Entropy Loss in Long-Distance DNA Looping

Andreas Hanke*[†] and Ralf Metzler[‡§]

*Institute for Theoretical Physics, University of Stuttgart, Stuttgart, Germany; [†]Department of Physics–Theoretical Physics, University of Oxford, Oxford, United Kingdom; [‡]NORDITA–Nordic Institute for Theoretical Physics, Copenhagen, Denmark; and [§]Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts USA

ABSTRACT   The entropy loss due to the formation of one or multiple loops in circular and linear DNA chains is calculated from a scaling approach in the limit of long chain segments. The analytical results allow us to obtain a fast estimate for the entropy loss for a given configuration. Numerical values obtained for some examples suggest that the entropy loss encountered in loop closure in typical genetic switches may become a relevant factor in comparison to both $k_B T$ and typical bond energies in biopolymers, which has to be overcome by the released bond energy between the looping contact sites.

INTRODUCTION

Gene expression in all organisms comprises the transcription of a certain gene on the DNA into messenger RNA through RNA polymerase starting from the promoter site, and its subsequent translation into a protein. The initiation of the transcription at a specific gene underlies a subtle cooperative scheme of transcription factors, which in turn is determined by a given set of boundary conditions such as the concentration of the transcription factors. Transcription factors often act cooperatively, and they are known to interact with each other over distances of several thousand basepairs (bp). This interaction is effected through DNA looping (Alberts et al., 1994; Blackwood and Kadonaga, 1998; Bolsover et al., 2001; Ptashne and Gunn, 2002; Révet et al., 1999; Snustad and Simmons, 2003), compare Fig. 1.

A typical example for DNA looping is found in the genetic switch which determines whether the replication of bacteriophage λ in *Escherichia coli* follows either the lysogenic or the lytic pathway (Ptashne, 1992; Ptashne and Gunn, 2002; Snustad and Simmons, 2003). A key component of this λ-switch is the λ-repressor which activates the expression of a gene that encodes the production of the λ-repressor itself. λ-repressor can bind to the three operator sites O_R which overlap the two promoter sites of the switch. λ-repressor binds cooperatively as a dimer, and typically under stable lysogenic conditions two such dimers on O_R form a tetramer, the next higher order of cooperativity, which is the main factor for the stability of the λ-switch against noise (Aurell and Sneppen, 2002; Aurell et al., 2002; Metzler, 2001). However, λ-repressor can also bind to the very similar operator O_L, which is located roughly 2300 bp away and not part of the λ-switch. It has been found that the two λ-repressor tetramers at O_L and O_R synergistically form an octamer through DNA looping. This higher-ordered oligomerization enhances the performance of the switch considerably (Amouyal et al., 1998; Bell and Lewis, 2001;

Bell et al., 2000; Blackwood and Kadonaga, 1998; Ptashne and Gunn, 2002; Révet et al., 1999; Semsey et al., 2002; Xu and Hoover, 2001). The specific binding along the tetramer-tetramer interface has recently been revealed through crystallographic structure determination (Bell and Lewis, 2001; Bell et al., 2000). Similar realizations of DNA looping also occur in linear DNA, naturally in the form of telomeres or in vitro in engineered DNA (De Bruin et al., 2001; Griffith et al., 1999; Zaman et al., 2002, compare Fig. 1). Multiple looping in large DNA molecules around a locus can be observed in vivo and can be induced in vitro by introducing of specific binding zones on the DNA, which leads to a considerable reduction of the gyration radius of the molecule such that it can be more easily transferred into, for example, mammalian cells (Montigny et al., 2001).

DNA looping often involves large loop sizes of several thousand bp. Therefore, the formation of these loops causes a non-negligible entropy loss which has to be overcome by the binding energy released at the bond formation on loop closure. In the present study, we quantify this entropy loss for such long DNA loops, taking into account self-avoiding effects due to both the monomer-monomer interaction within the loop and the additional effects due to the higher order contact points (vertices) at the loop closure site. The resulting numbers for typical systems suggest that the entropy loss is a relevant factor in the formation of DNA loops in comparison to the thermal energy and typical bond energies found in DNA nucleotides and other biopolymers, and it gives a lower bound for the bond-forming energy required to stabilize the loop.

Entropy loss due to loop formation was studied for the case of disconnected loops by Schellmann (1955), Flory (1956), and Semlyen (1997). In their seminal article, Poland and Scheraga (1965), and Wang and Uhlenbeck (1955), considered coupled Gaussian loops. To our knowledge the full effect of self-avoidance in the DNA looping network has not been considered before. Hereby, the contributions of non-trivial vertices turns out to be a relevant factor, and for multiple looping with a common locus actually become the dominating contribution. The analytical results presented here are derived from a scaling approach for general polymer
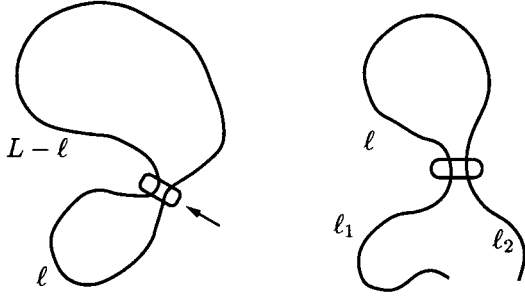

FIGURE 1 DNA looping in a circular (*left*) and linear DNA (*right*). The rounded boxes indicate the chemical bonds established between the transcription factors through looping at specific contact sites on the DNA double-helix, which are fairly distant from one another in terms of the arc length along the DNA. A telomere loop corresponds to the right configuration with vanishing ℓ_1 [or ℓ_2].

networks and provide the advantage that, on their basis, estimates for the entropy loss in a given DNA system can be computed in a straightforward manner. It should also be noted that the additional vertex effects studied herein may be crucial in the analytical treatment of the DNA looping *dynamics*, as the higher-order self-interaction at such vertices poses an additional barrier in the loop closure process (Lee et al., 1976; Merlitz et al., 1998). Our results for long DNA with large loops complement the investigations of the bending and twisting energies in small DNA plasmids (Coleman et al., 2000; Tobias et al., 2000). In the case of intermediate-sized DNA segments, both approaches may be combined.

In what follows, we calculate the scaling results for the entropy loss on looping for the three different cases: 1), looping in a circular DNA; 2), looping in a linear DNA; and 3), multiple looping in a circular DNA. In the Appendix, the general expressions for calculating the system entropy of an arbitrary polymer network are compiled so that the entropy loss for different configurations can be calculated according to the general procedure developed below.

LOOPING IN A CIRCULAR DNA CHAIN

As stated before, we consider the limit in which each segment of the looped DNA, e.g., both subloops created in the circular DNA upon looping, are long in comparison to the persistence length ℓ_p of the double-stranded DNA chain. (It can be assumed that this criterion is fulfilled if the segment contains more than ten persistence lengths.) In this long chain limit, we can neglect energetic effects due to bending or twisting, such that we treat the DNA as a flexible self-avoiding polymer. Therefore, we can employ results for the configuration number of a general polymer network, which we briefly review in the Appendix.

Before looping, the free energy of the circular DNA of total length L is given by

$$F_{circ} = H_0 - TS_{circ}, \tag{1}$$

where H_0 combines all binding enthalpies in the macromol-

ecule and the entropy $S_{circ} = k_B \ln \omega_{circ}$ is determined by the number of configurations (see De Gennes, 1979; see also the Appendix, this article)

$$\omega_{circ} = A_{circ} \mu^L L^{-3\nu}, \tag{2}$$

of a simply connected ring polymer of length L in units of the monomer length. The latter can be estimated by the persistence length ℓ_p of the polymer (~500 Å for double-stranded DNA corresponding to 100 bp; see Marko and Siggia, 1996). In Eq. 2, A_{circ} is a nonuniversal amplitude, μ is the support dependent connectivity constant, and $\nu \simeq 0.588$ (Guida and Zinn-Justin, 1998) is the Flory exponent. Thus, S_{circ} has the form

$$S_{circ} = k_B(\ln A_{circ} + L \ln \mu - 3\nu \ln L). \tag{3}$$

On looping, as sketched in Fig. 1 to the left, the circular DNA is divided into two subloops of lengths ℓ and $L - \ell$ by creation of a vertex at which four legs of the chain are bound together. For a self-avoiding chain, the number of configurations of the resulting figure-eight shape (Metzler et al., 2002a,b) is not simply the product of the configuration numbers of the two created loops, but has the more complicated form (Duplantier, 1986, 1989; Ohno and Binder, 1988; Schäfer et al., 1992; see also the Appendix, this article),

$$\omega_8 = A_8 \mu^L (L - \ell)^{-6\nu+\sigma_4} \mathcal{Y}_8\left(\frac{\ell}{L - \ell}\right). \tag{4}$$

In this expression, A_8 is a nonuniversal amplitude, $\mathcal{Y}_8$ is a universal scaling function, and $\sigma_4 \simeq -0.48$ is a universal exponent associated with the vertex with four outgoing legs. Note that in the Gaussian chain limit, the exponents σ_N vanish; as we are going to show, the inclusion of the additional effects due to the higher order vertex formation reflected by nonzero values for σ_N are non-negligible. Given the entropy $S_8 = k_B \ln \omega_8$ of the figure-eight configuration, the entropy loss suffered from creating this configuration out of the original circular DNA amounts to $|S_8 - S_{circ}|$. To proceed, we now evaluate the scaling function $\mathcal{Y}_8(x)$ in some special cases, and calculate typical numbers for the required entropy loss compensation. Two limiting cases can be distinguished.

Limiting case 1

If one of the loop sizes is much smaller than the other ($\ell \ll L - \ell$, say), the big loop of size $L - \ell$ will essentially behave like a free circular chain so that its contribution to ω_8 will scale like a regular ring polymer, i.e., like $(L - \ell)^{-3\nu}$. Consequently, we find the behavior $\mathcal{Y}_8(x) = ax^{-3\nu+\sigma_4}$ for $x \ll 1$, where a is a universal amplitude, and therefore (Hanke and Metzler, 2002; Metzler et al., 2002a,b),

$$\omega_8 = A_8 a \mu^L (L - \ell)^{-3\nu} \ell^{-3\nu + \sigma_4}. \tag{5}$$

In this case, the free energy difference between the initial circular and the looped states becomes

$$\Delta F = \Delta H_{\text{bond}} - T(S_8 - S_{\text{circ}}), \tag{6}$$

where ΔH_{bond} is the binding enthalpy at the loop closure site. The formation of the looping bond has to release a higher enthalpy than what is lost in entropy, i.e., $\Delta H_{\text{bond}} < T(S_8 - S_{\text{circ}})$. Collecting the different expressions, we thus find the condition

$$\Delta H_{\text{bond}} < k_B T \left[\ln \frac{A_8 a}{A_{\text{circ}}} + 3\nu \ln \frac{L}{\ell(L-\ell)} + \sigma_4 \ln \ell \right]. \tag{7}$$

In this expression (and in similar expressions below), the first term in the square brackets is nonuniversal and depends on details of the model (for a Gaussian random walk on a cubic lattice, $\ln(A_8 a / A_{\text{circ}}) = \ln(A_8 \mathcal{Y}_8(1)/A_{\text{circ}}) = \ln[3/(2\pi)] \simeq -0.74$, is small compared with the other contributions in Eqs. 7 and 9; for the cases discussed below, we give the corresponding numbers in Table 1), whereas the remaining contributions are universal (apart from the fact that L is measured in units of the nonuniversal monomer length).

To get an estimate for the magnitude of the entropy loss, consider the case of the λ-repressor loop in *E. coli*. With the size of the entire DNA of $\sim 3.5 \times 10^3$ kbp and the looping branch of ~ 2.3 kbp, the two loops correspond to 3.5×10^4 and 23 monomers, respectively (each monomer corresponds to a persistence length ℓ_p of 100 bp; see above). Neglecting the nonuniversal first term in brackets in Eq. 7, these numbers produce

$$\Delta H_{\text{bond}} < -7.0 \, k_B T = -17.5 \, \text{kJ/mol} = -4.2 \, \text{kcal/mol}; \tag{8}$$

here and in the following examples, we choose $T = 300$ K and make use of the gas constant, $R = 8.31 \, \text{JK}^{-1} \, \text{mol}^{-1}$, and the conversion factor 1 cal $= 4.2$ J (Abramowitz and Stegun, 1972). Eq. 8 gives a considerable minimal value for the required bond energy between the two looping sites. For comparison, the typical free energy for basepair formation in

DNA is 8 kcal/mol for AT pairs and 13 kcal/mol for GC pairs (Breslauer et al., 1986). Thus, even for the relatively small loop of 23 monomers, the required enthalpy release is non-negligible. Note that the relative contribution stemming from the σ_4 term in Eq. 7 amounts to $\sim 20\%$ of the required enthalpy. In the Conclusion section, we compare the results obtained in the text with the corresponding ones in the Gaussian chain limit, demonstrating the significant increase of $|\Delta H_{\text{bond}}|$ if self-avoidance is taken into consideration (see Table 1).

Limiting case 2

If the two created loops are of comparable size, i.e., $x = \ell / (L - \ell) \approx 1$, the corresponding value of the scaling function $\mathcal{Y}_8(x)$ is a finite number. For example, for $\ell = L/2$ one finds

$$\Delta H_{\text{bond}} < k_B T \left[\ln \frac{A_8 \mathcal{Y}_8(1)}{A_{\text{circ}}} + \sigma_4 \ln \frac{L}{2} - 3\nu \ln \frac{L}{4} \right]. \tag{9}$$

In a modified DNA with two loops of 2.3 kbp each, one finds a bond enthalpy requirement of

$$\Delta H_{\text{bond}} < -5.8 \, k_B T = -14.5 \, \text{kJ/mol} = -3.4 \, \text{kcal/mol}, \tag{10}$$

where we again neglect the nonuniversal first term in the square brackets. If both loops are of size 2×10^3 kbp each, the required bond enthalpy would increase to $\Delta H_{\text{bond}} < -12.5$ kcal/mol.

LOOPING IN A LINEAR DNA CHAIN

A linear chain of length L can assume

$$\omega_{\text{lin}} = A_{\text{lin}} \mu^L L^{\gamma - 1} \tag{11}$$

distinct configurations, where A_{lin} is a nonuniversal amplitude and $\gamma \simeq 1.16$ is a universal exponent (Caracciolo et al., 1998; Guida and Zinn-Justin, 1998). If looping occurs and produces the A-shape in Fig. 1 to the right, the configuration number is modified to

$$\omega_A = A_A \mu^L (L - \ell)^{\gamma - 1 - 3\nu + \sigma_4} \mathcal{Y}_A \left(\frac{\ell}{L - \ell}, \frac{\ell_1}{\ell_2} \right), \tag{12}$$

where ℓ is the size of the loop, and ℓ_1 and ℓ_2 are the sizes of the two loose end-segments, respectively.

We distinguish four different cases belonging to two groups: the configuration with $\ell_1 \approx \ell_2$, and the telomere configuration for which $\ell_1 = 0$ (or $\ell_2 = 0$).

Case 1

If $\ell_1 = \ell_2$, we find

$$\omega_A = A_A \mu^L (L - \ell)^{\gamma - 1 - 3\nu + \sigma_4} \mathcal{W}_A \left(\frac{\ell}{L - \ell} \right), \tag{13}$$

where $\mathcal{W}_A(x) = \mathcal{Y}_A(x, 1)$. If furthermore $\ell \ll L - \ell$, an anal-

TABLE 1

Equation	$-\Delta H_{\text{bond}}^{\text{SAW}}/(k_B T)$	$-\Delta H_{\text{bond}}^{\text{Gauss}}/(k_B T)$	$|\Delta H^{\text{NU}}|/(k_B T)$
7, 8	7.0	4.7	0.74
9, 10	5.8	3.7	0.74
15, 16	7.0	4.7	0.74
18, 19	8.6	5.7	0.3
22, 23	6.3	4.7	0.74
25, 26	15.8	11.6	0.74
28, 29	67.4	38.8	2.2

Comparison of the calculated minimum bond energies in the fully self-avoiding (*SAW*) and in the Gaussian ($\nu = 1/2$, $\gamma = 1$, and $\sigma_N = 0$) cases: the increase due to self-avoiding effects is distinct. We also compare to the (negligible) nonuniversal (*NU*) contributions (first term in square brackets in the cited equations for ΔH_{bond}), which are calculated for a Gaussian random walk on a cubic lattice.

ogous reasoning as in the limiting case 1 of the previous section, leads to

$$\omega_A = A_A b \mu^L (L - \ell)^{\gamma - 1} \ell^{-3\nu + \sigma_4}, \tag{14}$$

where b is a universal number. The fact that ℓ carries the same exponent as in Eq. 5 is due to the local effect of self-interaction for the small loop; in both cases, the small loop is connected to a 4-vertex.

For the binding enthalpy, we obtain the condition

$$\Delta H_{bond} < k_B T \left[\ln \frac{A_A b}{A_{lin}} + (\gamma - 1) \ln \frac{L - \ell}{L} - (3\nu - \sigma_4) \ln \ell \right]. \tag{15}$$

To obtain a numerical value, consider the λ-repressor loop of 23 monomers and the *E. coli* DNA length of 3.5×10^4 monomers, a configuration which can be obtained by cutting the *E. coli* DNA. Neglecting the (nonuniversal) first term in the square brackets, we find in this case

$$\Delta H_{bond} < -7.0 \, k_B T = -17.5 \, \text{kJ/mol} = -4.2 \, \text{kcal/mol}, \tag{16}$$

where the exact numerical value is slightly smaller than in Eq. 8.

Case 2

Conversely, if $\ell = \ell_1 = \ell_2$, the simpler expression

$$\omega_A = A_A \mathcal{W}_A \left(\frac{1}{2} \right) \mu^L \left(\frac{2L}{3} \right)^{\gamma - 1 - 3\nu + \sigma_4} \tag{17}$$

emanates, and the binding enthalpy has to fulfill

$$\Delta H_{bond} < k_B T \left[\ln \frac{A_A \mathcal{W}_A(1/2)}{A_{lin}} - (\gamma - 1) \ln \frac{2}{3} \right.$$
$$\left. - (3\nu - \sigma_4) \ln \frac{2L}{3} \right]. \tag{18}$$

Taking 23 monomers for each segment and neglecting the (nonuniversal) first term in the square brackets yields the condition

$$\Delta H_{bond} < -8.6 \, k_B T = -21.6 \, \text{kJ/mol} = -5.1 \, \text{kcal/mol} \tag{19}$$

for the binding energy. If the segments are larger by a factor of 100, this value gets modified to $\Delta H_{bond} < -11.3 \, \text{kcal/mol}$.

Cases 3 and 4

The next two cases belong to the telomere configuration corresponding to Fig. 1 (*right*) with $\ell_1 = 0$ and $\ell_2 = L - \ell$.

Case 3

This case involves a 3-vertex instead of a 4-vertex, and has only one loose end-segment. The number of configurations the telomere configuration can assume is

$$\omega_{telo} = A_{telo} \mu^L (L - \ell)^{-3\nu + \sigma_3 + \sigma_1} \mathcal{X}_{telo} \left(\frac{\ell}{L - \ell} \right), \tag{20}$$

where $\sigma_3 \simeq -0.18$ and $\sigma_1 = (\gamma - 1)/2 \simeq 0.08$ (see the Appendix). We note that in Eq. 20, the explicit occurrence of σ_1 is due to the existence of only one loose end. In the two cases (case 1 and case 2, this section) considered above, we encounter two loose ends, and the σ_1 can be incorporated into $\gamma = 1 + 2\sigma_1$.

Let us first calculate the entropy loss in the small loop limit $\ell \ll L - \ell$. Here, the linear chain part should essentially behave like a simple linear chain, which implies that $\mathcal{X}_{telo}(x) = c x^{-3\nu + \sigma_3 - \sigma_1}$ for $x \ll 1$, and thus

$$\omega_{telo} = A_{telo} c \mu^L (L - \ell)^{\gamma - 1} \ell^{-3\nu + \sigma_3 - \sigma_1}, \tag{21}$$

where c is a universal number.

The corresponding condition for the bond enthalpy reads

$$\Delta H_{bond} < k_B T \left[\ln \frac{A_{telo} c}{A_{lin}} + (\gamma - 1) \ln \frac{L - \ell}{L} \right.$$
$$\left. - \left(3\nu - \sigma_3 + \frac{\gamma - 1}{2} \right) \ln \ell \right]. \tag{22}$$

Taking a loop of 2.3 kbp in a chain of length 3500 kbp and neglecting the (nonuniversal) first term in the square brackets gives

$$\Delta H_{bond} < -6.3 \, k_B T = -15.8 \, \text{kJ/mol} = -3.8 \, \text{kcal/mol}. \tag{23}$$

For comparison, if the loop size is 230 kbp, this value is increased to $\Delta H_{bond} < -9.3$ kcal/mol.

Case 4

If the loop size and the linear chain segment are of equal size, $\ell = L/2$, the configuration number becomes

$$\omega_{telo} = A_{telo} \mathcal{X}_{telo}(1) \mu^L \left(\frac{L}{2} \right)^{-3\nu + \sigma_3 + \sigma_1}, \tag{24}$$

and we obtain the condition

$$\Delta H_{bond} < k_B T \left[\frac{A_{telo} \mathcal{X}_{telo}(1)}{A_{lin}} - (3\nu - \sigma_3) \ln \frac{L}{2} \right.$$
$$\left. - \frac{\gamma - 1}{2} \ln (2L) \right]. \tag{25}$$

Taking a chain length of 460 kbp and neglecting the (nonuniversal) first term in the square brackets we find

$$\Delta H_{bond} < -15.8 \, k_B T = -39.3 \, \text{kJ/mol} = -9.4 \, \text{kcal/mol}. \tag{26}$$

MULTIPLE LOOPING IN A CIRCULAR DNA CHAIN

Assume that m potential connector points are distributed evenly along a circular DNA chain of total length L. If these condense to form a common locus, a number m of loops of equal size are created which are held together at this locus, as sketched in Fig. 2 (Montigny et al., 2001). This creates, in the scaling limit, a high-order vertex where $2m$ legs are joined. The procedure for the configuration number for this locus configuration yields

$$\omega_{\text{locus}} = A_{\text{locus}} \mu^{\text{L}} \left(\frac{L}{m} \right)^{-3m\nu + \sigma_{2m}}, \qquad (27)$$

where the universal exponent σ_{2m} is associated with a vertex with $2m$ outgoing legs (see the Appendix). It should be noted that this result holds true only if the size of the locus is much smaller than the sizes of the created loops (Metzler et al., 2002a,b).

Due to the assumption that all m loops are of the same size, we immediately arrive at

$$\Delta H_{\text{bond}} < k_B T \left[\ln \frac{A_{\text{locus}}}{A_{\text{circ}}} + 3\nu(1-m) \ln L \right.$$
$$\left. + 3m\nu \ln m + \sigma_{2m} \ln \frac{L}{m} \right]. \qquad (28)$$

The absolute value of σ_{2m} increases rapidly with increasing m, and can be determined from Padé or Padé-Borel analysis as shown in Schäfer et al. (1992). We list the topological exponents up to order 8 in the Appendix. Taking a circular chain of 3500 kbp and $m = 4$, and neglecting the (non-universal) first term in the square brackets, we find that the entropy loss is fairly high (using $\sigma_8 = -2.4$),

$$\Delta H_{\text{bond}} < -67.4\,k_B T = -168\,\text{kJ/mol} = -40.1\,\text{kcal/mol}. \qquad (29)$$
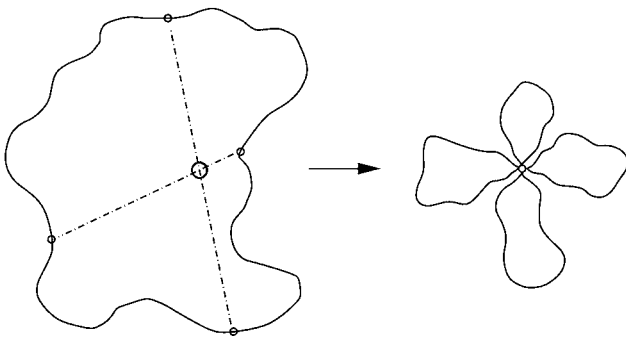

FIGURE 2 DNA loop condensation. The circles in the original DNA double-helix denote likely contact points. Formation of bonds between these contacts with one common agglomeration center, as indicated by the dashed lines, result in the locus configuration on the right. Note the reduction in the gyration radius during this process. A higher-order vertex is created at the locus point (Montigny et al., 2001).

In this case, the contribution due to the σ_8 term is as large as 50% of the total entropy loss.

CONCLUSIONS

We have presented an analytical method to estimate the entropy loss in different scenarios of DNA looping in the limit of long segments. This approach takes explicitly the self-avoidance and interacting nature of the formed loops and other segments into account, and considers the additional effect of vertex formation, i.e., the effective interaction between different segments at the point where they are joined. This is possible via the scaling theory for arbitrary polymer networks derived by Duplantier. The obtained numbers do not vary much, due to the logarithmic dependence on the segment sizes. However, they are all non-negligible, and therefore have to be compensated by the released bond energy on formation of the DNA loop. We noted that the entropy loss is of the same order or close to the bond melting energy required for splitting an AT or GC bond, i.e., a considerable amount. Moreover, it is to be expected that the vertex effect increases the characteristic bond formation times in analytical approaches which are based on the free energy.

In Table 1 we summarize the obtained estimates for the minimum bond energy ΔH_{bond} in the present self-avoiding model, and compare them with the corresponding results in the Gaussian limit. The differences are significant, clearly showing the much higher entropy loss in the self-avoiding case. In Table 1 we also show estimates for the corresponding nonuniversal contributions, which we neglected in the explicit calculation of ΔH_{bond}.

Our calculations are valid in the long chain limit. In units of the monomer size of a typical DNA double-helix persistence length $\ell_p \sim 100$ bp, a minimum number of at least 10 monomers is expected to be required to consider a segment in the final structure flexible. For shorter segments, additional effects due to bending and twisting energy are expected to become relevant. As the mentioned examples document, there are numerous systems, both in vivo and in vitro, in which the flexibility condition is easily fulfilled, and in which our estimation method for the entropy loss becomes fully applicable. The persistence length of single-stranded DNA and RNA is much shorter, typically taken to be of the order $\ell_p \sim 8$ bases. Thus, in single strand looping experiments the expected entropy loss will be considerably larger.

APPENDIX A: CONFIGURATION EXPONENTS FOR A GENERAL POLYMER NETWORK

A general polymer network $\mathcal{G}$ like the one depicted in Fig. 3 consists of vertices which are joined by $\mathcal{N}$ chain segments of lengths $s_1, \ldots, s_{\mathcal{N}}$. Their total length is $L = \sum_{i=1}^{\mathcal{N}} s_i$. In the scaling limit $s_i \gg 1$, the number of configurations of such a network is given by Duplantier (1986, 1989), Ohno and Binder (1988), and Schäfer et al. (1992) as
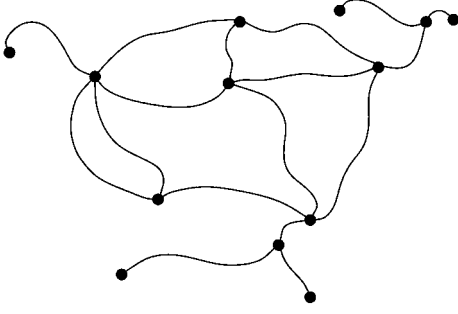
FIGURE 3 Polymer network $\mathcal{G}$ with vertices (●) of different order ($n_1 = 5$, $n_3 = 4$, $n_4 = 3$, and $n_5 = 1$).

$$\omega_{\mathcal{G}} = A_{\mathcal{G}} \mu^L s_{\mathcal{N}}^{\gamma_{\mathcal{G}}-1} \mathcal{Y}_{\mathcal{G}} \left(\frac{s_1}{s_{\mathcal{N}}}, \ldots, \frac{s_{\mathcal{N}-1}}{s_{\mathcal{N}}} \right), \qquad \text{(A1)}$$

where $A_{\mathcal{G}}$ is a nonuniversal amplitude, μ is the effective connectivity constant for self-avoiding walks, and $\mathcal{Y}_{\mathcal{G}}$ is a scaling function. The topology of the network is reflected in the configuration exponent

$$\gamma_{\mathcal{G}} = 1 - 3\nu\mathcal{L} + \sum_{N \geq 1} n_N \sigma_N. \qquad \text{(A2)}$$

$\mathcal{L} = \sum_{N \geq 1} (N-2) n_N / 2 + 1$ is the Euler number of independent loops, n_N is the number of N-vertices, and σ_N is an exponent connected to an N-vertex. Thus, Eq. A1 generalizes the familiar form $\omega \sim \mu^L L^{\gamma-1}$ of a linear polymer chain. The numerical values we use in the text are given in Table 2 for the topological exponents σ_N; furthermore, we employ $\nu = 0.588$ and $\gamma \simeq 1.16$ (Caracciolo et al., 1998; Guida and Zinn-Justin, 1998). We also make use of the relation $\gamma = 2\sigma_1 + 1$.

Note that in this work we consider the *entropy* of a given polymer network, in which enters the total number of physically distinct configurations. Two configurations are considered distinct if they cannot be superimposed by translation. In particular, the monomers of the chain are distinguishable. For a simple ring of length L this implies that two configurations are distinct even if they have the same trajectory, but differ from each other by a reptation (translation of the chain within the trajectory) by a noninteger multiple of L. The number of configurations of the simple ring is therefore (Duplantier, 1986, 1989),

$$\omega_{\text{circ}} = \tilde{\omega} L \sim L^{-3\nu}, \qquad \text{(A3)}$$

where $\tilde{\omega} \sim L^{-3\nu-1}$ is the number of configurations of a ring polymer with indistinguishable monomers. Likewise, ω_{circ} corresponds to the number of closed random walks of length L which start and end at a given point in space (compare also to Metzler et al., 2002a,b).

The number of configurations of a looped structure (with at least one vertex) is also given by Eq. A1. This is due to the fact that the established looping bond is chemically fixed within the chain, so that the chain cannot reptate within a given trajectory. For the same reason (and in contrast to references Hanke and Metzler, 2002; Metzler et al., 2002a,b), different loops cannot exchange length with each other.

REFERENCES

Abramowitz, M., and I. Stegun. 1972. Handbook of Mathematical Functions. Dover, New York.

Alberts, B., K. Roberts, D. Bray, J. Lewis, M. Raff, and J. D. Watson. 1994. The molecular biology of the cell. Garland, New York.

Amouyal, M., N. Perez, and S. Rolland. 1998. Adjacent cooperation of proteins on DNA are not representative of long-distance interactions. *C. R. Acad. Sci. Paris.* 321:877–881.

Aurell, E., and K. Sneppen. 2002. Epigenetics as a first exit problem. *Phys. Rev. Lett.* 88:048101.

Aurell, E., S. Brown, J. Johanson, and K. Sneppen. 2002. Stability puzzles in phage lambda. *Phys. Rev. E.* 65:051914.

Bell, C. E., and M. Lewis. 2001. Crystal structure of the lambda repressor C-terminal domain octamer. *J. Mol. Biol.* 314:1127–1136.

Bell, C. E., P. Frescura, A. Hochschild, and M. Lewis. 2000. Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding. *Cell.* 101:801–811.

Blackwood, E. M., and J. T. Kadonaga. 1998. Going the distance: a current view of enhancer action. *Science.* 281:60–63.

Bolsover, S. R., J. S. Hyams, S. Jones, E. A. Shephard, and H. A. White. 2001. From Genes to Cells. Wiley, New York.

Breslauer, K. J., R. Frank, H. Blocker, and L. A. Marky. 1986. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA.* 83:3746–3750.

Caracciolo, S., M. S. Causo, and A. Pelissetto. 1998. High-precision determination of the critical exponent gamma for self-avoiding walks. *Phys. Rev. E.* 57:R1215–R1218.

Coleman, B. D., D. Swigon, and I. Tobias. 2000. Elastic stability of DNA configurations. II. Supercoiled plasmids with self-contact. *Phys. Rev. E.* 61:759–770.

De Bruin, D., Z. Zaman, R. A. Liberatore, and M. Ptashne. 2001. Telomere looping permits gene activation by a downstream UAS in yeast. *Nature.* 409:109–113.

De Gennes, P.-G. 1979. Scaling concepts in polymer physics. Cornell University Press, Ithaca, New York.

Duplantier, B. 1986. Polymer network of fixed topology: renormalization, exact critical exponent gamma in two dimensions, and $d = 4$-epsilon. *Phys. Rev. Lett.* 57:941–945.

Duplantier, B. 1989. Statistical mechanics of polymer networks of any topology. *J. Stat. Phys.* 54:581–680.

Flory, P. J. 1956. Theory of elastic mechanisms in fibrous proteins. *J. Am. Chem. Soc.* 78:5222–5235.

Griffith, J. D., L. Comeau, S. Rosenfield, R. M. Stansel, A. Bianchi, H. Moss, and T. de Lange. 1999. Mammalian telomeres end in a large duplex loop. *Cell.* 97:503–514.

Guida, R., and J. Zinn-Justin. 1998. Critical exponents of the N-vector model. *J. Phys. A.* 31:8103–8121.

TABLE 2

N	σ_N Padé	σ_N Padé-Borel	σ_N Padé-Borel
3	−0.19	−0.19	−0.17
4	−0.48	−0.49	−0.47
5	−0.86	−0.87	−0.84
6	−1.33	−1.29	
7	−1.88	−1.75	
8	−2.51	−2.23	
9	−3.21	−2.71	

Topological exponents σ_N for network vertices of order N in three dimensions from various approximate techniques. The columns correspond to the first three columns in Table 1 in Schäfer et al. (1992), where the scaling relation $\gamma_F - 1 = \sigma_N + N\sigma_1$, with γ_F from Schäfer et al. (1992) and $\sigma_1 = (\gamma - 1)/2$ with the best available value $\gamma = 1.1575$ (Caracciolo et al., 1998) have been used. Note the large discrepancy between the different methods for increasing N.

Hanke, A., and R. Metzler. 2002. Towards the molecular workshop: entropy-driven designer molecules, entropy activation, and nanomechanical devices. *Chem. Phys. Lett.* 359:22–26.

Lee, C. S., R. E. Pyeritz, and C. A. Thomas, Jr. 1976. Kinetic studies on cyclization of T7 and T1 DNAs. *Biopolymers.* 15:671–686.

Marko, J. F., and E. D. Siggia. 1996. Stretching DNA. *Macromolecules.* 28:8759–8770.

Merlitz, H., K. Rippe, K. V. Klenin, and J. Langowski. 1998. Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation. *Biophys. J.* 74:773–779.

Metzler, R. 2001. The future is noisy: on the role of spatial fluctuations and the merits of cooperativity in genetic switching. *Phys. Rev. Lett.* 87: 068103.

Metzler, R., A. Hanke, P. G. Dommersnes, Y. Kantor, and M. Kardar. 2002a. Equilibrium shapes of flat knots. *Phys. Rev. Lett.* 88:188101.

Metzler, R., A. Hanke, P. G. Dommersnes, Y. Kantor, and M. Kardar. 2002b. Tightness of slip-linked polymer chains. *Phys. Rev. E.* 65:061103.

Montigny, W. J., C. R. Houchens, S. Illenye, J. Gilbert, E. Coonrod, Y.-C. Chang, and N. H. Heintz. 2001. Condensation by DNA looping facilitates transfer of large DNA molecules into mammalian cells. *Nucleic Acids Res.* 29:1982–1988.

Ohno, K., and K. Binder. 1988. Scaling theory of star polymers and general polymer networks in bulk and semi-infinite good solvents. *J. Phys. (Paris).* 49:1329–1351.

Poland, D. C., and H. A. Scheraga. 1965. Statistical mechanics of non-covalent bonds in polyamino acids. VIII. Covalent loops in protein. *Biopolymers.* 3:379–399.

Ptashne, M., and A. Gunn. 2002. Genes and signals. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Ptashne, M. 1992. A Genetic Switch. Cell Press and Blackwell Science, Cambridge, Massachusetts.

Révet, B., B. van Wilcken-Bergmann, H. Bessert, A. Barker, and B. Müller-Hill. 1999. Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. *Curr. Biol.* 9:151–154.

Schäfer, L., C. van Ferber, U. Lehr, and B. Duplantier. 1992. Renormalization of polymer networks and stars. *Nucl. Phys. B.* 374:473–495.

Schellmann, J. A. editor. 1955. The stability of hydrogen-bonded peptide structures in aqueous solution. *C. R. Trav. Lab. Carlsberg Ser. Chim.* 29:230–259.

Semlyen, J. A. 1997. Large Ring Molecules. John Wiley and Sons, Chichester, West Sussex, England.

Semsey, S., M. Geanacopoulos, D. E. A. Lewis, and S. Adhya. 2002. Operator-bound GalR dimers close DNA loops by direct interaction: tetramerization and inducer binding. *EMBO J.* 21:4349–4356.

Snustad, D. P., and M. J. Simmons. 2003. Principles of Genetics. John Wiley and Sons, New York.

Tobias, I., D. Swigon, and B. D. Coleman. 2000. Elastic stability of DNA configurations. I. General theory. *Phys. Rev. E.* 61:747–758.

Wang, M. C., and G. E. Uhlenbeck. 1955. On the theory of the Brownian motion II. *Rev. Mod. Phys.* 17:323–342.

Xu, H., and T. R. Hoover. 2001. Transcriptional regulation at a distance in bacteria. *Curr. Op. Microbiol.* 4:138–144.

Zaman, Z., C. Heid, and M. Ptashne. 2002. Telomere looping permits repression 'at a distance' in yeast. *Curr. Biol.* 12:930–933.