ELSEVIER

Biophysical Chemistry 52 (1994) 227–249

Biophysical Chemistry

A model for the binding of *E. coli* single-strand binding protein to supercoiled DNA

James B. Clendenning, J. Michael Schurr

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

Received 13 December 1993; accepted in revised form 2 March 1994

Abstract

A model is proposed for the binding of *E. coli* single strand binding protein (SSB) to supercoiled DNA. The basic tetrameric binding units of SSB are assumed to bind in pairs to the complementary single strands of a locally melted region. The cooperativity of the binding includes contributions from both protein–protein and base-pair stacking interactions. Each bound SSB tetramer is assumed to unwind $l = 34$ bp, which implies an unwinding angle of 3.27 turns. The resulting loss of superhelical strain is the essential driving force for binding SSB to supercoiled DNAs. All molecular parameters entering into the theory are estimated from available data, except for the composite binding constant (K_a), which is adjusted to best-fit the theory to the fluorescence quenching (FQ) and diffusion coefficient (D_0) data of Langowski et al. Very good fits are obtained with optimum values of K_a that are consistent with estimates from other data. This binding model predicts several noteworthy features. (1) SSB binds essentially always in a single contiguous stack on a supercoiled plasmid, and relative fluctuations in stack length are quite small, in agreement with results of electron microscopy studies. (2) The progressive loss of superhelical strain with increasing bound ligand decreases the affinity of the DNA for SSB. This anti-cooperativity offsets the cooperativity of the binding and causes apparent saturation of the binding at rather low binding ratios. Consequently, over the limited span of the measurements, the FQ data can also be satisfactorily fitted by a non-cooperative model comprising a small number of independent sites. (3) When SSB binds to a population of different topoisomers, the distribution of linking differences of the resulting complexes is extremely narrow. Thus, SSB acts to level any differences in superhelical strain in a population of topoisomers. Finally, the effects of restricting binding to a region comprising only part of the plasmid are assessed.

Keywords: E. coli; Supercoiled DNA; Binding

1. Introduction

E. coli single-strand binding protein (SSB) is a well-known helix-destabilizing protein that is able to denature linear double-strand DNA at low concentrations ($\leqslant 10$ mM) of 1:1 salt [1]. However, at NaCl concentrations of 50 mM or greater, SSB binds tightly and cooperatively to single-strand DNA, but shows no detectable binding to double-strand linear or nicked circular DNA [1–4]. SSB exists in solution as a stable tetramer of subunits with monomer molecular weight, $M_r = 18873$ [5]. Indeed, the tetramer is so stable that studies of wild-type monomer binding to single strand DNA are currently not feasible [6]. In complexes with homopolydeoxyribonucleotides,

SSDI 0301-4622(94)00036-J

SSB exhibits multiple binding modes that differ in binding site size (bases/SSB tetramer), cooperativity, and appearance in the electron microscope [7–10]. SSB performs important functions in DNA replication, repair, and recombination [8,11–14].

SSB binds to supercoiled DNAs at moderate salt concentrations, where binding to linear and nicked circular DNAs is not observed [4,15]. The binding of SSB to supercoiled DNAs exhibits several interesting features, for which no even qualitatively satisfactory theory has been presented.

(1) Binding of SSB to supercoiled pcDm506 was studied by crosslinking the bound protein with glutaraldehyde and examining the products by electron microscopy (EM) [4]. The SSB–DNA complexes exhibited an open configuration with very little evidence of residual supercoiling, in contrast to the naked native plasmids, which appeared to be highly supercoiled and exhibited a rosette type of tertiary structure. Within each SSB–plasmid complex there existed only *one* 'melted' region in which the separated individual strands complexed with SSB were visible. The length of the open region was reported to be nearly the same in all of the plasmids examined. Crosslinked SSB–pcDm506 complexes were also linearized by EcoRI before EM. Evidence for preferential binding to two different regions was discussed, although those preferences were evidently not absolute. These results, if not artifactual, imply three conclusions.

(i) SSB binds to supercoiled DNA in a highly cooperative manner *in a single stack at only one site*, though not necessarily at the same site on every DNA. Moreover, the dispersion of the stack lengths is quite small.

(ii) A substantial fraction of the superhelical turns are removed by SSB binding.

(iii) Certain regions of a supercoiled molecule may bind SSB more readily than others.

(2) The binding of SSB to supercoiled pBR322 DNA was studied by 'reverse titration' of the SSB fluorescence, which is quenched upon binding to DNA [15]. The experimental conditions were 50 mM NaCl, 20 mM phosphate, pH 7.4, at 40°C. The fluorescence quenching (FQ) data were fitted to several simple models, of which an *independent site* model with 10 independent sites and an intrinsic binding constant $K_0 = 2.6 \times 10^7\ M^{-1}$ gave the best fit [15]. This observation apparently contradicts the highly cooperative single-stack binding that is evident in EM. It was suggested that the cooperative binding of SSB might be significantly offset by the anticooperativity associated with the reduction in superhelical strain, thereby yielding apparent 'agreement' with an independent site model with a modest number of sites over the limited span of the data [15]. A primary objective here is to propose a specific model of SSB binding and to demonstrate quantitatively that the apparent contradiction can be resolved in the suggested manner.

(3) Dynamic light scattering (DLS) curves of D_{app} versus K^2 were measured for solutions containing the same DNA concentration, but with different numbers (v) of added SSB per plasmid [15]. The experimental conditions were 0.1 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.2, at 21°C. D_0 (the average value of D_{app} at $\theta = 30°$, 40°, 50° for $\lambda = 632.8$ nm) decreased sharply in the region $0 \leqslant v \leqslant 4$ and appeared to saturate at $v = 8$–10. The observed 18% decrease in D_0 far exceeds any possible contribution of SSB to the hydrodynamic drag, and was assigned to a global change in the tertiary structure of the plasmid. D_{plat} (the average value of D_{app} at $\theta = 100°$, 110°, and 120° at $\lambda = 351.1$ nm) showed corresponding changes in the opposite direction, and leveled off at unprecedentedly high values. Co-plots of D_{app} versus K^2 for $v = 0$, 3, and 10 displayed an apparent isosbestic point. From these observations, the following inferences were drawn.

(i) A change in tertiary structure occurs when SSB binds, and at large v the final structure is sufficiently unwound or reorganized that internal interference acts to significantly enhance D_{plat} [16].

(ii) The saturation behavior indicates that either the effective SSB binding constant decreases rapidly with increasing v, or that essentially no further changes in tertiary structure that significantly affect D_0 or D_{plat} occur beyond $v = 8$–10, or both.

(iii) The apparent isosbestic point suggests that

only two main kinds of hydrodynamic species exist in solution, namely free plasmids and complexes, and that either the binding itself or its effect on the tertiary structure is largely all-or-none.

(4) Time-resolved fluorescence polarization anisotropy (FPA) measurements of intercalated ethidium revealed that SSB binding caused a decrease in the torsion constant by a factor of 0.6 at $v = 15$, although it remained essentially uniform [15]. Experimental conditions were the same as in 3 above.

(iv) This observation implies that by $v = 15$ an extensive change in secondary structure has occurred to an alternate form with substantially reduced torsion constant. The observed uniformity argues that this decrease does *not* arise from several separate rigidity weaknesses at well dispersed SSB binding sites. However, if the SSB is bound in a single stack as implied by the EM studies, then no rigidity weakness, however small, at that point could significantly affect the overall torsional rigidity of so long a DNA. Hence, this decrease must be ascribed to a change in secondary structure of pBR322 that extends far beyond the SSB binding site. Similar decreases in the torsion constants of supercoiled pBR322 DNAs relaxed by ethidium or chloroquine were also observed under conditions where the torsion constants of the corresponding linearized species were unaffected by ethidium or chloroquine binding [17,18].

Any *complete* theory must account for conclusions (i)–(iv) above. However, our objectives here are more limited, namely to develop a model for SSB binding based on the standard model of supercoiled DNAs, in which the secondary structure remains simply strained B-helix except for a melted region at the site of SSB binding, and to examine its predictions in regard to the fluorescence binding curve, the curve of D_0 versus v, the cooperativity of the binding as manifested in the distribution of bound SSB, the apparent saturation of the binding, and finally the distribution of effective linking differences in the region $v = 0$ to $v = 10$. We also endeavor to understand the largely all-or-none effect of SSB binding on the DLS experiments. In this work we make no attempt to account for the particular tertiary structure responsible for the anomalously large value of D_{plat} or for the alternate secondary structure responsible for the reduced torsion constant, or for any effect of such structures on the twist energy parameter (E_T) that governs the free energy of supercoiling. Large uncertainties in the experimental input parameters preclude a precise quantitative prediction with no adjustable parameters. Instead we seek a qualitative and semi-quantitative understanding of the nature of the binding predicted by this simplest model.

The outline of the paper is as follows. We first describe the basic model for SSB binding and then present the pertinent statistical mechanical theory for SSB binding to supercoiled DNAs. Estimates of the relevant input parameters are then presented. The additional assumptions and theory needed to simulate the FQ and D_0 results are presented next, and the protocols used to adjust the composite binding constant to fit each of those experimental data sets are described. The best-fit results obtained under the assumption that the population of plasmids consists of a single topoisomer are presented first. Then, best-fit results obtained under the assumption that the population of plasmids consists of a distribution of topoisomers are discussed. The best-fit composite binding constant is compared with the value estimated from available data in the literature. The nature of the binding is examined, especially in regard to the apparent saturation at comparatively low binding levels, the number of stacks in a given topoisomer, and the distribution of final effective linking differences or superhelical strains. In addition, the effects of restricting binding to a particular region of more easily melted DNA are investigated. Finally, the potential functional significance of these aspects of SSB binding is discussed.

2. Model of SSB binding to DNA

The proposed model for SSB binding to DNA is depicted in Figs. 1 and 2. Strictly for convenience we imagine that the binding occurs in several stages.

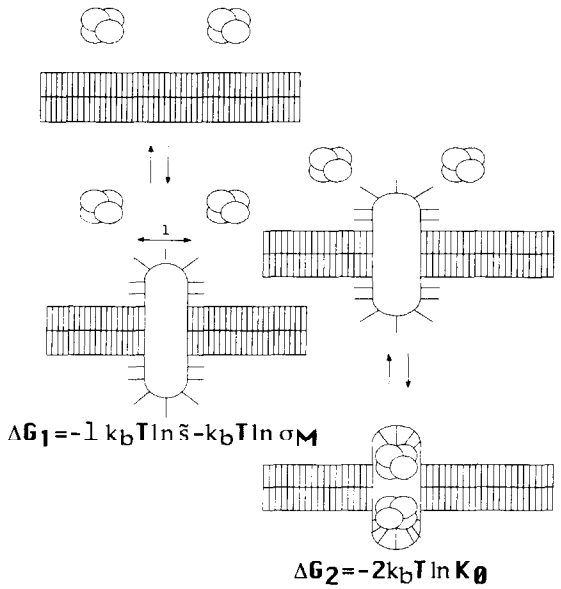

Fig. 1. Schematic illustration of the binding of a pair of SSB tetramers to a DNA molecule. SSB tetramers are required to bind in pairs to the complementary single-strands of a melted region of DNA whose minimum length is l base-pairs. The free energies for the opening and binding steps are indicated in the figure.

(1) Initially, free SSB and DNA exist separately in solution.

(2) A region of DNA containing l consecutive base-pairs melts to the open state. The standard free energy change for this step is

$$\Delta G_1 = -lk_BT \ln \tilde{s} - k_BT \ln \sigma_M, \quad (1)$$

where $\tilde{s}$ is the equilibrium constant per base-pair to form melted (or open) strands from DNA duplex, and σ_M is the cooperativity parameter for DNA melting.

(3) SSB tetramers (the basic binding units) bind in pairs, one to each complementary strand of the melted region. The standard free energy change for this step is

$$\Delta G_2 = -2k_BT \ln K_0, \quad (2)$$

where K_0 is the intrinsic equilibrium constant for binding a tetramer to one of the two complementary single-strands of the melted region.

(4) When two SSB pairs that are bound to separate melted regions of the DNA move together to form a single 'melted' region containing four SSB tetramers in a single stack, the standard free energy change is

$$\Delta G_3 = -2k_BT \ln \omega + k_BT \ln \sigma_M, \quad (3)$$

where ω is the tetramer–tetramer cooperativity parameter for binding to single-strand DNA. σ_M occurs here with a positive sign, due to the elimination of two junctions between duplex and open regions or, equivalently, to the gain of one extra base-pair stacking interaction, when the two melted regions fuse.

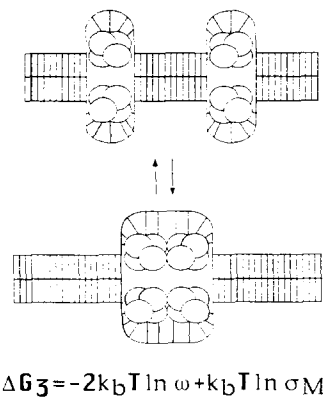

Fig. 2. Schematic illustration of cooperative interactions between two bound SSB pairs. The free energy change when non-adjacent bound pairs become nearest-neighbors arises from protein-protein interactions along the single-strands and the gain of an extra base-pair stacking interaction, as indicated formally in the figure.

We do not explicitly include the (unknown) effects of interactions between tetramers bound to the two different DNA strands. The standard free energy changes ascribed to binding and stacking of SSB–tetramer pairs can be written as

$$\Delta G_{bind} = \Delta G_1 + \Delta G_2 = -k_BT \ln K_a^2, \quad (4)$$

$$\Delta G_{stack} = \Delta G_3 = -k_BT \ln S^2, \quad (5)$$

wherein the composite constants

$$K_a = K_0 \tilde{s}^{l/2} \sigma_M^{1/2}, \quad (6)$$

$$S = \omega \sigma_M^{-1/2}, \quad (7)$$

have been introduced. K_a^2 is the effective equilibrium constant for binding a pair of SSB tetramers to relaxed duplex DNA in the absence of cooperative interactions and S^2 is the effective equilibrium constant for the coalescence of two SSB

binding regions to form a single stack of bound pairs of SSB tetramers.

This model does not admit preferential binding of SSB to specific regions of the plasmid. This possibility is considered subsequently.

The basic binding unit in our model is the SSB tetramer, which is also the prevailing species in solution. Henceforth, we adopt the convention that the symbol SSB always refers to the tetramer.

3. Theory for binding SSB to supercoiled DNA

Binding isotherms derived by the simple maximum term method are not applicable in the case of SSB binding to supercoiled DNAs for the following reasons. Such maximum term methods apply only when the average number of bound ligands is large and the uncomplexed DNAs constitute a negligible fraction of the total. The binding of SSB to duplex supercoiled DNAs is driven largely by its reduction of the prevailing superhelical strain. The large unwinding angle of a pair of SSBs (3.27 turns) causes a rapid decrease in affinity of the DNA for additional ligands, so the binding *apparently* saturates at relatively small numbers of bound SSB. Binding to any single topoisomer then exhibits substantial 'all-or-none' character, so that a significant concentration of uncomplexed topoisomers coexists with the 'fully' complexed topoisomers over most of the observed range. Thus, it is necessary to evaluate the grand partition function, as indicated below.

3.1. Concentrations of complexes and average numbers of bound SSB

The development here is closely related to previous treatments [18,19]. The major differences are that in the case of SSB binding the ligands bind in pairs, occupy many (l) consecutive lattice sites, and interact cooperatively with one another. One consequence of these differences is that the configurational degeneracy factors in the configuration sum (defined below) and the superhelical strain reduction factor in the partition function depend on the number (n_{2a}) of pairs of bound SSB, while the association factors ($c_a K_a$) and stacking factors containing S depend on the number ($n_a = 2n_{2a}$) of individual bound SSB.

Each topoisomer, denoted by $C(i)$, is regarded as a circular lattice of N identical sites with a particular linking difference in the absence of bound ligand, namely $\Delta L(i) = i - l_0$, where i is the linking number (number of turns of one strand around the other) and l_0 is the intrinsic twist. Although the linking number i is used to label the topoisomer, in fact only $\Delta L(i)$ is normally known from experiments. Since the superhelical strain free-energy depends only on $\Delta L(i)$, rather than i, the subsequent calculations of various properties of interest can be, and are, carried out using $\Delta L(i)$ without explicit knowledge of i.

Each pair of SSB molecules that binds to the lattice precludes binding to l adjacent sites. The reaction for $n_a = 2n_{2a}$ ligands, or n_{2a} *pairs* of ligands, to bind to fixed sites on complementary strands of the empty lattice $C(i)$ to form a particular complex denoted by $C_J(i, n_{2a})$, where J specifies the arrangement or configuration of bound ligands, is given by

$$C(i) + 2n_{2a}a = C_J(i, n_{2a}) \tag{8}$$

wherein a denotes an SSB molecule. The solution is assumed to be sufficiently dilute that each solute component obeys Henry's Law,

$$\mu_x = \mu_x^0 + k_B T \ln c_x \tag{9}$$

where c_x is the molarity of species x.

We assume that the difference in standard state chemical potentials for reaction (8) is equal to the sum of (1) the free energy change due to loss of superhelical strain, denoted by $\Delta A(\Delta L(i), n_{2a})$ (2) the standard free energy change upon binding n_{2a} pairs of SSB to an unstrained lattice, namely $n_{2a}\Delta G_{bind} = -n_{2a}k_B T \ln K_a^2$, and (3) the standard free energy change due to cooperative interactions between bound ligands, denoted by $p_J(n_{2a})\ \Delta G_{stack} = -p_J(n_{2a})k_B T \ln S^2$, where $p_J(n_{2a})$ is the number of nearest-neighbor contacts among the n_{2a} pairs of ligands in configuration J on the lattice. That is,

$$\begin{aligned}\mu^\circ_{C_J(i,\,n_{2a})} - \mu^\circ_{C(i)} - 2n_{2a}\mu^\circ_a \\ = \Delta A(\Delta L(i), n_{2a}) - n_{2a}k_B T \ln K_a^2 \\ - p_J(n_{2a})k_B T \ln S^2.\end{aligned} \tag{10}$$

The free energy change due to loss of superhelical strain was derived previously [18,19] and is given by

$$\Delta A(\Delta L(i), n_{2a}) = (k_B T E_T/N)\Phi^2\left(n_{2a}^2 - 2n_{2a}n_{2a}^*(i)\right), \quad (11)$$

wherein E_T is the twist energy parameter, Φ is the unwinding angle per pair of bound SSB in turns, and $n_{2a}^*(i) = -(i - l_0)/\Phi$. Eq. (11) is predicated on the assumption that the supercoiling free energy varies quadratically with linking difference from zero up to native values. Recent studies of ethidium binding in this laboratory demonstrate such behavior for p30δ DNA [20], but indicate considerably more complex behavior of pBR322 due to structural transitions induced by superhelical stress in that DNA [21]. Nevertheless, we assume ideal quadratic behavior for this calculation. A second important assumption entering into Eq. (11) is that the torsional rigidity of the complexed regions is *not* substantially *lower* than that of the duplex regions. Given the considerably enhanced mass per unit length of the complexed regions and the strong cooperative interactions among the stacked proteins therein, this is not unreasonable. Certainly, the EM pictures reveal no tendency of superhelical strain to accumulate in the complexed regions, as would be expected for a substantially softer torsion constant. If the complexed regions were significantly stiffer than normal duplex DNA, that would introduce no significant error into the results for any circumstance in which only a small fraction of the lattice is complexed, as is the case for SSB binding to supercoiled DNAs. This circumstance arises, because the *overall* bending constant, or torsion constant, for a linear sequence of springs is the reciprocal of the sum of the inverses of the individual spring constants. Thus, if the spring constants are initially essentially uniform, then a sufficiently large decrease in the spring constant for one or more individual springs could cause a substantial decrease in the overall bending constant, or torsion constant, but no increase in spring constant, however large, of *only a small fraction of the springs* could have an appreciable effect. In arriving at Eq. (11) we have also neglected several very small terms, including those arising from fluctuations in twisting and bending at constant *net* writhe (or *net* twist), fluctuations in *net* writhe (or *net* twist), and closure constraints for the circle [19].

At equilibrium for reaction (8),

$$\mu_{C(i)} + 2n_{2a}\mu_a = \mu_{C_J(i,n_{2a})} \quad (12)$$

After substituting Eq. (9) into (12), rearranging, and making use of Eq. (10), one has

$$\frac{c_{C_J(i,n_{2a})}}{c_{C(i)}} = \left[(c_a K_a)^2\right]^{n_{2a}} \times \exp\left[-\Delta A(\Delta L(i), n_{2a})/k_B T\right] \times S^{2p_J(n_{2a})}, \quad (13)$$

where c_a is the concentration of free SSB. Summing over all possible configurations, J, and then summing over all numbers $n_{2a} \geqslant 1$ of bound SSB pairs yields the ratio of the concentration of *all* complexes of species i with $n_{2a} \geqslant 1$, denoted by $c_{C(i,n_{2a}\geqslant 1)}$, to that of free $C(i)$,

$$\frac{c_{C(i,n_{2a}\geqslant 1)}}{c_{C(i)}} = \sum_{n_{2a}=1}^{N/l} \left[(c_a K_a)^2\right]^{n_{2a}} \times \exp\left[-\Delta A(\Delta L(i), n_{2a})/k_B T\right] \times \sum_J S^{2p_J(n_{2a})} = \mathrm{GPF}(i) \quad (14)$$

where the grand partition function GPF(i) is defined as the sum on the r.h.s. of the first line of Eq. (14). It can be written more compactly as

$$\mathrm{GPF}(i) = \sum_{n_{2a}=1}^{N/l} \left[(c_a K_a)^2\right]^{n_{2a}} \times \exp\left[-\Delta A(\Delta L(i), n_{2a})/k_B T\right] \times W(n_{2a}), \quad (15)$$

wherein the configuration sum, which is independent of $\Delta L(i)$ and superhelical strain energy, is defined formally by

$$W(n_{2a}) = \sum_J S^{2p_J(n_{2a})} \quad (16)$$

and given explicitly in Eqs. (21)–(25) below.

The conservation condition for topoisomer $C(i)$ is

$$c^{\mathrm{tot}}_{C(i)} = c_{C(i,n_{2a}\geqslant 1)} + c_{C(i)}, \tag{17}$$

where $c^{\mathrm{tot}}_{C(i)}$ is the total concentration of all species containing topoisomer i. $c^{\mathrm{tot}}_{C(i)}$ is given in terms of the total plasmid concentration c^{tot}_{C} by Eq. (27) below. Combining Eqs. (17) and (14) then yields the total concentration of complexes of SSB with $C(i)$,

$$c_{C(i,n_{2a}\geqslant 1)} = c^{\mathrm{tot}}_{C(i)}\mathrm{GPF}(i)/[1+\mathrm{GPF}(i)]. \tag{18}$$

The average number of bound SSB pairs *per complex* of topoisomer $C(i)$ is given by

$$\begin{aligned}\langle n_{2a}(C(i, n_{2a}\geqslant 1))\rangle &= (1/\mathrm{GPF}(i))\sum_{n_{2a}=1}^{N/l} n_{2a}\left[(c_a K_a)^2\right]^{n_{2a}} \\ &\quad\times\exp\left[-\Delta A(\Delta L(i), n_{2a})/k_B T\right] W(n_{2a})\end{aligned} \tag{19}$$

and the average number of nearest-neighbor pair–pair contacts *per complex* of topoisomer i is given by

$$\begin{aligned}\langle p_J(C(i, n_{2a}\geqslant 1))\rangle &= \langle p_J(n_{2a})\rangle_i \\ &= (S/2)[1/\mathrm{GPF}(i)]\partial\mathrm{GPF}(i)/\partial S.\end{aligned} \tag{20}$$

Eqs. (18)–(20) express the relevant quantities for each topoisomer in terms of its grand partition function GPF(i), which is evaluated directly from Eq. (15) with the help of Eqs. (11) and (21) (below).

3.2. The configuration sum in GPF

An expression for the configuration sum ($W(n_{2a})$) for ligand binding to circular lattices was presented by Reiter and Epstein [22]. This is adapted to the present model of pair binding to give

$$W(n_{2a}) = \sum_{c=c_{\min}}^{c_{\max}} \frac{f!}{c!(f-c)!}\mu\frac{(n_{2a}-1)!}{(n_{2a}-c)!(c-1)!} \times\left[S^2\right]^{n_{2a}-c}, \tag{21}$$

wherein

$$f = N - n_{2a}l \tag{22}$$

$$c_{\max} = \begin{cases} n_{2a}, & \text{for } f\geqslant n_{2a}, \\ f, & \text{for } f < n_{2a}\end{cases} \tag{23}$$

$$c_{\min} = \begin{cases} 0, & \text{for } n_{2a}=0, \\ 1, & \text{for } n_{2a}>0,\end{cases} \tag{24}$$

$$\mu = \begin{cases} N/f, & \text{for } f>0, \\ N/n_{2a}, & \text{for } f=0.\end{cases} \tag{25}$$

In Eqs. (21)–(25), f is the number of free lattice sites, and c is the island cluster index. An island cluster is defined as a group of contiguously bound ligands with at least one free lattice site on the right. The cluster index is simply the number of clusters on the lattice. The first combinatorial factor is the number of permutations of c island clusters and f free sites subject to the constraint that every cluster must have a free site on its right. Notice that the first combinatorial factor does not depend on the number of lattice sites covered by any particular cluster. All clusters are considered to be equivalent in the first factor, regardless of the number of ligands that comprise the cluster. The second combinatorial factory is the number of ways to choose c island clusters from n_{2a} ligands. The factor μ is the sliding degeneracy or the number or ways to form any unique pattern of clusters plus free sites on a circular lattice of length N. $n_{2a}-c$ is the number of pair–pair contacts. S^2 is the equilibrium constant for cooperative interactions per pair of SSB. The summation runs from the minimum to maximum number of island clusters with n_{2a} bound SSB pairs. We note that Eq. (21) is not defined for $f=0$. This circumstance does not arise here, because l is taken to be even, and $N=4363$, so N/l is non-integral.

The validity of Eqs. (21)–(25) was thoroughly checked by calculating GPF for non-supercoiled lattices according to Eq. (15) with $\Delta A=0$ and comparing with results of the matrix method. Perfect agreement is achieved at the level of numerical precision of the computations. In addition, a matrix formulation applicable for small values of Φ was developed from a Taylor series

expansion of ΔA about $\langle n_{2a}(C(i) \geq 1))\rangle$, with retention of just the constant and linear terms in n_{2a}. Very good agreement with results of Eq. (15) again confirms the validity of Eqs. (21)–(25).

3.3. The topoisomer distribution

The fractions of topoisomers with different linking differences in the native population are assumed to be given by

$$f_i = \frac{\exp\left\{-\frac{500}{4363}\left[\Delta L(i) - \Delta L(m)\right]^2\right\}}{\sum_{j=1}^{13} \exp\left\{-\frac{500}{4363}\left[\Delta L(j) - \Delta L(m)\right]^2\right\}}. \quad (26)$$

where $\Delta L(m)$ is the linking difference of the most probable topoisomer. The allowed linking differences are $\Delta L(i) = \Delta L(m) \pm j$, $j = 0, 1, \ldots, 6$. The width of this distribution is about twice that of a normal thermal distribution (for $E_T = 1000$), and is typical of native DNA preparations. The total concentration of the ith topoisomer is then

$$c_{C(i)}^{tot} = f_i c_C^{tot}, \quad (27)$$

where c_C^{tot} is the total concentration of supercoiled DNAs.

Each experimental condition to be analyzed is indexed by the subscript q, $q = 1, 2, \ldots, Q$, where Q is the total number of experiments of a particular kind (FQ or DLS). In each experiment, the total supercoiled DNA concentration ($(c_C^{tot})_q$), the total SSB concentration ($(c_a^{tot})_q$), and the ratio ($\upsilon_q = (c_a^{tot})_q/(c_C^{tot})_q$) are known. Making use of Eqs. (26) and (27) then yields all of the total topoisomer concentrations ($(c_{C(i)}^{tot})_q$) for each experiment. Calculation of all the GPF(i), $c_{C(i,n_{2a}\geq 1)}$, $\langle n_{2a}(C(i, n_{2a} \geq 1))\rangle$, and $\langle p_J(C(i, n_{2a} \geq 1))\rangle$ for each topoisomer in a given experiment by Eqs. (15) and (18)–(25) requires numerical values for several molecular parameters, namely N, K_a, S, $\Delta L(i)$, E_T, l, and Φ, which are discussed below, and also values for two additional prevailing experimental parameters, specifically the temperature (T) and free ligand concentration ($(c_a)_q$). Although T is precisely known, $(c_a)_q$ was not measured, so must be calculated from $(c_a^{tot})_q$ by taking account of conservation of SSB.

3.4. Conservation of SSB

The ratio of total added SSB per supercoiled DNA can be written as

$$\upsilon = c_a^{tot}/c_C^{tot} = \left[c_a + 2\sum_{i=m-6}^{m+6} c_{C(i,n_{2a}\geq 1)} \times \langle n_{2a}(C(i, n_{2a} \geq 1))\rangle\right]/c_C^{tot}, \quad (28)$$

where c_a^{tot} has been apportioned into its separate contributions of free and bound ligand in the second equality of Eq. (28). For each experiment, the free ligand concentration $(c_a)_q$ can be calculated from the known values of υ_q and $(c_C^{tot})_q$ and a given set of molecular parameters by iteratively 'solving' Eq. (28).

3.5. Molecular parameters

In order to evaluate the GPF(i) in Eq. (15) and other quantities of necessity or interest, numerical values must be specified for several molecular parameters, namely K_a, S, $\Delta L(i)$, E_T, N, l and Φ. Although N (4363 bp) is precisely known, considerable uncertainty surrounds the estimates of the other parameters. By far the largest uncertainty is associated with the composite binding constant K_a, since it is the product of three factors, including $\tilde{s}^{l/2}$. The high power to which $\tilde{s}$ is raised enormously amplifies any uncertainty in $\tilde{s}$. For this reason, we regard K_a as an adjustable parameter, and subsequently compare the best-fit value of K_a with its experimental estimate. The other parameters, which are listed in Table 1, are regarded as fixed input values and are estimated as follows.

3.5.1. Stacking constant (S)

According to Eq. (7), evaluation of S requires estimates of ω and σ_M. We take $\omega = 50 \pm 10$, independent of salt concentration or substrate identity, as suggested by studies of SSB binding to poly A and poly U over a range of salt concentrations [8]. We also assume that $\sigma_M = 4.5 \times 10^{-5}$, as employed in detailed DNA melting studies in

Table 1
Input parameters and best-fit composite binding constants

Parameter	Value
N	4363 (bp)
$S(20°) = S(40°)$	7.45×10^3
l	34 (bp)
Φ	3.27 (turns)
E_T	1000
$\Delta L(m)$ at 20°C	−25 (turns)
$\Delta L(m)$ at 40°C	−22.7 (turns)
$(K_a^{op})_{DLS}$ at 20°C	0.150 M^{-1}
$(K_a^{op})_{FQ}$ at 40°C	0.587 M^{-1}

0.2 M NaCl [23–25]. We obtain $S = \omega\sigma_M^{-1/2} = 50 \times 149 = 7.45 \times 10^3$. We further *assume* that S is essentially constant from 20 to 40°C.

3.5.2. Binding site width (l)

The apparent site-size (l) for SSB binding to single-strand DNA depends on the ionic environment and the SSB: single-strand-DNA binding ratio [9]. For SSB-poly dT below 10 mM NaCl, $l = 33 \pm 3$ bp. Above 0.2 M NaCl, $l = 65 \pm 5$ bp. Between these two salt concentrations l increases with increasing salt concentration [9]. At high binding ratios the lower value, $l = 33$ bp is favored [26]. The FQ and DLS studies that we are attempting to understand were performed in 50 mM and 100 mM NaCl, respectively. Both conditions favor $l < 65$ bp. The necessity for the DNA to melt in order to bind SSB can be expected to compress the site width, thus favoring smaller site sizes. We choose $l = 34$ bp, because it is well within the reported range (33 ± 3 bp), is consistent with the above observations, and is divisible by 2, so it is easily accommodated in the existing theory.

3.5.3. Unwinding angle Φ

The unwinding angle Φ determines the amount of strain reduction, and hence the decrease in supercoiling free energy, per binding event. This quantity is unknown and is experimentally difficult to obtain in the case of SSB, which does not bind to relaxed duplex DNA. We simply assume that $\Phi = l/10.4$, which is just the number of turns that must be melted to accommodate an SSB tetramer. With $l = 34$ bp, we obtain $\Phi = 3.27$ turns per pair of bound SSB. This is equivalent to assuming that there is no *intrinsic* helical structure within the melted binding region that contributes to the intrinsic twist of the DNA. Indeed, EM studies indicate that the strands in the complexed region are well separated and untwisted [1,4].

3.5.4. Twist energy parameter (E_T)

Experimental twist energy parameters for pBR322 vary widely. Topoisomer distribution experiments yield values near 1150 [21,27] and 1610 [28] while dye-binding measurements yield values over the range from 300 to 1800 [18,20]. A recent investigation indicates that dye-binding values for pBR322 depend on the time after cell-lysis and vary from one preparation to another, even when identical protocols are employed by the same experimenters [21]. These dye-binding studies on pBR322 generally indicate an increasing E_T with increasing bound dye, as the effective superhelix density simultaneously decreases from the native value toward zero [21]. However, in the case of p30δ (derivative of pBR322), no such variations of E_T with time after cell-lysis or with different preparations were detected by those same experimenters using the same protocols [20]. Moreover, for p30δ essentially the same value, $E_T = 1030 \pm 90$, was obtained from both topoisomer distribution and ethidium binding experiments, and it was found to be independent of superhelix density over the range from zero to native ($\sigma = -0.05$), as well as ethidium binding ratio over the range $r = 0$ to $r = 0.08$ dye/bp [20]. For the present calculations, we ignore the complications exhibited by real pBR322 DNAs and simply assume that $E_T = 1000$, which is a kind of consensus value obtained from topoisomer distribution experiments on several long DNAs [20,27,29,30]. We also assume that E_T remains constant, or equivalently that the supercoiling free energy varies quadratically with linking difference, for all superhelix densities from zero up to native values. Certainly, p30δ exhibits such behavior, even if pBR322 does not.

3.5.5. Linking difference ($\Delta L(i)$)

The linking differences of the topoisomers in the distribution are given by $\Delta L(i) = \Delta L(m) \pm j$,

$j = 0, 1, \ldots, 6$, where $\Delta L(m)$ is that of the most probable topoisomer. In practice, $\Delta L(m)$ varies from one preparation to another. We have isolated pBR322 with $\Delta L(m)$ as small as -20 and as large as -35 superhelical turns. Unfortunately, the most probable linking difference of the supercoiled pBR322 studied by Langowski et al. [15] was not measured. In this work, we adopt a representative value, $\Delta L(m) = -25.0$ turns in 0.1 M NaCl, 10 mM Tris at 20°C corresponding to a superhelix density, $\sigma = -0.06$. Selection of other values in the range -20 to -35 causes large changes in the best-fit composite binding constant K_a, but does not significantly alter the best-fit curve or affect any of the qualitative conclusions. Temperature and ionic strength affect the intrinsic twist, hence also the linking difference of a closed circular DNA. The DLS experiments were performed in 100 M NaCl, 10 mM Tris-HCl at 20°C, while the FQ studies were done in 50 mM NaCl, 20 mM potassium phosphate at 40°C. The intrinsic twist decreases with increasing temperature approximately according to $\Delta l_0/\Delta T = -(2.97 \times 10^{-5})N$ turns/degree, and increases with increasing ionic strength (X) approximately according to $\Delta\sigma/\Delta \mathrm{p}X \approx -(1/l_0)\,\Delta l_0/\Delta \mathrm{p}X = 4.47 \times 10^{-3}$, where $\sigma = (l - l_0)/l_0$ is the superhelix density and pX is $-\log_{10} X$ [31]. The intrinsic twist of the DNA in the FQ studies is estimated to be underwound compared to that of the DNA in the DLS studies by 2.59 turns due to the higher temperature and overwound by 0.30 turns due to the slightly higher ionic strength. Consequently, $\Delta L(m)$ in the DLS studies at 40°C is taken as -22.7 turns while that in the FQ studies at 20°C is taken as -25.0 turns.

4. Fitting the fluorescence quenching results

The fluorescence of SSB decreases markedly upon complexation with a single strand of DNA. In the reverse titration experiment under consideration, the fluorescence of a solution containing SSB at concentration c_a^{tot} in the absence of DNA is first measured to obtain f_0, the fluorescence of free SSB at concentration c_a^{tot}. Aliquots containing SSB also at concentration c_a^{tot} and a rather high DNA concentration are then added to raise the concentration of DNA *at constant* c_a^{tot}. The measured fluorescence of such a solution is denoted by f. Let $\gamma = f_b/f_0$ be the ratio of the fluorescence of bound SSB to that of free SSB at the same concentration. The value $\gamma = 0.26$ was estimated by extrapolating the curve of f/f_0 versus v^{-1} to large v^{-1}, under which condition the SSB is practically all bound. For each value of v in this titration at constant c_a^{tot}, the fluorescence intensity ratio is assumed to be given by

$$f/f_0 = \left[c_a + 2\gamma \sum_{i=m-6}^{m+6} c_{C(i,n_{2a}\geq 1)} \times \langle n_{2a}(C(i, n_{2a} \geq 1)) \rangle \right] / c_a^{tot}. \quad (29)$$

In all, Q fluorescence measurements were performed. For each fluorescence measurement (denoted by q), the quantities $(c_a^{tot})_q = 86$ nM, $(c_C^{tot})_q$, v_q, and $(f/f_0)_q$ are known. For a given set of molecular parameters, including a trial value of K_a, one can calculate all the $(c_a)_q$, $q = 1, \ldots, Q$, by iteratively solving Eq. (28). One can then compute other quantities of interest, in particular the theoretical value for $(f/f_0)_q$. The steps in the fitting algorithm are as follows:

(1) For each topoisomer in the distribution, n_{2a}^* is calculated from the assigned value of $\Delta L(i)$ according to $n_{2a}^* = -\Delta L(i)/\Phi$.

(2) An initial trial value, $K_a = 7.45 \times 10^{-4}$ M^{-1}, is selected. This corresponds to an early rough estimate, and has no effect on the final results.

(3) The q index of the measurement is set to $q = 1$.

(4) An initial trial value of the free ligand concentration is selected by setting $c_a^{trial} = (c_a^{tot})_q$.

(5) For each topoisomer in the distribution, GPF(i) is calculated by Eq. (15) using also Eqs. (11) and (21)–(25), $c_{C(i,n_{2a}\geq 1)}$ is calculated by Eqs. (18) and (27), and $\langle n_{2a}(C(i, n_{2a} \geq 1)) \rangle$ is calculated by Eq. (19). Finally, v_{trial} is calculated via the r.h.s. of Eq. (28). If $(v_{trial} - v_q) < -0.01\, v_q$, then c_a is a lower bound. Conversely, if $(v_{trial} - v_q) > 0.01\, v_q$, then c_a is an upper bound. The next c_a^{trial} is taken as one half the sum of the

prevailing upper and lower bounds, and this step 5 is repeated until $|v_{\text{trial}} - v_q| \geqslant 0.01\, v_q$, at which point $(c_a)_q = c_a^{\text{trial}}$.

(6) For each topoisomer, both $\langle p_f(C(i, n_{2a} \geqslant 1))\rangle$ and the effective or residual linking difference, $\Delta L(i)_{\text{eff}} = \Delta L(i) - \Phi\langle n_{2a}(C(i, n_{2a} \geqslant 1))\rangle = \Phi(n_{2a}^* - \langle n_{2a}(C(i, n_{2a} \geqslant 1))\rangle)$, are calculated for the prevailing v_q.

(7) The theoretical fluorescence ratio $(f/f_0)_q$ for the prevailing q-index is calculated via the r.h.s. of Eq. (29) and designated $(f/f_0)_q^{\text{th}}$.

(8) q is incremented by 1. If $q \leqslant Q$, return to step (4). If $q > Q$, proceed to step (9).

(9) The sum of squared deviations is calculated according to

$$(\text{SSD})_{\text{FQ}} = \sum_q \left[(f/f_0)_q^{\text{th}} - (f/f_0)^{\text{exp}}\right]^2. \qquad (30)$$

This quantity represents the total squared error for the prevailing choice of K_a.

(10) Steps (2)–(9) are repeated for different K_a values along a one-dimensional grid until the value of K_a that minimizes SSD is located.

(11) The entire procedure was repeated for different choices of $\Delta L(m)$. The results of this last step are presented in detail elsewhere [32], but are discussed only briefly here.

5. Fitting the D_0 results

In order to model the D_0 data of Langowski et al., [15] it is necessary to make some kind of assumption regarding the changes in D_0 caused by SSB binding. The rapid decline in D_0 with v implies that the structural change occurs already at very low levels of bound SSB. The apparent isosbestic point in plots of D_{app} versus K^2 for $v = 0$, 3 and 10 suggests that only two hydrodynamically distinct species, namely free DNA and complexes, exist in solution. Accordingly, we simply assume that all *free DNAs* exhibit the diffusion constant, $D_0^{\text{free}} = 4.54 \times 10^{-8}$ cm^2/s, and all *complexes* containing one or more bound SSB tetramer pairs exhibit the diffusion constant, $D_0^{\text{comp}} = 3.83 \times 10^{-8}$ cm^2/s. These values are selected to match the ends of the D_0 versus v curve at low and high v, respectively. At intermediate points, the average diffusion coefficient is assumed to be given by

$$D_0 = \sum_{i=m-6}^{m+6} \left[c_{C(i,n_{2a}\geqslant 1)} D_0^{\text{comp}} + c_{C(i)} D_0^{\text{free}}\right] / c_C^{\text{tot}}. \qquad (31)$$

Of the possible models that could be assumed, Eq. (31) predicts the earliest and most abrupt possible change in D_0 with increasing v. Reproducing the early onset and abruptness of the observed transition is one of the most stringent tests of possible models. A discussion of the implications of this assumption is postponed until after the results are examined in Section 6.

The steps in the fitting algorithm are as follows:

(1)–(6) These steps are the same as in fitting the FQ data. At this point one has calculated $c_{C(i,n_{2a}\geqslant 1)}$, $c_{C(i)}$, and other quantities for each topoisomer at the prevailing v_q.

(7) D_0 for this prevailing v_q is calculated via Eq. (31).

(8) This step is the same as in fitting the FQ data.

(9) The sum of squared deviations is calculated according to

$$(\text{SSD})_{D_0} = \sum_q \left[(D_0^{\text{th}})_q - (D_0^{\text{exp}})_q\right]^2 \times 10^{16}\ \text{cm}^4\ \text{s}^{-2}. \qquad (32)$$

(10)–(11) These steps are the same as in fitting the FQ data.

6. Results and discussion

6.1. Single topoisomer analysis

It is assumed for the moment that the binding behavior of the distribution can be modelled by that of a single topoisomer, whose linking difference is identical to the most probable linking number of the actual distribution, namely $\Delta L(m) = -25$ turns for the D_0 experiments at 20°C, and $\Delta L(m) = -22.7$ turns for the FQ experiments at 40°C. This assumption introduces no serious er-

rors into the analysis of ethidium binding to long DNAs [18,20]. It is of interest to examine the validity of this assumption in the case of SSB binding.

6.1.1. Analysis of the fluorescence quenching data

The best-fit theoretical fluorescence ratios are compared with the experimental values in Fig. 3. The optimum value of the composite binding constant, $K_a = 0.510\ M^{-1}$, yields a reasonably good fit with $(SSD)_{FQ} = 5.27 \times 10^{-3}$. Although the quality of the fit declines noticeably beyond $v^{-1} = 0.15$, the overall quality of the fit is comparable to that obtained previously for a model of 10 independent binding sites with no cooperative interactions whatsoever [15]. It is clear at this point that the reduction of superhelical strain almost completely offsets the highly cooperative binding to a very long lattice, which would otherwise saturate at (2) (4363/34) = 256 bound tetramers, in such a way that the predicted FQ results *over the limited span of the data* can be satisfactorily fitted by a non-cooperative model with a much smaller number of sites (10), as shown by Langowski et al. [15] However, the actual character of the binding implied by the present model differs radically from that of the non-cooperative model, as described further below.

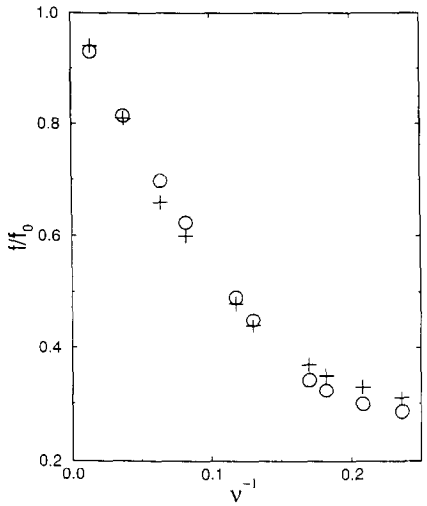

Fig. 3. Fluorescence intensity ratio f/f_0 versus v^{-1}, where v is the added SSB/plasmid ratio. (+) Experimental data of Langowski et al. [15] (○) Theoretical values calculated for a single topoisomer via Eqs. (14), (15), (18), (19), (28), and (29) together with the fixed parameters: $c_a^{tot} = 86$ nM, $\gamma = 0.26$, $l = 34$ bp, $\Phi = -3.269$ turns/bound SSB pair, $S = 7.45 \times 10^3$, $E_T = 1000$, $N = 4363$ bp, and the optimum value $K_a = 0.510\ M^{-1}$. The plasmid DNA is here represented by a single topoisomer with $\Delta L = -22.7$ turns.

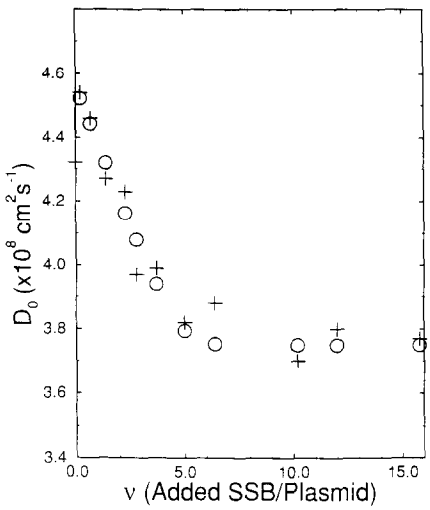

Fig. 4. Diffusion coefficient D_0 versus v, where v is the added SSB/plasmid. (+) Experimental data of Langowski et al. [15] (○) Theoretical values calculated for a single topoisomer via Eqs. (14), (15), (18), (19), (28), and (31) together with the fixed parameters: $c_C^{tot} = 16.3$ nM (plasmids), $D_0^{comp} = 3.83 \times 10^{-8}$ cm^2/s, $D_0^{free} = 4.54 \times 10^{-8}$ cm^2/s, $l = 34$ bp, $\Phi = -3.269$ turns/bound SSB pair, $S = 7.45 \times 10^3$, $E_T = 1000$, $N = 4363$ bp, and the optimum value $K_a = 0.015\ M^{-1}$. The plasmid distribution is here represented by a single topoisomer with $\Delta L = -25$ turns.

6.1.2. Analysis of the D_0 data

The best-fit values of D_0 are compared with the experimental values in Fig. 4. The optimum value of the composite binding constant, $K_a = 0.015\ M^{-1}$, yields a reasonably good fit with $(SSD)_{D_0} = 4.40 \times 10^{-2}$ cm^4 s^{-2}. This value of K_a, which applies at 20°C, is $\frac{1}{34}$ of that obtained from the FQ data, which applies at 40°C, so K_a evidently increases with increasing temperature which is opposite to the usual behavior observed for binding constants. However, this is not unexpected in view of the fact that K_a contains the equilibrium constant per base-pair to form melted strands from duplex DNA ($\bar{s}$) raised to the power

of $l/2 = 17$, as well as the equilibrium constant (K_0) for SSB binding to single-strand DNA. The opposing trends of $\tilde{s}^{l/2}$ and K_0 with temperature will cancel to some extent, but the former evidently wins out. The resulting inverted temperature dependence argues in favor of a model, such as that considered here, wherein the duplex must be melted to accommodate SSB binding.

6.2. Analysis for a distribution of topoisomers

It is now assumed that the distribution of topoisomers in the native population is given by Eq. (26) with $\Delta L(m) = -25$ for the D_0 data at 20°C and $\Delta L(m) = -22.7$ for the FQ data at 40°C. The simultaneous binding equilibria of all 13 topoisomers are considered. All topoisomers exhibit the same molecular parameters except for $\Delta L(i)$ and $n_{2a}^{*}(i) = -\Delta L(i)/\Phi$.

6.2.1. Analysis of the fluorescence quenching data

The best-fit theoretical fluorescence ratios are compared with the experimental values in Fig. 5. The optimum composite binding constant, $K_a = 0.587\ M^{-1}$, yields a good fit to the data with $(SSD)_{FQ} = 2.387 \times 10^{-3}$. This optimum K_a is about 15% larger than in the case of the single topoisomer analysis, and the fit is somewhat better, as $(SSD)_{FQ}$ is reduced by 2.2-fold.

6.2.2. Analysis of the D_0 data

The best-fit values of D_0 are compared with the experimental values in Fig. 6. The optimum composite binding constant, $K_a = 0.150\ M^{-1}$, yields a good fit to the data with $(SSD)_{D_0} = 4.09 \times 10^{-2}\ cm^4\ s^{-2}$. This optimum K_a is 10-fold larger than that obtained from the single-topoisomer analysis. This surprisingly large difference can be understood in the following way. The D_0 versus v curves in either case are remarkably insensitive to K_a from $v = 0$ to 2.0, provided K_a is sufficiently large, but are considerably more sensitive to K_a in the tail from $v = 2.5$ to 6.0. In all events, the binding is quasi all-or-none and proceeds more or less sequentially from the most supercoiled plasmids to the least, as described below. In the tail beyond $v = 2.5$, only the less supercoiled half of the topoisomer distribution

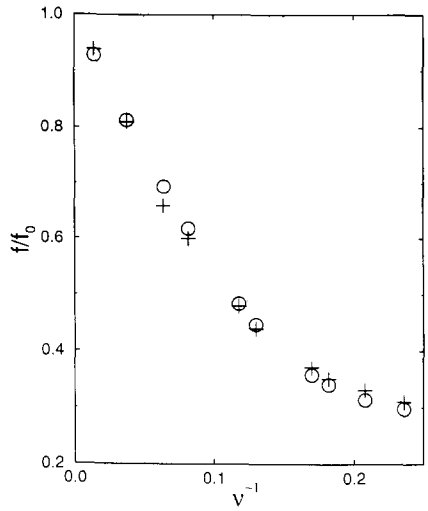

Fig. 5. Fluorescence intensity ratio f/f_0 versus v^{-1}, where v is the added SSB/plasmid ratio. (+) Experimental data of Langowski et al. [15] (○) Theoretical values calculated for a distribution of topoisomers given by Eq. (26) with $\Delta L(m) = -22.7$ turns via Eqs. (14), (15), (18), (19), (28), and (29) together with the same fixed parameters indicated in the legend of Fig. 3, except for the optimum K_a, which is now $K_a = 0.587\ M^{-1}$.

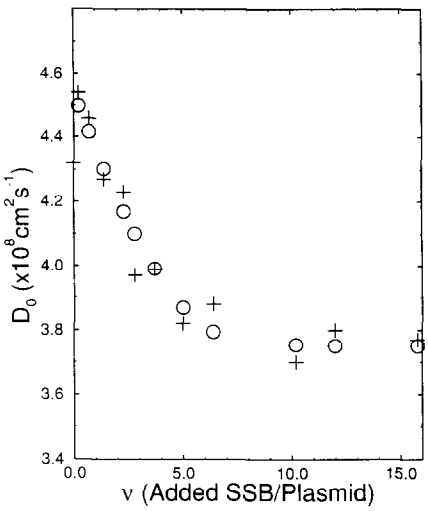

Fig. 6. Diffusion coefficient D_0 versus v^{-1}, where v is the added SSB/plasmid. (+) Experimental data of Langowski et al. [15] (○) Theoretical values calculated for a distribution of topoisomers given by Eq. (26) with $\Delta L(m) = -25$ turns via Eqs. (14), (15), (18), (19), (28), and (31) together with the same fixed parameters indicated in the legend of Fig. 4, except for the optimum K_a, which is now $K_a = 0.150\ M^{-1}$.

with $\Delta L(i) \leqslant -24$ remains uncomplexed, and a much larger (10-fold) value of K_a is required to induce those to bind than to induce the single topoisomer with $\Delta L(m) = -25$ to bind. The insensitivity of the D_0 versus v curve to K_a from $v = 0$ to 2 stems from the fact that most of the SSB is complexed under those conditions, so a very large increase in K_a has only a small effect on the average D_0 value. The crossover from insensitive (to K_a) to sensitive behavior occurs (in a surprisingly abrupt fashion) when v exceeds the average number of bound SSB per complex, so that the concentration of free SSB becomes significant compared to the bound SSB. From data presented in Fig. 8, this crossover occurs at about $v = 2.5$. Finally, we note that use of the full distribution in place of a single topoisomer reduced $(SSD)_{D_0}$ by only 7%.

The fluorescence quenching data are evidently somewhat more sensitive than the D_0 data to the width of the topoisomer distribution, since a much larger reduction in χ^2 was achieved in that case by using the full topoisomer distribution.

6.2.3. Effect of varying the linking difference

Varying $\Delta L(m)$ over the range from -20 to -35 turns had virtually no effect on the best-fit curves or on the SSD, but produced very large variations in the best-fit K_a. The empirical linear relations between $\ln K_a^{fit}$ and $\Delta L(m)$ are presented elsewhere [32].

6.3. Comparison of best-fit K_a with an independent estimate

K_a can be estimated according to Eq. (6), provided estimates of K_0, σ_M, and $\tilde{s}$ are at hand. As above, we take $\sigma_M = 4.5 \times 10^{-5}$. K_0 is the equilibrium constant for SSB binding to one of two complementary single strands of a melted region. It will differ from that (K_1) for binding to an *isolated* single-strand by a factor (R) due to the greater charge density surrounding the two strands. That is, $K_0 = K_1\ R$. To estimate K_1, we take the product $K_1\ \omega = 5.7 \times 10^{10}$, which is the midpoint of the reported range for 'natural' single-strand DNAs [2]. Combining this figure with the above estimate of $\omega = 50$ yields $K_1 = 1.1 \times 10^9\ M^{-1}$. The electrostatic enhancement factor is estimated in Appendix A to be $R = 500$. Thus we finally obtain $K_0 = 5.5 \times 10^{11}\ M^{-1}$. In Appendix B we estimate that $\tilde{s} = 0.213$, which corresponds to a free energy change, $\Delta G = -RT\ \ln\ \tilde{s} = 902$ cal/(mol bp), for melting at 293 K in 0.1 M NaCl. This value should be regarded as very approximate. The resulting estimate for the composite binding constant at 20°C is then

$$K_a^{est} = (5.5 \times 10^{11})(0.213)^{17}(4.5 \times 10^{-5})^{1/2} = 0.014^{-1}. \quad (33)$$

The optimum K_a at 20°C, namely $K_a = 0.150\ M^{-1}$, exceeds K_a^{est} by a factor of 10.7. This discrepancy is not significant for the following reason. If we ask by what factor (F) must $\tilde{s}$ be increased to bring K_a^{est} into line with the optimum K_a, we find $F = (10.7)^{1/17} = 1.15$. Thus, a mere 15% increase in $\tilde{s}$, or a decrease in ΔG_{bp} from 902 to 821 cal/(mol bp), is all that is required to bring the estimated and optimum K_a values into line. These changes are well within the considerable uncertainties in the estimates of $\tilde{s}$ and ΔG. Thus, the application of the present model to analyze the data of Langowski et al. [15] yields a best-fit K_a that is entirely consistent with other known data.

6.4. Nature of SSB binding to a single topoisomer

The average number of bound SSB pairs *per complex*, $\langle n_{2a}(C(i,\ n_{2a} \geqslant 1))\rangle$, and average number of pair–pair contacts, $\langle p_J(C(i,\ n_{2a} \geqslant 1))\rangle$, per complex are plotted versus $c_a/c^{tot}_{C(i)}$, the ratio of the *free* SSB concentration to *total* concentration of the topoisomer with $\Delta L(i) = -25$ turns, in Fig. 7. Perhaps the most striking aspect of these results is that the average number of pair–pair contacts is always very close to the average number of bound pairs *per complex* minus 1.0. This indicates that the SSB pairs essentially always bind contiguously in a single-stack on a given plasmid! This prediction agrees very well with the EM results [4].

The shape of the binding curve over the limited range presented in Fig. 7 suggests an apparent saturation of the binding at a rather small

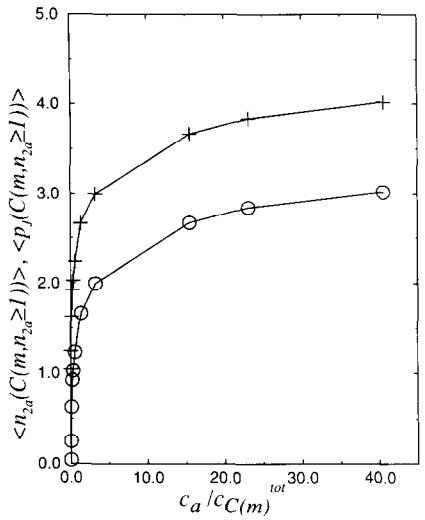

Fig. 7. Average number of bound SSB pairs $\langle n_{2a}(C(m, n_{2a} \geq 1))\rangle$, and average number of pair–pair contacts $\langle p_J(C(m, n_{2a} \geq 1))\rangle$ per complex of the topoisomer m versus $c_a / c^{tot}_{C(m)}$, where c_a is the free ligand concentration. (+) The average number of bound SSB pairs per complex formed by the topoisomer with $\Delta L(m) = -25$ turns according to Eq. (19). (○) The average number of pair–pair contacts according to Eq. (20). The solid lines are drawn merely to aid the eye. The other parameter values are 16.3 nM (plasmids), $l = 34$ bp, $\Phi = -3.269$ turns/bound SSB pair, $S = 7.45 \times 10^3$, $E_T = 1000$, $N = 4363$ bp, and $K_a = 0.150\ M^{-1}$. This K_a value best fits the D_0 data in Fig. 6. Because the ordinate is proportional to *free* ligand concentration, these results apply equally to a single isolated topoisomer with $\Delta L(m) = -25$ turns, or to the most probable topoisomer in the distribution of Eq. (26) with $\Delta L(m) = -25$ turns. Note that $\langle n_{2a}(C(m), n_{2a} \geq 1)\rangle$ always exceeds $\langle p_J(C(m), n_{2a} \geq 1)\rangle$ by 1.0 for each value of $c_a / c_{C(m)}$ investigated.

number of bound SSB per plasmid. However, calculations over a much wider range of concentration indicate that this apparent saturation is illusory, and that binding to the lattice can be driven upward by sufficiently high concentrations of free ligand. Of course, the restriction to free ligand concentrations in the experimentally accessible range eventually caps the binding at a rather modest level. The essential point is that over a limited span of concentrations, the decreasing affinity of DNA for SSB, due to its diminishing superhelical strain, acts to produce *apparent* saturation behavior at a level far below the actual saturation level of the lattice. This is manifested clearly in the fluorescence quenching data.

Relative fluctuations in the number of bound SSB pairs ($[(\langle n_{2a}(C(i, n_{2a} \geq 1))^2\rangle - \langle n_{2a}(C(i, n_{2a} \geq 1))\rangle^2)]^{1/2} / \langle n(C(i, n_{2a} \geq 1))\rangle$) are rather small, and exhibit periodic maxima at the midpoints of transitions from 1 to 2, 2 to 3, 3 to 4, etc., bound SSB pairs *per complex* (data not shown). This indicates that the binding is largely all-or-none in the sense that each molecule of topoisomer i has either no SSB (and therefore is not counted as a complex) or its normal complement *appropriate for the prevailing concentration of free ligand*. The size of this normal complement increases somewhat with increasing SSB concentration, but appreciable deviations from the normal complement at any fixed SSB concentration are evidently quite uncommon. This prediction also matches the EM results [4].

Because the abscissa in Fig. 7 is proportional to free ligand concentration, this figure and the results it displays are valid for either the isolated topoisomer or the same topoisomer in a distribution of topoisomers.

6.5. Nature of SSB binding to a distribution of topoisomers

The distribution employed to analyze the D_0 data consists of 13 different topoisomers. In Fig. 8, the average number of bound SSB pairs *per complex* is plotted versus v (total added SSB per plasmid) for three topoisomers in that distribution, namely those with $\Delta L(i) = -19$, -25, and -31 turns. By definition, the minimum number of bound SSB pairs *per complex* is 1.0. The onset of significant binding is then associated with increases above this minimum value. Evidently, increasingly larger v values are required to induce the onset of binding in plasmids with increasingly smaller linking differences. This reflects the prominent role played by supercoiling in directing the binding to the most strained plasmids. All of these curves show some tendency toward horizontal flattening, as the average number of bound SSB pairs increases past integral values, and toward steepening at half-integral values. This happens for the following reason. When the average number of bound SSB pairs per complex reaches an integral value, the next SSB molecules are

more often directed to assemble a new stack on a free plasmid, which offers the greatest superhelical strain reduction, than to enlarge the stack on an existing complex. Thus, the average number of bound SSB *per complex* changes relatively little with increased binding at these points. However, when the average number of bound SSB pairs *per complex* is half-integral, then addition to an existing stack competes more favorable with assembly of a new stack on a free plasmid, so the average number of bound SSB per complex rises more steeply at those points.

The average number of bound SSB pairs and average number of pair–pair contacts *per complex* are plotted versus v for one particular topoisomer in Fig. 9. The average number of pair–pair contacts is again always very close to the average number of bound SSB minus 1.0. Thus, the binding occurs essentially always in a single contiguous stack also in the topoisomer distribution. The

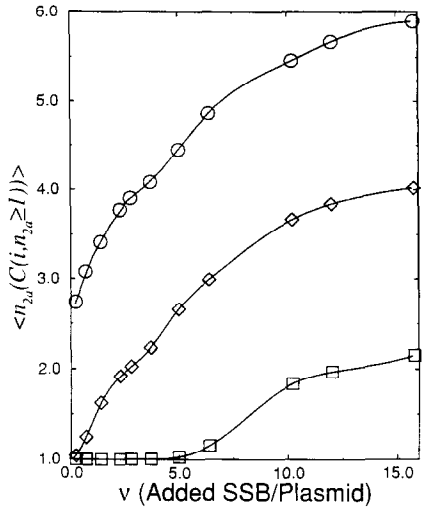

Fig. 8. Average number of bound SSB pairs per complex $\langle n_{2a}(C(i),\ n_{2a} \geq 1)\rangle$ versus added SSB per plasmid (v) for different molecules in the topoisomer distribution. Results are displayed for three different linking differences $\Delta L(i)$ in the population. (□) $\Delta L(i) = -19$ turns; (◇) $\Delta L(i) = -25$ turns, and (○) $\Delta L(i) = -31$ turns. The solid curves are drawn to guide the eye. The parameter values are: $c_C^{tot} = 16.3$ nM (plasmids), $l = 34$ bp, $\Phi = -3.269$ turns/bound SSB pair, $S = 7.45 \times 10^3$, $E_T = 1000$, $N = 4363$ bp, and $K_a = 0.150\ M^{-1}$. The topoisomer distribution is given by Eq. (26) with $\Delta L(m) = -25$ turns.

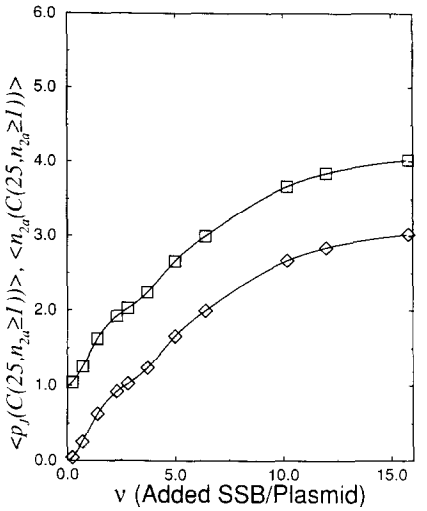

Fig. 9. Average number of bound SSB pairs $\langle n_{2a}(C(m),\ n_{2a} \geq 1)\rangle$ and average number of pair–pair contacts $\langle p_f(C(m),\ n_{2a} \geq 1)\rangle$ of the most probable topoisomer m versus added SSB per plasmid (v). Squares (□) denote the average number of bound SSB per complex according to Eq. (19) and diamonds (◇) denote the average number of pair–pair contacts per complex according to Eq. (20). These results apply for the most probable topoisomer in the distribution of Eq. (26) with $\Delta L(m) = -25$ turns. Other parameters are $c_C^{tot} = 16.3$ nM (plasmids), $l = 34$ bp, $\Phi = -3.269$ turns/bound SSB pair, $S = 7.45 \times 10^3$, $E_T = 1000$, $N = 4363$ bp, and $K_a = 0.150\ M^{-1}$. Note that $\langle n_{2a}(C(m),\ n_{2a} \geq 1)\rangle$ always exceeds $\langle p_f(C(m),\ n_{2a} \geq 1)\rangle$ by 1.0 for each value of v examined.

apparent saturation of the binding over the limited span of the data is again evident.

6.6. Narrowing of the distribution of linking differences

The intrinsic twist of a complex is given by $l_0 - n_{2a}\Phi$, consequently the linking difference can be defined generally to encompass both complexes and free topoisomers by setting

$$\Delta L(i) = i - (l_0 - n_{2a}\Phi). \tag{34}$$

In either case, the linking difference serves as an indicator of the prevailing superhelical strain. The relative concentrations of both free topoisomers and complexes are plotted versus negative mean linking difference $(-\langle \Delta L(i)\rangle = -[i - (l_0 - \langle n_{2a}(C(i,\ n_{2a} \geq 1))\rangle)\Phi])$ in Fig. 10 for three different added SSB/plasmid ratios, $v = 0$, 3, and

10. Especially striking is the remarkably narrow distribution of mean linking differences of the complexes. Binding of SSB evidently acts to almost completely level the superhelical strain throughout the entire population of complexes, regardless of the original linking differences of the free topoisomers. This implies that all of the complexes for a given value of υ will exhibit practically the same distribution of tertiary structures. With increasing υ, the consensus value of $-\langle \Delta L(i) \rangle$ declines, and all of these tertiary structure distributions are expected to change in unison.

6.7. The change in tertiary structure manifested in D_0

The change in D_0 reflects a change in tertiary structure, but what is the origin of that? At least three rather different scenarios are conceivable.

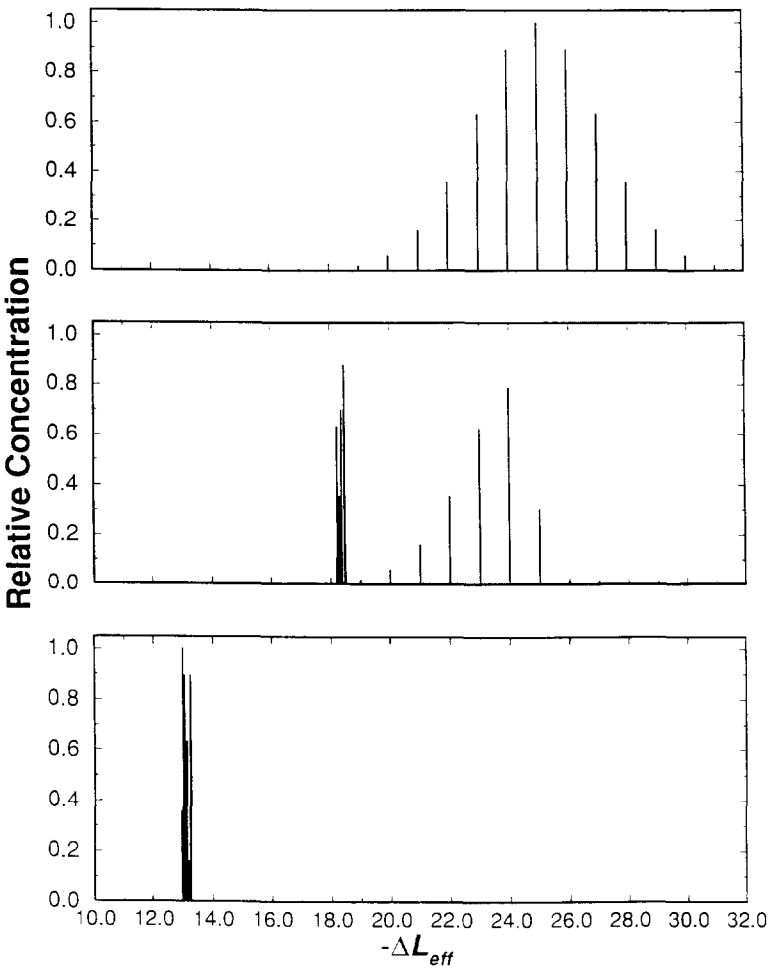

Fig. 10. Relative topoisomer concentration versus $-\Delta L_{\mathrm{eff}}$. The effective mean linking difference of a topoisomer in the presence of an unwinding ligand is defined by $\Delta L_{\mathrm{eff}} = \langle \Delta L(i) \rangle = [i - (l_0 - \langle n_{2a}(C(i),\ n_{2a} \geqslant 1) \rangle \Phi)]$. The top panel shows the initial distribution in the absence of ligand (i.e. for $\upsilon = 0.0$). This is the distribution given by Eq. (26) with $\Delta L(m) = -25$ turns. The middle panel shows the distribution for $\upsilon = 3.0$. About $\frac{1}{2}$ of the initial distribution has been converted to complexes which are centered near $\Delta L_{\mathrm{eff}} = -18$ turns. The bottom panel shows the distribution for $\upsilon = 10.0$. Essentially all of the DNAs exist as complexes with effective linking differences near $\Delta L_{\mathrm{eff}} = -13$ turns. Other parameters are $c_C^{\mathrm{tot}} = 16$ nM (plasmids), $l = 34$ bp, $\Phi = -3.269$ turns/bound SSB pair. $S = 7.45 \times 10^{-3}$, $E_{\mathrm{T}} = 1000$, $N = 4363$ bp, and $K_{\mathrm{a}} = 0.150\ \mathrm{M}^{-1}$.

In the first, D_0 is simply a function of linking difference, regardless of whether any SSB is bound or not. This scenario requires a substantial change in tertiary structure, when $-\langle \Delta L(i) \rangle$ drops below a particular threshold value. In order to account for the early onset and abruptness of the observed change in D_0, the threshold linking difference must be set no lower than about 5 turns below $\Delta L(m)$. However, evidence for an 18% step in D_0 that lies in or just below the range of native superhelix density is presently lacking [33]. In the second scenario, the site of the complex may provide a somewhat more flexible region that may stabilize a superhelix end or branch at that point and lead to a substantial reorganization of the tertiary structure, regardless of the linking difference. Evidence for this phenomenon is also presently lacking. In the third scenario, complex formation is accompanied by an extensive transition to an alternate secondary structure with reduced torsion constant, so that some of the writhe is absorbed into twist, the tertiary structure becomes more extended, and D_0 decreases. In fact, the torsion constant is found to be substantially reduced by SSB binding at $\upsilon = 15$, but whether it is similarly reduced at smaller υ values, when the negative mean linking difference is larger, is not known. Evidence that extensive changes in secondary structure can accompany ligand binding, as well as local sequence changes, is presented elsewhere [34]. Without additional information, the question of the origin of the tertiary structure change must be regarded as still open.

All three of the proposed scenarios would yield something akin to an isosbestic point in the DLS data, as observed [15].

6.8. Incomplete relaxation of superhelical strain

The present model and theory with the parameters adopted here relaxes only about half of the superhelical strain by $\upsilon = 10$. In contrast, the EM results indicate essentially complete loss of supercoils [4]. This disagreement might be attributed to various factors. A substantial decrease in torsion constant upon complex formation, as was actually observed, would result in the absorption of considerable writhe into twist, provided the bending constant either remained the same or increased. This would produce a substantial loss of visible supercoils. Another possibility is that the irreversible glutaraldehyde crosslinking procedure drives the binding to much higher levels than produced by the equilibrium binding constant K_a. Yet another possibility is that under the EM conditions the intrinsic twist is sufficiently reduced to remove the remaining superhelical turns. The existing discrepancy can only be resolved through further experimentation.

6.9. Effect of restricting binding to only part of the lattice

The EM data suggest that SSB binds preferentially to certain regions of the plasmid [4]. This is not surprising, since SSB would be expected to exhibit the highest affinity for the most easily melted regions, which are typically A–T rich. Although we cannot readily incorporate sequence-specific effects into our model, we can incorporate the possibility that the binding is restricted to a region of arbitrary size comprising only part of the larger circular lattice, as shown below.

If the restricted zone is N_r bp long, then the GPF(i) in Eq. (15) can be evaluated by (1) restricting the upper limit of the sum over n_{2a} to N_r/l, (2) evaluating $W(n_{2a})$ for a linear lattice of length N_r using the appropriate formulas [20,22] and (3) evaluating the free-energy change due to loss of superhelical strain according to Eq. (11) for the full circular lattice with $N = 4363$ bp, and $n^*_{2a} = -\Delta L(i)/\Phi$. Calculations for illustrative purposes are carried out for a single topoisomer with $\Delta L(i) = -24$ turns, $K_a = 10^{-4}$ M^{-1}, $S = 10^2$, $l = 34$ bp, $\Phi = 3.27$ turns, and $E_T = 1000$ at 20°C. The average number of bound SSB pairs *per complex* is plotted versus c_a in Fig. 11. Evidently, the binding *per complex* over this range of free ligand concentration is practically independent of the length of the restricted region from $N_r = 105$ to 2000 (and on up to 4363) bp, but is strongly perturbed when that length decreases to 100 bp. The reason is that the average number of bound SSB pairs per complex, which is essentially

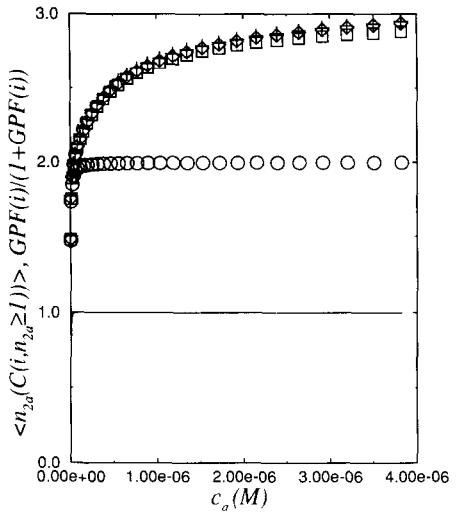

Fig. 11. Average number of bound ligands per complex $\langle n_{2a}(C(i), n_{2a} \geq 1) \rangle$ and $GPF(i)/[1+GPF(i)]$ versus free ligand concentration for various lengths N_r of the restricted binding zone in a 4363 bp topoisomer with $\Delta L(i) = -24$ turns. (○) $N_r = 100$; (□) $N_r = 105$; (◇) $N_r = 500$; and (+) $N_r = 2000$. The solid line indicates $GPF(i)/(1+GPF(i))$. In this case of restricted binding the calculation of $GPF(i)$ is modified, as described in the text. Note that $GPF(i)/[1+GPF(i)]$ is practically indistinguishable from 1.0 for all but the very smallest values of c_a. Hence, the average number bound per plasmid is equal to the average number bound per complex over most of the region displayed.

the average stack length, remains slightly less than 3.0 over this range of free ligand concentration. Since $3.0 \times 34 = 102$ bp, a stack of 3.0 bound SSB pairs can be accommodated in a restricted region containing $N_r \geq 105$ bp, but cannot be accommodated in one with only 100 bp. In the latter case, the binding is limited to 2.0 bound SSB pairs. Evidently, the length of the restricted region has little effect on the binding *per complex* up until it begins to limit the mean stack length. Although not visible in Fig. 11, restricting the length of the binding region in any way *does* affect the $GPF(i)$ and thereby alter the average number of bound SSB pairs *per plasmid*, which is defined by

$$\langle n_{2a}(i) \rangle = \langle n_{2a}(C(i, n_{2a} \geq 1)) \rangle c_{C(i, n_{2a} \geq 1)} / c^{tot}_{C(i)}$$
$$= \langle n_{2a}(C(i, n_{2a} \geq 1)) \rangle GPF(i) / (1 + GPF(i)). \quad (35)$$

Although the onset of binding *per plasmid* occurs at lower free ligand concentrations for longer restricted regions, the curves of $\langle n_{2a}(i) \rangle$ versus c_a for different lengths nonetheless all coalesce at larger free ligand concentrations, such that $GPF(i) \gg 1.0$ and $GPF(i)/(1 + GPF(i)) \approx 1.0$. Thus, except at extremely small c_a, the $\langle n_{2a}(C(i), n_{2a} \geq 1) \rangle$ in Fig. 11 may be identified with $\langle n_{2a}(i) \rangle$. Typically, $GPF(i) \gg 1.0$ for any free ligand concentration such that the average number of bound SSB pairs *per complex* ($\langle n_{2a}(C(i, n_{2a} \geq 1)) \rangle$) equals or exceeds 2.0, as illustrated in Fig. 11.

In summary, restricting SSB binding to a region comprising only part of the plasmid shifts the initial onset of binding *per plasmid* to a somewhat higher free ligand concentration, but has almost no effect on the binding at higher free ligand concentrations such that the average number of bound SSB pairs *per complex* exceeds 2.0, provided the length of the restricted region exceeds the maximum prevailing stack length. This maximum prevailing stack length generally rises with free ligand concentration, but eventually becomes limited by the length of the restricted region. At that point the binding *per complex* and *per plasmid*, which are practically identical, asymptotically approach the length of the restricted region instead of continuing their upward trajectory.

6.10. Potential functional significance

The unusual properties predicted for SSB binding to supercoiled DNAs by the present model suggest several potential biological functions for SSB. (1) The fact that SSB binds essentially always in a single stack would act to limit any open regions introduced or maintained by SSB to one per plasmid or one per superhelical domain. Thus, when SSB is involved in the initiation of replication, either unidirectional or bidirectional replication would be expected to begin always at only a single site in any given plasmid or superhelical domain, as appears to be the case [35]. Of course that site would also be localized to the most easily melted region, as observed [35]. If

for some reason synchronous initiation of replication in several different superhelical domains were required, then the action of SSB to level the effective superhelix density in different superhelical domains might be essential to insure comparable affinities of those domains for the (other) DNA replication proteins. (2) The apparent saturation of the binding over a limited span of free ligand concentration would insure that the amount of bound SSB is rather insensitive to fluctuations in SSB concentration. Under conditions where the effective superhelix density might be controlled by SSB, the fluctuations in effective superhelix density arising from fluctuations in bound SSB would then be effectively suppressed. Since the affinity of supercoiled DNAs for proteins with large unwinding angles is extremely sensitive to the prevailing superhelix density, this suppression of fluctuations in SSB binding and effective superhelix density might be important in some cases. (3) The remarkable narrowing of the distribution of linking differences and levelling of the superhelix densities in a population of topoisomers or superhelical domains by SSB binding would act to regularize the interactions of other unwinding proteins, such as (other) DNA replication proteins, with all members of the population. Properties (2) and (3) suggest that SSB (and other single strand binding proteins) could function not only to stably regulate the mean superhelical strain but also to strongly suppress fluctuations about that value. In the absence of some means to level the superhelix density, those topoisomers and superhelical domains with the higher superhelix densities would conscript the majority of the available DNA replication proteins, and would be reproduced at significantly higher rates. In cases where such differences between superhelical domains are required for the regulation of gene activity, the levelling action of SSB would have to be suppressed by either its absence or its inhibition. However, in cases where such differences between superhelical domains are detrimental to the orderly progression of events inside the cell, the action of SSB to level the superhelix density might be essential. Thus, SSB concentration (or activity) may be an important variable not only in controlling the mean superhelix density, but also in limiting the deviations of particular superhelical domains from that mean value.

Appendix A. Estimation of the electrostatic enhancement factor

SSB bears an electric charge, which is estimated below to be $n = +4.3$ charges. The value of K_1 measured for SSB binding to single-strand DNA contains the effects of electrostatic interactions between SSB and that single-strand. The binding of SSB to single-strand regions of a duplex DNA, as considered here, is expected to be greatly enhanced due to the much greater overall linear charge density and negative electrostatic potential of the duplex in comparison to a single strand. Thus, the binding constant K_1 needs to be corrected for electrostatic effects. This can be accomplished approximately by using a theory for the binding of multivalent ligands to polyelectrolytes due to Manning [36]. According to Manning's theory, the electrostatic binding constant for a ligand bearing n positive charges to a negatively charged polyelectrolyte with a structural charge spacing b is given by

$$K_{el} = V_p\left(\frac{1-\xi^{-1}}{V_p}\right)^n \tag{A.1}$$

wherein $\xi = L_B/b$.

$$L_B = \varepsilon_0^2/\varepsilon k_B T \tag{A.2}$$

is the Bjerrum length, and

$$V_P = 8\pi b^3(\xi - 1)N_{AV}/1000 \tag{A.3}$$

is the free volume occupied by the bound counterions. Eq. (A.3) follows from Manning's Eq. (29) in the limit that the multivalent ligand contributes negligibly to the overall charge compensation of the DNA. The binding of a positively charged ligand to duplex DNA is enhanced over binding to single-strand DNA by the factor

$$R = \frac{K_{el}^{ds}}{K_{el}^{ss}} = \frac{V_p^{ds}\left[\left(1-\xi_{ds}^{-1}\right)/V_p^{ds}\right]^n}{V_p^{ss}\left[\left(1-\xi_{ss}^{-1}\right)/V_p^{ss}\right]^n}, \tag{A.4}$$

where the labels ds and ss refer to double-strand and single-strand respectively. For a double-strand DNA, $b_{ds} = 1.7$ Å and $\xi_{ds} = 4.2$, whereas for a single-strand DNA, $b_{ss} = 4.0$ Å and $\xi_{ss} = 1.8$. The SSB charge is estimated from the salt dependence of the observed equilibrium constant. According to Manning, [36]

$$n = -\frac{\partial[\ln(K_0)]}{\partial[\ln(c_1)]}, \tag{A.5}$$

where c_1 is the concentration of univalent salt. From the data of Lohman et al. [28] and Krauss et al. [37] we find $n = 4.3$. When n and the values of b are substituted into Eq. (A.4) we find $R = 500$. Thus, SSB is expected to bind nearly 500 times more strongly to single strand regions of duplex DNA than to single strand DNA.

Appendix B. Estimation of the equilibrium constant for melting

Our objective is to estimate the equilibrium constant $\tilde{s}$ per base pair to form the melted state from duplex DNA at 20°C in 0.1 M NaCl. Unfortunately, the measured ΔH and ΔS for melting at T_M cannot be reliably extrapolated to 20°C. To avoid this long extrapolation, we make use of the enthalpy of melting of A_7U_7 at 298 K in 1.0 M NaCl, and correct that for the much smaller difference in temperature and the difference in salt concentration. At best, this procedure will suffer from a systematic error, since the resulting value of $\tilde{s}$ applies to an RNA rather than a DNA. In any case, we estimate $\Delta G = -RT \ln \tilde{s}$ by using the relation

$$\Delta G = \Delta H(293\ \text{K}, 0.1\ \text{M}) - (293)\Delta S(293\ \text{K}), \tag{B.1}$$

wherein the temperature and NaCl concentration in parentheses specify the conditions of our experiment. The available literature data apply at 298 K or T_M for 0.2 M or 1.0 M NaCl, so we first estimate $\Delta H(298, 0.2\ \text{M})$ and $\Delta S(298)$. We assume that the entropy change depends on T, but is essentially independent of salt concentration, and that the ratio of the enthalpies at two different salt concentrations is independent of temperature or kind of DNA, so

$$\Delta H(298\ \text{K}, 0.2\ \text{M}) = \Delta H(298, \text{K}, 1.0\ \text{M}) \times \frac{\Delta H(T_M, 0.2\ \text{M})}{\Delta H(T_M, 1.0\ \text{M})}, \tag{B.2}$$

where T_M is the melting temperature of pBR322 in 0.2 M NaCl. The problem now is to estimate the three quantities on the r.h.s. of (B.2). The enthalpy of melting for A_7U_7 at 298 K, 1.0 M NaCl is given in Table 1 of Wilcoxon and Schurr [38] whence $\Delta H(298\ \text{K}, 1.0\ \text{M}) = 7643$ cal/mol bp. For pBR322 we obtain $\Delta H(T_M, 0.2\ \text{M}) = T_M \Delta S(T_M, 0.2\ \text{M}) = 9000$ cal/mol bp by using the standard value, $\Delta S = 24.8$ cal/(K mol) bp for DNAs in 0.2 M NaCl in the melting region, and also the appropriate $T_M = 362.9$°C, which is estimated from the base-composition (50% GC) according to the formula of Marmur and Doty [39]. The corresponding enthalpy change in 1.0 M NaCl is estimated according to $\Delta H(T'_M, 1.0\ \text{M}) = T'_M\ \Delta S(T'_M)$, where T'_M is the melting temperature in 1.0 M NaCl. The increase in T_M on going from 0.2 M to 1.0 M NaCl is estimated to be $T'_M - T_M = 13.7$°C by using the relation $\partial(T_M/\partial \ln [\text{NaCl}]) = 8.5$ [40]. We also assume that, over this relatively small temperature range ΔH and ΔS remain constant, so $\Delta H(T_M, 1.0\ \text{M}) = \Delta H(T'_M, 1.0\ \text{M}) = T'_M\ \Delta S(T'_M) = T'_M \Delta S(T_M) = (362.9 + 13.7)\ (24.8) = 9340$ cal/mol bp. Collecting the preceding results into Eq. (B.2) yields $\Delta H(298\ \text{K}, 0.2\ \text{M}) = 7365$ cal/mol bp. Comparison with $\Delta H(298\ \text{K}, 1.0\ \text{M}) = 7643$ cal/mol bp shows that the correction for the difference in salt concentration is not large. We then calculate $\Delta S(298\ \text{K})$ by using the relation of Martin et al. [41] which applies to the data in Table 1 of Wilcoxon and Schurr [38].

$$\Delta G(298\ \text{K}, 1.0\ \text{M}) = \Delta H(298\ \text{K}, 1.0\ \text{M}) - 298[\Delta H(298\ \text{K}, 1.0\ \text{M})/353 + R \ln \beta], \tag{B.3}$$

where $R \ln \beta$ is the extra contribution that accompanies the pairing of the first two bases,

when separate strands form a duplex. We assume that $R \ln \beta$ contributes negligibly to ΔG(298 K, 1.0 M) for *long* DNAs. The second term in Eq. (B.3) represents the entropic contribution to ΔG(298 K, 1.0 M), namely $-298\ \Delta S(298\ \mathrm{K}) = -298\ \Delta H(298\ \mathrm{K}, 1.0\ \mathrm{M})/353$, whence $\Delta S(298\ \mathrm{K}) = 21.65$ cal/(K mol) bp. We again assume that ΔH and ΔS are independent of T over the small temperature range from 298 to 293 K, so $\Delta H(293\ \mathrm{K}, 0.2\ \mathrm{M}) = 7365$ cal/mol bp and $\Delta S(293\ \mathrm{K}) = 21.65$ cal/(K mol) bp. We now employ the analogous relation to (B.2), namely

$$\Delta H(293\ \mathrm{K}, 0.1\ \mathrm{M}) = \Delta H(293\ \mathrm{K}, 0.2\ \mathrm{M}) \times \frac{\Delta H(T_\mathrm{M}, 0.1\ \mathrm{M})}{\Delta H(T_\mathrm{M}, 0.2\ \mathrm{M})}, \tag{B.4}$$

to correct the enthalpy from 0.2 to 0.1 M NaCl. The only unknown quantity on the r.h.s. of Eq. (B.4) is $\Delta H(T_\mathrm{M}, 0.1\ \mathrm{M})$. As above, this is estimated according to $\Delta H(T_\mathrm{M}, 0.1\ \mathrm{M}) = T''_\mathrm{M} \times \Delta S(T_\mathrm{M})$, where T''_M is now the melting temperature in 0.1 M NaCl. From the relation cited above, we obtain $T''_\mathrm{M} - T_\mathrm{M} = -5.9$°C, and $\Delta H(T_\mathrm{M}, 0.1\ \mathrm{M}) = (362.9 - 5.9)(24.8) = 8854$ cal/mol bp. Use of the values estimated above in (B.4) yields $\Delta H(293\ \mathrm{K}, 0.1\ \mathrm{M}) = 7245$ cal/mol bp. Finally, $\Delta G(293\ \mathrm{K}, 0.1\ \mathrm{M}) = 7245 - 293(21.65) = 902$ cal/mol bp. Hence $\tilde{s} = \exp[-\Delta G(293\ \mathrm{K}, 0.1\ \mathrm{M})RT] = 0.213$.

Acknowledgement

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

References

[1] N. Sigal, H. Delius, T. Kornberg, T. Gefter and B. Alberts, Proc. Natl. Acad. Sci. USA 12 (1972) 3537–3541.
[2] W.T. Ruyechan and T.G. Wetmur, Biochemistry 25 (1975) 7799–7802.
[3] J.H. Weiner, L.L. Bertsch and A. Kornberg, J. Biol. Chem. 250 (1975) 1972–1980.
[4] G.C. Glikin, G. Gargiulo, L. Rena-Descalzi and A. Worcel, Nature 303 (1983) 770–774.
[5] A. Sancar, K.R. Williams and J.W. Chase, Proc. Natl. Acad. Sci. USA 78 (1981) 4274–4278.
[6] W. Bujalowski and T.M. Lohman, J. Mol. Biol. 217 (1991) 63–74.
[7] W. Bujalowski and T.M. Lohman, J. Mol. Biol. 197 (1987) 897–907.
[8] T.M. Lohman, L.B. Overman and S. Datta, J. Mol. Biol. 187 (1988) 603–615.
[9] T.M. Lohman and L.B. Overman, J. Biol. Chem. 260 (1985) 3594–3603.
[10] T.-F. Wei, W. Bujalowski and T.M. Lohman, Biochemistry 31 (1992) 6166–6174.
[11] R.R. Meyer, J. Glassberg and A. Kornberg, Proc. Natl. Acad. Sci. USA 76 (1979) 1702–1705.
[12] J. Glassberg, R.R. Meyer and A. Kornberg, J. Bacteriol. 140 (1979) 14–19.
[13] J.W. Chase and K.R. Williams, Ann. Rev. Biochem. 55 (1986) 103–136.
[14] K. Muniyappa, S.L. Shaner, S.S. Tsang and C.M. Radding, Proc. Natl. Acad. Sci. USA 81 (1984) 2757–2761.
[15] J. Langowski, A.S. Benight, B.S. Fujimoto and J.M. Schurr, Biochemistry 24 (1985) 4022–4028.
[16] P. Wu, L. Song and J.M. Schurr, Biopolymers 29 (1990) 1211–1232.
[17] P. Wu, B.S. Fujimoto, L. Song and J.M. Schurr, Biophys. Chem. 41 (1991) 217–236.
[18] P. Wu, L. Song, J.B. Clendenning, B.S. Fujimoto, A.S. Benight and J.M. Schurr, Biochemistry 27 (1988) 8128–8144.
[19] J.B. Clendenning and J.M. Schurr, Biopolymers 34 (1994) 849–868.
[20] J.B. Clendenning, A.N. Naimushin, B.S. Fujimoto, D.W. Stewart and J.M. Schurr, Biophys. Chem. 52 (1994) 191.
[21] A.N. Naimushin, J.B. Clendenning, U.S. Kim, L. Song, B.S. Fujimoto, D.W. Stewart and J.M. Schurr, Biophys. Chem. 52 (1994) 219.
[22] J. Reiter and I.B. Epstein, J. Phys. Chem. 91 (1989) 4813–4820.
[23] W. Hillen, T.C. Goodman, A.S. Benight, R.M. Wartell and R.D. Wells, J. Biol. Chem. 156 (1981) 2761–2766.
[24] R.M. Wartell and A.S. Benight, Phys. Rept. 126 (1985) 67–107.
[25] A.S. Benight, J.M. Schurr, P.F. Flynn, B.R. Reid and D.E. Wemmer, J. Mol. Biol. 200 (1988) 377–399.
[26] J.D. Griffith, L.D. Harris and J. Register, Cold Spring Harbor Symp. Quant. Biol. 49 (1984) 553–559.
[27] D.S. Horowitz and J.C. Wang, J. Mol. Biol. 173 (1984) 75–94.
[28] D. Shore and R.L. Baldwin, J. Mol. Biol. 170 (1983) 983–1007.
[29] R.E. Depew and J.C. Wang, Proc. Natl. Acad. Sci. USA 72 (1975) 4275–4279.
[30] D.E. Pulleyblank, M. Shure, D. Tang, J. Vinograd and H.P. Vosberg, Proc. Natl. Acad. Sci. USA 72 (1975) 4280–4284.

[31] W. Bauer, Ann. Rev. Biophys. Bioeng 1 (1978) 287–313.
[32] J.B. Clendenning, Ph.D. Thesis, University of Washington (1993).
[33] L. Song, Ph.D. Thesis, University of Washington (1989).
[34] U.S. Kim, B.S. Fujimoto, C.E. Furlong, C.E. Sundstrom, R. Humbert, D.C. Teller and J.M. Schurr, Biopolymers 33 (1993) 1725–1745.
[35] A. Kornberg, DNA replication, (Freeman, New York, 1980) pp. 352–357.
[36] G.S. Manning, Quart. Rev. Biophys. 11 (1978) 179–246.
[37] G. Krauss, H. Sindermann, U. Schomburg and G. Maas, Biochemistry 20 (1981) 5346–5342.
[38] J. Wilcoxon and J.M. Schurr, Biopolymers 22 (1983) 2273–2321.
[39] J. Marmur and P. Doty, J. Mol. Biol. 5 (1962) 109–118.
[40] M.T. Record Jr, Biopolymers 14 (1975) 2137–2158.
[41] C.F. Martin, O.C. Uhlenbeck and P. Doty, J. Mol. Biol. 57 (1971) 201–215.