THE ROLE OF WRITHE IN DNA CONDENSATION

T. BLACKSTONE[1*], P. McGUIRK[2*], C. LAING[3], M. VAZQUEZ[4], J. ROCA[5], J. ARSUAGA[4,6,+]

Abstract. It has long been known that extraction of DNA from P4 phages results in a large proportion of highly knotted nicked DNA circles. We showed that most of these knots are formed by the cyclization of the DNA inside the phage capsid. Since these DNA knots could be informative of the chromosome geometry inside the viral capsid, we further studied the distribution of knot families of up to 6 crossings, both experimentally and by computer simulations. Our study revealed an absence of achiral knots, which suggested that the arrangement of the DNA inside the capsid is writhe directed. We here perform a numerical study on how an excess of writhe affects different properties of a polymer chain. Our future goal is to use these results to further understand the geometry of DNA chains compacted inside biological structures.

1. Introduction and Biological Motivation

Bacteriophages are viruses that propagate in bacteria. Most bacteriophages consist of an icosahedral capsid that contains the nucleic acid (i.e. DNA or RNA), a tail with tail fibers used during the process of infection, and a molecular motor that connects the capsid and the tail called connector. All icosahedral bacteriophages with double-stranded DNA genomes are believed to pack their chromosomes in a similar manner [1]. Furthermore icosahedral bacteriophages are believed to hold similar DNA arrangements as those found in some animal viruses [2] and lipo-DNA complexes used in gene therapy [3].

In our research we have investigated the DNA packing in bacteriophage P4. Bacteriophage P4 has a linear double-stranded DNA genome, 10 to 11.5 Kb in length, flanked by 16 bp single-stranded ends that are complementary in sequence [4]. During phage morphogenesis a proteinic procapsid is first assembled. This is followed by the transfer of one copy of the viral genome inside it.

There are a number of models for the global organization of DNA inside phage capsids. These include the ball of string model [5], the coaxial spooling model [5, 6, 7, 8], the spiral-fold model [9], the folded toroidal model [10] and the liquid crystalline models [11]. Spooling and toroidal models suggest a global homogenous organization while

coaxial and liquid crystalline models, which are consistent with local optimal hexagonal packing, imply a less organized global picture [12]. These models need to account for the extreme conditions of the DNA inside the capsid. It is known that the DNA molecule and its associated water molecules fill the entire capsid volume and that DNA fibers reach concentrations of 800 mg/ml [13] and osmotic pressures of about 80 atmospheres [14]. In 1980 Liu and collaborators found that DNA extracted from P4 bacteriophage "capsids", that is phages with no tail was highly knotted [15, 16, 17]. We investigated these distributions of knots and found that most of these knots are formed inside the viral capsid [18]. By carefully analyzing distributions of up to 7 crossings we also found a lack of four crossing knots. Our studies also found that torus knots prevailed over other families of knots. By comparing these distributions with those found in computer simulations we demonstrated that the DNA packing is not random and concluded that it is chirally organized [19]. Such chiral organization is in agreement with our previous studies and those of others using molecular dynamics methods which also favor toroidal and spool-like arrangements for highly condensed DNA [20, 21, 22, 23, 24]. Calculations of optimal spool-like conformations of DNA in phage P4 estimated a writhe of -45 for the 10-kb DNA, which interestingly corresponds with the level of supercoiling density typically found in bacterial chromosomes.

If the DNA molecule has a large writhe inside the viral capsid, one would ask how such writhe contributes to the knotting probability as well as to the overall organization of the chromosome inside the capsid. Hence, here we compute several parameters to investigate the potential contribution of the writhe to the condensation of DNA. Quantification parameters for DNA condensation are difficult to establish. Traditionally biologists provided the linear condensation ratio (DNA length/capsid diameter). More recently, it has been proposed a volume condensation ratio, which compares the volume occupied by the DNA with that of a random coil of equal length [25]. But even this measure is somewhat gross and one would like to find others that account for different organizations of DNA fibers. As an extension of these ideas we here study the knot probability, average number of crossings and radius of gyration of polygons whose writhe has been fixed. These three measures are not independent but somehow reflect different aspects of DNA condensation. Next we briefly describe each of these measures.

Every achiral knot has average writhe zero [26]. The relationship between writhe and knot distributions has already been analyzed using the wormlike chain model [27]. In this study the authors increased

the writhe value and found that the probability of finding unknotted molecules rapidly decreases with increasing writhe values. Interestingly these studies suggested that this probability does not necessarily increase with DNA size. Consistent with our hypothesis they also observed a large fraction of toroidal knots. In the case of the wormlike chain it was shown that toroidal knots minimize the free energy, lower than other knot populations, because they have less torsion.

The more complex the knot is the smaller its radius of gyration. The radius of gyration estimates the radius of the smallest sphere that can contain a given knot. Therefore, more compact knots have smaller radius of gyration. The relationship between the radius of gyration and the knotting probability has been studied in [28]. This study led the authors of the paper to propose the concept of equilibrium length of knots.

The Average Crossing Number (ACN) is a geometric measure of knot complexity. Complex knots are those with high ACN. The importance of the ACN is that it can be detected in the laboratory since it has been correlated with knot migration in agarose gels.

Complex knots have a high average number of crossings. A number of studies have investigated this relationship. We will just mention here that knot migration in agarose gels was inferred from these studies [31].

2. Numerical Methods

The Polymer Model

In our previous investigations we considered the freely jointed polymer model for analyzing conformations inside phage capsids and we will continue using this model here. In the freely jointed model segments of equal length are joined together by hinges of infinite flexibility. The total number of segments N used in our simulations ranged from 14 to 200. Excluded volume effects were not taken into account.

Markov Chain Monte-Carlo

Ensembles of equilateral polygons were generated by the crankshaft algorithm. This algorithm generates configurations as follows: (i) two vertices of the chain are selected at random, thus dividing the polygon into two subchains, and (ii) one of the two subchains is selected at random (with equal probabilities for each subchain), and the selected subchain rotated through a random angle around the axis connecting the two vertices. The subchains are allowed to pass through each other during this rotation. This algorithm is known to generate an ergodic

Markov chain in the set of polygons of fixed length [32]. Biasing with respect an average writhing number was implemented using the Metropolis algorithm [33]. In this algorithm each chain is assigned an energy given by the square of the difference between a prefixed value of the writhe and the observed writhe. Probability of acceptance of chains along the Markov chain was implemented using the standard criterion Metropolis criterion. Configurations along the Markov chain are correlated by construction. Correlation along the Markov chains was accounted for by time-series analysis methods described in Madras and Slade [32] and implemented by S. Whittington.

Identification of Knotted Conformations
From each of the selected polygons in the Markov chain a random projection was taken. From each regular projection the corresponding knot diagram [33] and its associated Dowker code [34] were computed and simplified [35] by removing crossings that did not change the knot type (Reidemeister I and II moves; ref. 31). Once the diagram was simplified, the identification of knotted polygons was performed by computing the Alexander polynomial [35, 36, 37] evaluated at $t = -1$ [denoted by $\Delta(-1)$]. It is known that $\Delta(-1)$ does not identify all knotted chains and we expect some error introduced by this simplification, especially with high writhe and/or long polymer chains.

Computation of the Writhe and the ACN
The writhe is a geometric measure of self-entanglement which measures non-planarity of a spatial conformation. The writhe (Wr) of a closed curve C can be defined using the Gauss double integral. If $\gamma : S^1 \to \mathbb{R}^3$ is an arc-length parametrization of C, then the writhe is given by

$$Wr = \frac{1}{4\pi} \int\int_{S^1 \times S^1 \setminus D} \frac{\langle T\gamma(t), T\gamma(s), \gamma(t) - \gamma(s) \rangle}{\|\gamma(t) - \gamma(s)\|^3} \, \mathrm{d}t \, \mathrm{d}s$$

where D is the diagonal of $S^1 \times S^1$, $T_\gamma(t)$ denotes the unit tangent vector, $\|.\|$ is the 2-norm in $\mathbb{R}^3$ and $\langle ., ., . \rangle$ is the triple scalar product of three vectors. For the case of a right-handed crossing of the tangent vectors, one can think of the integrand as the measure of the solid angle created by those projections from which the tangent vectors appear to cross. If the crossing is a left-handed crossing, the integrand is equal to the measure of this solid angle multiplied by -1. For either type of crossing, the quantity is called the signed measure of the solid angle and is denoted $\Omega_\gamma(t, s)$. This interpretation of the integrand allows for an easy calculation of the writhe when C is a polygon of N segments.

In this case, the Gauss integral can be written as a double sum over the segments of the polynomial

$$Wr = 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{\Omega_{ij}}{4\pi} = \frac{1}{2\pi}\sum\sum_{1\leq i\leq j\leq N}\Omega_{ij}$$

where Ω_{ij} is the signed measure of the solid angle defined by the perspectives from which the ith and jth line segments appear to cross. Notice that $\Omega_{ij} = \Omega_{ji}$, thus the reason of why the multiplicative factor of two appears. Because the writhe is an average over all projections, we divide each solid angle by the total area of the unit sphere (4π) We followed the first method proposed by Klenin and Langowski [38] to calculate Ω_{ij}. Calculation of the ACN of a curve was done using the same algorithm, except that the handedness of the crossing was not taken into account. We then calculated the ACN by taking the absolute value of Ω_{ij} in every term of the summation used to calculate Wr.

Computation of the Radius of Gyration
The radius of gyration R_g is generally defined for a collection of N particles and is given by the rms distance from each particle to the average position of those particles. We calculate the radius of gyration of a polygon by calculating the radius of gyration of the vertices of that polygon.

$$R_g = \frac{1}{N}\sum_{i=0}^{N-1}\|r_i - \langle r\rangle\|^2$$

The radius of gyration is denoted by R_g, $\langle r\rangle$ denotes the average position of the N particles and $\|.\|$ is as before the 2-norm in $\mathbb{R}^3$.

3. Results

The effects of the writhe bias on the chain dimensions were analyzed by Markov-chain Monte Carlo simulations [30]. Symmetric Markov chains of freely jointed ideal equilateral polygons with a prefixed writhe were generated by importance sampling as explained in Numerical Methods. To test that our programs were working properly we first measured the distribution of writhes once a particular writhe value has been fixed. Results are shown in Figure 1. As expected we found a Boltzman distribution around the desired writhe value.

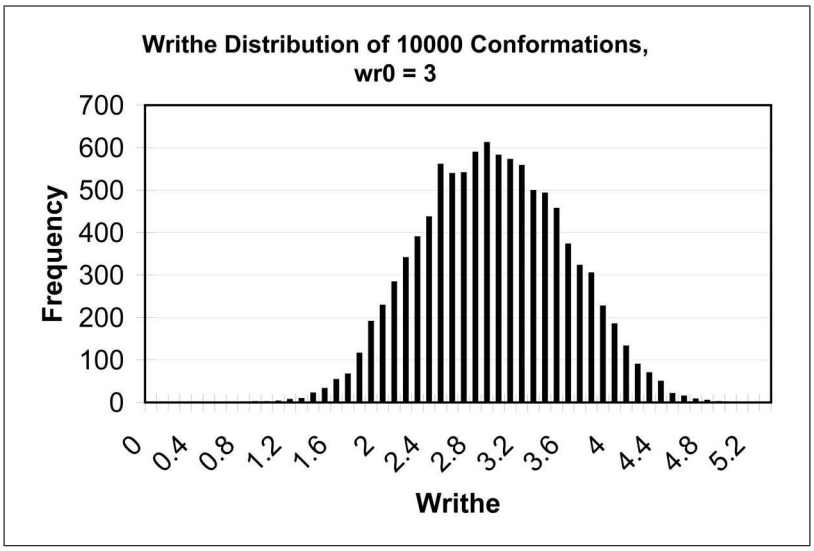

FIGURE 1. **Histogram of writhe values.** We used Metropolis Monte-Carlo simulations to generate populations of polygons around a pre-selected writhe. The histogram shows a population of polygons of length 100 with a pre-selected writhe of 3.

Next we investigated how the different measures proposed were affected by the writhe bias. First we investigated the knotting probability (Figure 2). Knotted conformations were detected by the Alexander polynomial evaluated at $t = -1$. For a fixed chain length and average writhe we computed the knotting probability by dividing the number of knotted conformations by the total number of conformations generated. For high fixed writhe and small polygons we observed that the knotting probability surprisingly decreased upon increasing the size of the polymer. This behavior had already been suggested by other authors [27]. Here we study this behavior systematically. We believe that a high writhe in an unknotted chain means an abundance of Reidemeister type I moves, which are much less likely to occur than within knotted conformations. Analyzing generated chains appears to support this hypothesis (Figure 3). The decreasing of the knotting probability disappeared once the chain length was large enough. This is presumably because once the chains are long enough Reidemeister type I moves are easier to introduce. For larger polygons the knotting probability increased with increasing length as expected [41]. Interestingly and in agreement with previous studies we observed that for any fixed length the knotting probability increased rapidly with increasing writhe.

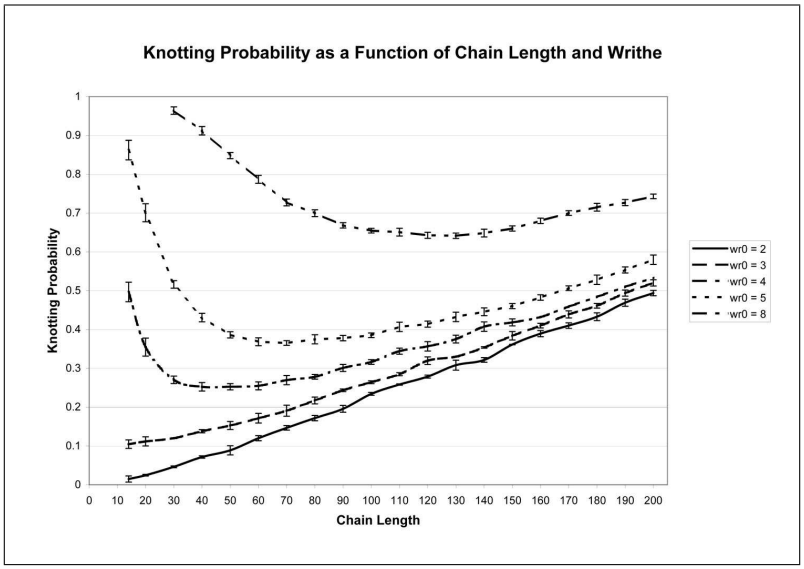

FIGURE 2. **Knotting probability for chains as a function of the writhe.** Comparison of knotting probability for different values of the writhe for chains from 14 to 200 segments. Writhe values ranged from 2 to 8. Interestingly knotting decreased first and increased for longer chains. Overall knotted molecules became highly probable with small deviations from zero writhe.

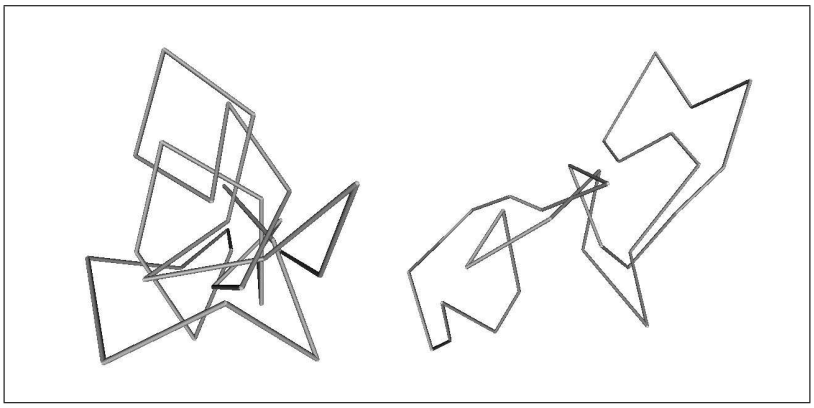

FIGURE 3. **Comparison of knotted and unknotted chains with high writhe.** Comparison of chain configurations between a knotted (left) and unknotted (right) chain, both with writhe 8. Each chain is comprised of 30 segments. Note the appearance of Reidemeister twists (type I moves) within the unknotted chain.

Next we investigated the geometrical complexity of the knotted configurations. We first investigated the ACN for chains of fixed length and fixed average writhe. For fixed chain length and writhe we computed

the ACN (Figure 4). As expected, the ACN increased with increasing writhe and/or chain length. To better understand the behavior of our samples, we subdivided the population between knotted and unknotted conformations and found that in general knotted conformations had a higher ACN than unknotted ones, although this difference was clearer for polygons with more than 50 segments. It can also be observed that for a given length and writhe knotted molecules had a higher ACN suggesting a higher degree of condensation.

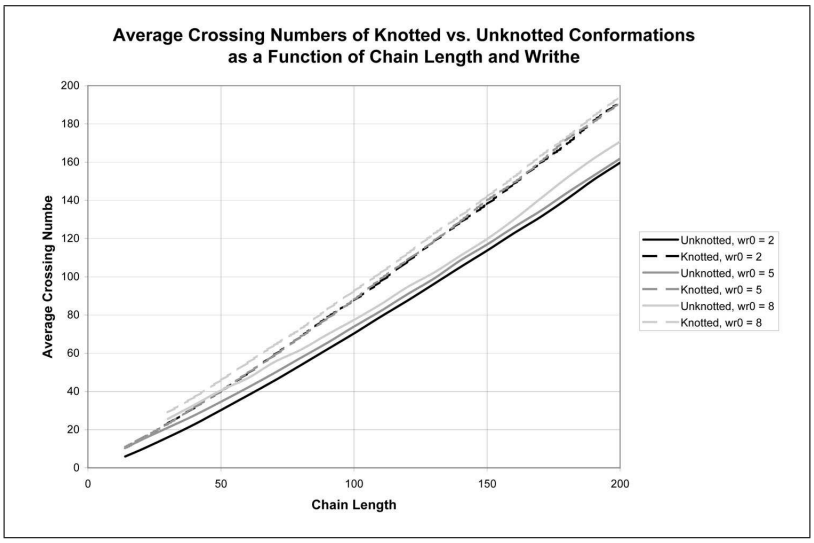

FIGURE 4. **Average crossing number as a function of the chain length and of the pre-selected writhe.** The graph shows the average crossing number for polygons between 14 and 200 segments and writhe values between 2 and 8. Unknotted molecules had fewer crossings than knotted molecules although this difference was negligible for short changes.

Finally we investigated the dependence of the radius of gyration on the writhe (Figure 5). Similar to the behavior of the ACN, for fixed writhe and increasing chain length we observed an increasing radius of gyration while for fixed length and increasing writhe we observed decreasing values of the radius of gyration indicating a higher degree of condensation. As in the previous case the population was subdivided in unknotted and knotted molecules. Overall knotted molecules were more compact that unknotted ones for the identical writhe values.

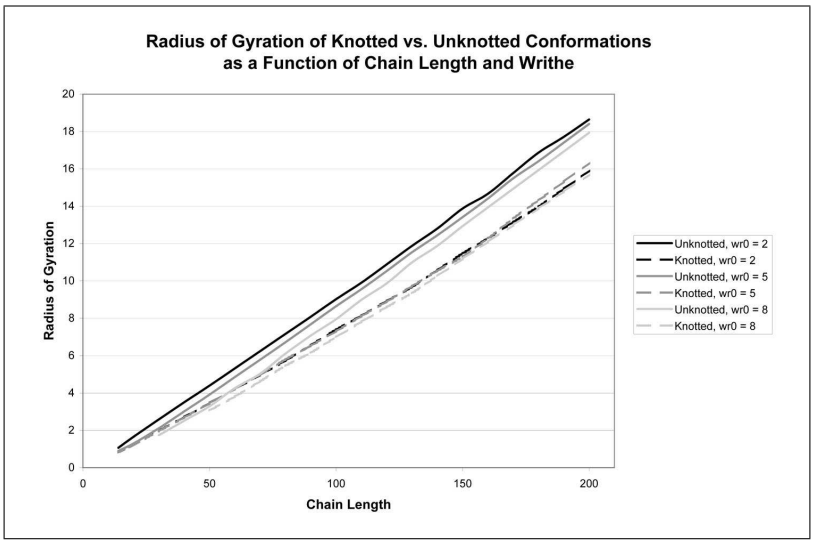

FIGURE 5. **Radius of Gyration as a function of the chain length and of the Writhe.** Polygons between 14 to 200 segments were analyzed for writhe values ranging between 2 and 8. For fixed length and writhe knotted molecules showed a smaller radius of gyration.

4. Conclusions and Future Directions

Inspired by our previous work on DNA packing in bacteriophages we have here investigated the role of the writhe in polymer condensation. Our results show that polymers chains under a strong writhe constraint have a higher probability of being knotted and that the resulting knots tend to be also very complex (i.e. high ACN) and more compact (i.e. radius of gyration). Nevertheless, in order to be able to make accurate scaling arguments our work needs to be extended to longer chains (up to 400-500 segments) and higher writhe values. This implies an important computational challenge which we will address in future publications. From this study we conclude that the writhe may be a primary contributor to the knot probability observed experimentally [17]; and therefore that it may be a determining factor to drive chromosomal organization. If that is the case we believe that knot distributions will be very characteristic and different from those found by confinement only. It will be necessary to verify this hypothesis with the experimentally observed knot distributions. Next we describe some of the challenges ahead of us.

i. Quantification of DNA Condensation and Organization
What are the correct mathematical expressions to define chromosome organization? Other measures that the ones studied here have been proposed to study the entanglement of fibers. Those include those proposed by Rogen and Fain [39] to study proteins which are more sophisticated and may help better characterize chromosome organization.

ii. Models of DNA in Confined Volumes
Biochemical and structural analyses have revealed that DNA inside the capsid is kept in its B-form [39, 40], that there are no specific DNA-protein interactions [7, 42] or correlation between DNA sequences and their spatial location inside the capsid, with the exception of the DNA ends in some viruses [43]. In our experiments we also found that DNA kinks induced by long AT repeats in the DNA sequence do not play a major role in the formation of DNA knots [18]. These observations justified our approach for considering curves confined to spherical volumes. The freely jointed chain is a basic model that needs to be improved to be able to make better predictions of the viral packing. How can this be achieved? First it is still unknown the degree of flexibility of the DNA chain under those conditions found inside the capsid. Therefore more sophisticated models of DNA that incorporate DNA bending in confined volumes are needed. Second, our models do not include volume exclusion. There are experimental observations that show domains of DNA fibers organized parallel to each other and with an average distance of 25Å, thus suggesting that the volume exclusion of the DNA also plays an important role [6, 11].

iii. Experimental Characterization of Knots
The observed knot distributions need to be better characterized experimentally. We are working in two approaches in this direction. On one hand we are developing electrophoresis and microscopy methods to visualize complex knots, on the other hand we are developing computer algorithms that simulate the unknotting of type II topoisomerases [46]. With these algorithms we can try to understand the probability of transitioning from one knot to another by one topoisomerase strand passage action [47, 48].

iv. DNA Knots in Vivo

Our results suggest that DNA knots should be common in chromosomes in other organisms. In fact DNA knots are found in a number in Escherichia coli cells harboring mutations in the GyrB or GyrA genes [44], in bacteriophage P2 [15, 16, 17], and cauliflower mosaic viruses [45]. The origin of these knots or their biological meaning remains to be identified.

5. Acknowledgements

We would like to thank S. Whittington for valuable comments in the manuscript and by providing the software to compute autocorrelations in the Markov Chains. This manuscript contains images rendered using Robert G. Scharein's KnotPlot program. This work was supported by grants from Program in Mathematics and Molecular Biology (PMMB) through a Burroughs Welcome Fund Interfaces Grant (C. Laing), the Ministerio Español de Ciencia y Tecnología, grant number BFU2005-01311 (J. Roca), from the Research Infrastructure in Minority Institutions Program, National Center on Minority Health and Health Disparities (NCMHD), NIH, grant number P20 MD000262 (M. Vazquez) and by (U56 CA096217) from the National Cancer Institute, NIH (J. Arsuaga).

References

[1] Earnshaw, W. C. and Casjens, S. R. (1980) *Cell* **21**, 319-331
[2] San Martin, C. and Burnett, R. (2003) *Curr. Top. Microbiol. Immunol.* **272**, 57-94.
[3] Schmutz, M., Durand, D., Debin, A., Palvadeau, Y., Eitienne, E. R. and Thierry, A. R. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 12293-12298.
[4] Wang, J. C., Martin, K. V. and Calendar, R. (1973) *Biochemistry* **12**, 2119-2123.
[5] Richards, K., Williams, R. and Calendar, R. (1973) *J. Mol. Biol.* **78**, 255-259.
[6] Earnshaw, W. C. and Harrison, S. (1977) *Nature* **268**, 598-602.
[7] Serwer, P. (1986) *J. Mol. Biol.* **190**, 509-512.
[8] Cerritelli, M., Cheng, N., Rosenberg, A., McPherson, C., Booy, F. and Steven, A. (1997) *Cell* **91**, 271-280.
[9] Black, L., Newcomb, W., Boring, J. and Brown, J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7960-7964.
[10] Hud, N. (1995) *Biophys. J.* **69**, 1355-1362.
[11] Lepault, J., Dubochet, J., Baschong, W. and Kellenberger, E. (1987) *EMBO J.* **6**, 1507-1512.
[12] Starostin, E. L. (2006) J. Phys.: Condens. Matter S187-S204
[13] Kellenberger, E., Carlemalm, E., Sechaud, J., Ryter, A. and Haller, G. (1986) in *Bacterial Chromatin*, eds. Gualerzi, C. and Pon, C. L. (Springer, Berlin), p. 11.
[14] Evilevitch A, Lavelle L, Knobler CM, Raspaud E, Gelbart WM. (2003) *Proc. Natl. Acad. Sci. USA* **100**:9292-5.
[15] Liu, L. F., Davis, J. L. and Calendar, R. (1981) *Nucleic Acids Res.* **9**, 3979-3989.
[16] Liu, L. F., Perkocha, L., Calendar, R. and Wang, J. C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5498-5502.
[17] Wolfson, J. S., McHugh, G. L., Hooper, D. C. and Swartz, M. N. (1985) *Nucleic Acids Res.* **13**, 6695-6702.

[18] Arsuaga, J., Vazquez, M., Tigueros, S., Sumners, D. W. and Roca, J. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 5373-5377.
[19] Arsuaga, J., Vazquez, M., McGuirk, P., Trigueros, S., Sumners, D., and Roca, J. (2005) *Proc. Natl. Acad. Sci. USA* **102**:9165-9.
[20] Maritan, A., Micheletti, C., Trovato, A. and Banavar, J. (2000) *Nature* **406**, 287-289.
[21] Arsuaga, J., Tan, R. K., Vazquez, M., Sumners, D. W. and Harvey, S. (2002) *Biophys. Chem.* **101**, 475-484.
[22] Tzill, S., Kindt, J. T., Gelbart, W. and Ben-Shaul, A. (2003) *Biophys. J.* **84**, 1616-1627.
[23] LaMarque, J. C., Le, T. L. and Harvey, S. C. (2004) *Biopolymers* **73**, 348-355.
[24] Kulic, I. M. , Andrienko, D. and Deserno, M. (2004) Twist-bend instability for toroidal DNA condensates. *Europhys. Lett.* **67**, 418-424 (2004)
[25] Holmes, V. and Cozzarelli, NR. (2000) *Proc. Natl. Acad. Sci. USA* **97**:1322-4.
[26] Janse Van Rensburg, E. J., Sumners, D. W., Whittington, S. G. (1998) in *Series of knots and everything*, eds. A. Stasiak, V. Katrich, L. H. Kauffman.
[27] Podtelezhnikov, AA., Cozzarelli, NR., Vologodskii, AV. (1999) *Proc. Natl. Acad. Sci. USA* **96**:12974-9. 28 Dobay A,
[28] Dobay, A., Dubochet J., Millet, K., Sottas, PE., Stasiak, A. (2003) *Proc. Natl. Acad. Sci. USA* **100**:5611-5615.
[29] Millet, K., (1994) in *Series of Knots and Everything*, eds. Sumners, D.W. and Millet, K.
[30] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953) *J. Chem. Phys.* **21**:1087-1092
[31] Vologodskii, AV., Crisona, NJ., Laurie, B., Pieranski, P., Katritch, V., Dubochet, J., Stasiak, A. (1998) *J. Mol. Biol.* **278**:1-3.
[32] Madras, N. and Slade, G. (1993) *The Self-Avoiding Walk* (Birkhäuser, Boston).
[33] Burde, G. and Zieschang, H. (1985) *Knots* (de Gruyter, Berlin), Vol. 5.
[34] Dowker, C. H. and Thistlethwaite, M. B. (1983) *Topology Appl.* **16**:19-31
[35] Arsuaga, J. (2000) *Ph.D. thesis* (Florida State Univ., Tallahassee, FL).
[36] Michels, J. P. J. and Wiegel, F. W. (1986) *Proc. R. Soc. London Ser. A* **403**:269-284
[37] Frank-Kamenetskii, M. D., Lukashin, A. V. and Vologodskii, A. V. (1975) *Nature (London)* **258**:398-402
[38] Klenin, K. and Langowski, J. (2000) *Biopolymers.* **54**:307-17.
[39] Rogen, P. and Fain, B. (2003) *Proc. Natl. Acad. Sci. USA* **100**:119-24.
[40] Aubrey, K., Casjens, S. and Thomas, G. (1992) *Biochemistry* **31**:11835-11842
[41] Diao, Y. (1995) *J. Knot Theor. Ramif.* **4** 189-196
[42] Hass, R., Murphy, R. F. and Cantor, C. R. (1982) *J. Mol. Biol.* **159**:71-92
[43] Chattoraj, D. K. and Inman, R. B. (1974) *J. Mol. Biol.* **87**:11-22
[44] Shishido, K. , Komiyama, N. and Ikawa, S. (1987) *J. Mol. Biol.* **195**:215-218
[45] Ménissier, J., de Murcia, G., Lebeurier, G. and Hirth, L. (1983) *EMBO J.* **2**, 1067-1071
[46] Trigueros, S., Arsuaga, J., Vazquez, ME., Sumners, DW., Roca, J. (2001) *Nucleic Acids Res.* **29**:E67-7.
[47] Hua, X., Baghavan, B., Nguyen, D., Arsuaga, J. and Vazquez, M. In press. *Topol. Appl.*
[48] Flammini, A., Maritan, A., Stasiak, A. (2004) *Biophys J.* **87**:2968-75.

[1]Computer Science Department, San Francisco State University, San Francisco, CA

[2]Physics Department, University of Wisconsin-Madison, Madison, WI

[3]Department of Mathematics, Florida State University, Tallahassee, FL

[4]Mathematics Department, San Francisco State University, San Francisco, CA

[5]Molecular Biology Department, CID-CSIC, Barcelona, Spain

[6]Center for Computing in the Life Sciences, San Francisco State University, San Francisco, CA

*Contributed equally to this work,

+To whom correspondence should be addressed