gamma_{2|w} - \Gamma_{1|w}$, is negative, which implies that the 2-state has less water and/or more EG in its neighborhood than does the 1-state. The present data lack sufficient precision to assess the cooperativity ($1/J$) of the transition, but when combined with other considerations, which argue that $-1.0 \leq \Delta\Gamma \leq 0$, they suggest that $1/J \geq 30.0$ bp, which in turn suggests an average domain size, $d \geq 31$ bp, at the midpoint of the transition.

Future studies of the effects of EG on DNAs with different sequences will likely require elaboration of the present model to incorporate sequence dependent variations of K_0, $\Delta\Gamma$, and J.

Acknowledgments

This work was supported in part by a grant R01 GM61685 from the National Institutes of Health.

References

1. Schurr JM, Delrow JJ, Fujimoto BS, Benight AS. Biopolymers 1997;44:283. [PubMed: 9591480]
2. Shibata JH, Wilcoxon J, Schurr JM V, Knauf V. Biochemistry 1984;23:1188. [PubMed: 6712943]
3. Langowski J, Benight AS, Fujimoto BS, Schurr JM, Schomburg U. Biochemistry 1985;24:4022. [PubMed: 2996588]
4. Wu PG, Song L, Clendenning JB, Fujimoto BS, Benight AS, Schurr JM. Biochemistry 1988;27:8128. [PubMed: 3233199]
5. Wu P, Schurr JM. Biopolymers 1989;28:1695. [PubMed: 2597725]
6. Song L, Fujimoto BS, Wu JM, Thomas JC, Shibata JH, Schurr JM. J Mol Biol 1990a;214:307. [PubMed: 2370668]
7. Wu PG, Fujimoto BS, Song L, Schurr JM. Biophys Chem 1991;41:217. [PubMed: 1777574]
8. Schurr, JM.; Fujimoto, BS.; Wu, P-G.; Song, L. Topics in fluorescence spectroscopy. In: Lakowicz, JR., editor. Biochemical applications. Vol. 3. Plenum Press; New York: 1992. p. 137
9. Kim US, Fujimoto BS, Furlong CE, Sundstrom JA, Humbert R, Teller DC, Schurr JM. Biopolymers 1993;33:1725. [PubMed: 8241430]
10. Clendenning JB, Schurr JM. Biopolymers 1994;34:849. [PubMed: 8054468]
11. Clendenning JB, Naimushin AN, Fujimoto BS, Stewart DW, Schurr JM. Biophys Chem 1994;52:191. [PubMed: 7999972]
12. Naimushin AN, Clendenning JB, Kim US, Song L, Fujimoto BS, Stewart DW, Schurr JM. Biophys Chem 1994;52:219. [PubMed: 7999973]

13. Heath PJ, Clendenning JB, Fujimoto BS, Schurr JM. J Mol Biol 1996;260:718. [PubMed: 8709150]
14. Delrow JJ, Heath PJ, Schurr JM. Biophys J 1997;73:2688. [PubMed: 9370462]
15. Delrow JJ, Heath PJ, Fujimoto BS, Schurr JM. Biopolymers 1998;45:503. [PubMed: 9577231]
16. Naimushin AN, Fujimoto BS, Schurr JM. Biophys J 2000;78:1498. [PubMed: 10692335]
17. Rangel DP, Sucato CA, Spink CH, Fujimoto BS, Schurr JM. Biopolymers 2004;75:291. [PubMed: 15386272]
18. Suh D, Sheardy RD, Chaires JB. Biochemistry 1991;30:8722. [PubMed: 1888733]
19. Riccelli PV, Vallone PM, Kashin I, Faldasz BD, Lane MJ, Benight AS. Biochemistry 1999;38:11197. [PubMed: 10460177]
20. Owczarzy R, Vallone PM, Goldstein RF, Benight AS. Biopolymers 1999;52:29. [PubMed: 10737861]
21. Vallone PM, Benight AS. Biochemistry 2000;39:7835. [PubMed: 10869190]
22. Qu X, Ren J, Riccelli PV, Benight AS, Chaires JB. Biochemistry 2003;42:11960. [PubMed: 14556627]
23. Mandell K, Vallone PM, Owczarzy R, Riccelli PV, Benight AS. Biopolymers 2005;82:199. [PubMed: 16345003]
24. Fujimoto BS, Brewood GP, Schurr JM. Biophys J 2006;91:4166. [PubMed: 16963514]
25. Rangel DP, Brewood GP, Fujimoto BS, Schurr JM. Biopolymers 2007;85:222. [PubMed: 17111396]
26. Chitra R, Smith PE. J Phys Chem B 2001;105:11513.
27. Schurr JM, Rangel DP, Aragon SR. Biophys J 2005;89:2258. [PubMed: 16055532]
28. White JH. Am J Math 1969;91:693.
29. Fuller FB. Proc Nat Acad Sci USA 1971;68:815. [PubMed: 5279522]
30. Depew RE, Wang JC. Proc Natl Acad Sci USA 1975;72:4275. [PubMed: 172901]
31. Pulleyblank DE, Shure M, Tang J, Vinograd J, Vosberg HP. Proc Natl Acad Sci USA 1975;72:4280. [PubMed: 1060106]
32. Horowitz D, Wang JC. J Mol Biol 1984;173:75. [PubMed: 6321743]
33. Shore D, Baldwin RL. J Mol Biol 1983;170:983. [PubMed: 6644817]
34. Clendenning JB, Naimushin AN, Fujimoto BS, Stewart DW, Schurr JM. Biophys Chem 1994;52:191. [PubMed: 7999972]
35. Gebe JA, Allison SA, Clendenning JB, Schurr JM. Biophys J 1995;68:619. [PubMed: 7696514]
36. Sucato CA, Rangel DP, Aspleaf D, Fujimoto BS, Schurr JM. Biophys J 2004;86:3079. [PubMed: 15111422]
37. Rangel DP, Fujimoto BS, Schurr JM. Biophys Chem. 2007 submitted.
38. di Mauro E, Caserta M, Negri R, Carnevalli F. J Biol Chem 1985;260:152. [PubMed: 2981201]
39. Flick, E. Industrial Solvents Handbook. William Andrew Publishing; Noyes: 1998.
40. Weast, RC., editor. CRC Handbook of Chemistry and Physics. Boca Raton, Florida: CRC Press, Inc; 1980.
41. Brewood, GP. PhD Thesis. University of Washington; 2006.
42. Stanley C, Ran DC. Biophys J 2006;91:912. [PubMed: 16714350]
43. Spink CH, Chaires JB. Biochemistry 1999;38:496. [PubMed: 9890933]
44. Preisler RS, Chen HH, Colombo MF, Choe Y, Short BJ Jr, Rau DC. Biochemistry 1995;34:14400. [PubMed: 7578044]
45. Tunis-Schneider MJB, Maestre MF. J Mol Biol 1970;52:521. [PubMed: 4923749]
46. Nelson RG, Johnson WC Jr. Biochem Biophys Res Commun 1970;41:211. [PubMed: 5466495]
47. Hanlon S, Brudno S, Wu TT, Wolf B. Biochemistry 1975;14:1648. [PubMed: 1168489]
48. Baase WC, Johnson WC Jr. Nucleic Acids Res 1979;6:797. [PubMed: 424316]
49. Tinoco, I., Jr; Sauer, K.; Wang, JC.; Puglisi, JD. Physical Chemistry, Principles and Applications in the Life Sciences. Prentice-Hall; Upper Saddle River, N. J: 2002. p. 575
50. Wolf B, Hanlon S. Biochemistry 1975;14:1661. [PubMed: 1168490]
51. Vologodskaia M, Vologodski AV. J Mol Biol 2002;317:205. [PubMed: 11902837]

52. Hill, TL. An Introduction to Staistical Thermodynamics. Addison-Wesley Publishing Co., Inc; Reading, Massachusetts: 1962.

Appendix A. Effect of torsional strain on a conformational equilibrium

We consider a chain of N subunits, each of which can undergo a 1□ 2 transition. Both the equilibrium twist, φ_{kj}^0 (rad), and the torsion elastic constant, α_{kj} (erg), of the kth torsion spring *kj* between the kth and (k+1)th subunits depend upon the state, j (j=1 or 2), of the kth subunit. In the absence of twisting torques, the intrinsic twist of the chain (turns) is given by,

$$l_0=\left[\left(N-n_2^0\right)\varphi_1^0+n_2^0\varphi_2^0\right]/2\pi \tag{A1}$$

where φ_1^0 and φ_2^0 are the intrinsic twists (radians) of the springs in states 1 and 2, respectively, and n_2^0 is the number of subunits of the *unperturbed* chain that are in the 2-state. In the presence of a net torsional strain, the only stable configuration of the chain is the minimum energy configuration, wherein the net torque on every subunit vanishes, which requires that all 1-springs exhibit the same displacement, $\delta\varphi_1$, and that all 2-springs exhibit the same displacement, $\delta\varphi_2$, and furthermore that $\delta\varphi_1 = (\alpha_2/\alpha_1)\,\delta\varphi_2$. In this case, the net twist (turns) can be written as,

$$t=\left[(N-n_2)\left(\varphi_1^0+\delta\varphi_1\right)+n_2\left(\varphi_2^0+\delta\varphi_2\right)\right]/2\pi \tag{A2}$$

where n_2 is the number of springs in state 2. In general, $n_2 \neq n_2^0$, because the imposed torsional strain shifts the 1□ 2 equilibrium, although this effect may be extremely slight in some cases. The relevant question is whether the net twisting strain suffices to *significantly* shift the 1□ 2 equilibrium. Combining equations (A1) and (A2) yields the net twisting strain (radians),

$$\begin{aligned}2\pi(\Delta t)=2\pi(t-l_0)=&\left(n_2-n_2^0\right)\Delta\varphi^0+((N\\&-n_2)\alpha_2\\&/\alpha_1+n_2)\delta\varphi_2\end{aligned} \tag{A3}$$

where $\Delta\varphi^0 \equiv \varphi_2^0 - \varphi_1^0$ is the difference in intrinsic twist per spring between the 2 and 1 states. Equation (A3) can be rearranged to yield

$$\begin{aligned}\delta\varphi_2=&\left(2\pi\Delta t-\left(n_2-n_2^0\right)\Delta\varphi^0\right)\\&/((N-n_2)\alpha_2\\&/\alpha_1+n_2)\end{aligned} \tag{A4}$$

The total (minimum) strain energy arising from Δt is (with $\delta\varphi_1 = (\alpha_2/\alpha_1)\,\delta\varphi_2$),

$$U(n_2,\Delta t)=n_2(\alpha_2/2)\delta\varphi_2{}^2+(N-n_2)(\alpha_2/\alpha_1)(\alpha_2/2)\delta\varphi_2{}^2$$
$$=(1/2)\left\{\frac{1}{(N-n_2)(1/\alpha_1)+n_2/\alpha 2}\right\}\left(2\pi\Delta t-\left(n_2-n_2^0\right)\left(\Delta\varphi^0\right)\right)^2 \quad \text{(A5)}$$

The second line in equation (A5) follows from the first by using equation (A4) for $\delta\varphi_2$.

In order to assess the effect of Δt on the equilibrium value of n_2, we must incorporate $U(n_2, \Delta t)$ into the formalism for reckoning the mean value, $\bar{n}_2$. Specifically, we begin with the conformational partition function χ defined by equation (12) in the text. In the presence of net twist, Δt, this would have to be modified to,

$$\chi \equiv \sum_{n_2\geq 0} B^{n_2} e^{-U(n_2,\Delta t)/kT} \sum_{I(n_2)} J^{n_J} \quad \text{(A6)}$$

where $U(n_2, \Delta t)$ has been explicitly incorporated. In fact, each kind of spring exhibits not only its minimum energy value for a given net twist, but also thermal fluctuations about that value.[35] In a circular DNA with a fixed writhe and a given net twist, t, the fluctuations of individual springs are not entirely independent, but are constrained to sum to zero. The leading contribution of this constraint to the free energy *per spring* is ~ $(kT/2N)\ln N$,[4] which is negligibly small for $N \geq 1000$ bp, and is ignored here. Hence, these fluctuations in the torsion springs may be regarded as effectively independent in sufficiently large circles. The contribution of thermal fluctuations in twist of each spring to the configuration integral and free energy of its associated subunit, relative to that of a 1-spring, is assumed to be already included in B, so only the minimum energy arising from the net twist, $U(n_2, \Delta t)$, must be incorporated into χ. Because $U(n_2, \Delta t)$ is not linear in n_2, the matrix method cannot be employed directly. However, because $U(n_2, \Delta t)$ is independent of the arrangement of the n_2 2-states among the subunits, the configuration sum,

$$W(n_2) \equiv \sum_{I(n_2)} J^{n_J} \quad \text{(A7)}$$

is independent of, and unaffected by, $U(n_2, \Delta t)$, and must take the same value for a given n_2, regardless of the value of Δt or $U(n_2, \Delta t)$.

For a sufficiently long chain, it is permissible to replace $\ln \chi$ by ln(max term), where maxterm is the largest term in the sum.[52] Thus,

$$\ln\chi=\ln\left[B^{n_2}e^{-U(n_2,\Delta t)/kT}W(n_2)\right]_{\max} \quad \text{(A8)}$$

where the subscript indicates the maximum value, which is assumed to occur, when n_2 takes its mean value in the presence of Δt. From equation (13) we also have for the average value,

$$\bar{n}_2=\left(1/c^{tot}\right)\sum_{n_2\geq 0} n_2 c_{n_2}=B(\partial\ln\chi/\partial B) \quad \text{(A9)}$$

which remains valid, despite the incorporation of $\exp[-U(n_2, \Delta t)/kT]$ into χ. The condition for the maximum term in equation (A8) is,

$$(\partial/\partial n_2)(n_2 \ln B - U(n_2, \Delta t)/kT + \ln W(n_2)) = 0 \tag{A10}$$

which becomes

$$\ln B - (\partial U(n_2, \Delta t)/\partial n_2)/kT + \partial \ln W(n_2)/\partial n_2 = 0 \tag{A11}$$

The solution of equation (A11) for n_2 gives the mean value, $\bar{n}_2$.

The last term in equation (A11) can be evaluated by considering first the case, when $U(n_2, \Delta t) = 0$ for all n_2, so the second term can be omitted. Then after exponentiating both sides and multiplying by $f_2 = n_2/N$, we obtain f_2 as the solution of

$$f_2 = BF(n_2) \tag{A12}$$

where

$$F(n_2) \equiv f_2 \exp[\partial \ln W(n_2)/\partial n_2] \tag{A13}$$

It is clear from equation (A7) that $W(n_2)$ depends only upon n_2 and J, but not upon B or $U(n_2, \Delta t)$, hence $F(n_2)$ also depends only upon n_2 and J.

When $U(n_2, \Delta t) = 0$, the quantity $F(n_2)$ can be found in the following way. The characteristic equation for the transfer matrix, **M**, in equation (16) of the main text is

$$\lambda^2 - \lambda(B+1) + B\left(1 - J^2\right) = 0 \tag{A14}$$

Implicit differentiation leads to

$$\partial\lambda/\partial B = \left(\lambda + J^2 - 1\right)/(2\lambda - (B+1)) \tag{A15}$$

By using the relation, $f_2 = (B/\lambda)\partial\lambda/\partial B$, which follows from equations (13) and (15) in the main text, B can be expressed in terms of λ and f_2 as,

$$B = \lambda f_2(2\lambda - 1)/\left(\lambda(1+f_2) + J^2 - 1\right) \tag{A16}$$

By substituting (A16) into (A14), it is possible to eliminate B, and express λ in terms of f_2 and J as,

$$\lambda=\left\{\left(1-J^2\right)(1-2f_2)+1+\left[\left(\left(1-J^2\right)(1-2f_2)+1\right)^2-4\left(1-J^2\right)(1-f_2)^2\right]^{1/2}\right\} /(2(1-f_2)) \tag{A17}$$

Solving (A16) for f_2 yields,

$$f_2=B\left(\lambda(1+f_2)+J^2-1\right)/(\lambda(2\lambda-1)) \tag{A18}$$

Comparison of (A18) with (A12) implies that

$$F(n_2)=\left(\lambda(1+f_2)+J^2-1\right)/(\lambda(2\lambda-1)) \tag{A19}$$

where λ is given in terms of f_2 and J by (A17). Thus, in the present large N limit, $F(n_2)$ is a function only of $f_2 = n_2/N$ and J. Although $F(n_2)$ in equation (A19) was derived for the case $U(n_2, \Delta t) = 0$, it must also apply for arbitrary $U(n_2, \Delta t)$, as noted above.

When both sides of (A11) are exponentiated and multiplied by f_2, and use is made of (A13), we obtain finally,

$$f_2=B\cdot\exp\left[-(\partial U(n_2,\Delta t)/\partial n_2)/kT\right]\cdot F(n_2) \tag{A20}$$

From equation (A5), we find,

$$(\partial U(n_2,\Delta t)/\partial n_2)= (1/2)\left\{\frac{-(1/\alpha_2-1/\alpha_1)\left(2\pi\Delta t-\left(n_2-n_2^0\right)\Delta\varphi^0\right)^2}{\left[(N-n_2)/\alpha_1+n_2/\alpha_2\right]^2}-\frac{2\Delta\varphi^0\left(\left(2\pi\Delta t-\left(n_2-n_2^0\right)\Delta\varphi^0\right)\right)}{\left[(N-n_2)/\alpha_1+n_2/\alpha_2\right]}\right\} \tag{A21}$$

When equation (A19) for $F(n_2)$, (A17) for λ, and (A21) are substituted into equation (A20), there results a transcendental equation, which can be solved numerically for $n_2 = \bar{n}_2$. The n_2^0 in (A21) is the corresponding solution of equation (A12), when $U(n_2, \Delta t) = 0$. Note that, when $\Delta t = 0$, both $U(n_2, \Delta t)$ and $\partial U(n_2, \Delta t)/\partial n_2$ vanish for $n_2=n_2^0$, as expected.

The effect of the twisting strain present in a supercoiled DNA to shift n_2 away from n_2^0 can be estimated in the following way. The linking difference, $\Delta l = l - l_0$, is distributed between net twist, $\Delta t = t - l_0$, and writhe, w (turns), according to $\Delta l = \Delta t + w$. Under standard conditions, Δt accounts for about 1/3 of Δl and w accounts for about 2/3 of Δl. The most prominent topoisomers in a thermally equilibrated population lie in the range $\Delta l = -2$ to $+2$. Under standard conditions (with $\Delta l = +2$), one would expect $\Delta t \approx 2/3$ turn, whence $2\pi\Delta t = 4\pi/3$ radians. The spring constants of the 1- and 2-states in Topo I buffer are estimated to be $\alpha_1 = 6.08\times10^{-12}$ erg and $\alpha_2 = 11.99\times10^{-12}$ erg, respectively (c.f. Table 3). It is also found 0 in the present study

that $(l_{20} - l_{10}) \cong -0.84$ turns, so $\Delta\varphi^0 = 2\pi(l_{20} - l_{10})/4932 = -0.0011$ radians/bp. We assume that B=1.0, so that, when $\Delta t = 0$, the solution is $n_2=n_2^0=N/2$, as is evident from equation (18) in the main text. We consider this choice (B=1.0), because many properties, such as the effective elastic constant, are maximally sensitive to perturbation of the chemical equilibrium at its midpoint. We now consider a small non-vanishing value of Δt, and imagine that equation (A20) is to be solved by iteration, beginning with the choice, $n_2=n_2^0=N/2$ on the right-hand side. With this choice, then, on the right-hand side one has, $f_2 = 1/2$, λ =1+ J (from (A17)), and $F(n_2)$ = 1/2 (from (A19), so that (A20) now reads

$$f_2=(1/2)\exp\,[\,-(\partial u(n_2,\Delta t)/\partial n_2)/kT\,] \tag{A22}$$

where the exponential factor is to be evaluated at $n_2=n_2^0=N/2$. Numerical evaluation of the exponent yields,

$$f_2=(1/2)\exp[\,-0.00022\,] \tag{A23}$$

The second term in braces in equation (A21) yields a positive contribution that is about 4.0 times that of the first term for Δl = +2. The right-hand side of equation (A23) differs negligibly from the input value, $f_2 = 1/2$, which indicates that the populations are not significantly affected by the net twisting strain, Δt, in this example. If the effective equilibrium constant, B, were increased from 1.0 to exp[+0.00022], then equation (A20) would be precisely satisfied by f_2 = 1/2. Thus, the effect of the twisting strain in a topoisomer with Δl = +2, when the transition is near its midpoint, is equivalent to a 1.00022-fold decrease in the effective equilibrium constant, B, for the 1 □ 2 transition. The positive twisting strain evidently favors very slightly the 1-state, primarily because of its greater intrinsic twist, and to a lesser extent because of its smaller torsion elastic constant.

We have no information regarding the difference in curvature between states 2 and 1. Nevertheless, the net bending strain required to close the 4932 bp DNA into a circle and add 4/3 of a writhe turn is also extremely slight, and is believed not to significantly shift f_2 away from f_2^0.

Because neither the twisting nor bending strain significantly shifts the 1 □ 2 equilibrium, no significant error is encountered in the present study by assuming that $f_2=f_2^0$.

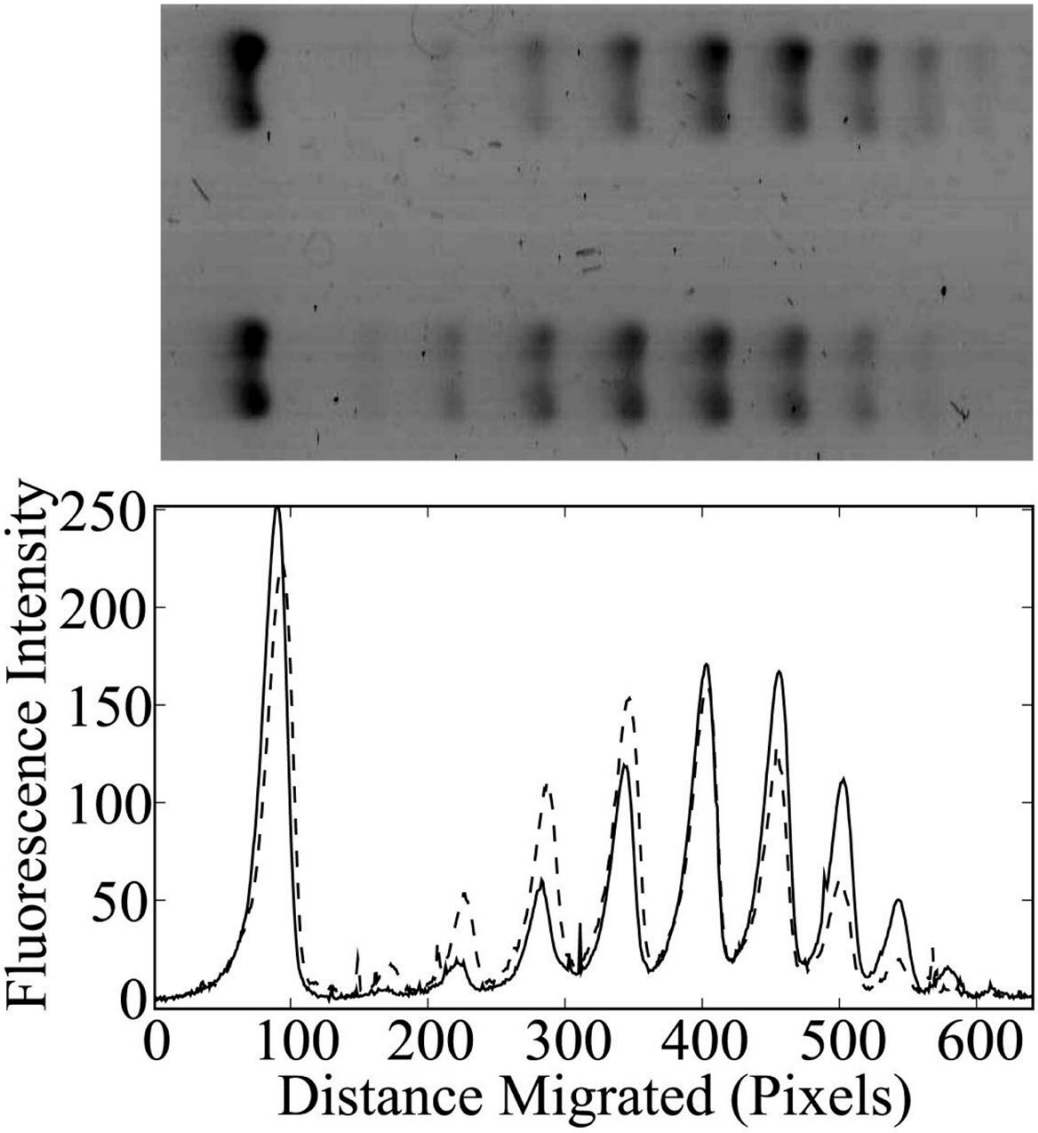

Figure 1.

(Top) Fluorescence scanner image of two adjacent lanes of gel, wherein DNA migrated from left to right. Upper and lower lanes contain p30δ DNAs that were relaxed at 37 °C by Topo I in buffers containing, respectively, 0 and 40 w/v% EG. (Bottom) Lane traces showing the band intensities as a function of position (pixels) along the lane. The solid and dashed curves apply for, respectively, 0 and 40 w/v% EG. The bands at ~80 pixels represent nicked circular DNA, and all other bands represent topoisomers of covalently closed circular DNAs with small positive linking differences *under conditions prevailing in the gel*. Those farther to the right have a more positive linking difference. Note that the topoisomer distribution obtained in 40 w/v% EG is shifted leftward by nearly one turn from that obtained in 0 w/v% EG. Hence, the intrinsic twist in 40 w/v% EG is less than that in 0 w/v% EG by about one turn.

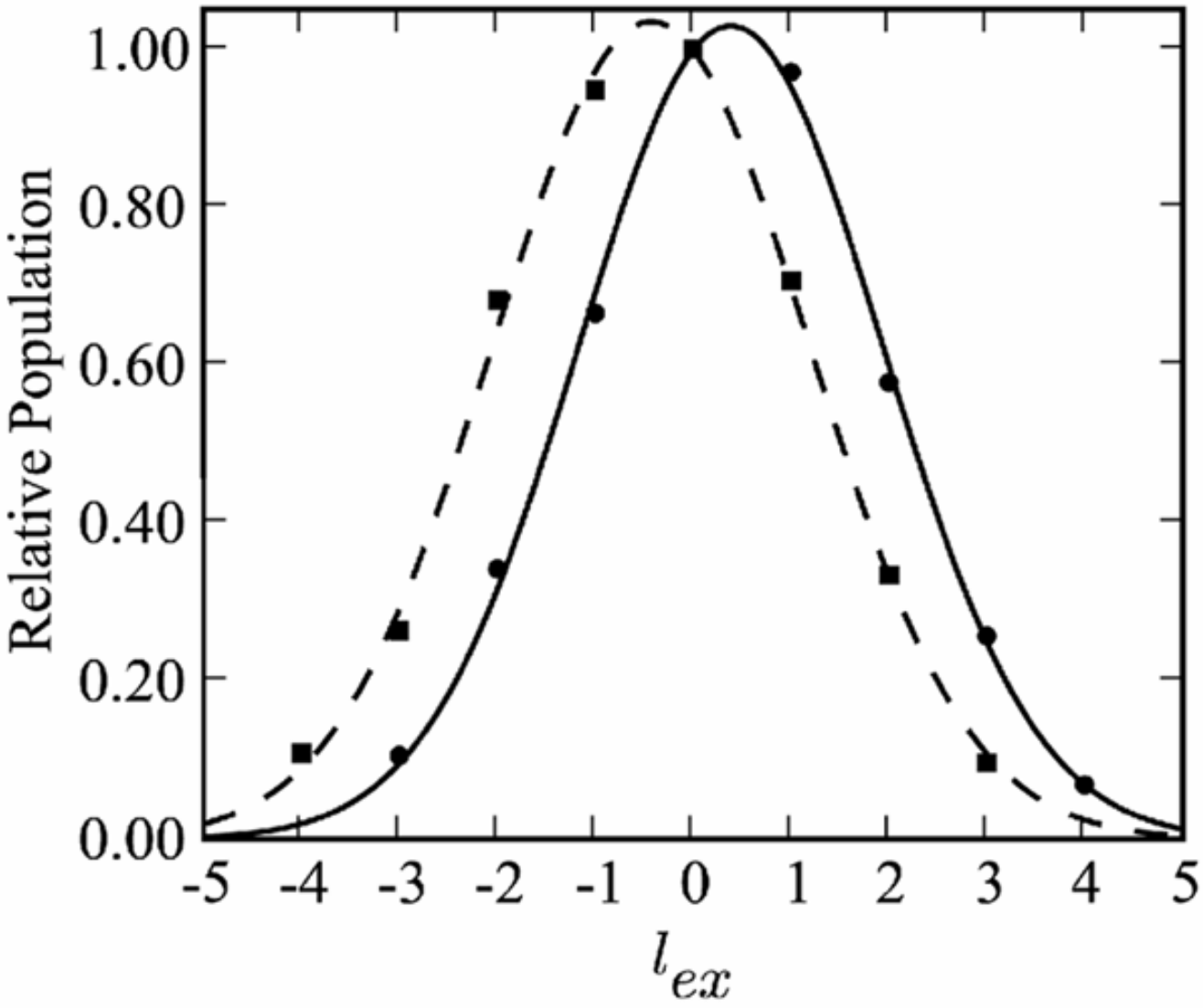

Figure 2.

Relative topoisomer population vs. excess linking number, l_{ex}. The relative population of a given topoisomer is the ratio of its integrated fluorescence intensity (area under the corresponding peak in Figure 1) to that of the most abundant topoisomer in the same lane. The solid circles correspond to 0 w/v% EG, and the solid squares to 40 w/v% EG. The solid and dashed curves represent best fits of equation (33) to the data in each case, with E_T and Δl_0 as adjustable parameters. The best-fit parameters are collected in Table 1.

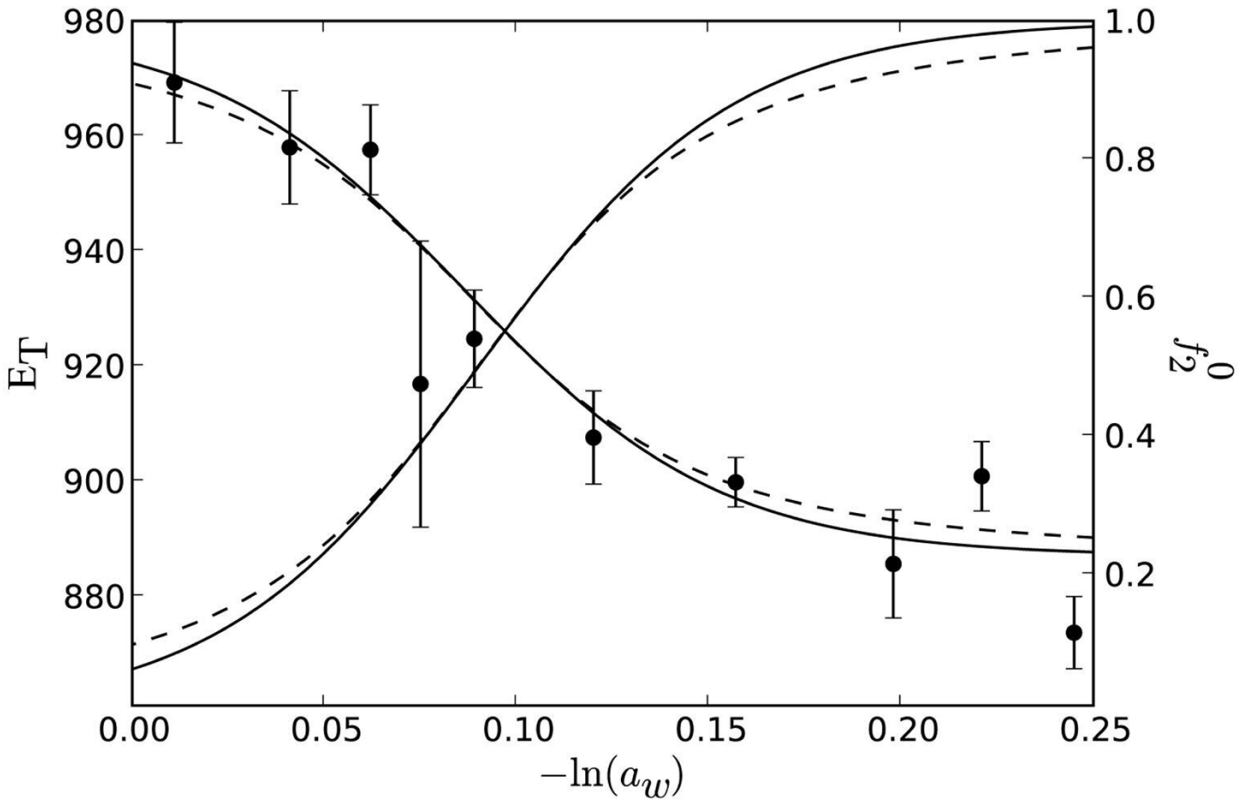

Figure 3.

Left- hand ordinate: E_T vs. −ln (a_w) for p30δ DNAs with small linking differences at 37 °C in buffers containing various concentrations of ethylene glycol. Solid circles are the present measurements on a single sample of p30δ DNA. The *solid* curve is a fit of equations (29) and (18) with *assumed* $J = 1.0$ (*non-cooperative* model) to the experimental data. The adjustable parameters are E_{T_1}, E_{T_2}, ΔΓ, and K_0, and the optimum values are presented in Table 2. The *dashed* curve is the *prediction* of a *corresponding cooperative* model with an *assumed* ΔΓ= −1.0 and the same values of E_{T_1}, E_{T_2}, midpoint value, $(-\ln a_w)_{1/2}$, and midpoint slope $(\partial E_T/\partial \ln a_w)_{1/2}$ (c.f. Table 2), as were found for the non-cooperative model. The values of all parameters in equations (18) and (29) for this corresponding cooperative model are presented in Table *Right-hand ordinate*: f_2^0 vs. −ln (a_w). The solid curve is calculated from equation (18) using the best-fit parameters obtained for the non-cooperative model ($J = 1.0$) from Table 2. The dashed curve is calculated from equation (18) using the implied parameters of the corresponding cooperative model (ΔΓ= −1.0) from Table 2.

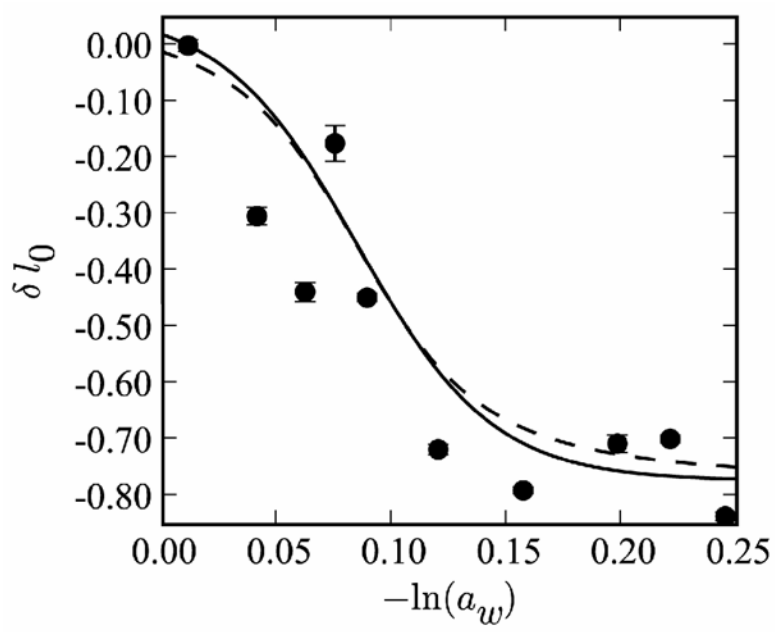

Figure 4.

δl_0 vs. −ln (a_w) for p30δ DNAs with small linking differences at 37 °C in buffers containing various concentrations of ethylene glycol. Solid circles are the experimental values of δl_0 (the change in intrinsic twist from its value in 0 w/v% EG), and are determined as described in the main text. The solid and dashed curves are predicted for, respectively, the non-cooperative (J = 1.0) and corresponding cooperative ($\Delta\Gamma$= −1.0) models using the appropriate f_2^0 values in Figure 3 and a suitable estimate of the difference in intrinsic twist between the 1 and 2 states, namely $l_{02} - l_{01}$= −0.836. The experimental data are evidently compatible with the f_2^0-values in Figure 4, which were determined by fitting equations (18) and (29) to the E_T vs. −ln (a_w) data in Figure 3.

Table 1

Water activities and experimental results for different w/v% EG in the Topo I buffer.

w/v%	a_w	$-\ln(a_w)$	Δl_0(turn)[a]	Δ(NINT)	$\Delta\Delta l_0$[b]	δl_0(turn)[c]	E_T[d]	σ_{ET}[e]
0	0.9891	0.011	−0.416	0	0	0	969	10.5
10	0.9598	0.041	−0.113	0	0.303	−0.303	958	9.9
15	0.9399	0.062	0.0212	0	0.438	−0.438	958	7.8
17.5	0.9277	0.075	−0.242	0	0.174	−0.174	917	24.9
20	0.9148	0.089	−0.0134	0	0.448	−0.448	925	8.5
25	0.8690	0.120	0.301	0	0.717	−0.717	908	8.1
30	0.8547	0.157	0.373	0	0.790	−0.790	900	4.3
35	0.8204	0.198	0.150	0	0.707	−0.707	886	9.4
37.5	0.8017	0.221	0.282	−1	−0.302	−0.698	901	6.0
40	0.7827	0.245	0.419	−1	−0.164	−0.836	874	6.3

[a] Δl_0 is the linking difference, $l - l_0$, of the most populous topoisomer.

[b] $\Delta\Delta l_0$ is the difference in Δl_0 between DNA with and without osmolyte, given by Eq. (34).

[c] δl_0 is the difference in l_0 between DNA with and without osmolyte, given by Eq. (35).

[d] E_T is twist energy parameter in Eqs. (25), (29), and (33).

[e] The σ_{E_T} is the standard deviation *of the mean* of E_T.

Table 2

Comparison of best-fit parameters of the non-cooperative model with the implied parameters of a corresponding cooperative model.

Model	E_{T_1}[a]	E_{T_2}[b]	J	ΔΓ	K_0	$-(\ln a_w)_{1/2}$	$\left(\partial E_T^{eff} / \partial \ln a_w\right)_{1/2}$
non-cooperative	979	887	1.0[c]	−30.0	0.067	0.0901	688
cooperative[a]	979	887	1/30.0	−1.0[d]	0.914	0.0901[e]	688[e]

[a] E_T-value of pure 1-state.

[b] E_T-value of pure 2-state.

[c] Assumed value of J for the non-cooperative model.

[d] Assumed value of ΔΓ for this particular corresponding model.

[e] Assumed value required to qualify as a corresponding model.

Table 3

Absolute, relative, and best-fit theoretical values of the CD.

	y=(w/v% EG)/100.0				
	0.0	**0.1**	**0.2**	**0.3**	**0.4**
CD_{272}	2.61	2.17	2.02	1.81	1.50
CD_{245}	−2.70	−3.24	−3.21	−3.33	−3.49
$S(y)^a$	1.0	0.83	0.77	0.69	0.57
$S(y)^{th}$ [b]	0.99	0.845	0.76	0.70	0.56

[a] $S(y) = CD_{272}(y)/CD_{272}(0)$ is the relative CD_{272} at y divided by that at $y = 0$.

[b] $S(y)^{th}$ is computed from equation (38) using the best-fit values, a = 1.7479, b = 0.9639, d = 0.2983 determined by a least-squares fit of equation (38) to the $S(y)$ data.

Table 4

Measured and estimated values of the torsion elastic constant in different buffers.[a]

					Non Cooperative			Cooperative		
w/v% EG	buffer	$-\ln a_w$	T	$\alpha^{eff}\times10^{12}$	f_2^0	$\alpha_1\times10^{12}$	$\alpha_2\times10^{12}$	f_2^0	$\alpha_1\times10^{12}$	$\alpha_2\times10^{12}$
0	TE	0.004	40	4.99	_[b]	-	-	_[b]	-	-
20	TE	0.0829	40	6.32	_[b]	-	-	_[b]	-	-
0	TE	0.004	37	5.16	0.070	4.97	10.70	0.105	4.85	11.40
20	TE	0.0829	37	6.53	0.446	4.97	10.70	0.448	4.85	11.40
0	Topo I	0.011	37	6.35	0.085	6.08	11.99	0.118	5.95	12.54
20	Topo I	0.089	37	8.03	0.492	6.08	11.99	0.493	5.95	12.54

[a] Values in TE buffer at 40°C were measured. All other values were derived from these values by applying various corrections as described in the text.

[b] f_2^0-values at 40 °C cannot be estimated without knowledge of the temperature dependence of K_0, $\Delta\Gamma$, and J.