LBL-35488
UC-408

Lawrence Berkeley Laboratory

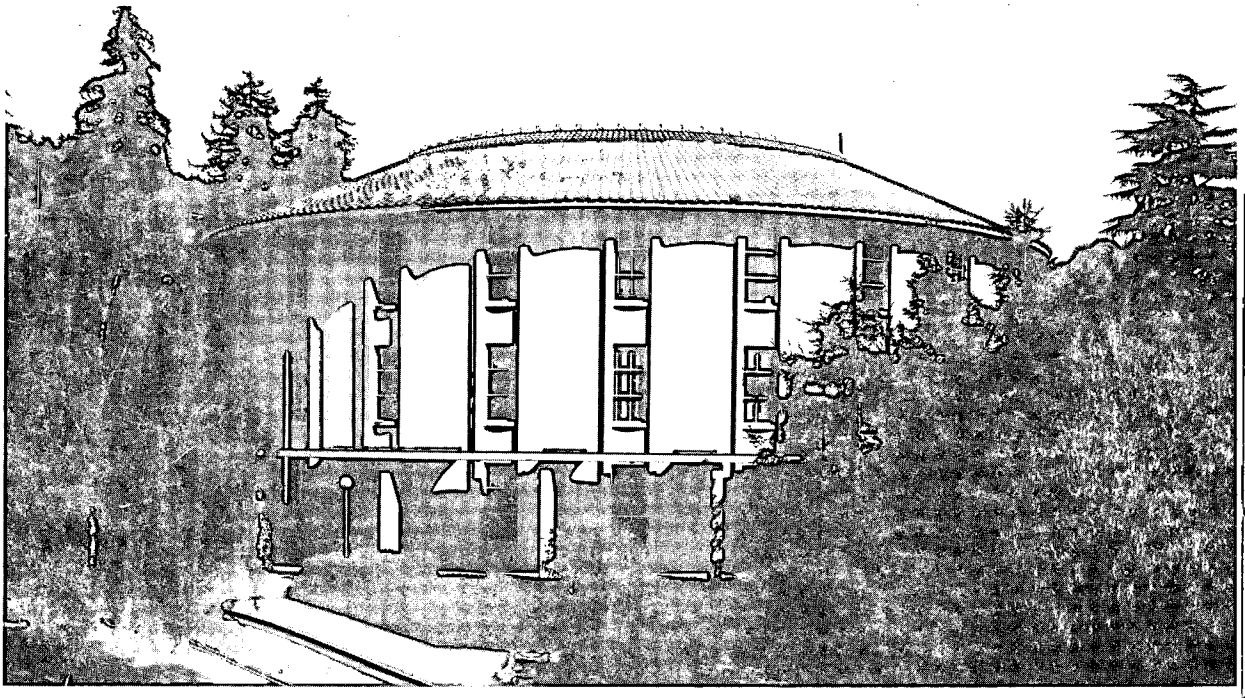
UNIVERSITY OF CALIFORNIA

STRUCTURAL BIOLOGY DIVISION

Theoretical Study of the Conformation and Energy of Supercoiled DNA

N.G. Hunt
(Ph.D. Thesis)

January 1992



REFERENCE COPY
Does Not
Circulate
Bldg. 50 Library.
LBL-35488
Copy 1

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-35488
UC-408

Theoretical Study of the Conformation and Energy of Supercoiled DNA

Nathaniel George Hunt
Ph.D. Thesis

Department of Physics
University of California

and

Division of Structural Biology
Lawrence Berkeley Laboratory
University of California
Berkeley, California

January 1992

This work was supported by the Director, Office of Energy Research, Office of Health Effects Research, of the U.S, Department of Energy under Contract No. DE-AC03-76SF00098, NIH, Grant GM41911and NSF graduate fellowship.

Theoretical study of the conformation and energy of supercoiled DNA

by

Nathaniel G. Hunt

Abstract

The two sugar-phosphate backbones of the DNA molecule wind about each other in helical paths. For circular DNA molecules (plasmids), or for linear pieces of DNA with the ends anchored, the two strands have a well-defined linking number, Lk. If Lk differs from the equilibrium linking number Lk_0, the molecule is supercoiled. The linking difference $\Delta Lk = Lk - Lk_0$ is partitioned between torsional deformation of the DNA, or twist (ΔTw), and a winding of the DNA axis about itself known as writhe (Wr).

In this dissertation, the conformation and energy of supercoiled DNA are examined by treating DNA as an elastic cylinder. Finite-length and entropic effects are ignored, and all extensive quantities (e.g. writhe, bend energy) are treated as linear densities (writhe per unit length, bend energy per unit length). Two classes of conformation are considered: the plectonemic or interwound form, in which the axis of the DNA double helix winds about itself in a double superhelix, and the toroidal shape in which the axis is wrapped around a torus. Minimum energy conformations are found. For biologically relevant values of specific linking difference, the plectonemic conformation is energetically favored over toroidal conformations. For plectonemic DNA, the superhelical pitch angle α is in the range $45^o < \alpha \leq 90^o$. For low values of specific linking difference $|\sigma|$ ($\sigma = \Delta Lk / Lk_0$), most linking difference is in writhe. As $|\sigma|$ increases, a greater proportion of linking difference is in twist. Interaction between DNA strands is treated first as a hard-body excluded volume and then as a screened electrostatic repulsion. Ionic strength is found to have a large effect, resulting in significantly greater torsional stress in supercoiled DNA at low ionic strength. Results

are compared with electron microscopy data and measurements of supercoiling-induced DNA conformational transitions. Good agreement with experimental results is found for DNA in monovalent salt solutions.

Approved by

Prof. Alan Portis
Department of Physics
University of California, Berkeley

Acknowledgments

This dissertation owes a lot to the friends and colleagues in Chemical Biodynamics who taught me what biology I know. John Hearst and Dave Cook took a special interest in the problem of DNA supercoiling and offered invaluable insight and encouragement. Paul Selvin helped get me into the whole biology business and so bears primary responsibility for any catastrophes which may result. Mel Klein was supportive and encouraging, both with my EPR work and with the DNA supercoiling. Pete Spielmann put heroic efforts into preparing DNA samples for the EPR. When plied with sufficient quantities of caffeine, Robert Van Buskirk has lent his insight into supercoiling and related problems.

Many people helped me avoid science when necessary. Paul was again a prime culprit, both here and in Central America. Robert, with his boundless energy and dedication, is one of those compañeros who justify my faith in human nature. Bill Scott and Tariq Rana have become friends of the sort which cannot be lost. Other political co-conspirators have been Ed, Eva, Alixe, Eve, Jessie, Ray, Ben and Cathy. Marc, Karina, Marc, Storrs, Mike, Lois and Sunita were crucial in getting me out of the lab and into cafés, the mountains, to the beach or in front of a chess-board, as required.

Finally, I would like to give a special thanks to Dove, who has caused me so much trouble over the past year, and to my parents, Sarah, Bob, Rivers, Priscilla, Jennifer and Lee.

This work was supported by the NIH, Grant GM41911; the DOE Office of Energy Research, Office of Health and Environmental Research, Contract DE AC03 - 76SF00098; and a 3 - year NSF graduate fellowship.

Theoretical study of the conformation and energy of supercoiled DNA

Nathaniel G. Hunt

Table of Contents

Abstract	1
Acknowledgments	ii
Table of Contents	iii
List of Figures	iv
Chapters:	
I. Introduction: DNA - a physicist's overview	1
II. Elastic model of supercoiling with hard-body excluded volume	23
III. Elastic and electrostatic model of supercoiling	52
IV. Conclusion and future work	81
Appendix A: Evaluation of Wr	86
Appendix B: Fortran code for plectonemic helix	92
Glossary	100

List of Figures

Chapter I:

Figure 1. Toroidal conformation 6

Figure 2. Plectonemic conformation 7

Figure 3. Cruciform structure 11

Chapter II:

Figure 4. Bend, twist, and total energies per unit length 27

Figure 5. Total energy per unit length vs. α for different σ 30

Figure 6. Total energy per unit length vs. α for different R_{ex} 31

Figure 7. Energy cost of writhe and twist 33

Figure 8. Bend, twist and total energies vs. σ 34

Figure 9. Linking difference in twist and writhe vs. σ 35

Figure 10. Proportion of linking difference in writhe vs. σ 36

Figure 11. Total energies for toroidal and plectonemic conformations vs. σ 41

Figure 12. Total plectonemic energy and parabolic expression vs. σ 47

Chapter III:

Figure 13. Total, bend, twist and electrostatic energies vs. σ; I = 150 mM 59

Figure 14. Total energies vs. σ; I = 50, 150 & 500 mM 60

Figure 15. Superhelix radius vs. σ; I = 50, 150 & 500 mM 61

Figure 16. Proportion of ΔLk in Wr vs. σ; I = 50, 150 & 500 mM 62

Figure 17. Total energy for plectonemic and toroidal conformations vs. σ 64

Figure 18. Hard-body, electrostatic, and Monte Carlo energies vs. σ; low salt 66

Figure 19. Hard-body, electrostatic, and Monte Carlo energies vs. σ; high salt 67

Figure 20. Hard-body, electrostatic, and Monte Carlo stresses vs. σ; low salt 69

Figure 21. Hard-body, electrostatic, and Monte Carlo energies vs. σ; high salt 70

Figure 22. Superhelix radius - theory and electron microscopy data, no Mg^{2+} 73

Figure 23. Superhelix radius - theory and electron microscopy data, with Mg^{2+} 74

Figure 24. Probability of cruciform formation - theory and experiment 76

Appendix A:

Figure 25. Writhe of a single helix - infinite and finite-length cases 90

Figure 26. Writhe of a plectonemic helix - infinite and finite-length cases 91

I. Introduction: DNA - a physicist's overview

Of all the remarkable properties of the DNA molecule, this dissertation focuses on a small number. These are its topology, its elastic bend and torsion constants, and the fact that it is charged. Slightly more detail is included only when considering the conformational transitions which result from large torsional stress on the DNA . Using this simple model, the conformation and energy of supercoiled DNA are examined.

For the benefit of readers with a limited biological background, there is a brief glossary of biological terms on page 100.

The biological interest in supercoiling is substantial. Some of the specific biological issues related to DNA supercoiling are outlined below, in section I.B. There are also more general questions about supercoiling which the work presented here may help elucidate. Why is DNA supercoiling such a ubiquitous phenomenon? Does it play a regulatory role in DNA expression? Does it serve to store energy to drive DNA transcription? The results derived here do not answer these questions, but they provide a framework for considering the supercoiling phenomenon.

Some specific examples of issues which can be understood based on this thesis and which previously caused some confusion are whether supercoiled DNA is toroidal or plectonemic, the likely shape of supercoiled DNA (i.e. superhelical pitch angle and radius), the partitioning of supercoiling-induced deformations between twist and writhe, whether the energy of supercoiling is purely a quadratic function of linking difference, why B-Z transitions occur in supercoiled DNA at low salt concentration while in linear DNA they require high salt, the salt-concentration dependence of supercoiling-induced cruciform formation, and the salt-concentration dependence of enzymatically-induced supercoiling.

A. Topology and DNA

Any two unknotted, disjoint, closed space curves have a linking number, Lk. Roughly speaking, Lk is the number of times the two curves wrap around each other. More rigorously, if the two curves are given by parameterizations **r** (u_1), **s** (u_2), then

$$Lk = \frac{1}{4\pi} \int_{C_1, C_2} du_1\, du_2 \frac{[\vec{r}(u_1) - \vec{s}(u_2)] \cdot \left[\frac{d\vec{r}(u_1)}{du_1} \times \frac{d\vec{s}(u_2)}{du_2}\right]}{|\vec{r}(u_1) - \vec{s}(u_2)|^3} \qquad (1.1)$$

Lk can be either positive, negative or zero and is an integer. The sign of Lk can be determined by assigning orientations to the curves. Such an assignment is implicit in the parameterization of the curves in equation (1.1). All other quantities in this thesis are independent of parameterization. For the most part, I will not play close attention to the sign of Lk since most results derived here are independent of this sign.

In relating equation (1.1) to DNA, we can start by considering the two sugar-phosphate backbones of a DNA molecule. They wrap around each other in helical paths with a pitch of 10.4 base pairs (b.p.) or 35.4 Å. This pitch is commonly called the helical repeat, h_0. A closed circular piece of DNA (a plasmid) thus has a natural linking number, Lk_0, which is the number of base pairs divided by 10.4. Plasmids vary greatly in size, roughly from 10^3 to 10^5 b.p., so that typical values for Lk_0 range from 10^2 to 10^4.

We could also choose the two curves in equation (1.1) to be one of the sugar-phosphate backbones of the DNA and the other to be the axis of the DNA double helix. This choice is preferable, as discussed below, and gives the same value of Lk.

In 1969, the mathematician James White (White, 1969) found that Lk can be expressed as the sum of two quantities, which Fuller (1971) later named twist (Tw) and writhe (Wr):

$$Lk = Tw + Wr \tag{1.2}$$

Roughly speaking, the twist of a curve A about a curve B (Tw_{AB}) is the number of times a vector from B to A goes around B. Tw_{AB} does not in general equal Tw_{BA}, and twist does not have to be an integer. To define twist rigorously, a correspondence between the curves A and B must be defined, i.e. for each point **r** (u) on curve A there is a corresponding point **s** (u) on curve B. This correspondence is easiest to visualize if the two points are close to each other, but this does not have to be the case. Let the vector from **s** (u) to **r** (u) be called **z** (u). For well-behaved curves and choices of the correspondence between them, **z** (u) has a component in the plane perpendicular to the tangent to **s** (u). The unit vector parallel to this planar component of **z** (u) is ν_{AB}. Tw_{AB} is then

$$Tw_{AB} = \frac{1}{2\pi} \int_B \left[\vec{T}_B \times \vec{\nu}_{BA} \right] \cdot d\vec{\nu}_{BA} \tag{1.3}$$

where $\mathbf{T}_B$ is the unit tangent to B. For a more detailed discussion, see the paper by White and Bauer (1986).

Writhe is a property of a single space curve and is given by

$$Wr = \frac{1}{4\pi} \int_{C_1, C_2} du_1\, du_2 \frac{\left[\vec{r}(u_1) - \vec{r}(u_2)\right] \cdot \left[\frac{d\vec{r}(u_1)}{du_1} \times \frac{d\vec{r}(u_2)}{du_2}\right]}{\left|\vec{r}(u_1) - \vec{r}(u_2)\right|^3} \tag{1.4}$$

The resemblance to equation (1.1) is not coincidental. A simple derivation of equations (1.2) - (1.4) from equation (1.1) is given by Frank-Kamenetskii and Vologodskii (1981). The derivation, which Frank-Kamenetskii and Vologodskii attribute to V.V. Anshelevich, depends upon expressing the two curves in equation (1.1) as **r** (u) and **r** (u) + ε **a** (u) , where **a** (u) is a unit vector normal to the tangent vector **r**' (u). The parameter ε must be small enough that Lk does not change if ε is decreased (i.e. in the terminology of differential geometry, **r** (u) and **r** (u) + ε **a** (u) form a ribbon or strip (Fuller, 1971)). The region of integration in equation (1.1) can then be divided into two parts: a small neighborhood around the line where $u_1 = u_2$, and the remaining region of integration. It is not difficult to show that, in the first region, equation (1.1) reduces to equation (1.3), while the contribution from the remaining region yields equation (1.4). The paper by Frank-Kamenetskii and Vologodskii (1981) is considerably easier to follow than White's original paper, which treated generalizations of equations (1.1) and (1.2) in many dimensions.

One property of Wr which can be deduced from equation (1.4) is that the writhe of a curve lying in a plane is zero. This can be seen from the numerator of equation (1.4). If **r** (u) lies in a plane, then the tangent vectors lie in the same plane, and the cross-product is perpendicular to the plane. Therefore the dot product is zero.

For the most part, the only forms of equation (1.4) used in this dissertation are the results for infinite single helices (Figure 1) and infinite plectonemic or interwound helices (Figure 2). These are, for the single helix,

$$Wr = N (1 - \sin \alpha) \tag{1.5}$$

while for the plectonemic helix,

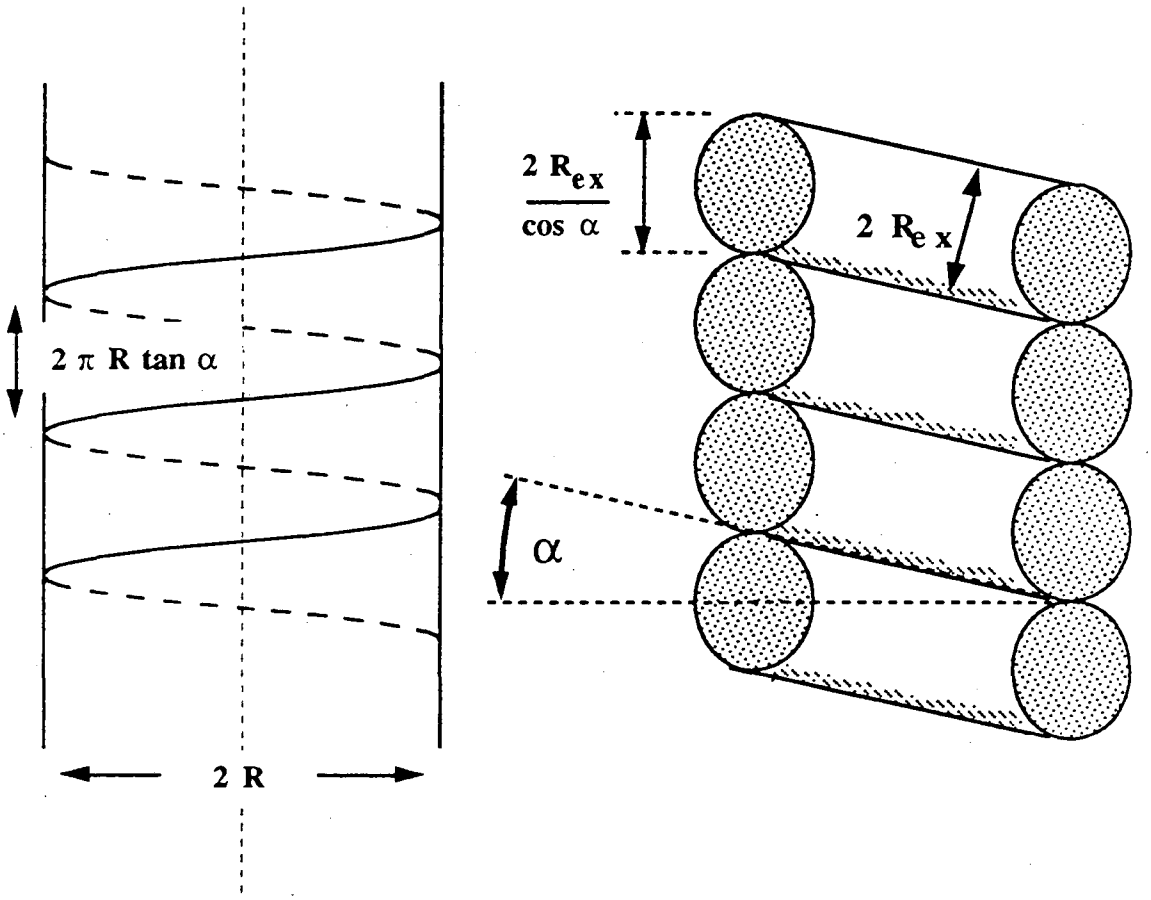
$$Wr = N \sin \alpha \tag{1.6}$$

where N is the number of helical turns and α is the pitch angle of the helix (Fuller, 1971). As usual, I pay little attention to the sign of writhe. Unless otherwise stated, Wr and α can be assumed positive. It is, however, worth noting that the handedness of the single helix and the plectonemic helix is opposite - i.e. a right-handed single helix has the same sign of writhe as a left-handed plectonemic helix.

Fuller (1971) derived the expressions for the writhe of single and plectonemic helices, equations (1.5) and (1.6), using geometrical arguments. He did this by considering the case in which the two curves are given by $\mathbf{r}$ (u) and $\mathbf{r}$ (u) + ε $\mathbf{a}$ (u) (i.e. the two curves define a strip) where $\mathbf{r}$ (u) describes either a single or a plectonemic helix and $\mathbf{a}$ (u) is the principal normal to $\mathbf{r}$ (u). In this case, the twist of the second curve about the first is known to be equal to the geometric torsion of the curve. (In the general case, the geometric torsion of a curve is not equal to the twist of a pair of curves. Nor should the geometric torsion be confused with a physical torsional deformation. The geometric torsion of a space curve is defined in Appendix I.) The geometric torsion τ of a helix with pitch angle α and radius R is given by

$$\tau = \frac{\sin \alpha \cos \alpha}{R} \tag{1.7}$$

Each turn of the helix has length $2 \pi R / \cos \alpha$, the total twist (in turns, rather than radians) of N turns of a helical strip is $N \sin \alpha$. In the case of the single helix, each turn of the helix contributes a unit to Lk, whereas in the plectonemic case, the contributions to Lk from each strand cancel. Thus, equations (1.5) and (1.6) follow by applying equation (1.2). Again, I have not paid attention to arithmetic signs.



(a) (b)

Figure 1. Long-length limit of toroidal DNA. a) Helical path of the axis of the DNA double helix. R - superhelix radius; α - pitch angle of superhelix; $2 \pi R \tan \alpha$ - pitch of superhelix. b) Cross-section showing excluded volume radius of the DNA. R_{ex} - excluded volume radius.

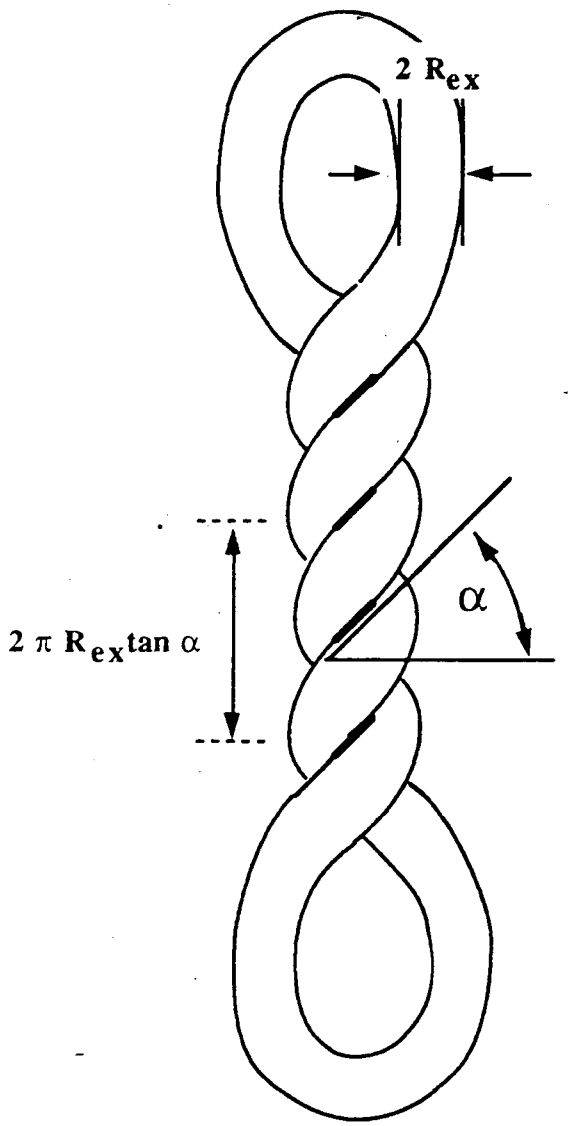


Figure 2. Finite plectonemic helix with end loops. Each strand represents the DNA double helix. α - superhelix pitch angle; R_{ex} - excluded volume radius; 2 π R_{ex} tan α - superhelix pitch.

I have verified equations (1.5) and (1.6) by numerical integration of the general writhe expression, equation (1.4), for the cases of single and plectonemic helices (see Appendix I). It seems likely that the integration can be carried out analytically, given the simplicity of the results, but I have been unable to do this.

Before applying these topological quantities to DNA, we re-write equation (1.2) in a more precise form:

$$Lk = Tw_{AB} + Wr_B \tag{1.8}$$

In words, the linking number of two curves A and B is the twist of A about B plus the writhe of B. Alternatively, Lk could be expressed as the sum of Tw_{BA} and Wr_A. In applying equation (1.8) to DNA, there are several possible choices for the curves A and B. One choice would be to take A and B to be the two sugar-phosphate backbones of the DNA double helix. Another choice is to take A to be either one of the sugar-phosphate backbones and B to be the axis of the DNA double helix. This second possibility turns out to be the most convenient choice for considering the energetics and conformation of supercoiled DNA. With this choice, a linear or planar piece of DNA has no writhe, and the twist of a relaxed piece of DNA, Tw_0, is just the number of base pairs divided by 10.4, i.e. Lk_0. With these definitions, we write

$$\Delta Lk = \Delta Tw + Wr \tag{1.9}$$

where $\Delta Lk = Lk - Lk_0$ and $\Delta Tw = Tw - Tw_0$.

A useful quantity is the specific linking difference σ:

$$\sigma = \Delta Lk / Lk_0 \tag{1.10}$$

σ is also called the superhelix density, but this is sometimes taken to imply writhe only. The term specific linking difference emphasizes that σ is the density of ΔLk, regardless of how it is partitioned into ΔTw and Wr.

An alternative topological description of DNA, known as surface linking theory, has proved useful for analyzing nuclease digestion of DNA bound in nucleosomes. The shortcomings of surface linking theory for plectonemically supercoiled DNA are described in section II.E.1.

B. DNA supercoiling - biological relevance

Supercoiled DNA was discovered by Vinograd *et al.* in 1965 (Vinograd *et al.*, 1965). Their work, described in a recent review by Lebowitz (1990), relied largely on sedimentation and electron microscopy studies of a plasmid. The principal observations were that the same plasmid could have substantially different sedimentation coefficients, and that these coefficients were correlated with the appearance of the plasmids on electron micrographs. The plasmids which ran more slowly appeared spread out while those which ran more quickly appeared coiled up with themselves, in what is now called a branched plectonemic structure.

Although supercoiling was discovered in plasmids and is often studied in them, any piece of DNA which is topologically restricted can become supercoiled. Such topological anchoring can be provided by attachment to the nuclear membrane (Paulson and Laemmli, 1977) or by large protein complexes bound to DNA, such as RNA polymerase and associated mRNA and ribosomes.

DNA supercoiling has been shown to be relevant to a number of biological phenomena. Plasmids isolated from bacteria are supercoiled, typically with specific linking difference $\sigma = -0.06$. Liu and Wang (1987) demonstrated that transcription of DNA can generate positively and negatively supercoiled domains of DNA on opposite

sides of a transcribing RNA polymerase. Baker *et al.* (1986) showed that an intermediate in the replication of a plasmid was highly underwound. Studies by Nash (1990) and Bliska and Cozzarelli (1987) have demonstrated the central role supercoiling plays in DNA recombination. Sekiguchi and Kmiec (1989) have studied the role of DNA supercoiling in determining the transcriptional activity of chromatin.

The specific observations mentioned in the previous paragraph can be thought of as subsidiary to the larger question of the biological role of DNA supercoiling. A phenomenon as ubiquitous as DNA supercoiling must serve some essential biological purpose or purposes. Supercoiling may serve merely to facilitate DNA packing, as in chromatin. In chromatin, it is known that the wrapping of DNA around histone proteins gives the DNA net writhe, which can be made to manifest itself as supercoiling when the histones are removed. Supercoiling may also serve to regulate genetic expression by its influence on the rate of DNA transcription. While negative supercoiling (underwinding) enhances DNA transcription *in vitro*, whether it regulates transcription *in vivo* remains uncertain (D.N. Cook, personal communication).

One phenomenon related to DNA supercoiling which proves to be a useful test and application of the theory developed here is DNA conformational transitions. When DNA is supercoiled to a sufficient degree, it is prone to transitions which relieve the resulting stress, which is primarily torsional. The two conformational transitions which we will examine are cruciform formation and B-Z transitions. Both occur for $\sigma < 0$, i.e. for DNA which is underwound. Cruciform formation can occur at DNA sequences which are inverted repeats. An example of an inverted repeat and the cruciform structure it can form is shown in Figure 3. Such a structure adds (N / 10.4) turns to the remaining DNA, where N is the number of base pairs in the cruciform. This relieves stress induced by supercoiling when $\sigma < 0$ (i.e. when the DNA is underwound). In the

GATT AAAGGATCTT CT TGAGAT CCTTT TT TT
CTAATT T CCTAGAAGAACT CTAGGAAAAAAA

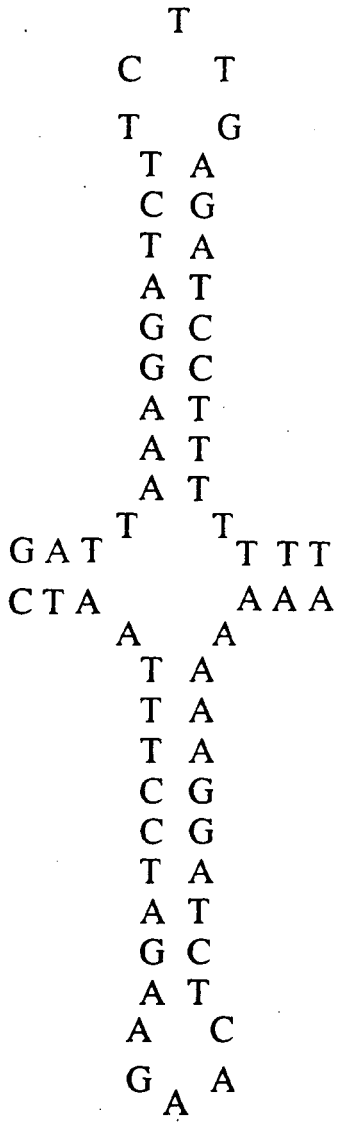


Figure 3. Top: DNA sequence containing an inverted repeat (underlined). Bottom: The cruciform structure which can be formed by the inverted repeat.

B-Z transition, a section of the right-handed DNA double helix flips to form a left handed double helix. This adds approximately (N / 10.4 + N / 12.0) turns to the remaining DNA, where 12.0 is the helical repeat of Z DNA (Rich *et al.*, 1984). As in the case of cruciform formation, this process relieves stress induced by supercoiling.

C. Some experimental techniques

Of the two techniques used in early studies of DNA supercoiling, sedimentation and electron microscopy, the latter is still widely used. Sedimentation has been largely superseded by gel electrophoresis, in which DNA, being charged, is pulled through a gel by an electric field. As in the case of sedimentation, the rate at which DNA plasmids move through the gel is a function of their linking difference. With this technique, plasmids which differ only in their values of Lk, known as topoisomers, can be separated.

One important type of experiment has been measurements of thermal distributions of topoisomers. In these experiments, plasmids are put in solution with topoisomerase I, an enzyme which relaxes superhelical stress by making and then closing single-strand nicks in DNA. The resulting distribution of topoisomers is determined by gel electrophoresis and is Gaussian, implying that the free energy of supercoiling is quadratic in σ. The consensus value is

$$E_{super} = 10\, RT\, N\, \sigma^2 \qquad (1.11)$$

where R is the gas constant and N the number of base pairs in the plasmid (Depew and Wang, 1975; Pulleyblank *et al.*, 1975). By comparing experimental results with the results of Monte Carlo simulations (see section I.D), it is possible to determine values of

the DNA bend and torsion constants, β_{bend} and β_{twist} (Shore and Baldwin, 1983; Horowitz and Wang, 1984; Shimada and Yamakawa, 1988).

A related experiment is the ring-closure technique. Linear pieces of DNA are put in solution with DNA ligase, an enzyme which ligates the pieces together. It also makes plasmids out of linear pieces of DNA. By examining the relative reaction rates, it is possible to determine the energy of supercoiling, as in the topoisomer distribution experiment. Shimada and Yamakawa (1985, 1984) obtained values of β_{bend} and β_{twist} from DNA ligase experiments by comparison with theoretical calculations based on the earlier work of Jacobson and Stockmeyer (1950) on the conformational entropy of simple polymers with excluded volume. Shimada and Yamakawa's theoretical treatment of the writhe of polymer conformations involves approximations which are only valid for short pieces of DNA and low values of Wr, as they note.

The analysis presented in this thesis depends upon interactions between DNA double helices. Results are taken from two experimental techniques which measure the effective DNA radius for interhelical interactions, known as the excluded radius. These are osmotic pressure measurements of the second virial coefficient of concentrated DNA solutions, and measurements of the sedimentation behavior of short linear pieces of DNA.

In the osmotic pressure measurements (Yarmola *et al.*, 1985), DNA in solution is equilibrated with a DNA-free solution through a membrane which allows exchange of solvent and salt. The second virial coefficient is extracted by varying the DNA concentration. This coefficient is affected by the fact that the concentrations of the salt anions and cations will differ on different sides of the membrane. This is known as the Donnan effect and arises because DNA is itself a polyanion. By comparing the second virial coefficient with the cluster integral theory developed by McMillan and Mayer, an effective excluded radius for the DNA can be defined. The theory is involved, and an

introductory discussion and references can be found in Stigter (1987). A basic introduction to the Donnan effect is given in Richards (1980).

Brian *et al.* (1981) studied the sedimentation behavior of concentrated DNA solutions in an ultracentrifuge. The concentration of DNA as a function of position (and hence as a function of pressure) within the ultracentrifuge is determined by measuring the absorbance at 302 nm. The theory is similar to that for the osmotic pressure case mentioned in the previous paragraph. Brian *et al.* give an introductory treatment of the theory and provide references.

D. Conformation and energy - previous theoretical work

Fuller (1971) first applied equation (1.8) to DNA. In his 1971 paper he clearly outlined much of the math and physics of supercoiled DNA conformation which is given in this dissertation. He derived the equations for the writhe of single and plectonemic helices, equations (1.5) and (1.6), using the geometric arguments given above. Fuller also noted that, for sufficiently large ΔLk, the single helical conformation must become energetically favored over the plectonemic form, since it can accommodate more writhe per length. This statement is made quantitative in section II.D.

Benham (1977, 1978, 1979, 1983) and Le Bret (1978, 1979, 1980, 1984) independently investigated the equations of equilibrium for an elastic rod under topological constraint. Benham and LeBret both found equilibrium solutions of the elastic equations which approximate toroids. However, they were also able to show that such conformations are mechanically unstable. The essential problem with this approach is the difficulty of incorporating helix-helix interactions (excluded volume effects). Benham and LeBret did not attempt to include these effects. This would not be a problem if the minimum-energy conformations did not involve excluded volume effects. However, plectonemic conformations can yield arbitrarily large values of

writhe for arbitrarily small cost in bending energy if excluded volume effects are ignored (Fuller, 1971). This fact, demonstrated in section II.C, shows that excluded volume must play a central role in determining minimum-energy conformations. Benham (1982) nevertheless outlined the application of supercoiling to DNA conformational transitions (see chapter III). Le Bret (1984) developed a nice description of the behavior of a closed circle under increasing topological stress and the initial formation of the plectonemic conformation.

The theory of elasticity analysis followed by Benham and LeBret involved looking for solutions to the coupled equations

$$\vec{M} = \beta_{bend}\, \vec{t} \times \frac{d\vec{t}}{ds} + \beta_{twist}\; 2\,\pi\, \rho\Delta Tw\, \vec{t}$$
$$\frac{d\vec{M}}{ds} = \vec{F} \times \vec{t} \tag{1.12}$$

where **M** is the torsional stress, **F** is the force in the DNA (**F** is zero for DNA in the absence of excluded volume interactions), $\rho\Delta Tw$ is the twist per unit length, **t** is the unit tangent vector, and s is the arc-length. A different but equivalent formulation of the problem is to write a static Hamiltonian for the energy per unit length:

$$H = \frac{\beta_{bend}}{2}\kappa^2 + \frac{\beta_{twist}}{2}(2\pi)^2\,(\rho\Delta Tw)^2 + H_{excl} \tag{1.13}$$

where κ is the curvature of the DNA backbone and H_{excl} represents the excluded volume interaction. The previous theoretical studies discussed above ignored H_{excl}, while those discussed below take H_{excl} to be a hard-body excluded radius of the DNA. I use an excluded radius in chapter II, while in chapter III H_{excl} is a screened electrostatic interaction between DNA strands.

One conclusion which can be drawn from equations (1.12) or equation (1.13) is that minimum-energy or stationary solutions will always have constant twist density $\rho\Delta Tw$. This result is left implicit in several of the earlier theoretical studies.

Three studies previous to my work have focussed on an elastic treatment of the plectonemic supercoil. Tanaka and Takahashi (1985) published minimum-energy conformations for finite-length loops, i.e. plectonemic helices with end loops (Figure 2). Tsuru and Wadati (Tsuru and Wadati, 1986; Wadati and Tsuru, 1986) did the same, with an improved model of the end loops. Neither study examined the resulting conformations in detail, and the inclusion of end loops, while clearly of some inherent interest, precluded the general conclusions we draw (see chapter II). Camerini - Otero and Felsenfeld (1978) considered a section of plectonemic supercoil and ignored end effects, but they committed several errors in their analysis. They erroneously equated the results of their model with the empirical quadratic dependence of the energy of supercoiling on linking difference. They also introduced a Lagrange multiplier constraint which they did not include consistently in their analysis. One result of these errors was their conclusion that the pitch angle α of the plectonemic helix lies in the range $45^o < \alpha < 54^o$. In chapter II we show the range to be $45^o < \alpha \leq 90^o$.

One theoretical approach to DNA supercoiling which has proved useful is Monte Carlo simulations. The basic idea of the Monte Carlo technique, as proposed by Metropolis *et al.* (1953) and as applied to DNA supercoiling, is to simulate an ergodic walk through phase space, calculating the internal energy at each conformation considered. The $(n + 1)^{th}$ conformation is generated from the nth conformation by a random re-arrangement. If the energy calculated for the $(n + 1)^{th}$ conformation is less than that for the nth conformation, the step is allowed. If the energy increases, the step is allowed with probability $\exp[(E (n) - E (n + 1)) / k_BT]$, where $E (n + 1)$ and $E (n)$ are the energy of the $(n + 1)^{th}$ and n^{th} conformation, respectively. The principal

advantages of the technique are its flexibility and the fact that entropic effects are included.

The application of the Monte Carlo technique to DNA supercoiling can be thought of as a simulation of the topoisomer distribution experiment (see section I.C). Random closed chains are generated. The internal energy of each conformation (as opposed to the free energy) is taken to be the bend energy. Since the conformations used consist of straight segments rather than a continuously flexible rod, the bend energy is the sum of the bend energies at each of the joints. The steps to new conformations usually consist of random rotations of a small number of the chain segments. Excluded volume effects are enforced by rejecting any steps which would bring any portion of two segments within an excluded volume diameter of each other. Effects of twist can be easily included, assuming that, for nicked plasmids (i.e. plasmids which are not topologically constrained due to nicks in the sugar-phosphate backbones) twist and writhe are statistically independent. The twist energy for the Monte Carlo study is identical to that given in equation (1.13).

One problem with Monte Carlo simulations had been an inability to examine high-writhe conformations, since these occur rarely. This problem has recently been overcome by biasing the walk through conformation space in favor of high-writhe conformations (Klenin *et al.*, 1991).

In spite of the recent success of the Monte Carlo treatment, I believe there is still plenty of room for the approach outlined in this dissertation. A major drawback of the Monte Carlo technique is the amount of computer time required. This in turn limits the molecular length and level of supercoiling which can be simulated. For the same reason, the treatment of the electrostatic interaction has been limited to a simple hard-body excluded volume. All of these limitations are overcome in the approach developed here. There are also subtle numerical and statistical pitfalls in the Monte Carlo

technique, so it is helpful to have independent theoretical approaches. A comparison of the results of this thesis and the Monte Carlo results is given in section III.D.3.

E. Entropic Contributions to the Free Energy of Supercoiling

For the most part, entropic effects are not included in this dissertation. In this section, we survey the approaches which have been taken to the conformational entropy of supercoiling.

The problem of conformational degrees of freedom can be thought of formally as the evaluation of the partition function (Shimada and Yamakawa, 1985)

$$Z(k_n, \{\vec{r}\}) = \int D\{\vec{r}\} \exp\left[-\beta U\{\vec{r}\}\right] \tag{1.14}$$

where the integral is a Wiener functional integral (Freed, 1987) over closed chains $\{\mathbf{r}\}$, k_n denotes a given family of knotted curves (usually k_0, unknotted curves), and $U\{\mathbf{r}\}$ is the sum of bend and twist energies for the conformation $\{\mathbf{r}\}$ and the given linking number. In practice, equation (1.14) seems intractable when topological constraints and excluded volume are included.

There have been a number of approaches to evaluating entropic contributions to the free energy of supercoiling. Most of these are intended to help interpret the results of topoisomer distribution and ring-closure experiments (see section I.C). The studies can be divided into analytical analyses and Monte Carlo models.

The most successful analytical model of topoisomer and ring-closure data is that of Shimada and Yamakwa (Shimada and Yamakwa, 1985, 1984; Yamakawa, 1984). Their approach, based on the earlier work of Jacobson and Stockmeyer (1950), was to consider only small deviations from equilibrium conformations derived from the theory

of elasticity by LeBret (1984). Since LeBret considered only figure-8 and toroidal conformations, Shimada and Yamakawa's analytical results may be questioned for $Wr \geq 2$, but this is the upper limit of the values of Wr which they considered. They did in fact revise their conclusions in a later Monte Carlo study (Shimada and Yamakawa, 1988).

A recent analytical study is that of Hearst and Hunt (1991). In that paper, the plectonemic helix with end-loops is approximated by the leading terms of a spatial Fourier transform. Numerical calculations of Wr are carried out for these helices and for other combinations of simple sines and cosines. It is found that the plectonemic helices have greater writhe per bend energy than other simple space curves. By comparing spatial thermal r.m.s. deviations with the dimensions of the lowest energy plectonemic conformations, it is concluded that, at sufficiently high $|\sigma|$, the entropic degrees of freedom are greatly restricted and the molecule cannot deviate significantly from a plectonemic helix.

The analysis of Hearst and Hunt (1991) has a closer bearing on the work presented in this dissertation than the other treatments of entropy in the literature, since it is applicable to the cases of high $|\sigma|$ and long molecular length. At a number of points, the analysis depends on linear approximations to non-linear expressions, such as the expressions for contour length, bend energy and writhe.

There have been quite a number of Monte Carlo treatments of the free energy of supercoiling (Jacobson, 1969; Vologodskii *et al.*, 1979; LeBret, 1980; Chen, 1981; Iwata, 1983; Levene and Crothers, 1986a,b). The most successful of these have been the analysis of ring-closure data by Shimada and Yamakawa (1988) and the recent study of Klenin *et al.* (1991). In the latter, the problem of simulating large values of Wr has been largely overcome by biasing the Monte Carlo walks to favor high-writhe conformations. As noted above and in section III.E, the method still suffers from

requirements for large amounts of computer time and consequent limits on the molecular length, values of $|\sigma|$, and treatments of inter-strand interactions which can be modelled.

References

Baker, T.A., Sekimuzu, K., Funnell, B.E., and Kornberg, A. (1986) *Cell* **45**, 53

Benham, C.J. (1977) *Proc. Nat. Acad. Sci. U.S.A.* **74**, 2397

Benham, C.J. (1978) *J. Mol. Biol.* **123**, 361

Benham, C.J. (1979) *Biopolymers* **18**, 609

Benham, C.J. (1982) *Biopolymers* **21**, 679

Benham, C.J. (1983) *Biopolymers* **22**, 2477

Bliska, J.R. and Cozzarelli, N.R. (1987) *J. Mol. Biol.* **194**, 205

Brian, A.A., Frisch, H.L., and Lerman, L.S. (1981) *Biopolymers* **20**, 1305

Camerini-Otero, R.D. and Felsenfeld, G. (1978) *Proc. Nat. Acad. Sci. U.S.A.* **75**, 1708

Chen, Y.-D. (1981) *J. Chem. Phys.* **75**, 2447

Depew, R.E. and Wang, J.C. (1975) *Proc. Nat. Acad. Sci. U.S.A.* **72**, 4275

Frank-Kamenetskii, M.D. and Vologodskii, A.V. (1981) *Sov. Phys. Usp.* **24**, 679

Freed, K.F. (1987) *Renormalization Group Theory of Macromolecules*, J. Wiley and Sons, New York

Fuller, F.B. (1971) *Proc. Nat. Acad. Sci. U.S.A.* **68**, 815

Horowitz, D.S. and Wang, J.C. (1984) *J. Mol. Biol.* **173**, 75

Hearst, J.E. and Hunt, N.G. (1991) *J. Chem. Phys.* **95**, 9322

Iwata, K. (1983) *J. Chem. Phys.* **78**, 2778

Jacobson, H. (1969) *Macromolecules* **2**, 650

Jacobson, H. and Stockmeyer, W.H. (1950) *J. Chem. Phys.* **18**, 1600

Klenin, K.V., Vologodskii, A.V., Anshelevich, V.V., Dykhne, A.M., and M.D. Frank-Kamenetskii (1991) *J. Mol. Biol.* **217**, 413

Lebowitz, J. (1990) *TIBS* **15**, 202

LeBret, M. (1978) *Biopolymers* **17**, 1939

LeBret, M. (1979) *Biopolymers* **18**, 1709

LeBret, M. (1980) *Biopolymers* **19**, 619

LeBret, M. (1984) *Biopolymers* **23**, 1835

Levene, S.D. and Crothers, D.M. (1986a) *J. Mol. Biol.* **189**, 61

Levene, S.D. and Crothers, D.M. (1986b) *J. Mol. Biol.* **189**, 72

Liu, L.F. and Wang, J.C. (1987) *Proc. Nat. Acad. Sci. U.S.A.* **84**, 7024

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) *J. Chem. Phys.* **21**, 1087

Nash, H.A. (1990) *TIBS* **15**, 222

Paulson, J.R. and Laemmli, J.K. (1977) *Cell* **12**, 817

Pulleyblank, D.E., Shure, M., Tang, D., Vinograd, T., and Vosberg, H.-P. (1975) *Proc. Nat. Acad. Sci. U.S.A.* **72**, 4280

Rich, A., Nordheim, A., and Wang, A.H.-J. (1984) *Ann. Rev. Biochem.* **53**, 791

Richards, E.G. (1980) *An Introduction to the Physical Properties of Large Molecules in Solution* Cambridge University Press, Cambridge

Sekiguchi, J.M. and Kmiec, E.B. (1989) *Mol. Gen. Genet.* **220**, 73

Shimada, J. and Yamakawa, H. (1984) *Macromolecules* **17**, 689

Shimada, J. and Yamakawa, H. (1985)*J. Mol. Biol.* **184**, 319

Shimada, J., and Yamakawa, H. (1988) *Biopolymers* **27**, 657

Shore, D., and Baldwin, R.L. (1983) *J. Mol. Biol.* **170**, 983

Stigter, D. (1987) *Cell Biophysics* 11, 139

Tanaka, F., and Takahashi, T. (1985) *J. Chem. Phys.* **83**, 6017

Tsuru, H. and Wadati, M. (1986) *Biopolymers* **25**, 2083

Vinograd, J., Lebowitz, J., Radloff, R., Watson, R., and Laipis, P. (1965) *Proc. Nat. Acad. Sci. U.S.A.* **53**, 1104

Vologodskii, A.V., Anshelevich, V.V., Lakashin, A.V. and Frank-Kamenetskii, M.D. (1979) *Nature* **280**, 294

Wadati, M. and Tsuru, H. (1986) *Physica D* **21**, 213

White, J.H. (1969) *Am. J. Math.* **91**, 693

White, J.H. and Bauer, W.R. (1986) *J. Mol. Biol.* **189**, 329

Yamakawa, H. (1984) *Ann. Rev. Phys. Chem.* **35**, 23

Yarmola, E.G., Zarudnaya, M.I., and Lazurkin, Yu. S. (1985) *J. Biomol. Str. and Dynamics* **2**, 981

II. Elastic Model of Supercoiling with Hard-Body Excluded Volume

A. Introduction

In this chapter we present elastic models of plectonemic and toroidal supercoiling. The electrostatic interaction is treated as a hard-body excluded volume. This material is mostly covered in the paper by Hunt and Hearst (1991).

All extensive quantities (e.g. writhe, bend energy) are treated as linear densities (writhe per unit length, bend energy per unit length). Minimum energy conformations are found. For low values of specific linking difference, most linking difference is in writhe. As specific linking difference increases, a greater proportion of linking difference is in twist. Ionic strength effects are discussed, and it is found that variation of excluded volume with ionic strength has a large effect, resulting in significantly greater torsional stress on supercoiled DNA at low ionic strength. For biologically relevant levels of specific linking difference, the plectonemic conformation is energetically favored over the toroidal form.

Many of the results in this chapter can be summarized by two main points. The first is that the writhe per unit length, ρWr, for a plectonemic helix has a maximum value, ρWr_{max}. For low values of specific linking difference $|\sigma|$ ($\sigma = \Delta Lk / Lk_0$), most linking difference can be taken up in writhe at very little energetic cost. At greater values of $|\sigma|$, where the writhe density is a substantial fraction of ρWr_{max}, linking difference goes increasingly into twist, at greater energetic cost. The energy of supercoiling which results is not simply a quadratic function of specific linking difference, as has often been assumed based on extrapolations of measurements made over limited ranges of specific linking difference (Depew and Wang, 1975; Pulleyblank *et al.*, 1975; Horowitz and Wang, 1984).

The second main point is the importance of DNA excluded volume. Osmotic pressure measurements of Donnan equilibria (Yarmola *et al.*, 1985), measurements of

the sedimentation behavior of short pieces of DNA (Brian *et al.*, 1981), and a theoretical treatment of the counter-ion atmosphere of DNA (Stigter, 1987) all show that the DNA excluded volume radius R_{ex} is a strong function of ionic strength, varying from about 15Å at high ionic strength to over 100Å at low ionic strength . ρWr_{max} is $1/(4 \pi R_{ex})$, so that six times more linking difference can be taken up by writhe at high ionic strength than at low ionic strength. Thus at low ionic strength, linking difference starts going into twist at much lower values of $|\sigma|$, resulting in significantly greater torsional stress. Since torsional stress can induce DNA conformational transitions (Benham, 1987; Peck and Wang, 1983; Gellert *et al.*, 1979; Lilley, 1980; Panayotatos and Wells, 1981; Lee and Bauer, 1985; Lyubchenko and Shlyakhtenko, 1988) this effect may account for some of the ionic strength dependence of transitions induced under superhelical stress at low ionic strength (Singleton *et al.*, 1982; Nordheim and Rich, 1983; Singleton, 1983; Kozyavkin *et al.*, 1990).

B. Formulation

1. Bend energy, twist energy, writhe and twist

Giving a DNA molecule twist or writhe requires energy. In this chapter we calculate the bending and twisting energies of two DNA conformations - the interwound form, or plectonemic helix (Figure 2); and the toroidal or solenoidal form (Figure 1). In both cases we consider the limit of long molecules and ignore end effects. All extensive quantities (e.g. writhe, bend energy) are treated as linear densities (writhe per length, bend energy per length, etc.). DNA is treated as an elastic rod characterized by a torsional rigidity β_{twist} , a bending rigidity β_{bend} , and an excluded volume radius R_{ex}.

Changing the twist from its equilibrium value has an energy cost per unit length

$$\rho E_{twist} = \frac{\beta_{twist}}{2} (2\pi)^2 \ (\rho \Delta Tw)^2 \tag{2.1}$$

where ρEt_{wist} is the twist energy per unit length and $\rho \Delta Tw$ is the deviation from the equilibrium twist per unit length. The factor of $(2\pi)^2$ converts twist from turns to radians, as required by the usual definition of β_{twist}.

The axis of a DNA molecule with net writhe curves in three dimensions. The bending energy per unit length, ρE_{bend}, of a DNA molecule is

$$\rho E_{bend} = \frac{\beta_{bend}}{2} \kappa^2 \tag{2.2}$$

where β_{bend} is the bending constant and κ is the curvature of the molecular axis. Changing the writhe of a DNA molecule changes the bend energy of the molecule, but there is no general relation between writhe and bend energy such as the relation equation (2.1) between twist and twist energy. The energy cost of adding or subtracting writhe depends on the molecule's specific geometry and is different for the plectonemic and toroidal conformations.

2. Writhe per unit length

The writhe of a space curve **r**(u), where u is some parameterization of the curve, is given by equation (1.4) (White, 1969; Vologodskii *et al.*, 1979). Writhe is a global property of a curve. If u is the arc-length rather than an arbitrary parameterization, we can define a local quantity, the writhe per unit length, by

$$\rho Wr\,(u_0) = \frac{1}{4\pi} \int_{C'} du' \; \frac{(\vec{r}(u_0) - \vec{r}\,(u'))\left(\frac{d\vec{r}\,(u)}{du}\right)_{u_0} \times \frac{d\vec{r}\,(u')}{du'}}{(\vec{r}\,(u_0) - \vec{r}\,(u'))^3} \tag{2.3}$$

For the infinite helices considered in this chapter, ρWr is independent of position on the curve, since the helices are spatially homogeneous modulo a translation and rotation, neither of which affect the writhe integral.

The curve **r**(u) is taken to be the axis of the DNA double helix. For a discussion of other possible choices, see section II.E.1.

The sign of the writhe depends on the handedness of the helix. The writhe of a plectonemic helix is positive if the helix is left-handed, while the writhe of a single helix is positive if it is right-handed. In this chapter we take α , σ and ρWr as positive. The results for $\sigma < 0$ are identical to those for $|\sigma|$, the only difference being the handedness of the helix.

One turn of a helix of radius R has length $2\pi R / \cos\alpha$. Using equation (1.5) and equation (1.6), the writhe per unit length for a plectonemic helix is

$$\rho Wr = \frac{\sin\alpha \cos\alpha}{2\pi R} \tag{2.4}$$

while for a single helix

$$\rho Wr = \frac{(1 - \sin\alpha)\cos\alpha}{2\pi R} \tag{2.5}$$

Equations (2.4) and (2.5) are exact for infinite helices, but only approximate for finite helices. The end effects can be calculated by comparing equations (2.4) and (2.5) with the exact results given by equation (2.3). Such a comparison is carried out in

Appendix I. For helices of more than two turns, end effects are negligible. Therefore the results of this chapter can be applied to finite helices without altering the expressions for ρWr.

C. Energetics of the Plectonemic Helix

Fuller (1971) pointed out that a plectonemic helix has a high value of writhe per bend energy. In fact, the energy cost can be made arbitrarily small if the two strands are allowed arbitrarily close - i.e. if excluded volume effects (and end effects) are ignored (see below). The plectonemic helix was therefore a promising candidate for the tertiary structure of superhelical DNA. One can also arrive at the interwound or plectonemic helix as a plausible structure by twisting a string or piece of tubing (LeBret, 1984).

For the plectonemic helix, R is taken to be the DNA excluded volume radius, R_{ex}. The curvature κ of a helix with radius R_{ex} is

$$\kappa = \frac{\cos^2\alpha}{R_{ex}} \tag{2.6}$$

so the bend energy per unit length ρE_{bend} is

$$\rho E_{bend} = \frac{\beta_{bend} \cos^4\alpha}{2 R_{ex}^2} \tag{2.7}$$

Using equation (1.2) in the form

$$\rho \Delta Tw = \frac{\sigma}{h_0} - \rho Wr \tag{2.8}$$

where h_0 is the equilibrium helical repeat (of the DNA double helix, not of the superhelix) in units of length, the total energy per unit length, ρE_{tot}, is, for the plectonemic helix

$$\rho E_{tot} = \rho E_{bend} + \rho E_{twist} = \frac{\beta_{bend} \cos^4\alpha}{2 R_{ex}^2} + \frac{\beta_{twist}}{2} (2\pi)^2 \left[\frac{\sigma}{h_0} - \frac{\sin\alpha \cos\alpha}{2 \pi R_{ex}}\right]^2 \qquad (2.9)$$

ρE_{tot} is plotted in Figures 4 - 6 as a function of α for different values of σ and R_{ex}.

The energy minimum for the plectonemic helix always lies in the pitch angle range

$$45° < \alpha_{opt} \leq 90° \qquad (2.10)$$

where α_{opt} denotes the pitch angle corresponding to minimum total energy. This is because the writhe per unit length, ρWr (given in equation 2.4), has its maximum $\rho Wr_{max} = 1/(4 \pi R_{ex})$ at $\alpha = 45^o$, whereas the writhe per bend energy

$$\frac{Wr}{E_{bend}} = \frac{\rho Wr}{\rho E_{bend}} = \frac{R_{ex} \sin\alpha}{\pi \beta_{bend} \cos^3\alpha} \qquad (2.11)$$

becomes infinite as $\alpha \rightarrow 90^o$. For low values of superhelix density σ, α_{opt} will be near 90°, where writhe is energetically cheap. As σ increases, α_{opt} decreases towards 45° where writhe is more costly, but ρWr is greater. Any conformation with $\alpha < 45^o$ (Torbet and DiCapua, 1989) is not a minimum energy conformation, since there is then a conformation with pitch angle $\alpha' = 90^o - \alpha$ which has the same ρWr (and therefore the same twist energy) but with lower bend energy.

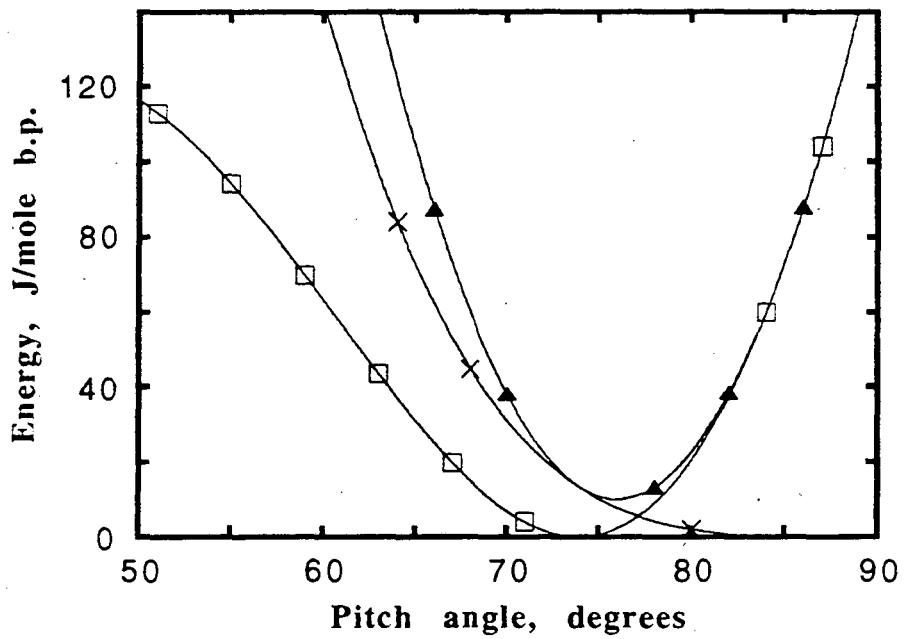


Figure 4. ρE_{bend} (x's), ρE_{twist} (squares) and ρE_{tot} (triangles) as functions of superhelix pitch angle. ρE_{twist} is zero when the writhe density ρWr equals the linking difference per unit length, σ / h_0. $R_{ex} = 30Å$, $\sigma = 0.05$, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm.

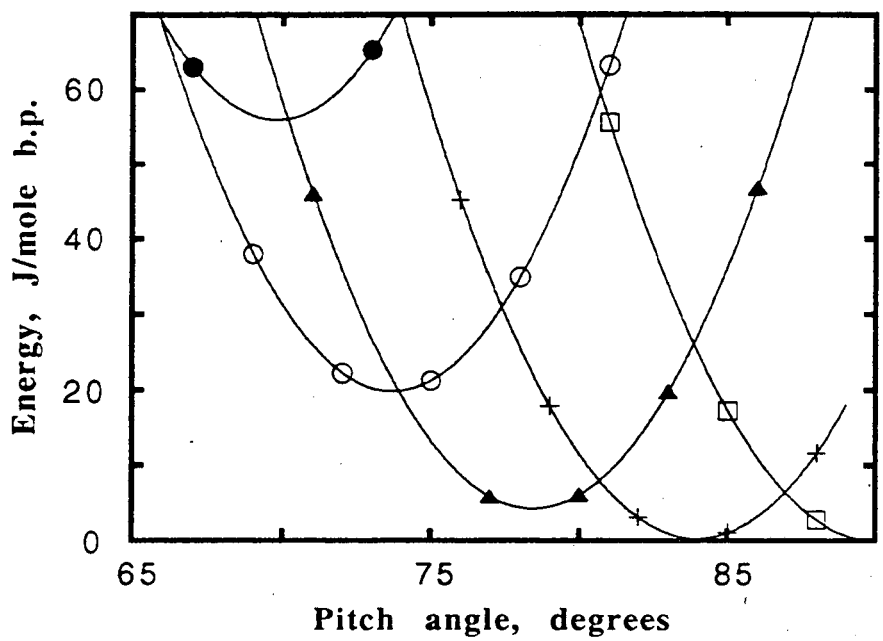


Figure 5. ρE_{tot} as a function of pitch angle for different values of σ. Squares - $\sigma = 0$; +'s - $\sigma = 0.02$; triangles - $\sigma = 0.04$; open circles - $\sigma = 0.06$; solid circles - $\sigma = 0.08$. The optimum pitch angle decreases towards the limiting value of 45° as σ increases. For all curves, $R_{ex} = 30$Å, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm.

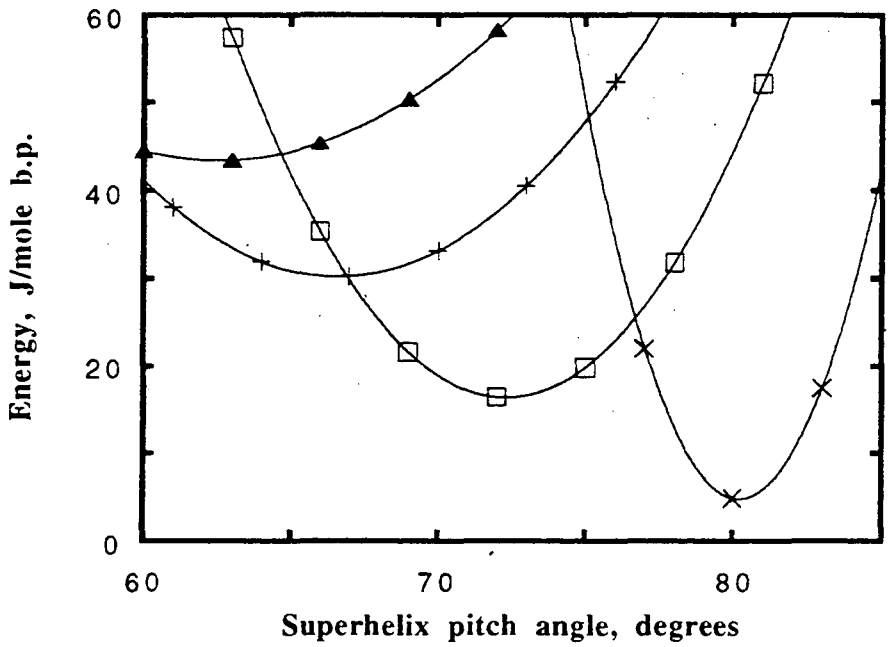


Figure 6. ρE_{tot} as a function of superhelix pitch angle for different values of the excluded volume radius R_{ex}. x's - R_{ex} = 20Å; squares - R_{ex} = 40Å; +'s - R_{ex} = 60Å; triangles - R_{ex} = 80Å. For all curves, $\sigma = 0.05$, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm.

We can now demonstrate the fact that, in the absence of excluded volume, the plectonemic conformation can give unbounded values of writhe for a finite cost in bend energy. The bend energy density is given by equation (2.7). If, as R is decreased, α is increased so that $\cos^2\alpha / R$ stays constant, then ρE_{bend} does not change. In the same limit, the writhe density given by equation (2.4) increases without bound.

The model can be parameterized in terms of ρWr in place of the pitch angle α. Using equation (2.4) we have

$$\cos^2\alpha = \frac{1 - \left[1 - 16\,\pi^2 R_{ex}^2 \rho Wr^2\right]^{1/2}}{2} \tag{2.12}$$

so that, substituting into equation (2.7),

$$\rho E_{bend} = \beta_{bend}\left[\frac{1 - \left[1 - 16\,\pi^2 R_{ex}^2 \rho Wr^2\right]^{1/2}}{4\,R_{ex}^2} - 2\,\pi^2 \rho Wr^2\right] \tag{2.13}$$

Equation (2.13) can be thought of as the energy cost of putting linking difference into writhe. Equation (2.13) is plotted with the corresponding expression for twist, equation (2.1), in Figure (7).

Minimum energy conformations can be found by minimizing the total energy ρE_{tot} with respect to pitch angle α. Setting $d\rho E_{tot} / d\alpha$ to zero gives

$$2\,\pi\,R_{ex}\frac{\sigma}{h_0} = \frac{2\,c\cos^3\alpha\sin\alpha}{\sin^2\alpha - \cos^2\alpha} + \sin\alpha\cos\alpha \tag{2.14}$$

where $c = \beta_{bend} / \beta_{twist}$. Equation (2.14) is an expression for the pitch angle α_{opt} which minimizes ρE_{tot} for a given R_{ex} and σ. Using equation (2.14), the values of ρWr, ρTw, $\rho Wr\, h_0 / \sigma$ and ρE_{tot} which correspond to α_{opt} are plotted as functions of σ in Figures 8 - 10.

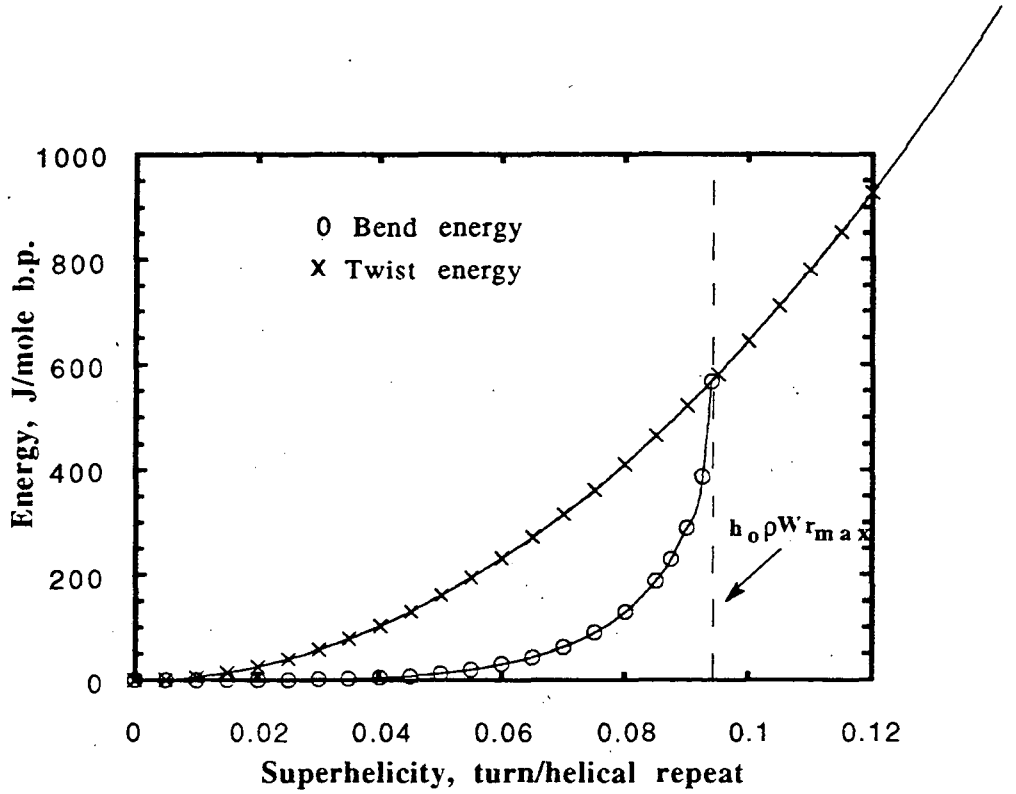


Figure 7. Energy cost of writhe (circles) and twist (x's). The twist cost is given by equation (2.1), the writhe cost by equation (2.13). The curve for writhe does not go beyond the value of maximum writhe density, or $h_0 \rho W r_{max} = 0.094$ in this case. At the point $\rho W r = \rho W r_{max}$, equation (2.13) has a finite maximum, $\beta_{bend}/(8 R_{ex}^2)$, but the slope of the curve is infinite. This figure should not be confused with Figure 8, which gives the distribution of twist and bend energies for energy-minimum conformations, or with Figure 9, which gives the distribution of twist and writhe for the same conformations. $R_{ex} = 30$Å, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm.

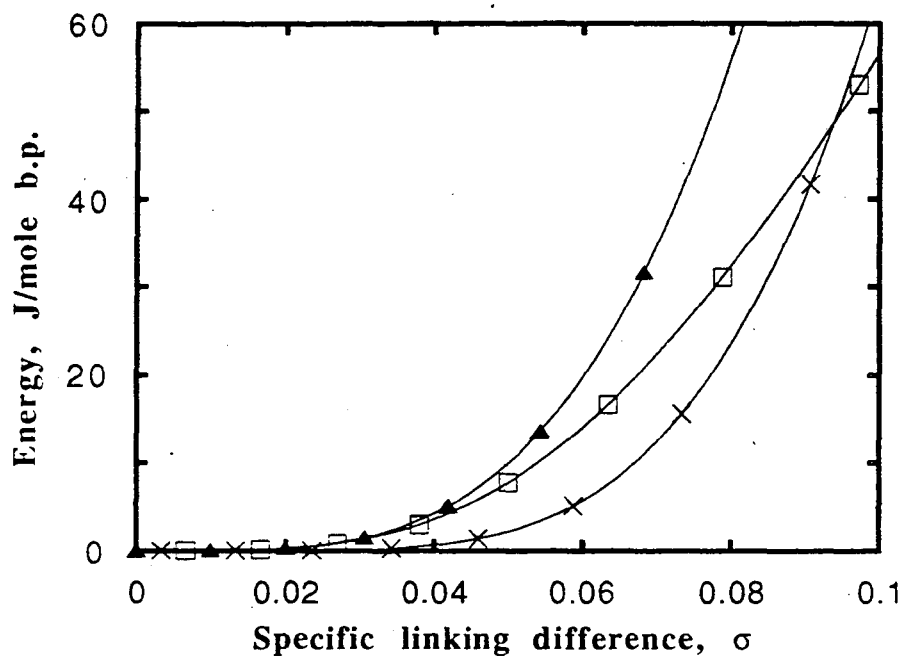


Figure 8. Bend (squares), twist (x's) and total (triangles) energies as functions of σ for minimum energy conformations. As σ increases, a greater proportion of superhelicity goes into twist, and twist energy dominates. $R_{ex} = 30Å$, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm.

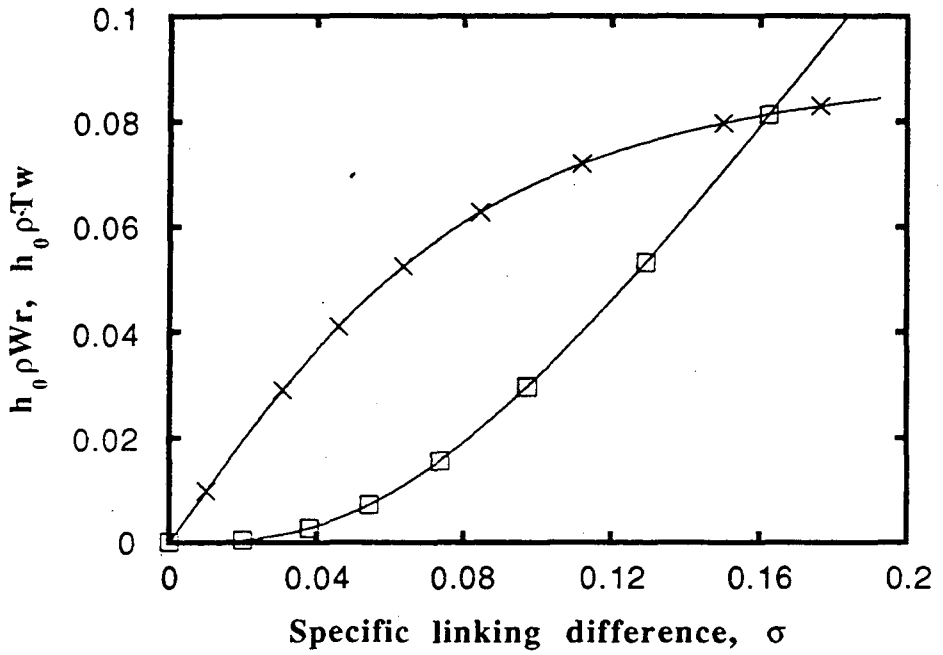


Figure 9. Superhelicity in twist ($h_0\ \rho\Delta Tw$) (squares) and writhe ($h_0\ \rho Wr$) (x's) as functions of σ, for minimum energy conformations. The sum of the two quantities is σ. ρWr approaches ρWr_{max} as σ approaches infinity. $R_{ex} = 30Å$, $\beta_{bend} = \beta_{twist}$.

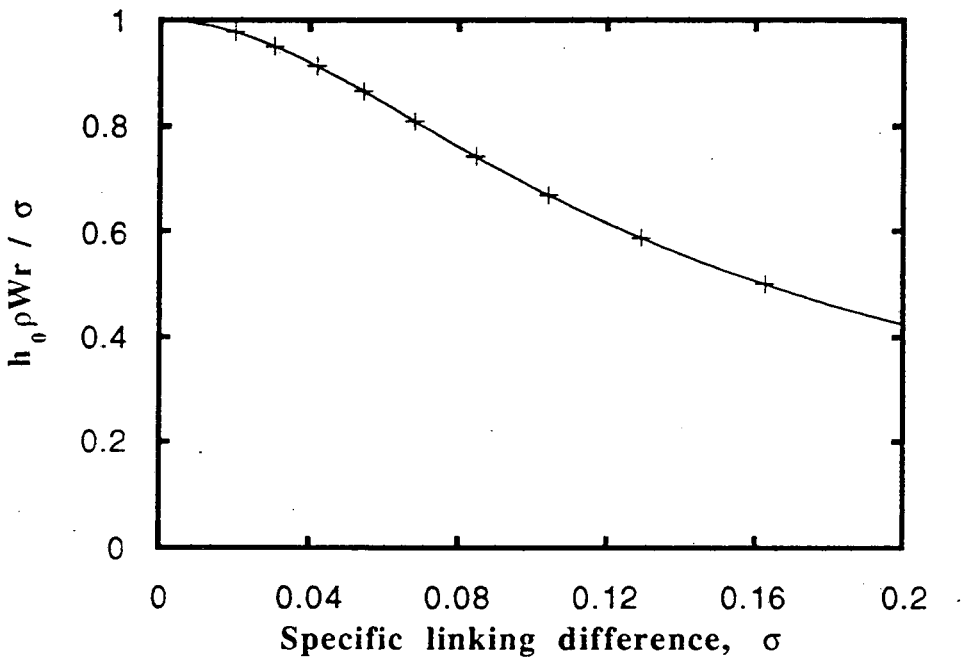


Figure 10. Proportion of superhelicity in writhe, $h_0 \rho Wr / \sigma$, as a function of σ, for minimum energy conformations. $R_{ex} = 30Å$, $\beta_{bend} = \beta_{twist}$.

If $\beta_{bend} = \beta_{twist} = \beta$ (c = 1), the expressions above have a simpler form and optimum values can be given as explicit functions of σ. For example, the values of ρE_{tot}, ρWr and α for minimum energy conformations are

$$\rho E_{tot} = \beta\left[\frac{[h_0^2 + 16\,\pi^2 R_{ex}^2\,\sigma^2]^{1/2} - h_0}{4\,R_{ex}^2[h_0^2 + 16\,\pi^2 R_{ex}^2\,\sigma^2]^{1/2}} + \frac{2\,\pi^2\sigma^2}{h_0^2}\right] - \beta\left[\frac{4\,\pi^2\,\sigma^2}{h_0\,[h_0^2 + 16\,\pi^2 R_{ex}^2\,\sigma^2]^{1/2}}\right] \tag{2.15}$$

$$\rho Wr = \frac{\sigma}{[h_0^2 + 16\,\pi^2 R_{ex}^2\,\sigma^2]^{1/2}} \tag{2.16}$$

$$\alpha_{opt} = \frac{\pi}{2} - \frac{1}{2}\,\mathrm{Tan}^{-1}\left[\frac{4\,\pi\,R_{ex}\,\sigma}{h_0}\right] \tag{2.17}$$

Equations 2.15 - 2.17 do not differ much qualitatively from optimized values obtained using equation (2.14) for c in the range of its experimental values (roughly between 0.67 and 1.5). They can therefore be used to get an idea of the approximate behavior of ρE_{tot} and ρWr.

D. Comparison of plectonemic and toroidal structures

An alternate structure to the plectonemic helix is a toroidal structure. In the limit of infinite length, the toroidal structure becomes equivalent to a single helix (Figure 1). While most experimental evidence favors the plectonemic form for DNA in solution (Boles *et al.*, 1990; Spengler *et al.*, 1985) there have been some reports of evidence for the toroidal structure (Brady *et al.*, 1987, 1983). Fuller (1971) noted that, at

sufficiently high specific linking difference, the single helix must become energetically favored over the plectonemic helix, since ρWr has a greater maximum value for the single helix than for the plectonemic helix. By calculating the energy per unit length of a single helix and comparing the results with those obtained above for the plectonemic helix, we will show that the plectonemic helix is energetically favored over the single helix for realistic values of σ and excluded volume radius R_{ex}.

For the single helix, R is not equal to the excluded volume radius R_{ex}. We will momentarily disregard excluded volume effects while finding the minimum energy conformation for the single helix. With ρWr given in equation (2.5), ρE_{bend} from equation (2.7) (with R in place of R_{ex}), and ρE_{twist} again given by equation (2.1), the total energy per unit length, ρE_{tot}, for the single helix is

$$\rho E_{tot} = \frac{\beta_{bend} \cos^4\alpha}{2 R^2} + \frac{\beta_{twist}}{2} (2\pi)^2 \left[\frac{\sigma}{h_0} - \frac{(1 - \sin\alpha) \cos\alpha}{2 \pi R} \right]^2 \tag{2.18}$$

For fixed R, setting $d\rho E_{tot} / d\alpha$ to zero gives

$$2 \pi R \frac{\sigma}{h_0} = \frac{2 c \cos^3\alpha \sin\alpha}{\cos^2\alpha + \sin\alpha - \sin^2\alpha} + \cos\alpha - \sin\alpha \cos\alpha \tag{2.19}$$

Fixing α and setting $d\rho E_{tot} / dR$ to zero gives

$$2 \pi R \frac{\sigma}{h_0} = \frac{c \cos^3\alpha}{1 - \sin\alpha} + \cos\alpha - \sin\alpha \cos\alpha \tag{2.20}$$

Equations (2.19) and (2.20) give equal values only at the non-physical point $\alpha = 90^o$ and $R = 0$. Therefore no energy minima for both R and α exist inside the interval $0^o < \alpha < 90^o$, and any such minima must occur at either $\alpha = 0^o$ or 90°. At the limit

$\alpha \rightarrow 0^o$, equation (2.20) becomes

$$R = \frac{(1+c)\,h_0}{2\,\pi\,\sigma} \tag{2.21}$$

Substituting this into equation (2.18) at $\alpha = 0^o$ gives

$$\rho E_{tot} = \frac{2\,\beta_{bend}\,\pi^2\,\sigma^2}{(1+c)^2\,h_0^2} + \frac{\beta_{twist}}{2}(2\pi)^2\left[\frac{\sigma\,c}{h_0\,(1+c)}\right]^2 \tag{2.22}$$

At $\alpha = 90^o$, all of the linking number difference is in twist, so that

$$\rho E_{tot} = \frac{\beta_{twist}}{2}(2\pi)^2\left[\frac{\sigma}{h_0}\right]^2 \tag{2.23}$$

Equation (2.23) is greater than equation (2.22) for any $c > 0$, so the global minimum in energy is given by equation (2.22) with $\alpha = 0^o$ and R given by equation (2.21).

For $c = 1$, the total energy of the single helix without excluded volume, equation (2.22), can be easily compared with the total energy of the plectonemic helix, equation (2.15). The plectonemic helix has lower total energy until σ / h_0 exceeds $2^{3/2}\rho W r_{max}$, generally quite a large value of $|\sigma|$. When excluded volume effects are included, as outlined below, the transition to the toroidal form takes place at an even greater value of $|\sigma|$.

The finite excluded volume radius of the DNA double helix puts a lower limit on the pitch. The pitch of a helix is defined as the displacement per turn and is $2\,\pi\,R\tan\alpha$. In Figure 2 it is shown that the minimum pitch with an excluded radius R_{ex} is $2\,R_{ex} / \cos\alpha$. Setting the two pitch expressions equal gives

$$\sin\alpha = \frac{R_{ex}}{\pi R} \; ; \quad \cos\alpha = \left[1 - \frac{R_{ex}^2}{\pi^2 R^2}\right]^{1/2} \tag{2.24}$$

Equations (2.23) can be substituted into equation (2.18) to get the energy at the smallest allowable pitch angle when excluded volume is taken into account. The resulting expression can be minimized with respect to R to get the global energy minimum. We did this numerically and plotted the results in Figure 11, together with the corresponding results for the plectonemic case. As Fuller predicted, the toroidal form becomes energetically favorable at sufficiently high $|\sigma|$, but the values of $|\sigma|$ at which this occurs are much greater than what is physiologically accessible.

E. Discussion

1. Alternate topological descriptions

Equation (1.7) relates a property of two curves, linking number (or linking number difference), to another property of two curves, the twist of one about the other, and a property of one of the curves, writhe. White and Bauer pointed out that, in applying equation (1.7) to DNA, there are several possible choices for the curves which do not give the same values for twist and writhe (White and Bauer, 1986). The usual choice is to take one curve to be the axis of the DNA double helix and the other to be either of the sugar-phosphate backbones. In this case, the writhe is the writhe of the helix axis and the twist is the twist of the backbone about the axis (which is not the same as the twist of the axis about the backbone). White and Bauer (1986) suggested that the correct choice for analyzing ring closure experiments (and, by extension, any study of DNA energetics) would be to take the two curves to be the two sugar-phosphate backbones of the DNA. While this choice is perfectly self-consistent topologically, it does not lead to correct values of twist to apply to equation (2). It is not a convenient

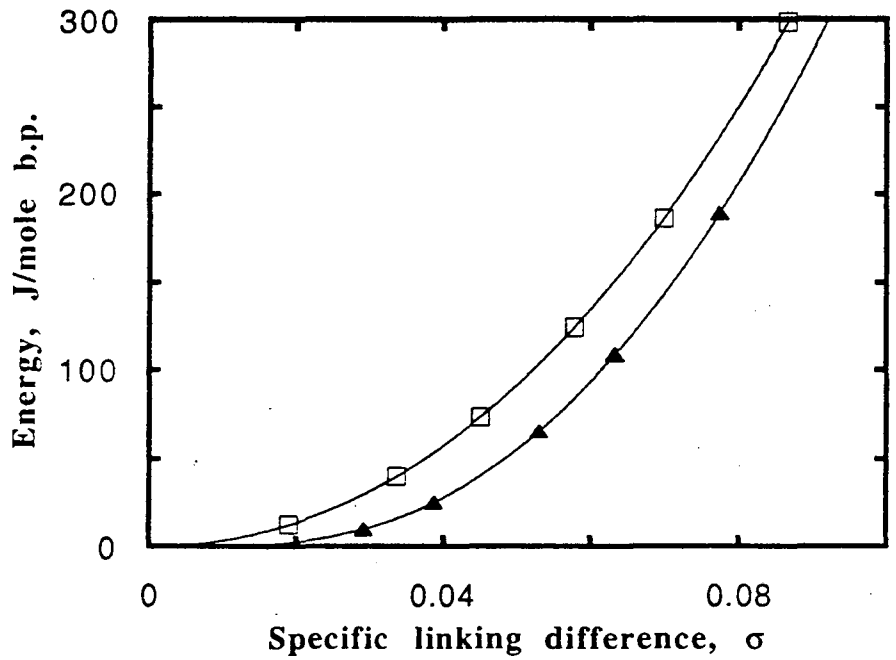


Figure 11. Total energies of plectonemic (triangles) and toroidal (squares) conformations as functions of σ, for minimum energy conformations. R_{ex} = 100Å, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm. For smaller values of R_{ex}, the difference between the two curves is greater. The two curves cross at $\sigma \approx 0.15$, $\rho E_{tot} \approx 1000$ J / mole b.p.

choice for bend energy and writhe, either. For example, the writhe of a straight piece of DNA would be non-zero by this definition, since the backbones are helical. The usual choice of the DNA axis and one of the backbones gives the right twist and curvature for equations (2) and (3), respectively (Landau and Lifshitz, 1959).

Another alternative to equation (1.7) was derived by White, Cozzarelli and Bauer (1988). In what was later called surface linking theory (Richardson *et al.*, 1988) linking number was expressed as the sum of two quantities, winding number and surface linking number. Both of these quantities are defined with respect to a surface which the DNA axis lies on. The winding number was shown to be the experimentally measured quantity in nuclease studies of helical repeat of DNA bound in nucleosomes. However, surface linking theory is not as useful for describing the conformation of plectonemic DNA. The surface linking number for plectonemic DNA is zero, so that the winding number equals the linking number (White *et al.*, 1988). Since this is true for any superhelical pitch angle, surface linking theory does not distinguish between conformations with differing values of twist and writhe. Twist and writhe are the relevant quantities for treating elastic energies of DNA, since twist energy is a function of twist and bend energy can be related to writhe.

2. Conformation of supercoiled DNA

The model of the plectonemic helix presented here makes several quantitative predictions about the conformation of superhelical DNA. At low specific linking difference $|\sigma|$, the linking difference is almost entirely in writhe. As $|\sigma|$ increases, a greater proportion of linking difference is in twist. There is an absolute limit on the writhe per unit length, ρWr_{max}, which is $1/(4\pi R_{ex})$; therefore the limit on the amount of specific linking difference $|\sigma|$ which can be taken up by writhe is $(h_0 \rho Wr_{max})$. This limit is approached as $|\sigma|$ approaches infinity; (for biologically relevant values of $|\sigma|$,

ρWr doesn't exceed 0.7 ρWr_{max}). A comparison of the plectonemic and toroidal structures shows that the plectonemic structure is energetically favored.

3. Range of superhelical pitch angle differs from an earlier model

Several previous models of DNA supercoiling have included some of the results in this chapter, at least implicitly (Wadati and Tsuru, 1986; Tsuru and Wadati, 1986; Tanaka and Takahashi, 1985; LeBret, 1979, 1984, 1988; Hao and Olson, 1989a,b; Olson and Cicariello, 1986; Benham, 1977, 1979, 1983, 1987). In particular, the paper by Camerini-Otero and Felsenfeld (1978) is very similar to this work in that it ignores end effects to derive properties of supercoiling in the limit of long pieces of DNA. Camerini-Otero and Felsenfeld came to different conclusions from this study, most notably in the statement that the superhelical pitch angle lies in the range $45^o < \alpha \leq 59^o$, whereas we conclude that the upper limit on α is 90°. The differing conclusions stem from two related errors in Camerini-Otero and Felsenfeld's work. One is that a constraint on two parameters of their model, their equation (7), is mistaken for an equation for the torsional constant (at that time unmeasured) and is not applied consistently as a constraint. The second error is that the free energy expression derived in their model is equated with an empirical quadratic dependence of free energy on ΔLk. While such a relation does hold over a restricted range of ΔLk (Depew and Wang, 1975; Pulleyblank *et al.*, 1975; Horowitz and Wang, 1984) it is only consistent with the model of Camerini-Otero and Felsenfeld if the torsional constant and/or the superhelical radius are made to vary in a non-physical manner.

A simple physical argument against the upper limit on α given by Camerini-Otero and Felsenfeld is that, in the limit where $\sigma = 0$, a conformation with $\alpha = 90^o$ has zero bend and torsional energy, whereas a conformation at any other pitch angle (e.g. 59°) has finite bend and twist energies. This energy is plotted in the $\sigma = 0$ curve of Figure 5.

4. Comparison with experimental results

a. Electron microscopy

The most basic prediction of this chapter, that supercoiled DNA in solution is plectonemic, agrees with the branched plectonemic structures observed by electron microscopy (Boles *et al.*, 1990; Adrian *et al..*, 1990). The most extensive electron microscopy study of DNA supercoiling is that of Boles *et al..* (1990). Comparison of the model of this chapter with their data on a more quantitative level is complicated by the fact that the observed superhelical radius is a function of σ, whereas we held the superhelix radius constant.

Two possible extensions to the model presented here might explain the variation of superhelical radius with σ. The first would be to replace the hard-body model of the DNA excluded volume with a screened electrostatic repulsion. This is carried out in the next chapter. Then the two strands of the plectonemic helix approach each other until the outward force per length balances the inward force per length caused by supercoiling. This results in a superhelix radius which decreases with increasing $|\sigma|$. The resulting changes in superhelix radius are on the order of the Debye length (Atkins, 1990).

A second modification to the model presented here would be to include entropic effects. Since there are more possible conformations at greater superhelical radius, the entropic contribution to free energy would tend to give larger values for superhelical radius than what results from the enthalpic contribution alone. As $|\sigma|$ increases, the dependence of the enthalpy on superhelical radius becomes sharper, resulting again in a superhelical radius which decreases with increasing $|\sigma|$. This entropic effect on superhelical radius would be greater for small $|\sigma|$. The justification for ignoring entropic effects at large values of $|\sigma|$ is presented in the paper by Hearst and Hunt (1991).

It is possible that the observed dependence of superhelix radius on σ may be partly an experimental artifact. If the interaction of the DNA with the polylysine matrix used in Boles *et al.* (1990) results in forces which disturb the solution conformation of the DNA, the resulting distortions would be greater for smaller values of $|\sigma|$, since the bend and twist energies vary more slowly with superhelical radius at small $|\sigma|$.

This work, together with data on R_{ex} (Yarmola *et al.*, 1985; Brian *et al.*, 1981; Stigter, 1987) suggests a strong dependence of superhelical parameters on salt concentration. Thus it is surprising that Boles *et al.* do not report a difference in results for two buffers of substantially different cationic concentrations. Adrian *et al.* (1990), using cryo-electron microscopy, report a strong dependence of superhelical diameter on salt concentration. Their results for superhelical diameter are somewhat smaller than values of the excluded volume obtained by osmotic pressure measurements (Yarmola *et al.*, 1985; Brian *et al.*, 1981; Stigter, 1987).

b. Energy of supercoiling

The energy of supercoiling inferred from ethidium binding studies and the distribution of linking numbers resulting from ligation of linear pieces of DNA is given in equation (1.10). In terms of energy per length it is

$$\rho E_{tot} = 10\ RT\ \sigma^2 / 3.4\ \text{Å} \tag{2.25}$$

where 3.4 Å is the length of DNA per base-pair. This quantity gives larger energies than our model except at very high values of $|\sigma|$, as shown in Figure 12. The lack of agreement is not too surprising, since equation (2.25) is based on measurements made at low $|\sigma|$, where entropic effects and end effects are large (Depew and Wang, 1975; Pulleyblank *et al.*, 1975; Horowitz and Wang, 1984).

c. Salt concentration and conformational transitions

The model parameters β_{bend}, β_{twist} and R_{ex} are all functions of ionic strength. While β_{bend} and β_{twist} may vary by as much as a factor of two in the range $1mM \leq [Na^+] \leq 1M$ (Hagerman, 1988) the excluded volume radius R_{ex} varies by a factor of five over the same range. The ionic strength dependence of R_{ex} has effects which are generally greater than the effects of the ionic strength dependences of β_{bend} and β_{twist}. At low ionic strength, R_{ex} is about 100Å (Yarmola *et al.*, 1985; Brian *et al.*, 1981; Stigter, 1987) which gives $(h_0 \rho Wr_{max}) = 0.028$. Therefore linking difference starts going into twist at low values of $|\sigma|$, resulting in large values of twist, twist energy, total energy and superhelical stress. Conversely, at high salt concentrations, R_{ex} is smaller, ρWr_{max} is larger and more linking difference goes into writhe, resulting in lower twist, twist and total energies, and lower superhelical stress.

Experimental data on the salt-concentration dependence of conformational transitions agrees qualitatively with our work. B-Z transitions (Singleton *et al.*, 1982; Nordheim and Rich, 1983) and cruciform formation (Singleton, 1983; Kozyavkin *et al.*, 1990) are both induced by lower levels of specific linking difference at low salt concentration than at higher salt concentration.

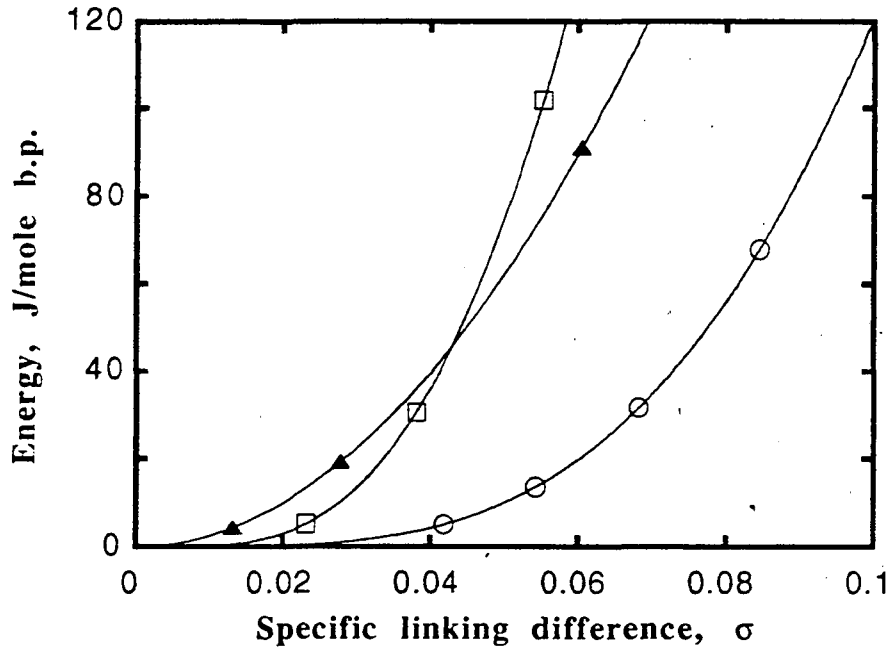


Figure 12. Total energy of plectonemic conformations compared with the parabolic energy of supercoiling, Eq. (2.24). Triangles - Eq. (2.24); squares - plectonemic helix with R_{ex} = 100Å, $\beta_{bend} = 2 \times 10^{-19}$ erg-cm, $\beta_{twist} = 3 \times 10^{-19}$ erg-cm; circles - plectonemic helix with R_{ex} = 30Å, $\beta_{bend} = \beta_{twist} = 2 \times 10^{-19}$ erg-cm.

5. Range of validity of the model

The models presented in this chapter of plectonemic and toroidal supercoiling are valid in the limit of long DNA molecules. The axis of a finite toroidal superhelix is circular, rather than straight, which results in greater bend energy. A finite plectonemic helix has loops at the end which contribute to the bend energy while contributing little to the writhe. Since the infinite length models ignore these effects, they are not valid for pieces of DNA shorter than a few persistence lengths.

This work does not entirely rule out toroidal structures for all lengths of DNA. It is possible that the finite length effects mentioned above result in a region of DNA length and σ which would favor the toroidal structure.

As noted above, these models do not take entropic effects into account (Hearst and Hunt, 1991). In general, entropic effects are of greatest importance when the structure is loose, i.e. when ρE_{tot} varies relatively slowly with R and α. This is the case for small values of $|\sigma|$. A more exact statement on the range of σ for which entropic effects are significant would require calculating a partition function for all conformations near the enthalpy-minimum ones obtained above, which is beyond the scope of this chapter. Empirically, one can say that the model is valid for $|\sigma|$ values high enough that interwound structures are observed experimentally.

F. Future work

Several aspects of the present model and its differences with experimental data suggest that a more detailed treatment of electrostatic interaction between strands, as given in the next chapter, is needed (Hunt, Klein and Hearst, in preparation). The energy of supercoiling is then raised by including this energy, bringing the total energy into better agreement with Eq. (2.24). An electrostatic term in the total energy also

results in the superhelical radius being a function of $|\sigma|$, as observed in electron microscopy data. Finally, the sensitivity of the results of the model to the value of R_{ex} suggests that electrostatic repulsion between and within DNA strands could have significant effects.

References

Adrian, M., ten Heggeler-Bordier, B., Wahli,W., Stasiak, A.Z., Stasiak, A. and Dubochet, J. (1990) *EMBO J.* **9**, 4551

Atkins, P.W. (1990) *Physical Chemistry* (Freeman, New York)

Benham, C.J. (1977) *Proc. Nat. Acad. Sci. USA* **74**, 2397

Benham, C.J. (1979) *Biopolymers* **18**, 609

Benham, C.J. (1983) *Biopolymers* **22**, 2477

Benham, C.J. (1987a) *Cell Biophysics* **10**, 193

Benham, C.J. (1987b) *Biopolymers* **26**, 9

Boles, T.C., White, J.H., and Cozzarelli, N.R. (1990) *J. Mol. Biol.* **213**, 931

Brady, G.W., Satkowski, M., Foos, D., and Benham, C.J. (1987) *J. Mol. Biol.* **195**, 185

Brady, G.W., Fein, D.B., Lambertson, V., Grassian, V., Foos, D., and Benham, C.J. (1983) *Proc. Nat. Acad. Sci. U.S.A.* **80**, 741

Brian, A.A., Frisch, H.L. and Lerman, L.S. (1981) *Biopolymers* **20**, 1305

Camerini-Otero, R.D. and Felsenfeld, G. (1978) *Proc. Nat. Acad. Sci. USA* **75**, 1708

Depew, R.E. and Wang, J.C. (1975) *Proc. Nat. Acad. Sci. USA* **72**, 4275

Fuller, F.B. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 815

Gellert, M., Mizuuchi,K., O'Dea, M.H., Ohmori, H.and Tomizawa,J. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 35

Hagerman, P.J. (1988) Ann. Rev. Biophys. Biophys. Chem. **17**, 265

Hao, M.-H. and Olson, W.K. (1989a) *Biopolymers* **28**, 873

Hao, M.-H. and Olson, W.K. (1989b) *J. Biomol. Str. and Dynamics* **7**, 661

Hearst, J.E. and Hunt, N.G. (1991) J. Chem. Phys. in press

Horowitz, D.S. and Wang, J.C. (1984) *J. Mol. Biol.* **173**, 75

Hsieh, T.-S. and Wang, J.C. (1975) *Biochemistry* **14**, 527

Hunt, N.G. and Hearst, J.E. (1991) *J. Chem. Phys.* **95**, 9329

Hunt, N.G., Klein, M.P., and Hearst, J. E. Manuscript in preparation

Kozyavkin, S.A., Slesarev, A.I., Malkhosyan, S.R. and Panyutin, I.G. (1990) ***Eur. J. Biochem.*** **191**, 105

Landau, L.D. and Lifshitz, E.M. (1959) *Theory of Elasticity* (Pergamon, London)

LeBret, M. (1979) *Biopolymers* **18**, 1709

LeBret, M. (1984) *Biopolymers* **23**, 1835

LeBret, M. (1988) *J. Mol. Biol.* **200**, 267

Lee, F.S. and Bauer,W.R. (1985) *Nucl. Acids Res.* **13**, 1665

Lilley, D.M.J. (1980) *Proc. Nat. Acad. Sci. USA* **77**, 6468

Lyubchenko, Yu. L. and Shlyakhtenko, L.S. (1988) *Nucl. Acids Res.* **16**, 3269

Nordheim, A. and Rich, A. (1983) *Proc. Nat. Acad. Sci. USA* **80**, 1821

Olson, W. K. and Cicariello, J. (1986) *Ann. N.Y. Acad. Sci.* **482**, 69

Peck ,L.J. and Wang, J.C. (1983) *Proc. Nat. Acad. Sci. U.S.A* . **80**, 6206

Panayotatos, N. and Wells, R.D. (1981) *Nature* **289**, 466

Pulleyblank, D. E., Shure, M., Tang, D., Vinograd, J. and Vosberg, H.-P. (1975) *Proc. Nat. Acad. Sci. U.S.A.* **72**, 4280

Richardson, S.M.H., Boles, T.C., and Cozzarelli, N.R. (1988) Nucl. Acids Res. **16**, 6607

Singleton, C.K., Klysik, J., Stirdivant, S.M. and Wells, R.D. (1982) *Nature* **299**, 312

Singleton, C.K. (1983) *J. Biol. Chem.* **258**, 7661

Spengler, S.J., Stasiak, A., and Cozzarelli, N.R. (1985) *Cell* **42**, 325

Stigter, D. (1987) *Cell Biophysics* **11**, 139

Tanaka, F. and Takahashi, H. (1985) *J. Chem. Phys.* **83**, 6017

Torbet, J. and DiCapua, E. (1989) *EMBO J.* **8**, 4351

Tsuru, H. and Wadati, M. (1986) *Biopolymers* **25**, 2083

Vologodskii, A.V., Anshelevich, V.V., Lakashin,A.V., and Frank-Kamenetskii, M.D. (1979) *Nature* **280**, 294

Wadati, M. and Tsuru, H. (1986) *Physica D* **21**, 213

White, J.H. (1969) *Am. J. Math.* **91**, 693

White, J.H. and Bauer,W.R. (1986) *J. Mol. Biol.* **189**, 329

White, J.H., Cozzarelli, N.R., and Bauer, W.R. (1988) *Science* **241**, 323

Yarmola, E.G., Zarudnaya,M.I., and Lazurkin, Yu.S. (1985) *J. Biomol. Str. and Dynamics* **2**, 981

III. Elastic and Electrostatic Model of Supercoiling

A. Introduction

The approach taken in this chapter is an extension of the work of the previous chapter and a recent publication (Hunt and Hearst, 1991 - referred to as paper I). DNA is modelled as an elastic cylinder with torsion and bending constants derived from experimental data. In chapter II, the electrostatic interaction between DNA strands was only taken into account indirectly, by treating DNA as a cylinder with a hard-body excluded volume. Here we model DNA as a uniformly charged cylinder and explicitly include the electrostatic interaction. This improvement allows us to predict the excluded volume radius of the DNA from first principles and the properties of the model.

The treatment outlined here does not take entropy into account, as do Monte Carlo studies in which free energies are calculated. However, the computational requirements are much less. This allows us to explicitly include the electrostatic interaction, which would significantly increase the computational requirements in a Monte Carlo study, and facilitates comparison with experimental data on conformational transitions. Unlike Monte Carlo models, the present treatment presents no difficulties in the limits of long DNA length and large linking difference.

B. Previous work

The treatment of DNA energy and conformation is identical to that used in chapter II with the addition of the electrostatic interaction. Here we briefly outline the approach of chapter II and list the major conclusions.

DNA was modelled as an elastic cylinder characterized by a bending constant β_{bend}, a torsional constant β_{twist}, and an excluded volume radius R_{ex}. End effects were ignored and two types of conformation were considered - a plectonemic or

interwound form, and a toroidal form (which, in the limit of no end effects or infinite length, is a single helix). All extensive quantities (e.g. bend energy, writhe) were treated as linear densities (bend energy per unit length, writhe per unit length).

After finding the conformations which minimize the sum of bend and twist energies for a given superhelix density and excluded volume radius, the following main conclusions were reached:

(1) The plectonemic conformation is energetically favored over the toroidal conformation for biologically relevant levels of supercoiling.

(2) The writhe per unit length, ρWr, has a maximum value for the plectonemic conformation, ρWr_{max}, which is inversely proportional to R_{ex}: $\rho Wr_{max} = 1 / (4\pi R_{ex})$. For low values of specific linking difference $|\sigma|$ ($|\sigma| = |\Delta Lk/Lk_0|$), supercoiling is accommodated in writhe at little energetic cost. As $|\sigma|$ increases and ρWr becomes a significant fraction of ρWr_{max}, an increasing proportion of supercoiling goes into twist, at increasing energetic cost.

(3) The excluded volume radius R_{ex}, obtained from theory (Stigter, 1987), osmotic pressure measurements (Yarmola *et al.*, 1985) and sedimentation data (Brian *et al.*, 1981) is a strong function of ionic strength, varying from as little as 15Å at ionic strength, I, of 1 M to over 100Å at I of 1 mM. Therefore ρWr_{max} also varies by more than a factor of six over this range of ionic strength. Thus at low ionic strength, where R_{ex} is large and ρWr_{max} small, little linking difference can be put into writhe. Linking difference goes into twist at low values of $|\sigma|$, resulting in greater torsional stress on supercoiled DNA at low ionic strength. This prediction of the model is in qualitative agreement with experiments measuring the ionic strength dependence of secondary structure transitions in supercoiled DNA (Kozyavkin *et al.*, 1990; Singleton, 1983; Singleton *et al.*, 1982).

The importance of the excluded volume parameter in chapter II prompted us to treat the electrostatic repulsion between double helical DNA strands in greater detail. The use of a hard-body model assumes that a close approach of DNA strands entails no energetic cost, except indirectly by limiting the fraction of $|\sigma|$ which can be accommodated in writhe. It is clear, however, that electrostatic repulsion between duplex strands does lead to some energetic cost as the molecule becomes compacted. In our current model, there is no longer an infinitely stiff excluded volume radius R_{ex} which defines the superhelix radius, but the above conclusions still hold qualitatively when the electrostatic interaction is included and are a useful basis for thinking about the energy and conformation of supercoiled DNA. Entropic effects change this picture somewhat at low values of $|\sigma|$, as discussed in section III.E.

C. Methods of Calculation

1. Electrostatic energy

We followed the procedure developed by Stigter (1975, 1987), with modifications for the geometry of the plectonemic and toroidal conformations. DNA is modelled as a uniformly charged cylinder of radius 12Å in an ionic atmosphere. The potential as a function of distance from the DNA is calculated using the Poisson-Boltzmann (P-B) equation in cylindrical coordinates, which is solved numerically for given concentrations of ions in solution. In the limit of small electrical potential, the P-B equation reduces to the simpler Debye-Hückel (D-H) equation. Since the D-H equation is more tractable than the P-B equation (in particular, the D-H equation is linear), the D-H solution which matches the numerical P-B solution in the region of low potential is found. This D-H solution can be thought of as the D-H potential generated by a line charge with some charge per unit length ν located at the axis of the cylindrical coordinates. ν may be less than or greater than the charge per length of the charged cylinder, depending on the

solution concentrations of ions. Once ν is found, the electrostatic interaction per unit length between the two strands is approximated by

$$\rho E_{electro} = \frac{\nu^2}{\varepsilon} \int_{-\infty}^{\infty} \frac{\exp[-r(s)/r_D]}{r(s)} \, ds \tag{3.1}$$

where the integration is carried out over one of the strands of the plectonemic helix, **r** (s) is the distance from a fixed point on one strand to the point corresponding to s on the other strand, r_D is the Debye length for the given ionic strength, and ε is the dielectric constant of water.

The point of these mathematical gymnastics is that the two strands of the plectonemic helix are a few Debye lengths apart, i.e. they are well into the region where the P-B equation reduces to the D-H equation. This prediction of the model is supported by electron microscopy measurements of supercoiled DNA, so we do not have to worry that our reasoning is circular. Since the D-H equation is linear, summing the potential contribution from each differential element of the opposing strand (integrating as in equation (3.1)) should be a good approximation to the full non-linear P-B treatment of the interaction between the two strands. For each conformation considered, the electrostatic interaction can be calculated from the well-behaved one-dimensional integral, equation (3.1), rather than by solving the non-linear P-B equation in three dimensions.

The electrostatic energy calculation is somewhat different for the toroidal or helical conformation. In the toroidal conformation, it was shown in chapter II that the most favorable conformation is a superhelix of low pitch angle and radius given approximately by

$$R = \frac{(1 + c)\, h_0}{2 \pi \sigma} \tag{3.2}$$

where c is $\beta_{bend} / \beta_{twist}$ and h_0 is the helical repeat of DNA in units of length (35.4 Å). The exact value of the optimum radius depends on excluded volume effects (see chapter II). Whereas in the plectonemic case, the electrostatic interaction is between two double helical DNA strands, in the toroidal case the electrostatic interaction is primarily between a given point and the points one full turn in either direction. One could calculate the electrostatic interaction of a point with points in the immediate vicinity and subtract the same quantity for a straight segment, but we assume that this local contribution is included in the bend constant. In no case did points two or more turns away from each other contribute significantly to the electrostatic energy.

2. Minimum energy conformations

For each value of σ and each ionic strength considered, the superhelix pitch angle α and radius R are varied to find minimum energy conformations. The total energy per unit length, ρE_{tot}, is the sum

$$\rho E_{tot} = \rho E_{electro} + \rho E_{bend} + \rho E_{twist} \tag{3.3}$$

where $\rho E_{electro}$, the electrostatic energy per unit length, is given by equation (3.1). ρE_{bend} and ρE_{twist}, the bend and twist energies per unit length, were derived in chapter II and are

$$\rho E_{bend} = \frac{\beta_{bend} \cos^4\alpha}{2 R^2} \tag{3.4}$$

plectonemic:

$$\rho E_{twist} = \frac{\beta_{twist}}{2} (2\pi)^2 \left[\frac{\sigma}{h_0} - \frac{\sin\alpha \cos\alpha}{2\pi R}\right]^2 \tag{3.5}$$

toroidal: $$\rho E_{twist} = \frac{\beta_{twist}}{2} (2\pi)^2 \left[\frac{\sigma}{h_0} - \frac{(1 - \sin\alpha)\cos\alpha}{2\pi R} \right]^2 \qquad (3.6)$$

In this paper we use the values $\beta_{bend} = 2.8 \times 10^{-19}$ erg-cm and $\beta_{twist} = 3.1 \times 10^{-19}$ erg-cm. These are the values obtained by Shimada and Yamakawa (1988) from their analysis of the topoisomer distribution data of Shore and Baldwin (1983) and of Horowitz and Wang (1984).

3. Calculation of fraction of topoisomers with cruciform

For each value of ΔLk, the energy of supercoiling for the pA03 plasmid is calculated for forms with and without the cruciform. The cruciform increases (towards zero) the effective linking difference by its length divided by the helical repeat. End effects are ignored and the energy is taken as the length of the plasmid multiplied by the energy per unit length for the given specific linking difference and ionic concentrations. The calculated fraction of plasmids with cruciforms, p^+, is then

$$p^+ = \frac{\exp\left[-(H^+_{super} + \Delta H_{cruc}) / RT \right]}{\exp\left[-H^0_{super} / RT \right] + \exp\left[-(H^+_{super} + \Delta H_{cruc}) / RT \right]} \qquad (3.7)$$

where H^+_{super} and H^0_{super} are the calculated supercoiling energy of the plasmid with and without the cruciform, respectively. ΔH_{cruc} is the energy cost of forming a cruciform and is obtained from the best fit to the experimental data. This procedure is carried out both using supercoiling energies derived from our model and using the usual quadratic expression (Depew and Wang, 1975; Pulleyblank *et al.*, 1975; Horowitz and Wang, 1984)

$$H_{super} = 10\,RT\,N\,\sigma^2 \tag{3.8}$$

where N is the number of base pairs.

D. Results

1. Plectonemic helix

The bend, twist, electrostatic and total energies per mole b.p. for the plectonemic conformation are shown in Figure 13 for I = 150 mM uniunivalent salt. The bend energy is dominant in the range of physiological interest, although twist energy dominates for $\sigma \geq 0.13$. The contribution of the electrostatic interaction is mostly indirect - the electrostatic repulsion between strands holds them apart and limits the amount of linking difference which goes into writhe.

The effect of varying salt concentration is shown in Figures 14 - 16. The energies and superhelical radii are both strong functions of salt concentration. At low salt concentrations the strands are held farther apart by the electrostatic interaction between strands, which is not shielded as much as it is at higher salt concentrations. As $|\sigma|$ increases, R and α decrease, increasing the writhe per length.

All linking difference is in writhe in the limit $|\sigma| \to 0$, as shown in Figure 16. The proportion of linking difference in writhe, $h_0\,\rho Wr\,/\,|\sigma|$, decreases with increasing $|\sigma|$ as the energy costs of bending the DNA and pushing against the repulsion between strands increase. At lower salt concentrations, the proportion in writhe drops more rapidly,

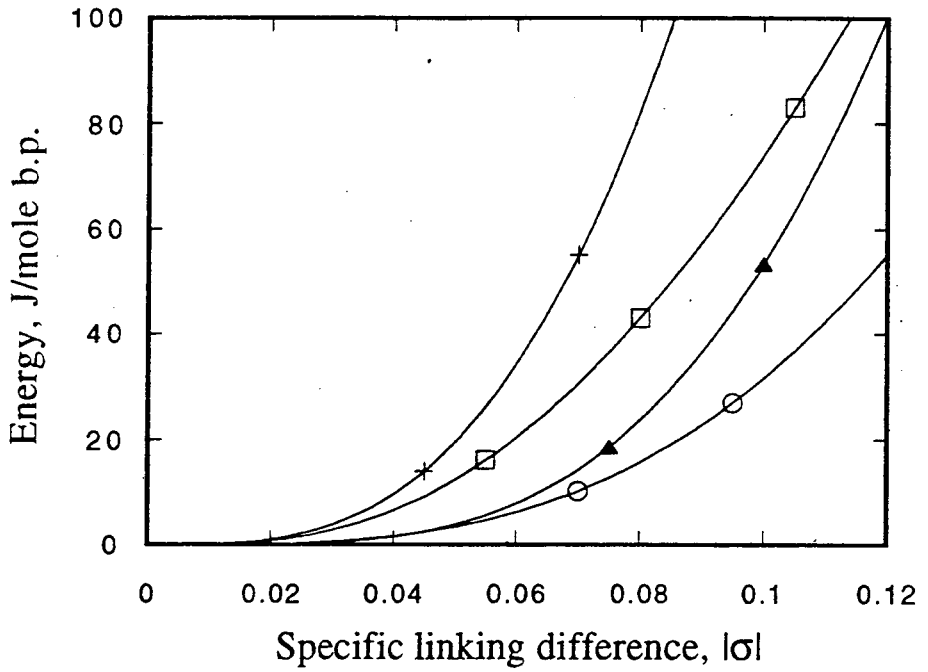


Figure 13. Energies per mole b.p. for optimum plectnemic helix. I = 150 mM.
+'s - ρE_{tot}; squares - ρE_{bend}; triangles - ρE_{twist}; circles - $\rho E_{electro}$

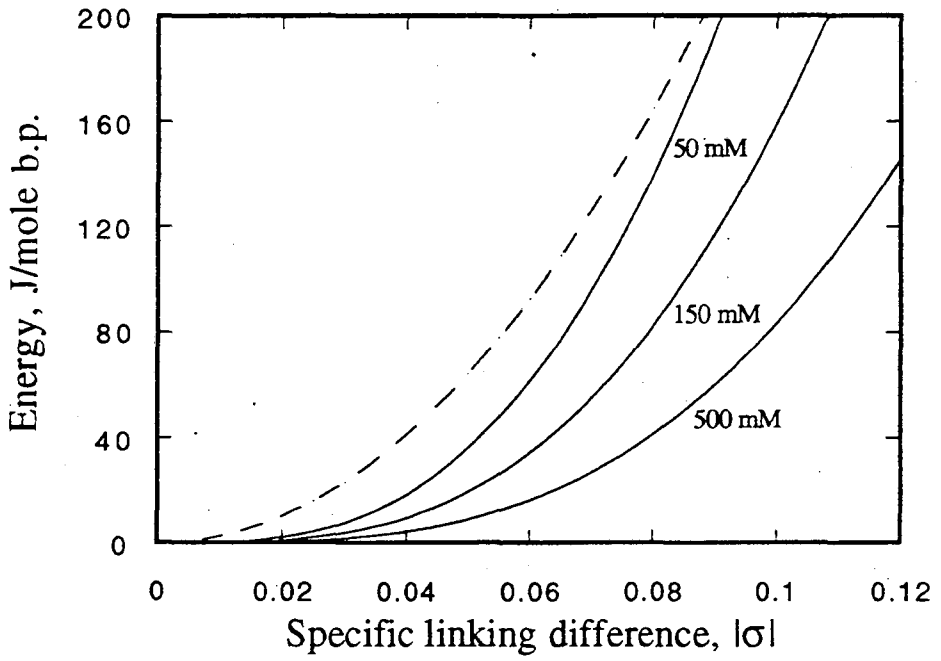


Figure 14. Total energies at different ionic strengths and the quadratic expression, equation 3.8. Dashed curve - equation (3.8); solid curves, from top to bottom: I = 50 mM, I = 150 mM, I = 500 mM.

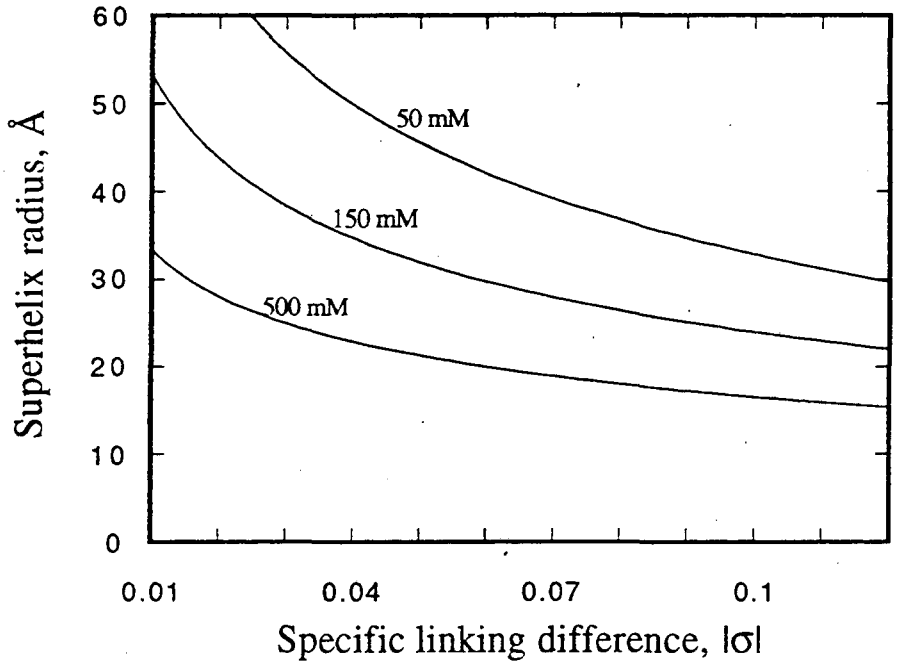


Figure 15. Superhelix radius at different salt concentrations. Curves, from highest to lowest: I = 50 mM, I = 150 mM, I = 500 mM.

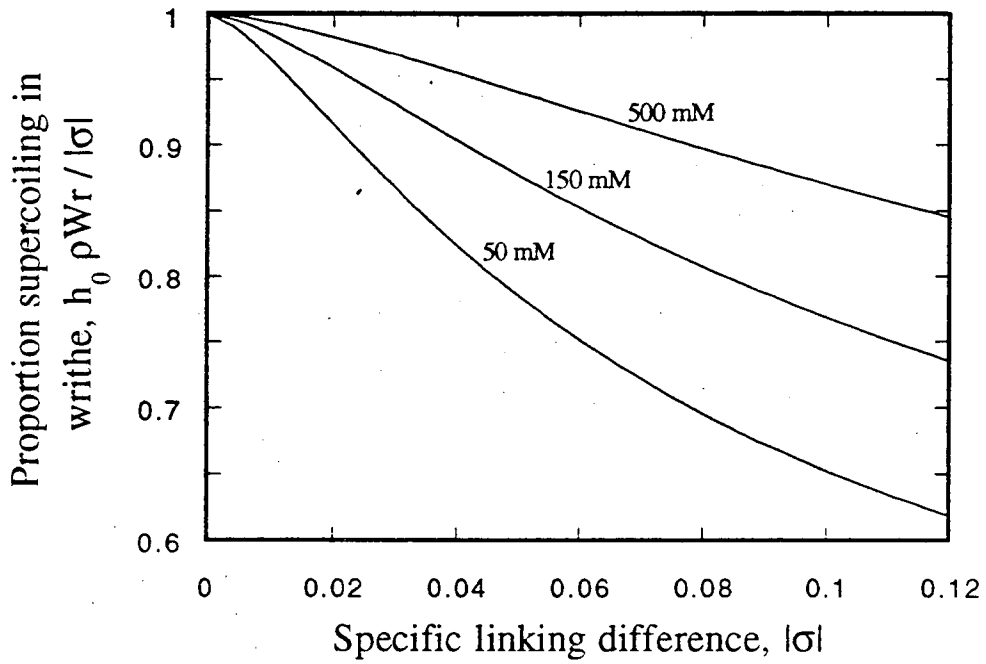


Figure 16. Proportion of linking difference in writhe, $h_0 \rho Wr / |\sigma|$. Curves, from highest to lowest: I = 500 mM, I = 150 mM, I = 50 mM.

since the strands are held farther apart. The conformational entropy of the DNA evidently alters this behavior at low $|\sigma|$ (see section III.E.).

2. Comparison between plectonemic and toroidal conformations

The total energies for the plectonemic and toroidal conformations are shown in Figure 17 for I = 50 mM. In all cases considered, the plectonemic conformation is energetically favored over the toroidal conformation for biologically relevant values of $|\sigma|$. In contrast to the plectonemic case, the energy of the toroidal conformation is not a strong function of salt concentration. The present study therefore does not completely rule out the possibility of a toroidal conformation at very low ionic strength and high levels of supercoiling. The treatment of the electrostatic interaction presented here does not work well at ionic strengths below about 10mM, so we have not thoroughly investigated the low ionic strength region. At sufficiently high $|\sigma|$, the toroidal conformation becomes energetically favored, since it can attain higher values of ρWr than the plectonemic conformation. However, large values of $|\sigma|$ and low ionic strength are conditions which put maximal torsional stress on DNA, inducing conformational transitions such as B-Z transistions and cruciform formation. We conclude that toroidal conformations are not likely to occur.

Perhaps the best evidence for the toroidal conformation is small-angle x-ray scattering data (Brady *et al.*, 1987, 1983). We note that most of the toroidal shapes proposed to fit this data seem unlikely when the elastic energy of supercoiling is considered. As shown in paper I, minimum energy toroidal conformations have small pitch angles and radii which decrease with increasing $|\sigma|$. The optimal radius for the toroidal conformation at 50 mM ionic strength (see Figure 17) and $|\sigma| = 0.05$ is 223Å, while the optimal pitch angle is 4.9^{o}. The values proposed to fit the x-ray scattering

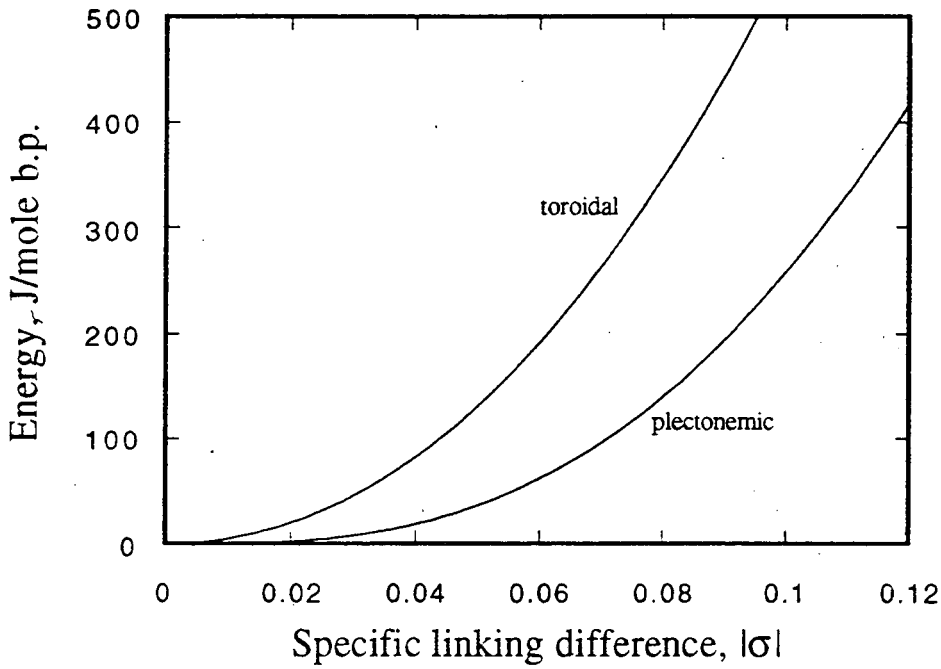


Figure 17. Comparison of energy for plectonemic and toroidal configurations, I = 50 mM. Upper curve - toroidal configuration; lower curve - plectonemic configuration.

data vary with ionic strength, from 15° to 55° for the pitch angle and from 67Å to 146Å for the radius.

Another way of making a similar point is to compare the twist energy of the proposed toroidal conformations with expressions for the total energy of supercoiling. If two-thirds of linking difference goes into twist, the smallest proportion proposed for the toroidal conformations, the twist energy alone is nearly twice the usual quadratic expression, equation (3.8), itself probably an overestimate for $|\sigma| \geq 0.02$.

3. Comparison with results from chapter II and a Monte Carlo study

In order to compare the results of developed in this chapter with models which use a hard-body excluded volume for DNA, a value for the hard-body excluded radius must be chosen to compare with the soft-body electrostatic results for a given salt concentration. The comparison is somewhat arbitrary since, for the electrostatic treatment, the superhelix radius varies as a function of σ, as shown in Figure 15. Nevertheless, a logical correspondence between excluded radius and salt concentration is the effective excluded radius of the DNA for the given salt concentration. The experimental techniques which yield values of the effective excluded radius are outlined in section I.C. Stigter's theoretical treatment (1987) yields similar values.

Figures 18 an 19 compare the results of the Monte Carlo study of Klenin *et al.* (1991) with the electrostatic model of this chapter and the hard-body excluded volume model of chapter II. The Monte Carlo results shown are for the longest plasmid simulated by Klenin *et al.*, 2650 base pairs.

Two principal conclusions can be reached by examining Figures 18 and 19. In comparing the hard-body results of chapter II with the soft-body results of this chapter, Figure 19 shows that the soft-body model results in significantly lower energy than the

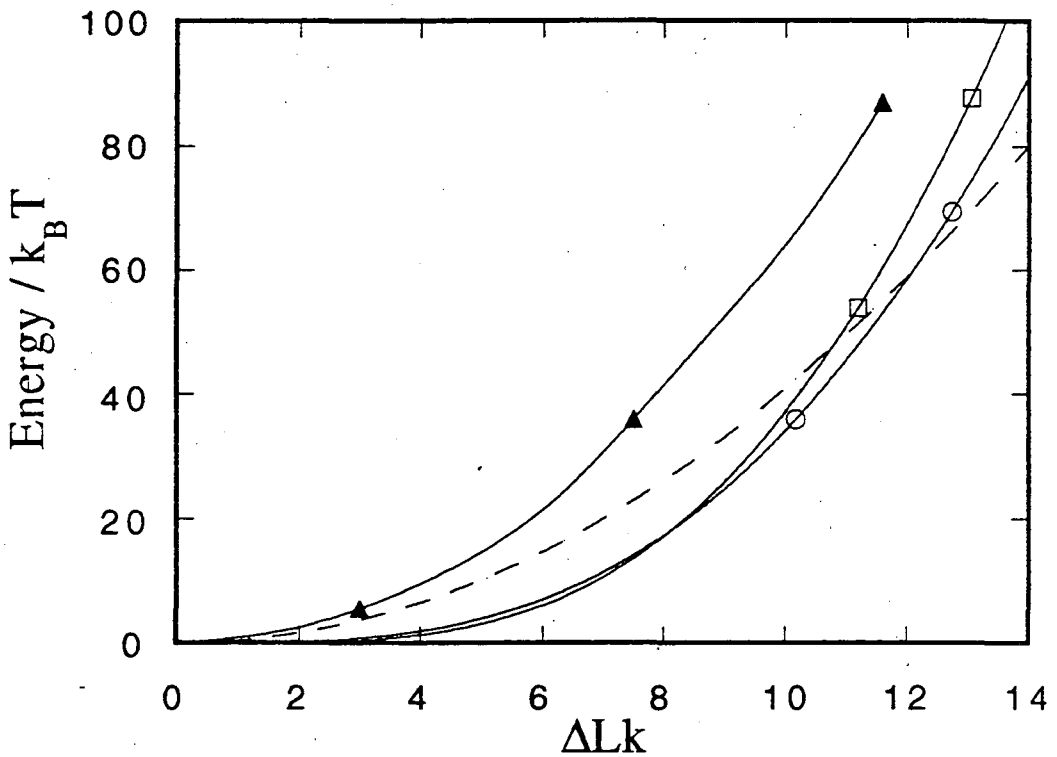


Figure 18. Circles - electrostatic model as developed in this chapter for ionic strength of 20 mM; Squares - hard-body excluded volume as developed in chapter II with an excluded volume radius of 50Å; Triangles - results of Monte Carlo calculation of Klenin et al. for an excluded volume radius of 50Å. All calculations used β_{bend} = 2 e-19 erg-cm, β_{twist} = 3 e-19 erg-cm. Stigter (1987) gives an effective DNA excluded radius of 112Å for 20 mM salt.

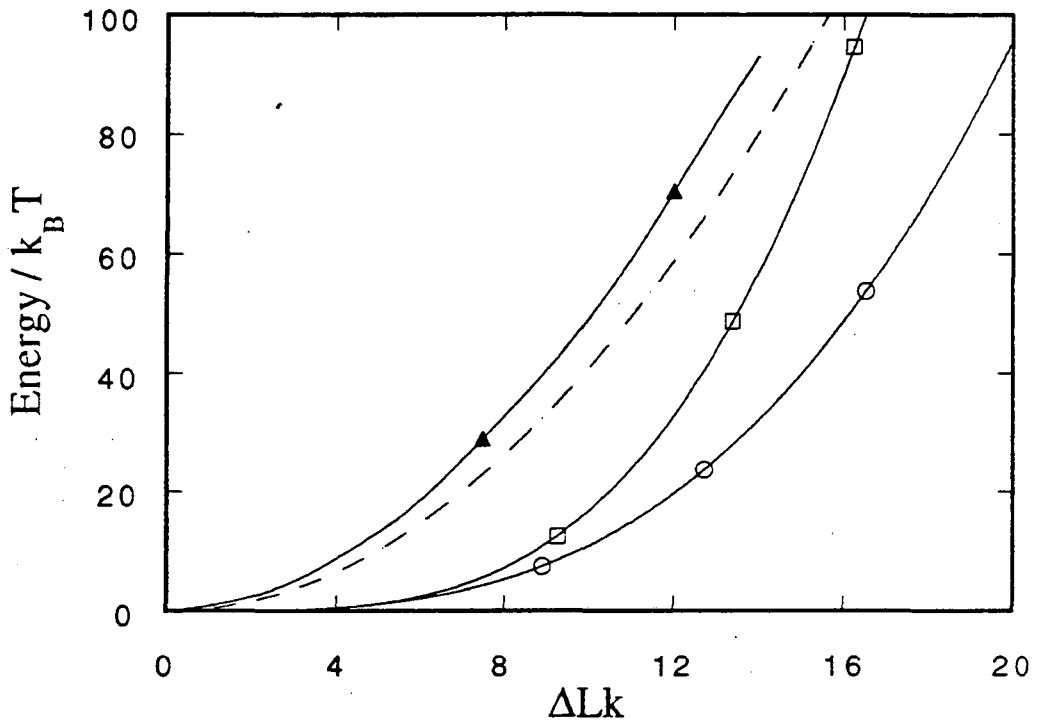


Figure 19. Circles - electrostatic model of this chapter for ionic strength of 100 mM; Squares - hard-body excluded volume model of chapter II with an excluded volume radius of 30Å; Triangles - Monte Carlo results of Klenin et al. for an excluded radius of 30Å. Bend and twist constants as in Figure 18. Stigter (1987) gives an effective DNA radius of 56Å for 100 mM salt.

hard-body model at physiological $|\sigma|$ ($|\sigma| = 0.06$ corresponds to $\Delta Lk = 15$ for a 2650 base pair plasmid) and physiological salt concentration (roughly 150 mM). As shown in Figure 18, the difference between the two models is much less at lower salt concentrations where the electrostatic interaction between DNA strands is not so thoroughly shielded.

The second principal conclusion is that the difference between the Monte Carlo model and our results is significant at low $|s|$, but much less so at physiological levels of supercoiling ($\Delta Lk = 15$). This result is in agreement with the fact that our models do not take entropy or end effects into account. Both loss of conformational entropy and finite length effects should make their greatest contributions to the free energy of supercoiling at low $|\sigma|$, as discussed in section II.E.4.a.

The relevant quantity for supercoiling-induced conformational transitions is the derivative of the energy with respect to ΔLk, (i.e. the supercoiling-induced stress). This quantity is plotted in Figures 20 and 21. These figures make more apparent the relatively close agreement of the chapter II results and the Monte Carlo results.

Work extending the results of this thesis to include finite-length effects is underway. A rough estimate of the magnitude of the end effects for the relatively short plasmid simulated by Klenin et al. can be made as follows. The end loops of the plectonemic helix (see Figure 2) are approximated as semi-circles. The writhe contributions of the end-loops are ignored. The length of the end loops is subtracted from the total molecular length to yield the length of the plectonemic region. In effect, the specific linking difference of the plasmid is increased roughly as

$$\sigma_{eff} = \left[\frac{h_0 \rho W r}{\sigma}\right] \frac{L_{tot}}{L_{tot} - L_{ends}} \sigma \qquad (3.9)$$

where σ_{eff} is the effective specific linking difference of the plasmid with end loops,

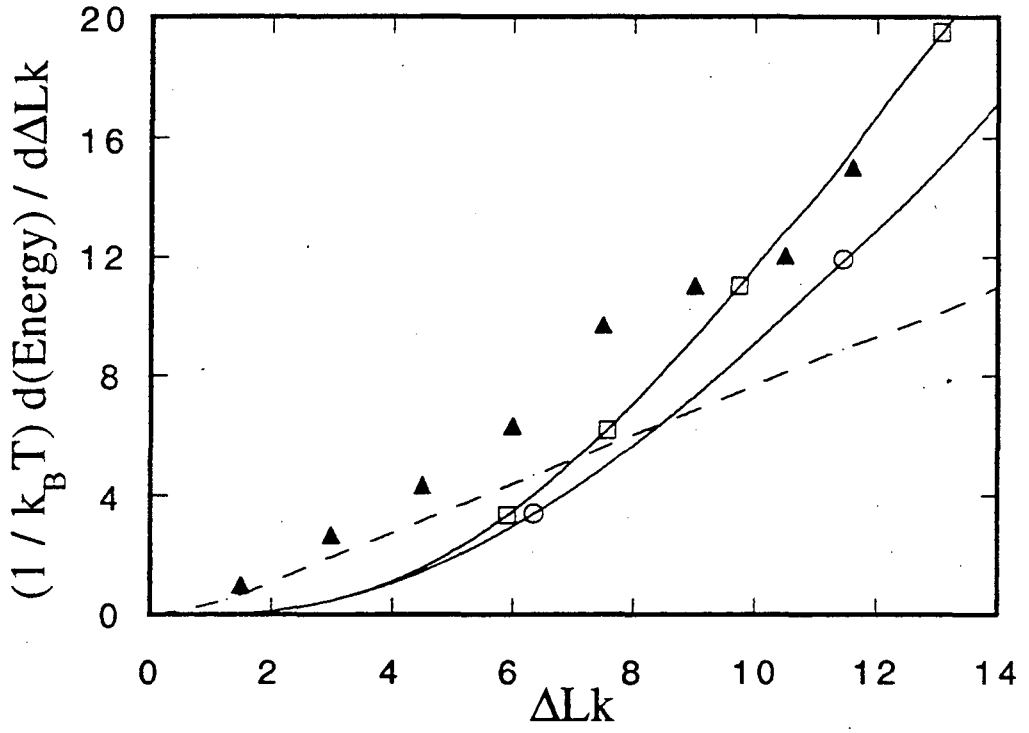


Figure 20. Supercoiling - induced stress, as calculated by the model of this chapter, the Monte Carlo study of Klenin et al., and equation (3.8). Symbols, bend and twist constants, and ionic strength the same as in Figure 18.

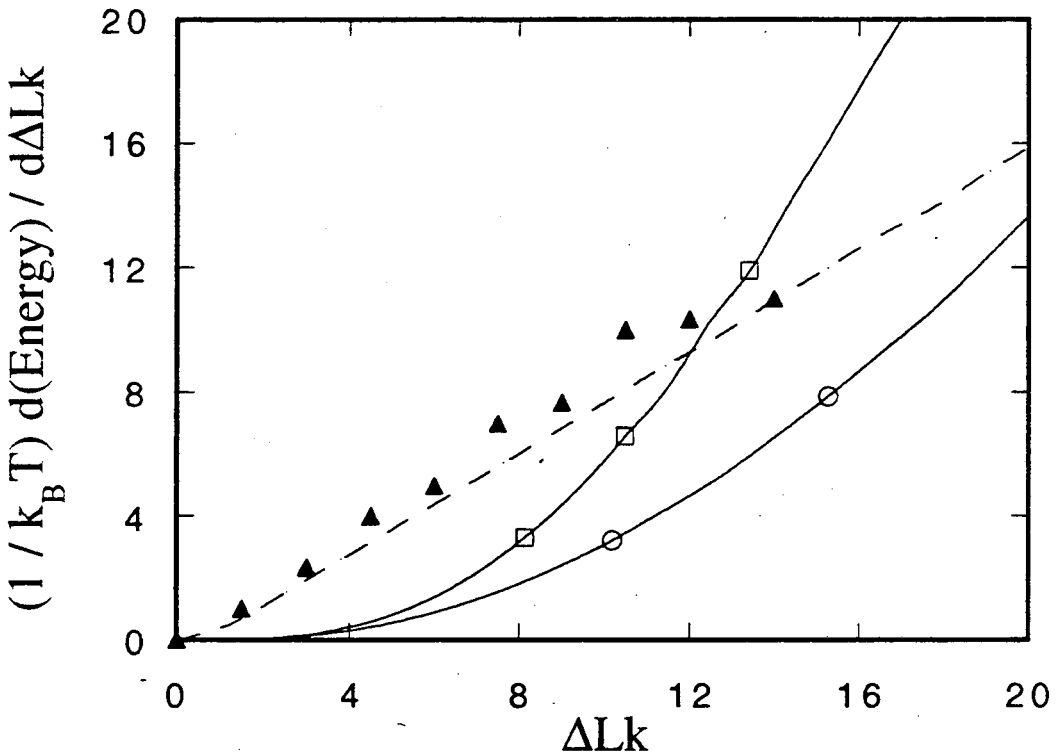


Figure 21. Supercoiling - induced stress for the model of this chapter, the Monte Carlo study of Klenin et al., and equation (3.8). Symbols, bend and twist constants, and ionic strength the same as in Figure 19.

$h_0 \rho Wr / \sigma$ is the proportion of linking difference in writhe (usually about 0.7), L_{tot} is the total length of the plasmid, and L_{ends} is the length of the end loops. A first order approximation for the end effects equates the increase in energy caused by the increase in effective specific linking difference with the bend energy of the end loops:

$$(L_{tot} - L_{ends})\, \rho E_{tot}(\sigma_{eff}) \approx L_{tot}\, \rho E_{tot}(\sigma) + \frac{2\pi^2 \beta_{bend}}{L_{ends}} \tag{3.10}$$

where $\rho E_{tot}(\sigma)$ and $\rho E_{tot}(\sigma_{eff})$ are the energy densities as given by equation (3.3) at specific linking differences σ and σ_{eff}, respectively; and the last term is the bend energy of the end loops. For a 2650 base pair plasmid at a ΔLk of 15, equation (3.10) is approximately satisfied by $L_{ends} = L_{tot} / 20$. This yields a bend energy for the end caps of about 20 kT, a substantial fraction of the difference between the hard-body model of chapter II and the Monte Carlo result. Therefore, until a model incorporating end effects is developed, one cannot say with certainty what fraction of the difference between the chapter II and Monte Carlo results is due to entropic effects as opposed to finite-length effects for this small plasmid.

4. Electron microscopy data

The most extensive electron microscopy study of DNA supercoiling is that of Boles *et al.* (1990). Values for the superhelical radius of plectonemically supercoiled DNA were obtained from electron micrographs by two methods - either by measuring the area of the plectonemic superhelix or by counting the number of superhelical branches and nodes. Plasmids in either TE solution (10 mM-Tris·HCl (pH 8.0), 1 mM-EDTA) or Mg solution (20 mM-Tris·HCl (pH 8.0), 50 mM-NaCl, 10 mM-$MgCl_2$) were spread on a polylysine-coated grid. They were then washed in a solution of 0.1 M-ammonium

acetate, stained with 5% (w/v) uranyl acetate and washed again with 0.01 M-ammonium acetate. Comparisons of the data with the calculated values of superhelix radius for each spreading solution are shown in Figures 22 and 23.

A recent cryo-electron microscopy study (Adrian *et al.*, 1990) gives superhelix radii at two different ionic strengths, also shown in Figures 18 and 19. This technique does not suffer from possible complications introduced by electrostatic interactions with a charged grid or exposure of the DNA to solutions of differing ionic concentrations.

The experimentally observed values for solutions without Mg^{2+} are in reasonable agreement with our model. In the case of the data of Boles *et al.*, however, this may be partly coincidental. The fact that their superhelical radii are basically independent of the ionic strength of the spreading solution suggests that the observed DNA conformations reflect some combination of the staining solution, wash solution, and interaction with the charged grid. The disagreement between our results and the high-salt point of Adrian *et al.* probably reflects a failure of our P-B treatment for solutions of DNA and Mg^{2+}, as discussed in section III.E.

5. Ionic strength dependence of cruciform formation

The electrostatic repulsion between strands is a strong function of ionic strength. As discussed above and shown in Figures 14 - 16, this causes the energy of supercoiling to be significantly greater at low ionic strength, mainly by decreasing the proportion of linking difference which goes into writhe. The superhelical stress on the DNA, which is $d\rho E_{tot} / d|\sigma|$, is also greater at low ionic strength. Conformational transitions such as cruciform formation and B-Z transitions relieve stress induced by supercoiling. Therefore conformational transitions induced by superhelical stress should occur at lower values of $|\sigma|$ at low ionic strength. Such a dependence on ionic

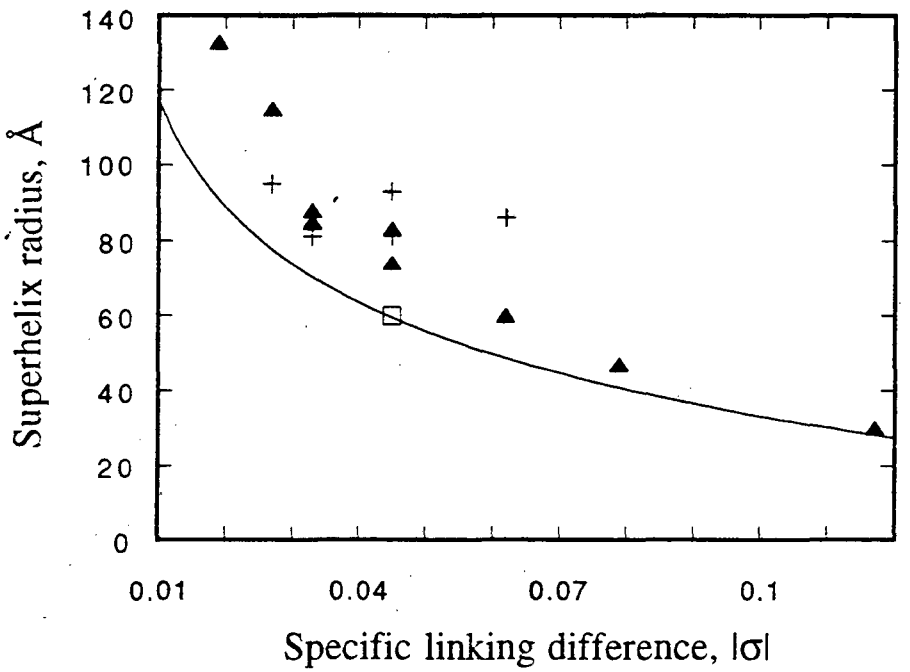


Figure 22. Superhelix radius from electron microscopy data and from optimum theoretical conformation. Triangles - data of Boles *et al.* with radius calculated by counting nodes and branches; +'s - data of Boles *et al.* with radius calculated form the area of the superhelix on the electron micrograph; square - data of Adrian *et al.* The sample of Adrian *et al.* was frozen in a buffer of ionic strength I = 14 mM. The points of Boles *et al.* were spread on a polylysine-coated grid in a buffer of the same ionic strength, then stained and washed. Curve - theoretical radius calculated with I = 14 mM.

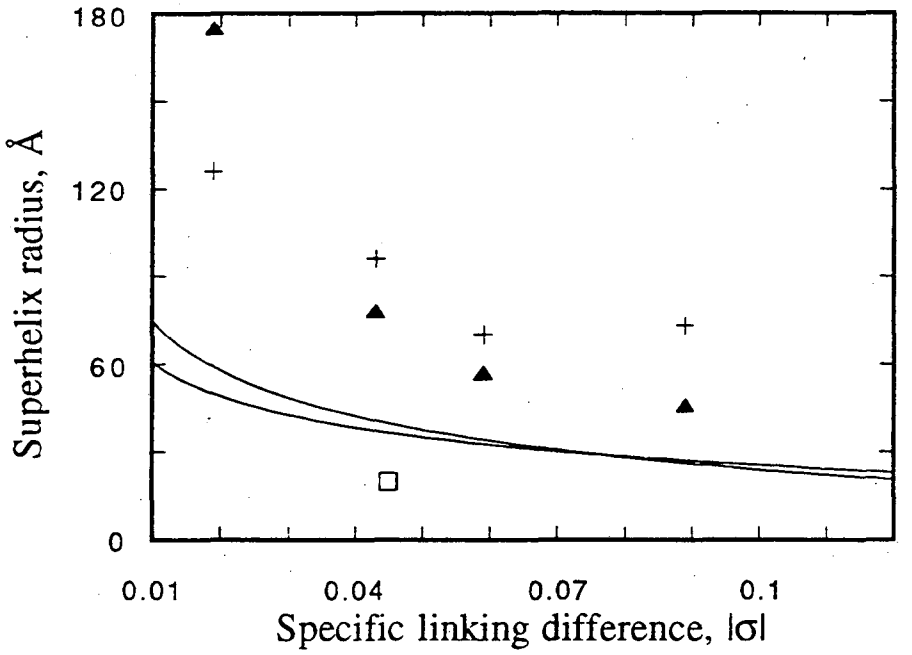


Figure 23. Superhelix radius from electron microscopy data and from optimum theoretical conformations, for solutions containing Mg^{2+} . Triangles - data of Boles *et al.* with radius calculated by counting nodes and branches; +'s - data of Boles *et al.* with radius calculated from the area of the superhelix on the micrograph; square - data of Adrian *et al.* The sample of Adrian *et al.* was frozen in a buffer with 14 mM monovalent salt and 10 mM $MgCl_2$. The samples of Boles *et al.* were spread on a polylysine-coated grid in a buffer of 70 mM monovalent salt and 10 mM $MgCl_2$, then stained and washed. Upper curve (on left) - calculated radius for ionic solution of Adrian *et al.*; lower curve (on left) - calculated radius for ionic solution of Boles *et al.*

strength is observed in supercoiling-induced B to Z transitions (Wells *et al.*, 1982) and cruciform formation (Singleton, 1983; Kozyavkin *et al.*, 1990).

Kozyavkin *et al.* (1990) carried out a systematic study of the dependence of cruciform formation in the pA03 plasmid on linking difference at different ionic concentrations. A comparison of some of their results with the model presented here and with the usual quadratic expression for the energy of supercoiling is shown in Figure 24. As Kozyavkin *et al.* noted, the width of transitions deduced from the quadratic expression (equation (3.9)) and equation (3.8) is roughly constant. Our treatment gives sharper transitions at low ionic strength, in agreement with the experimental data. The energy cost of forming a cruciform, ΔH_{cruc}, is derived from fitting the data. The best-fit values using equation (3.8) are $\Delta H_{cruc} = 6.6 \times 10^4$ J / mole at ionic strength I = 20mM and $\Delta H_{cruc} = 8.11 \times 10^4$ J / mole at I = 87mM. The corresponding values using our model are 6.7×10^4 J / mole and 5.7×10^4 J / mole, respectively.

Kozyavkin *et al.* also published data for higher ionic strengths. Neither our model nor equation (3.8) could be made to fit this data. As in the electron microscopy case, we believe this is because the solutions contained Mg^{2+}.

E. Discussion

The model we present here allows calculations of the energy and conformation of supercoiling in terms of bend, twist and electrostatic energy. Good agreement is found with experimental data for DNA in monovalent salt solutions.

One of two areas of disagreement with experiment we have found is for DNA solutions containing Mg^{2+}. The cryo-electron microscopy data of Adrian *et al.* show the two strands of plectonemically supercoiled DNA within 40Å of each other, about

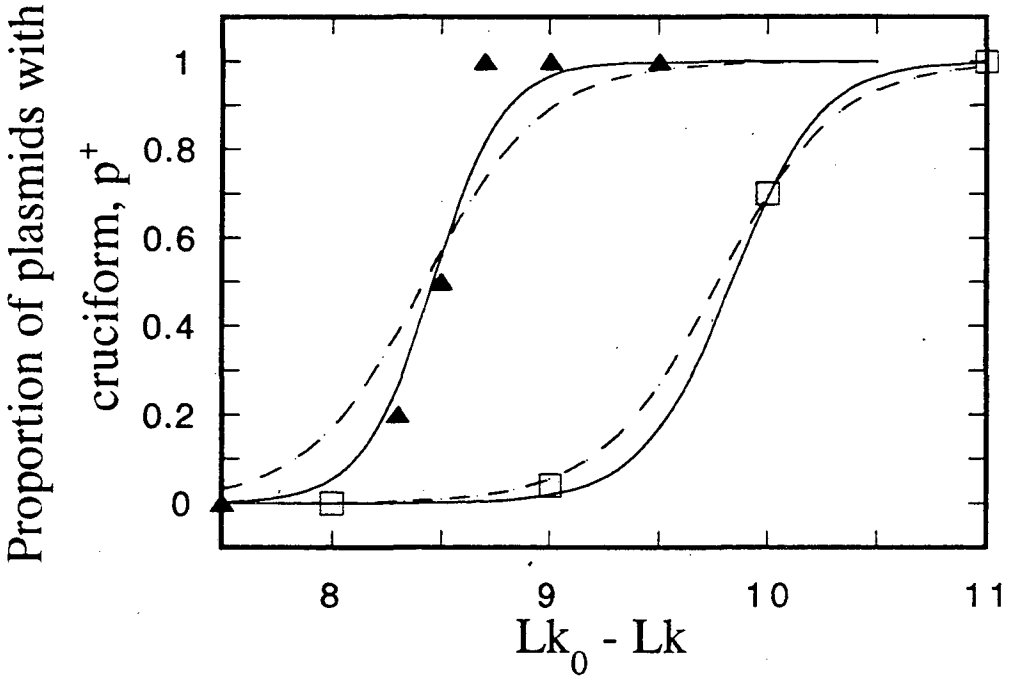


Figure 24. Proportion of plasmids with cruciform. Triangles - data of Kozyavkin *et al.* for I = 20 mM; squares - data of Kozyavkin *et al.* for I = 87 mM; solid curves - calculated best fits using the model of this paper; dashed curves - calculated best fits using equation 3.8.

half the distance predicted by the P-B treatment we used. The data of Kozyavkin *et al.*. for solutions containing Mg^{2+} (not shown) is also roughly consistent with the two strands coming much closer than our model predicts, raising the writhe per length and lowering the energy of supercoiling.

We conclude that the P-B treatment used here is inadequate to describe the interaction of Mg^{2+} with DNA. The P-B treatment outlined here does not allow for specific binding of cations to DNA. Such binding, even if weak, can have significant effects for divalent cations since the cationic concentration at the DNA surface is quite high (Anderson & Record, 1982) This suggests that Mg^{2+} has a greater affinity for DNA than the P-B treatment outlined here predicts.

The other area of disagreement with experiment is at low values of $|\sigma|$. For $|\sigma| \leq 0.02$ the energy of supercoiling given by equation (3.8) is several times greater than the energies given by our model, although the absolute difference is not so large. There is little reason to doubt the validity of equation (3.8) at low $|\sigma|$. Equation (3.8) was derived from thermal distributions of topoisomers, which only includes the low $|\sigma|$ region. The Monte Carlo study of Klenin *et al.* (1991) faithfully reproduces equation (3.8) at low $|\sigma|$.

A related discrepancy is the proportion of linking difference in writhe, $h_0 \rho Wr / \sigma$, at low $|\sigma|$. As shown in Figure 16, our model puts all the linking difference into writhe as $|\sigma| \to 0$, with the proportion monotonically decreasing with increasing $|\sigma|$. Both the Monte Carlo study of Klenin *et al.* and the electron microscopy data of Boles *et al.* give a figure of about 0.7 for $h_0 \rho Wr / \sigma$, roughly independent of $|\sigma|$.

Two factors left out of our model which might explain these discrepancies are finite-length effects and entropic effects. Simple calculations as outlined in section III.D.3 show that for moderate length plasmids (e.g. 10^4 b.p. or longer), the energy of forming end loops on a plectonemic supercoil can be made quite small. Thus we believe entropic effects must be responsible for the discrepancies with electron microscopy data

and with equation (3.8), while end effects also contribute for the short plasmid simulated by Klenin *et al.* The picture which emerges is of a relatively large loss of conformational entropy at low $|\sigma|$ as a loose plectonemic superhelix forms. At higher $|\sigma|$, as the superhelix adds turns and its radius shrinks, the loss of entropy is less significant. This agrees with a recent study which shows that entropic effects should be small at physiological levels of supercoiling (Hearst and Hunt, 1991).

Since our model excludes entropic effects, we expect and find it to be more accurate with increasing $|\sigma|$. The high - $|\sigma|$ regime is beyond the range from which equation (3.8) was derived and still presents some difficulties for Monte Carlo studies. While comparison with equation (3.8) and the Monte Carlo model of Klenin *et al.* at low $|\sigma|$ let us infer the behavior of the entropy loss of supercoiling, at high $|\sigma|$ and for large molecules our model has advantages over the Monte Carlo approach. The longest molecule studied by Klenin *et al.* was 2650 b.p., and the largest value of $|\sigma|$ used was 0.06. Our model gives behavior in the long-length limit and high $|\sigma|$ presents no difficulties. At sufficiently high $|\sigma|$, the approach outlined here shows that the proportion of linking difference in writhe must start to decrease, behavior which did not clearly emerge from the Monte Carlo study.

Another advantage is that the explicit treatment of electrostatic effects should give more accurate results than using an ionic-strength dependent excluded volume, the approach taken in chapter II and the study of Klenin *et al.*. As noted in section II.D.3 and shown in Figure 19, the soft-body electrostatic interaction between strands gives significantly different results from a hard-body model at physiological levels of $|s|$ and salt concentration. Modelling such effects with Monte Carlo calculations would take orders of magnitude more computer time (or computer speed) than the calculations already carried out. We expect the two approaches will continue to play complementary roles.

Perhaps the most promising application of this approach to DNA supercoiling is to supercoiling-induced conformational transitions. Since these transitions involve molecules at relatively high $|\sigma|$, entropic effects should play a small role, as may be inferred from Figures 20 and 21.

References

Adrian, M., Heggeler-Bordier, B., Wahli, W., Stasiak, A.Z., Stasiak, A. & Dubochet, J. (1990) *EMBO J.* **9**, 4551-4554

Anderson, C.F. & Record, M.T. (1982) *Ann. Rev. Phys. Chem.* **33**, 191-222

Boles, T.C., White, J.H. & Cozzarelli, N.R. (1990) *J. Mol. Biol.* **213**, 931-951

Brady, G.W., Fein, D.B., Lambertson, H., Grassian, V. & Benham, C.J. (1983) *Proc. Nat. Acad. Sci. U.S.A.* **80**, 741-744

Brady, G.W., Satkowski, M., Foos, D. & Benham, C.J. (1987) *J. Mol. Biol.* **195**, 185-191

Brian, A.A., Frisch, H.L. & Lerman, L.S. (1981) *Biopolymers* **20**, 1305-1328

Depew, R.E. & Wang, J.C. (1975) *Proc. Nat. Acad. Sci. U.S.A.* **72**, 4275-4279

Hearst, J.E. & Hunt, N.G. (1991) *J. Chem. Phys.* **95**, 9322

Horowitz, D.S. & Wang, J.C. (1984) *J. Mol. Biol.* **173**, 75 - 91

Hunt, N.G. & Hearst, J.E. (1991) *J. Chem. Phys.* **95**, 9329

Klenin, K.V., Vologodskii, A.V., Anshelevich, V.V., Dykhne, A.M. & Frank-Kamenetskii, M.D. (1991) *J. Mol. Biol.* **217**, 413-419

Kozyavkin, S.A., Slesarev, A.I., Malkhosyan, S.R. & Panyutin, I.G. (1990) *Eur. J. Biochem.* **191**, 105-113

Pulleyblank, D.E., Shure, M., Tang, D., Vinograd, J. & Vosberg, H.-P. (1975) *Proc. Nat. Acad. Sci. U.S.A.* **72**, 4280-4284

Shimada, J. & Yamakawa, H. (1988) *Biopolymers* **27**, 657-673

Shore, D. & Baldwin, R.L. (1983) *J. Mol. Biol.* **170**, 983-1007

Singleton, C.K., Klysik, J., Stirdivant, S.M., & Wells, R.D. (1982) *Nature* **299**, 312 - 316

Singleton, C.K. (1983) *J. Biol. Chem.* **258**, 7661 - 7668

Stigter, D. (1975) *J. Colloid Interface Sci.* **53**, 296-306

Stigter, D. (1987) *Cell Biophysics* **11**, 139-158

Yarmola, E.G., Zarudnaya, M.I., & Lazurkin, Yu. S. (1985) *J. Biomol. Str. and Dynamics* **2**, 981-993

IV. Conclusion and Future Work

A. Conclusion

The approach to the energy and conformation of supercoiled DNA presented in this dissertation has proven to be useful. By ignoring end effects and conformational entropy and by treating extensive quantities (e.g. writhe, bend energy) as linear densities (writhe per unit length, bend energy per unit length) we are able to reach several conclusions of biological interest. Among them are:

1. The plectonemic superhelix is energetically favored over toroidal conformations for biologically relevant values of σ. The pitch angle of the plectonemic superhelix approaches 90° in the limit $|\sigma| \to 0$ and approaches 45° as $|\sigma|$ approaches infinity. Optimal conformations for the single helix are described and have low pitch angles, limited by excluded volume (or electrostatic repulsion).

2. Since the electrostatic repulsion between strands is a strong function of ionic strength, the electrostatic energy and the writhe per unit length are also strong functions of ionic strength. At low ionic strength, the strands are held farther apart, resulting in a greater fraction of linking difference in twist and higher energy of supercoiling.

3. The ionic strength dependence of the energy of supercoiling implies that the stress on a supercoiled DNA molecule is also a strong function of ionic strength. Therefore conformational transitions which relieve stress caused by supercoiling should occur at lower values of $|\sigma|$ at low ionic strength. Such a pattern is observed for cruciform formation (Singleton, 1983; Kozyavkin *et al.*, 1990), B-Z transitions (Singleton *et al.*, 1982) and DNA melting (Lyubchenko and Shlyakhtenko, 1988).

4. By comparing the predictions of the model presented here with a Monte Carlo study (Klenin *et al.*, 1991), we infer that there is a relatively large loss of entropy associated with the initial formation of a loose plectonemic helix at low $|\sigma|$. As $|\sigma|$ increases, the superhelix adds turns and its radius shrinks, and the loss of entropy

becomes smaller. This is in accord with the study of Hearst and Hunt (1991), where this result was derived based on the polymer dynamics model of Harris and Hearst (1966). It follows that the effects of ignoring entropy should be particularly small for molecules under large superhelical stress, as in the case of supercoiling-induced conformational transitions.

5. The predictions of this model are in agreement with electron microscopy data on the conformation of supercoiled DNA for DNA in monovalent salt solutions. The model also agrees well with data on supercoiling-induced cruciform formation for DNA in monovalent salt solutions.

B. Future work

The three main areas which seem most promising for further work are a more detailed examination of the energies of supercoiling-induced conformational transitions, a refinement of the treatment of electrostatic interactions, and application of the ideas developed here to chromatin structure. Preliminary thoughts on these three areas are outlined below.

1. Conformational transitions

The initial work presented in chapter III modelling the data of Kozyavkin *et al.* illustrates the potential of the model to extract values of the energy of conformational transitions from experimental data. There is quite a bit of this data published, and re-examining it using this model should provide revised values for the energy of cruciform formation. The same type of modelling should be possible for B-Z transitions, although we have not yet tried this.

2. Refinement of the treatment of the electrostatic interaction

The treatment of the electrostatic interaction outlined in chapter III seems to be adequate for ionic strengths greater than 10 mM. Below 10 mM, problems with the behavior of the numerical solutions appear. These problems must be examined in detail if the theory is to be applied in this region of very low ionic strength.

One can question the approximation used for the electrostatic interaction between strands, equation (3.1). The D-H equation is valid in the region where the potential ϕ is small, i.e. where $e\,\phi / k_B T \ll 1$ (where e is the electron charge). In this region, the non-linear P-B equation reduces to the linear D-H equation. Since the two strands of the plectonemic helix are sufficiently far apart that each is in the D-H region of the other's potential field, equation (3.1) is plausible. But it remains to be demonstrated rigorously that equation (3.1) is a good approximation to the full three-dimensional P-B equation for the geometry of the superhelices examined.

Preliminary work on a full P-B treatment has started, in collaboration with Robert Van Buskirk. By expressing the problem in helical coordinates we have taken advantage of the helical symmetry of the problem to reduce it to two dimensions. If various numerical difficulties can be overcome, we should be able to check the validity of equation (3.1).

The results obtained in chapter III using equation. 3.1 seem to be fairly robust against changes in the potential. This can be understood by considering the exponential nature of the screened electrostatic potentials. If the potential between two strands is changed by a factor of e, the two strands move roughly one Debye length closer together, which does not alter the results that much, since the strands are several Debye lengths apart. In contrast, changing the ionic conditions changes the Debye length, which can alter the separation between strands and the energy of supercoiling significantly.

An adequate treatment of the electrostatic interaction for solutions containing Mg^{2+} is desirable, both for comparison with experimental results and because Mg^{2+} is present *in vivo*. The P-B treatment outlined here is not very stable numerically for solutions containing Mg^{2+}. A closer examination of the numerical behavior of the P-B treatment may yield some ideas for possible modifications.

One possible modification to the P-B treatment used would be to allow for specific binding of Mg^{2+} to the DNA. At the high concentrations of Mg^{2+} near the DNA predicted by the P-B equation, even weak binding can have significant effects (Anderson and Record, 1982).

It is possible that the failure of the P-B treatment is due to correlations between Mg^{2+} cations. Such an effect has been examined by Oosawa (1971) and more recently by Guldbrand *et al.* (1986). Their work may offer a basis for treating the effects of such correlations. Yet another possibility is a cluster expansion treatment of the type developed by Mayer (1950). Such an approach to DNA solutions has been developed by Manning (1978).

3. Possible applications to chromatin structure

DNA in eukaryotic cells is wrapped around proteins known as histones. The complex of the histone proteins and the strand of DNA wrapped around them twice is known called a nucleosome. Nucleosomes in turn form a higher order structure known as the 30 nm fiber. The geometry of the 30 nm fiber is not certain, but the leading model has the nucleosomes wrapped in a solenoid.

There have been some theoretical studies of the writhe of DNA in nucleosomes (Zivanovic *et al.*, 1988; LeBret, 1988) . I believe there is more to be said on this subject based on numerical calculations of the writhe integral (see Appendix) using the geometry of the nucleosome. Calculations of writhe using the proposed geometry of

the 30 nm fiber would also be of interest. A related calculation of interest would be a comparison of the energy of supercoiling of a plectonemic helix of DNA with the energy of supercoiling of DNA in nucleosomes and in the 30 nm fiber.

References

Guldbrand, L., Nilsson, L.G., and Nordenskiold, L. (1986) *J. Chem. Phys.* **85**, 6686

Harris, R.A. and Hearst, J.E. (1966) *J. Chem. Phys.* **44**, 2595

Hearst, J.E. and Hunt, N.G. (1991) *J. Chem. Phys.* **95**, 9322

Klenin, K.V., Vologodskii, A.V., Anshelevich, V.V., Dykhne, A.M. and Frank-Kamenetskii, M.D. (1991) *J. Mol. Biol.* **217**, 413

Kozyavkin, S.A., Slesarev, A.I., Malkhosyan, S.R. and Panyutin, I.G. (1990) *Eur. J. Biochem.* **191**, 105

Lyubchenko, Yu.L. and Shlyakhtenko, L.S. (1988) *Nucl. Acids Res.* **16**, 3269

Manning, G.S. (1978) *Q. Revs. Biophys.* **11**, 2

Mayer, J.E. (1950) *J. Chem. Phys.* **18**, 1426

Oosawa, F. (1971) *Polyelectrolytes* M. Dekker, New York

Singleton, C.K., Klysik, J., Stirdivant, S.M., and Wells, R.D. (1982) *Nature* **299**, 312

Singleton, C.K. (1983) *J. Biol. Chem.* **258**, 7661

Appendix A: Evaluation of Wr

1. Convergence and numerical considerations

The writhe integral

$$Wr = \frac{1}{4\pi} \int_{C',\, C''} du'\, du'' \frac{[\vec{r}(u') - \vec{r}(u'')] \cdot \left[\frac{d\vec{r}(u')}{du'} \times \frac{d\vec{r}(u'')}{du''}\right]}{|\vec{r}(u') - \vec{r}(u'')|^3} \tag{A.1}$$

can be evaluated using standard numerical integration routines for any given space curve **r** (u). The only complication in carrying out this integration occurs on the diagonal where u' = u". At these points, the denominator of the integrand in equation (A.1) goes to zero. The integrand itself is well-behaved on the diagonal and goes to zero for curves with continuous derivatives. To show this, we introduce the formulas of Frenet. A nice introduction to them can be found in Struik (1988).

A space curve **r** (s) parametrized by the arc-length s defines a local coordinate system at any point **r** (s_0). The axes of this coordinate system are the unit tangent vector **t** :

$$\mathbf{t} = d\mathbf{r} / ds\,, \tag{A.2}$$

the unit normal **n** :

$$\vec{n} = \frac{\frac{d\vec{t}}{ds}}{\left|\frac{d\vec{t}}{ds}\right|} = \frac{1}{\kappa}\frac{d\vec{t}}{ds} \tag{A.3}$$

where κ is the curvature, and the unit binormal **b** :

$$\vec{b} = \vec{t} \times \vec{n} \tag{A.4}$$

These three unit vectors and their derivatives are related by the formulas of Frenet:

$$\begin{aligned} \frac{d\vec{t}}{ds} &= \kappa \vec{n} \\ \frac{d\vec{n}}{ds} &= -\kappa \vec{t} + \tau \vec{b} \\ \frac{d\vec{b}}{ds} &= -\tau \vec{n} \end{aligned} \tag{A.5}$$

where τ is the torsion of the curve, given by

$$\tau = \left| \frac{d\vec{b}}{ds} \right| \tag{A.6}$$

In general, the geometric torsion τ of a curve has no relation to the physical torsion on DNA discussed in this dissertation.

With the formulas of Frenet, (A.5), the integrand in equation (A.1) can be Taylor expanded about the diagonal. For example, **r** (s + h) is

$$\begin{aligned} \vec{r}(s+h) &= \vec{r}(s) + h\frac{d\vec{r}}{ds} + \frac{h^2}{2}\frac{d^2\vec{r}}{ds^2} + \frac{h^3}{6}\frac{d^3\vec{r}}{ds^3} + \frac{h^4}{24}\frac{d^4\vec{r}}{ds^4} + O(h^5) \\ &= \vec{r}(s) + h\vec{t} + \frac{h^2}{2}\kappa\vec{n} + \frac{h^3}{6}\left[-\kappa^2\vec{t} + \kappa\tau\vec{b} + \kappa'\vec{n}\right] + \\ &\frac{h^4}{24}\left[-3\kappa\kappa'\vec{t} - \kappa^3\vec{n} - \kappa\tau^2\vec{n} + \kappa''\vec{n} + 2\kappa'\tau\vec{b} + \kappa\tau'\vec{b}\right] + O(h^5) \end{aligned} \tag{A.7}$$

where ' and " denote first and second derivatives with respect to s, respectively. Expanding the integrand in this manner and keeping lowest order terms gives

$$\frac{h}{12}\kappa^2\tau + \frac{h^2}{24}\left[2\kappa\kappa'\tau + \kappa^2\tau'\right] + O(h^3) \tag{A.8}$$

Therefore the integrand goes to zero linearly with h.

Although the integrand is well-behaved mathematically, it is troublesome from a numerical standpoint. Finite-precision calculations of the integrand behave erratically near the diagonal. I chose to get around this problem by excising a narrow strip surrounding the diagonal from the region of integration. Given knowledge about the values of κ and τ and their derivatives, equation (A.8) can be used to estimate the error introduced by excising the strip around the diagonal. The h^2 term is the lowest order error term, since the term proportional to h makes equal and opposite contributions on each side of the excised diagonal.

2. Evaluation of Wr for finite single and plectonemic helices

A helix of radius R and pitch angle a can be parametrized as

$$\begin{aligned} x(u) &= R \cos u \\ y(u) &= R \sin u \\ z(u) &= R\, u \tan \alpha \end{aligned} \tag{A.9}$$

Substituting into equation (1.4) gives, for a single helix,

$$Wr = \frac{\tan \alpha}{4\pi} \int du_1\, du_2 \, \frac{2 - 2\cos(u_2 - u_1) - (u_2 - u_1)\sin(u_2 - u_1)}{\left[2 - 2\cos(u_2 - u_1) + (u_2 - u_1)^2 \tan^2\alpha\right]^{3/2}} \tag{A.10}$$

Since the integrand is simply a function of the difference $u_2 - u_1$, equation (A.10) can be re-written as a one-dimensional integral:

$$Wr = \frac{\tan \alpha}{2\pi} \int_0^{u_{max}} du \, \frac{2 - 2\cos u - u \sin u}{\left[2 - 2\cos u + u^2 \tan^2\alpha\right]^{3/2}} \left[u_{max} - u\right] \tag{A.11}$$

This exact result for the finite single helix is plotted in Figure 25, along with the infinite-length result, equation (1.5). The diagonal has become an integrable singularity at $u = 0$.

The plectonemic helix can be treated in a similar way with the addition of a cross-term between strands. This gives

$$Wr = \frac{\tan\alpha}{2\pi}\int_0^{u_{max}} du \; \frac{2 - 2\cos u - u\sin u}{[2 - 2\cos u + u^2\tan^2\alpha]^{3/2}}\,[u_{max} - u] - \frac{2 + 2\cos u + u\sin u}{[2 + 2\cos u + u^2\tan^2\alpha]^{3/2}}\,[u_{max} - u] \tag{A.12}$$

Equation (A.12) is plotted in Figure 26, along with the infinite-length result, equation (1.6).

References

Struik, D.J. (1988) *Lectures on Classical Differential Geometry* Dover, New York

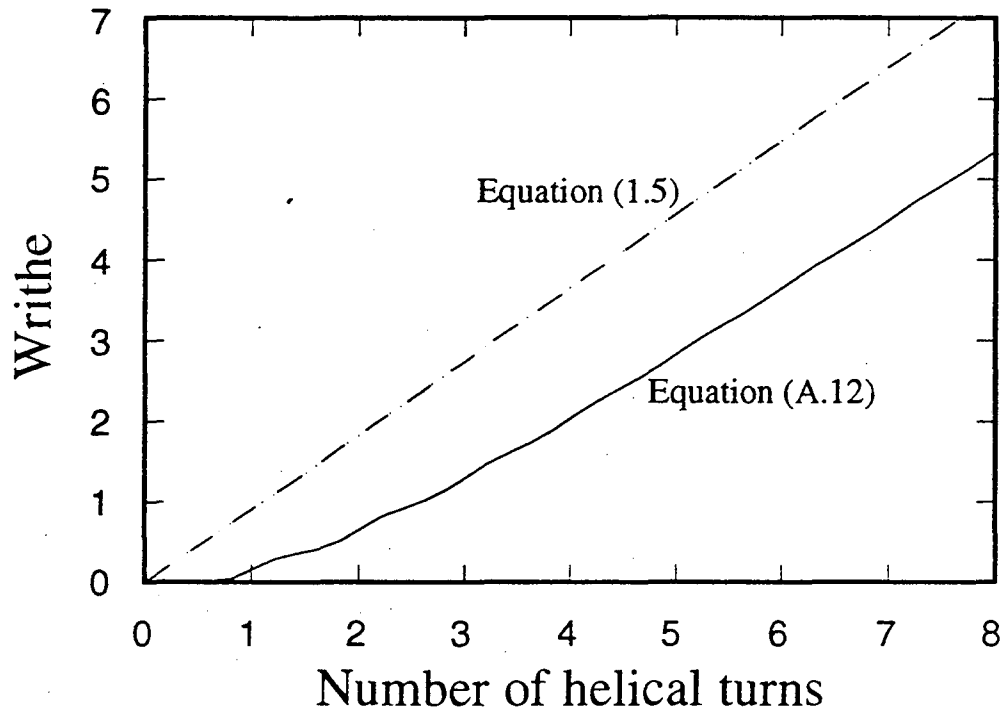


Figure 25. Writhe of a finite single helix with pitch angle $\alpha = 5^{o}$. Dashed curve - infinite length expression, equation (1.5); Solid curve - exact result for finite helix, equation (A.12).

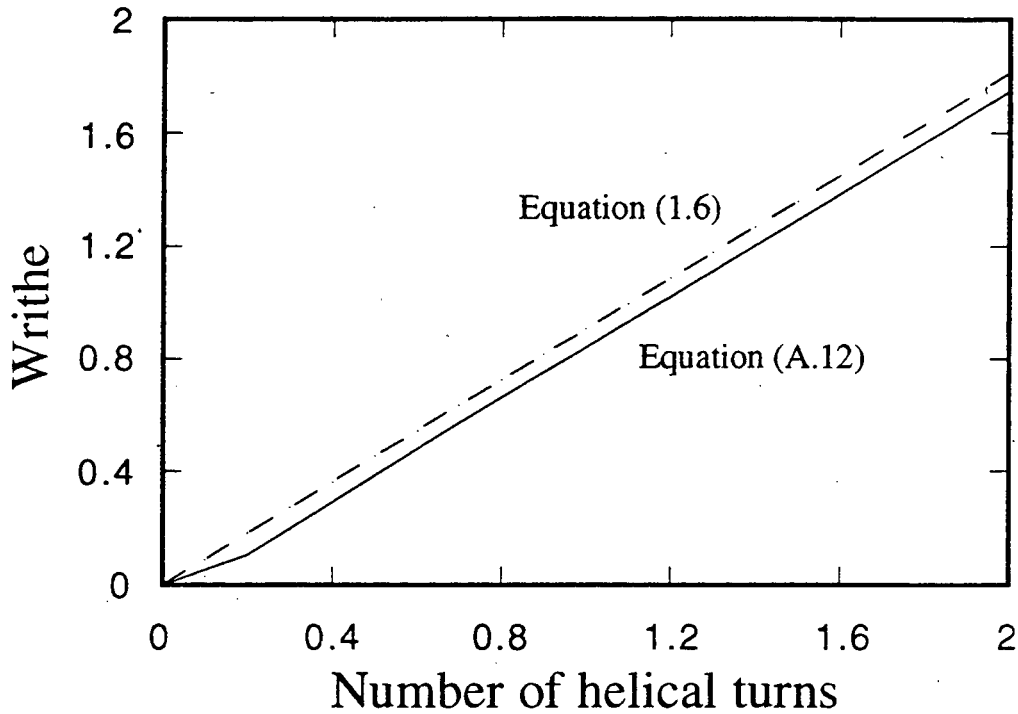


Figure 26. Writhe of a finite plectonemic helix with pitch angle 65o. Dashed curve - infinite length expression, equation (1.6); solid curve - exact result, equation (A.12). The difference between the two curves is about 0.02.

Appendix B: Fortran Code for Plectonemic Helix

The following fortran program calculates optimum plectonemic configurations as outlined in chapter III.

```
c       program pscrn.for
c       finds optimum plectonemic configurations for
c       given salt concentrations and for sigma between 0 and 0.2.
c       calculates potential from a helical dna
c       molecule by numerical integration
c       of the P-B equation and then matching
c       boundary conditions with the modified
c       Bessel function solution to the Debye-Huckel
c       equation
c       must be linked with the imsl library, e.g. $ link pscrn, imsl/lib

        integer mxparm, neq, interv, nout, maxsub
        parameter (mxparm=50,neq=2,maxsub=1000)
        integer ido, neval, nsubin, n, iparam(7),maxfcn
        real*8 pi,esq,kt,epsilon,na,kappa,xzero,dydxzero
        real*8 k0xz,k1xz,z,y(neq),zend,tol,param(mxparm)
        real*8 diff, zinit,dydxzout, diffold, dz, xend
        real*8 eperl, r, xinit, k1xi, k0xi, k0xe, bessend
        real*8 beta, kone, kzero, gamma, const, fcn
        real*8 qfperlsq, alpha, integral, msd, fz
        real*8 bound, errabs, errel, errest, f, result
        real*8 alist(maxsub),blist(maxsub),rlist(maxsub)
        real*8 elist(maxsub), iord(maxsub)
        real*8 sigma, ebend, etwist, eelect, etot
        real*8 bbend, btwist, h0, arg, xguess(2), xscale(2)
        real*8 fscale, rparam(7), x(2), gvalue, g, grtn
        real*8 etwonly, wrperl, wrperh0, twperl, twperh0
        real*8 prcntwr, prcnttw, delta, mg, ion
        real*8 ftol, s, fr, diffb, qp, kr
        real*8 above, below, dabove, dbelow
        integer seed


        common alpha, kappa, r, sigma, qfperlsq, bbend
        common btwist, h0, pi
        common mg, ion, na, epsilon

        external dbsk1, dbsk0, fcn, sset, divpbs, f, dq2agi
        external grtn, dumpol

        open(10,file='screen.dat',status = 'new')
        open(11,file='sigma.dat',status = 'new')
        open(12,file='r.dat',status = 'new')
        open(13,file='alpha.dat',status = 'new')
        open(14,file='etot.dat',status = 'new')
        open(15,file='eelect.dat',status = 'new')
```

```
      open(16,file='etwist.dat',status = 'new')
      open(17,file='ebend.dat',status = 'new')
      open(18,file='wrperl.dat',status = 'new')
      open(19,file='wrperh0.dat',status = 'new')
      open(20,file='twperl.dat',status = 'new')
      open(21,file='twperh0.dat',status = 'new')
      open(22,file='prcntwr.dat',status = 'new')
      open(23,file='prcnttw.dat',status = 'new')

      seed = 654721
      pi = 3.1415926535
c
c     esq is the electron charge squared, in erg-cm
c     kt is in ergs
c     epsilon is the dielectric constant of water
c
      esq = 2.3e-19
      kt = 4.05e-14
      epsilon = 80.
      bbend = 2.8e-19
      btwist = 3.1e-19
      h0 = 3.54e-7
c
c     na is salt concentration in M/l
c     mg is divalent cation concentration
c     ion is ionic strength
c     and is used to calculate the Debye length, 1/kappa
c
      na = 0.5
      mg = 0.0
      ion = na + 3. * mg
      kappa = (8. * pi * esq * ion * 1.e-3 * 6.e23/(epsilon * kt))
      kappa = dsqrt(kappa)
c
c     xzero is kappa times the dna radius, 12 Angstroms
c     dydxzero is the desired slope at xzero from Gauss'
c     law and qp * e per phosphate residue
c
      xzero = kappa * 1.2e-7
      qp = 0.73
      dydxzero = -4. * qp * esq/(3.4e-8 * xzero * epsilon * kt)
      zend = dlog(xzero)
c
c     k0xz is K0(xzero), the zeroeth order modified Bessel
c     function evaluated at xzero
c     k1xz is K1(xzero), the first order modified
c     Bessel function evaluated at xzero
c
      k0xz = dbsk0(xzero)
      k1xz = dbsk1(xzero)

      call sset(mxparm, 0.0, param, 1)
      param(10) = 0.0
```

```
      tol = 1.e-4

      write(10,100)
100     format(' zinit  xzero  dydxzero  dydxout yxzout diff')

c
c     set initial conditions
c
      z = 2.319 - 0.4 * ion**2
      if (z .gt. 2.310) z = 2.310
      if (ion .lt. 0.05) z = 2.293 + 0.5 * ion
      if (ion .lt. 0.02) z = 2.34919 - 7.47e-3/dsqrt(ion)
      if (ion .lt. 0.009) z = 2.2214 + 0.552 * dsqrt(ion)
      if (ion .lt. 0.007)z=2.39718 + 2.62111e-2 * dlog(ion)
      if (ion .lt. 0.004)z=2.38313+2.34826e-2 * dlog(ion)
      if (mg .ne. 0.) z = z - 0.044 - 3.* mg/2.
      zinit = z
      xinit = dexp(zinit)
      k1xi = dbsk1(xinit)
      k0xi = dbsk0(xinit)
      y(1) = 1.e-4
      y(2) = -1.e-4 * xinit * k1xi/k0xi
      ido = 1
      above = 3.
      below = 0.
      dabove = 10. * dydxzero
      dbelow = -10. * dydxzero
      call divpbs(ido,neq,fcn,z,zend,tol,param,y)
      dydxzout = y(2) * xzero
      diff = dydxzout - dydxzero
      if (diff .lt. 0.) then
       if (diff .gt. dabove) then
        dabove = diff
        above = zinit
       end if
      else
       if (diff .lt. dbelow) then
        dbelow = diff
        below = zinit
       end if
      end if

      dz = 0.001
      if (ion .lt. 0.02) dz = 1.e-4
      zinit = zinit + dz
      z = zinit
      ido = 3
      call divpbs(ido,neq,fcn,z,zend,tol,param,y)
      do 200 i=1,200
      xinit = dexp(zinit)
      k1xi = dbsk1(xinit)
      k0xi = dbsk0(xinit)
      y(1) = 1.e-4
```

```
      y(2) = -1.e-4 * xinit * k1xi/k0xi
      ido = 1
      call divpbs(ido,neq,fcn,z,zend,tol,param,y)
      dydxzout = y(2) * xzero
      write(10,110)zinit,xzero,dydxzero,dydxzout,y(1),diff
110     format(' ',f12.3,f12.3,f12.3,f12.3,e12.3,e12.3)
      diffold = diff
      diff = dydxzout - dydxzero
      delta = dabs(diffold - diff)
      if (diff .lt. 1.e-13) go to 201
      if (delta .lt. 1.e-14) go to 201
      dz = 0.5 * diff * dz/(diffold - diff)
      if (i .lt. 5) dz = dz/4.
      if (dz .gt. 1.e-1) dz = 1.e-1
      if (ion .lt. 0.02) then
       if (dz .gt. 1.e-2) dz = 1.e-2
      end if
      if (ion .lt. 0.005) then
       if (dz .gt. 5.e-4) dz = 5.e-4
       if (dz .lt. -5.e-4) dz = -5.e-4
      end if
      if (mg .ne. 0.) dz = dz/5.

      if (diff .lt. 0.) then
       if (diff .gt. dabove) then
        dabove = diff
        above = zinit
       end if
      else
       if (diff .lt. dbelow) then
        dbelow = diff
        below = zinit
       end if
      end if
      diffb = above - below
      zinit = zinit + dz
      if (above .eq. 3.) go to 195
      if (below .eq. 0.) go to 195
      if (zinit .gt. above) then
       fr = ran(seed)
       zinit = above * fr + below * (1. - fr)
      end if
      if (zinit .lt. below) then
       fr = ran(seed)
       zinit = above * fr + below * (1. - fr)
      end if
195     z = zinit
      ido = 3
      call divpbs(ido,neq,fcn,z,zend,tol,param,y)
200     continue
c
c     now we have the solution characterized by
c     zinit and the y's at z=zinit
```

```
c       now generate the potential as a f(x) in
c       J/mole b.p. based on this solution
c       or, alternatively, we can calculate Stigter's
c       parameters beta and gamma to get an effective
c       charge, less than the actual charge, to generate
c       the Debye-Huckel potential in the far-field region
c
201       xinit = dexp(zinit)
        kzero = dbsk0(xzero)
        kone = dbsk1(xzero)
        beta = -1. * y(2) * kzero / (y(1) * kone)
        k0xi = dbsk0(xinit)
        const = 1.e-4/k0xi
        gamma = y(1)/(const * kzero)
        write(10,240)
        qfperlsq = 2. * 0.73/(3.4e-8 * beta * gamma * xzero * kone)
        qfperlsq = qfperlsq**2 * esq
240       format(' ')
c
c       now find optimum configurations
c       start by using a guess for r and the optimal
c       alpha corresponding to this r and the given sigma
c
c
c       scale kappa by 1.e-8 to make integration and minimization
c       routines work better (kappa now in inverse Angstroms)
c       put r in Angstroms
c
        kappa = kappa * 1.e-8

        do 1000 i=1,40
        sigma = i * 0.005
        r = 23. + 1.85/dsqrt(ion*sigma)
        r = r - mg * sigma * 1500.
        arg = 4. * pi * r * 1.e-8 * sigma/h0
        alpha = pi/2. - (datan(arg))/2.
        alpha = 180. * alpha/pi
        xguess(1) = r
        xguess(2) = alpha
        n = 2
c       xscale(1) = 1.0
c       xscale(2) = 1.0
c       fscale = 1.0
c       iparam(1) = 0
        s = 0.025
        ftol = 1.e-8
        maxfcn = 2000
        call dumpol (grtn, n, xguess, s, ftol, maxfcn,
     &  x, gvalue)
        r = x(1)
        alpha = x(2)
        etot = gvalue
        interv = 1
```

```
      errabs = 0.
      errel = 1.e-3
      call dq2agi (f, bound, interv, errabs, errel, result,
     &  errest, maxsub, neval, nsubin, alist, blist, rlist,
     &  elist, iord)
      integral = result
      eelect = qfperlsq * integral/epsilon
      eelect = eelect * 3.4 * 6.02e8
      ebend = bbend * dcosd(alpha)**4/(2. * (r * 1.e-8)**2)
      ebend = ebend * 3.4 * 6.02e8
      etwist = ((sigma/h0) - dsind(alpha) * dcosd(alpha)/(2.*pi*r*1.e-8))
      etwist = (etwist * 2. * pi)**2 * btwist/2.
      etwist = etwist * 3.4 * 6.02e8
      etotchk = eelect + etwist + ebend
      etwonly = (2.*pi*sigma/h0)**2
      etwonly = etwonly*btwist/2
      etwonly = etwonly * 3.4 * 6.02e8
      wrperl = dsind(alpha)*dcosd(alpha)/(2.*pi*r*1.e-8)
      twperl = sigma/h0 - wrperl
      wrperh0 = wrperl * h0
      twperh0 = twperl * h0
      prcntwr = wrperh0/sigma
      prcnttw = twperh0/sigma
      kr = kappa * r
      write(10,494)xguess(1),r,xguess(2),alpha
494     format(' rg ',e12.3,' r ',e12.3,' ag ',e12.3,' a ',e12.3)
      write(10,495)etwonly,etotchk,s
495     format(' etwonly = 'e12.3,' etotchk = ',e12.3,' s = ',e12.3)
      write(10,510)sigma,etot,kr
510     format(' sigma ',e12.3,' etot ',e12.3,' kr ',e12.3)
      write(10,511)eelect,etwist,ebend
511     format(' eelect ',e12.3,' etwist ',e12.3,' ebend ',e12.3)
      write(10,512)
512     format(' ')
      write(11,600)sigma
      write(12,600)r
      write(13,600)alpha
      write(14,600)etot
      write(15,600)eelect
      write(16,600)etwist
      write(17,600)ebend
      write(18,600)wrperl
      write(19,600)wrperh0
      write(20,600)twperl
      write(21,600)twperh0
      write(22,600)prcntwr
      write(23,600)prcnttw
600     format(' ',e12.4)

1000  continue
      end
```

```
      subroutine grtn (n, x, g)

      external f, dq2agi
      common alpha, kappa, r, sigma, qfperlsq, bbend
      common btwist, h0, pi
      common mg, ion, na, epsilon
      real*8 alpha, kappa, r, sigma, qfperlsq, bbend
      real*8 btwist, h0, pi
      real*8 mg, ion, na, epsilon
      integer n, interv, maxsub, neval, nsubin
      parameter (maxsub=1000)
      integer iord(maxsub)
      real*8 bound, x(2), g, errabs, errel
      real*8 f, result, errest, alist(maxsub)
      real*8 blist(maxsub), rlist(maxsub), elist(maxsub)
      real*8 integral, eelect, ebend, etwist

      interv = 1
      bound = 0.
      errabs = 0.
      errel = 1.e-3
      r = x(1)
      alpha = x(2)
      call dq2agi (f, bound, interv, errabs, errel, result,
     &  errest, maxsub, neval, nsubin, alist, blist, rlist,
     &  elist, iord)
      integral = result
      eelect = qfperlsq * integral/epsilon
      eelect = eelect * 3.4 * 6.02e8
      ebend = bbend * dcosd(alpha)**4/(2. * (r * 1.e-8)**2)
      ebend = ebend * 3.4 * 6.02e8
      etwist = ((sigma/h0) - dsind(alpha) * dcosd(alpha)/(2.e-8*pi*r))
      etwist = (etwist * 2. * pi)**2 * btwist/2.
      etwist = etwist * 3.4 * 6.02e8
      g = eelect + etwist + ebend

      return
      end

      subroutine fcn (neq, z,  y, yprime)
      integer neq
      real*8 z, y(neq), yprime(neq)
      common alpha, kappa, r, sigma, qfperlsq, bbend
      common btwist, h0, pi
      common mg, ion, na, epsilon
      real*8 alpha, kappa, r, sigma, qfperlsq, bbend
      real*8 btwist, h0, pi
      real*8 mg, ion, na, epsilon

      yprime(2) = 0.
      yprime(1) = y(2)
      if (mg .eq. 0.) go to 15
      yprime(2) = (mg/ion)*(dexp(2.*y(1))-dexp(-1.*y(1)))
```

```
15    yprime(2) = yprime(2) + (na/ion) * dsinh(y(1))
      yprime(2) = dexp(2.*z) * yprime(2)
      return
      end

      real*8 function f (s)
      real*8 s, angle, rad
      real*8 alpha, kappa, r, sigma, qfperlsq, bbend
      real*8 btwist, h0, pi, mg, ion, na, epsilon
      common alpha, kappa, r, sigma, qfperlsq, bbend
      common btwist, h0, pi
      common mg, ion, na, epsilon

      angle = s * dcosd(alpha)/r
      rad = dsqrt(2.*r**2 * (1. + dcos(angle)) + s**2 * (dsind(alpha))**2)
      f = dexp(-1. * kappa * rad)/rad

      return
      end
```

Glossary

Base-pair: a pair of nucleotide bases in DNA , e.g. A-T or G-C . The spacing between base pairs is 3.4 Å in B-DNA. Abbreviated as b.p.

B-DNA: the usual form of double-helical DNA found in solution. The double helix is right-handed with a pitch or helical repeat of 10.5 base-pairs, or 3.4 Å.

Chromatin: the complex of DNA and associated proteins which makes up chromosomes.

Cruciform: a section of DNA in which each strand of the double-helix forms base-pairs with itself rather than with the opposing strand (see Figure 3). This can only occur for sequences of DNA which contain inverted repeats.

DNA ligase: an enzyme which covalently joins the ends of double-stranded DNA molecules.

DNA recombination: a process in which one strand in a double-stranded DNA molecule is exchanged with a third strand of single-stranded DNA.

Helical repeat: the pitch of the DNA double-helix, abbreviated as h_0. Helical repeat is expressed either in terms of base-pairs (10.5 for B-DNA) or in terms of length (35.4 Å for B-DNA).

mRNA: messenger RNA, the molecule which is created during the process of DNA transcription. The sequence of the mRNA molecule created during transcription is complementary to the sequence of the transcribed DNA.

Nuclease: an enzyme which cleaves the sugar-phosphate backbone of DNA.

Nucleosome: the basic structural unit of chromatin. A complex of about 150 base-pairs of DNA wrapped twice around proteins known as histones.

Plasmid: a covalently closed circular molecule of double-stranded DNA.

Plectonemic: having an interwound shape, as shown in Figure 2.

Ribosome: a complex of proteins and nucleic acids which translates the sequence of an mRNA molecule by synthesizing a protein molecule with the corresponding sequence.

RNA polymerase: a large enzyme which translates DNA to mRNA by polymerising and mRNA molecule with a sequence complementary to the DNA molecule.

Topoisomer: a plasmid characterized by a set value of the linking number, Lk. Topoisomers are plasmids which may differ from each other only by their values of Lk.

Topoisomerase: an enzyme which changes the linking number of a plasmid or of any topologically constrained piece of DNA.

Z-DNA: a form of double-stranded DNA which has a left-handed double helix instead of the usual right-handed one. Z-DNA has been observed under conditions of superhelical stress and also in very high salt solutions.

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720