



---

**The helical repeat of underwound DNA in solution**

---

S.M.H.Richardson, T.C.Boles\* and N.R.Cozzarelli

---

Department of Molecular Biology, University of California, Berkeley, CA 94720, USA

---

Received April 7, 1988

---

**ABSTRACT**

Closed circular DNA was relaxed with a topoisomerase in the presence of varying concentrations of the intercalating dye, ethidium bromide, to create underwound, planar DNA rings. We directly determined the helical repeat of these DNA molecules by the Gaussian center method and found that it varied as a simple predicted function of the degree of underwinding and the helical repeat of relaxed, dye-free DNA. We discuss these results in light of a recent mathematical treatment of DNA structure which predicts that the helical repeat of supercoiled DNA molecules in solution obeys the same function.

**INTRODUCTION**

The helical repeat of DNA is the number of base pairs per helical turn of the duplex. It is a fundamental structural property of DNA that is tightly correlated with the geometries of the backbone and bases and also determines the phasing relationship between *cis*-interacting sequences. In solution, the generally accepted value for helical repeat is around 10.5, as obtained from a variety of methods (1-5). This value is known to vary significantly as a function of base sequence (6, 7) and ambient conditions (8). A recent mathematical treatment of DNA structure that we shall call surface linking theory predicts that the helical repeat also varies with DNA supercoiling and in a surprisingly simple way (9). This is important because DNA in its natural state is negatively supercoiled, and many DNA binding proteins even require supercoiled substrates for activity.

This paper represents a first step in testing the proposal that the helical repeat changes with supercoiling. Negatively supercoiled DNA is underwound; that is, its linking number is less than that of relaxed circular DNA of the same length. We have experimentally determined the helical repeat for one special class of underwound DNA. Closed circular DNA was relaxed with topoisomerase I in the presence of ethidium bromide to create underwound molecules in which the axis of the DNA duplex is nearly planar, or writhelless. This is equivalent to adding to negatively supercoiled molecules a sufficient amount of intercalator to remove all writhe.

We have used the Gaussian center method introduced by Wang (1) to measure the helical repeat of such DNA. The method, described in detail below, depends on the change in the equilibrium distribution of topoisomer linking number about the average due to insertion of a small

piece of DNA. In our experiments, the presence of ethidium during relaxation maintains underwinding at a chosen level during repeated breakage and religation events by a topoisomerase. Since the data consist of intensities of topoisomer bands which are fixed when the relaxation reaction reaches equilibrium, length corrections and effects on electrophoretic mobility due to supercoiling or intercalator are immaterial.

Another feature of the approach is that we can predict quantitatively the change in helical repeat as a function of underwinding. For the special case where writhe is on average zero, both surface linking theory (9) and classical linking theory (10) predict the same equation, so we have confidence in the validity of the test.

Our results show that the helical repeat of such underwound DNA varies precisely as predicted and depends only on the degree of underwinding and the helical repeat of DNA relaxed without ethidium. According to the surface linking treatment of supercoiling, the helical repeat of supercoiled DNA in solution varies in exactly the same way, irrespective of writhe.

#### Surface linking theory and the helical repeat

We present first the relationship between helical repeat and underwinding as developed in the surface linking theory of White *et al.* (9). Central to the description of the helical repeat,  $h$ , of closed circular DNA in solution is the concept that the path of the duplex axis,  $A$ , lies on a virtual surface, just as the DNA of the nucleosome core particle lies on the actual surface of the histone octamer (11). The helical repeat can be defined rigorously with reference to this surface using two vectors,  $\mathbf{n}$  and  $\mathbf{s}$ , that start from the axis and lie in a plane.  $\mathbf{n}$ , the surface normal, always extends perpendicularly from the surface as it is moved along the axis.  $\mathbf{s}$  extends to one strand of the duplex; it rotates through  $\mathbf{n}$  as the plane containing them is moved along  $A$ . Because the DNA is closed circular,  $\mathbf{s}$  will have rotated an integral number of times through  $\mathbf{n}$  on its return to the starting point; this integer is the winding number,  $\Phi$ . The average helical repeat of DNA,  $h$ , is defined as the number of base pairs in the molecule,  $N$ , divided by  $\Phi$ . This rotating vector definition of  $h$  is equivalent to the biochemical  $h$  deduced from the periodicity of DNase I cleavage of DNA bound to surfaces (3).

The linking number between the two strands of the DNA duplex,  $Lk$ , is the sum of two integers,  $\Phi$  and  $SLk$ , the surface linking number.  $SLk$  describes the component of  $Lk$  determined by the shape of the virtual surface on which the DNA lies. Thus,

$$Lk = SLk + \Phi \quad (1)$$

$SLk$  can be defined as the linking number between the duplex axis  $A$  and a curve  $A_\epsilon$  displaced from  $A$  by a constant distance,  $\epsilon$ , along the normal to the surface,  $\mathbf{n}$ . Fully relaxed circular DNA completely without writhe lies on average in a plane. Therefore,  $SLk$  is zero, because  $A_\epsilon$  and  $A$  lie in parallel planes and are unlinked. We designate by a subscript zero the parameters for such a DNA relaxed in the absence of intercalator. Thus,

$$Lk_0 = \Phi_0 \quad (2)$$

Because negatively supercoiled DNA is underwound, its Lk is less than Lk<sub>0</sub>. The linking difference, ΔLk, is the expression of underwinding:

$$\Delta Lk = Lk - Lk_0 \quad (3)$$

Similarly, the specific linking difference, σ, is a commonly used size-independent descriptor of underwinding:

$$\sigma = \frac{\Delta Lk}{Lk_0} \quad (4)$$

For DNA relaxed in the presence of ethidium, SLk is zero, just as for DNA relaxed in the absence of dye, and thus Lk = Φ in this case. Substituting this relation and Eq. 2 into Eq. 4 gives:

$$\sigma = \frac{\Phi - \Phi_0}{\Phi_0} \quad (5)$$

Rewriting,

$$\sigma = \frac{\frac{N}{h}}{\frac{N}{h_0}} - 1 \quad (6)$$

Solving for h,

$$h = \frac{h_0}{\sigma + 1} \quad (7)$$

Thus, for molecules whose SLk is zero, helical repeat depends only on h<sub>0</sub> and σ.

Classical linking theory (10), which divides Lk into writhe and twist, results in the same equation if DNA relaxed with or without intercalator is assumed to be writhless. For this case, Φ is equal to twist, which in turn is equal to Lk and thus Eq. 7 can be derived similarly.

A rearrangement of Eq. 7 which gives additional insight shows that the ratio of helical repeats,  $\frac{h_0}{h}$ , is equal to 1 + σ; this is the ratio of Lk for dye-bound and dye-free relaxed states.

## MATERIALS AND METHODS

### DNA

All plasmids are deletion derivatives of pUC18 and are designated pUC18-N, where N is the number of base pairs removed. Plasmids pUC18-2 and pUC18-13 were constructed by Acc I cleavage, limited S1 nuclease digestion and ligation; pUC18-28 by EcoRI cleavage, S1 treatment and ligation; and pUC18-47 by EcoRI and HindIII cleavage, DNA polymerase I (Klenow fragment) filling of the recessed 3' ends and ligation using standard methods (12). Constructions were transformed into *Escherichia coli* strain DH5α (BRL) and DNA was prepared by an alkaline lysis procedure (12). Deletion lengths and sequences were determined by chemical cleavage (13) or dideoxy chain termination nucleotide sequencing (14).

### Relaxation reactions

Pairs of plasmids (50 μg/ml) were incubated for 2 hours at 20°C with excess wheat germ topoisomerase I (15) in 50 mM Tris-HCl, pH 7.6, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 50 μg/ml bovine

serum albumin and 0, 2.9, or 5.8  $\mu\text{g/ml}$  ethidium bromide (Sigma). Reactions were stopped by addition of sodium dodecyl sulphate to 1% and phenol extraction.

#### Gel electrophoresis and quantitation

After precipitation with ethanol, DNA was resuspended in electrophoresis buffer—90 mM Tris-borate, 2.5 mM EDTA, pH 8.3 (TBE) containing 0, 1.2, or 7  $\mu\text{g/ml}$  chloroquine phosphate (Sigma). DNA samples relaxed in the absence of ethidium bromide were applied to 0.5% agarose/2% polyacrylamide TBE composite vertical gels to resolve the topoisomer distributions. These gels (14 x 16 x 0.3 cm) were run at room temperature at 90 volts for 21 hours. All other DNA samples were electrophoresed in 0.8% agarose horizontal TBE gels (13 x 27 x 0.5 cm) with chloroquine in gel and buffer to resolve the more supercoiled distributions. Electrophoresis was performed at 90 volts for 24 hours at room temperature with recirculation of the buffer. After electrophoresis, gels were irradiated with short wavelength ultraviolet light (UV) for 2 minutes to nick the closed circular DNA prior to staining for 30 minutes in 0.5  $\mu\text{g/ml}$  ethidium bromide and destaining for 1 hour in distilled water. Gels were photographed under short wavelength UV illumination and the photographic negatives (Polaroid type 665) were scanned with a Zeineh soft laser densitometer. The linear range of film response was determined from densitometry of coelectrophoresed titrations of linearized pUC18, and intensity data were collected only within this range. Intensities were taken from peak heights; integration of the traced areas gave identical results. The position, in units of Lk, of the Gaussian center of each topoisomer distribution was calculated from the fluorescence intensities of the three most abundant topoisomers using the algorithm of Kolb and Buc (16). The specific linking difference,  $\sigma$ , was determined by the band-counting method (17), using the Gaussian centers of each distribution to give the mean linking differences with respect to the centers of the distributions relaxed in the absence of ethidium bromide, Lk<sub>0</sub>.

#### The Gaussian center method

The principle of the method as developed by Wang (1) is outlined in Fig. 1. Experimentally, pairs of plasmids differing by 26, 34, or 45 base pairs were incubated in the same topoisomerase I relaxation mixture. The Gaussian center (G) and its shift ( $\alpha$ ) in units of Lk from the position of the most abundant topoisomer in each distribution were determined. The value of  $\alpha$  for the shorter plasmid is subtracted from  $\alpha$  for the larger plasmid to find the change in  $\alpha$  caused by the insert. This value is a measure of the rotation,  $\omega$  (in units of fractional turns), required to close the ends after the addition of the insert. Calculation of  $h$  for the insert depends on the positions of the ends prior to ring closure because this is what is changed by addition of B base pairs. This is simply  $1 - \omega$ , which is equal to the non-integral residual of the insert size divided by the helical repeat,  $(\frac{B}{h})_{\text{mod}1}$ . The number of integral turns in the insert is unambiguous for the short insert sizes we have used; e.g., 2 for the 26 base pair insert.

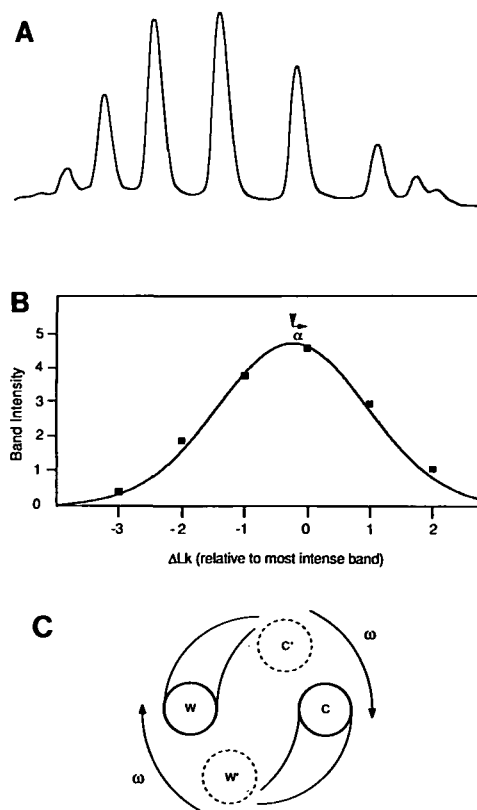


Figure 1. Explanation of the Gaussian center method. A. Densitometer tracing of a topoisomer distribution of relaxed pUC18-13. Each topoisomer differs from its neighbor by one unit of Lk. B. Measurements of the amounts of DNA in each band in part A were quantitated and plotted as a function of Lk. The Gaussian curve was calculated from the data by the method of least squares. The arrowhead marks the Gaussian center and represents the mean linking number of the population. In general, the center of the Gaussian distribution will not coincide with the most abundant topoisomer. This arises because Lk can only assume integral values in closed circular DNA, but the plasmid size need not be an integral multiple of  $h$ . The shift ( $\alpha$ ) in units of Lk between the Gaussian center and the most abundant topoisomer is a direct reflection of the smallest angle of rotation,  $\omega$  (in units of fractional turns), between the most probable, but unproductive, backbone geometry and a productive one. C. Schematic illustrating the effect of small insertions on  $\omega$ . Shown are the backbones of two linear duplex ends, viewed down the helix axis, about to be joined in a ring closure reaction in the plane of the page. The closure of a double-strand break rather than a nick is illustrated for clarity. Because the length of the linear DNA is not an exact multiple of  $h$ , the end coming down from above the page (backbones W', C') must be rotated relative to the end receding from the page (W, C) by the angle  $\omega$  to achieve proper alignment. Addition of base pairs at W' and C' changes the rotation necessary for alignment.  $\alpha$  is a direct measure of  $\omega$  in units of fractional turns. Insertion of DNA lengths that are integral multiples of  $h$  will have no effect on  $\omega$  or  $\alpha$ . For insertion of non-integral multiples of  $h$ ,  $\omega$  and  $\alpha$  will vary as the non-integral residual of the insert length divided by the helical repeat.

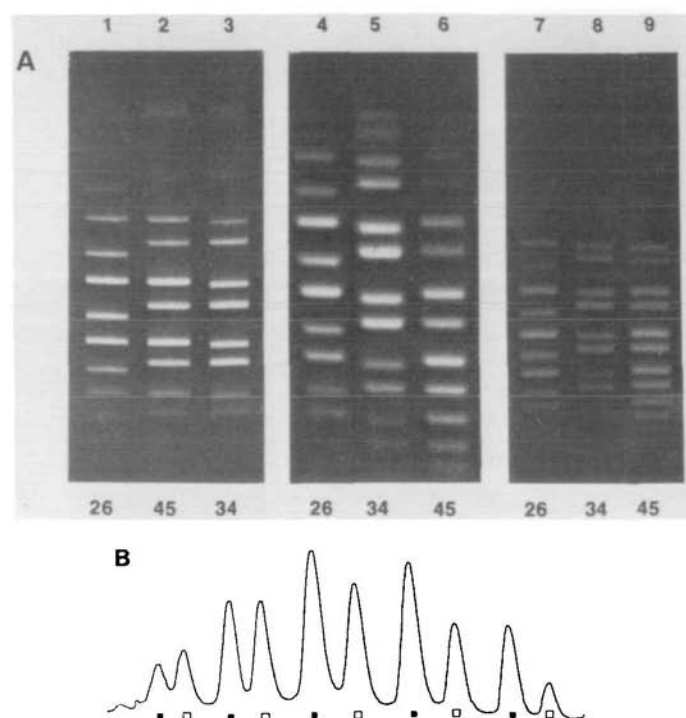


Figure 2. A. Mixed topoisomer distributions resolved by gel electrophoresis. Pairs of plasmids differing in length by insert sizes of 26, 34, and 45 base pairs were relaxed in the same reaction mixture as described in Materials and Methods. The panel shows a composite of three gels containing 0 (lanes 1-3), 1.2 (lanes 4-6), or 7 (lanes 7-9)  $\mu\text{g/ml}$  chloroquine, resolving paired distributions around  $\sigma = 0$ ,  $-0.03$  and  $-0.06$ , respectively. Under these conditions,  $\Delta Lk$  is increasing in the direction of electrophoresis (top to bottom). B. Laser densitometer scan of lane 4 of part A showing pUC18-2 (closed symbols) and pUC18-28 (open symbols) at  $\sigma = -0.03$ ; these plasmids differ by 26 base pairs. The direction of electrophoresis is from right (less positive topoisomers) to left (more positive).  $\alpha$  is  $+0.280$  for the pUC18-2 distribution and  $-0.153$  for the pUC18-28 distribution. Subtracting  $\alpha$  of the parental plasmid from the insert-containing one gives  $\omega = -0.153 - (+0.288) = -0.441$ . By convention (1), negative values of  $\omega$  are converted to positive values to reflect the equivalent right-handed rotation and  $\omega(+)=1 + -0.441 = 0.559$ . The nonintegral residual of insert size (B) divided by  $h, (\frac{B}{h})_{\text{mod}1}$ , is  $1 - \omega$  or  $+0.441$  which corresponds to a helical repeat of  $\frac{26}{2.441} = 10.65$ .

## RESULTS

The DNA sequences chosen for this study are all parts of the polylinker multiple cloning site of pUC18. Sequenced deletions were made by single or multiple restriction, filling or trimming, and religation. In pairwise comparisons the larger plasmid was considered to be an

Table 1. Summary of helical repeat measurements

Insert size (base pairs)	Specific linking difference ( $\sigma$ )*	Helical repeat measured (base pairs/turn)	Standard error of the mean	Number of measurements	Helical repeat predicted** (base pairs/turn)
26	0.0	10.48	0.020	4	10.5
34	0.0	10.44	0.015	4	10.5
45	0.0	10.55	0.016	6	10.5
26	-0.034	10.84	0.075	13	10.87
34	-0.035	10.92	0.040	14	10.88
45	-0.031	10.95	0.033	16	10.84
26	-0.061	11.07	0.104	14	11.18
34	-0.063	11.26	0.118	11	11.2
45	-0.057	11.15	0.032	12	11.13

\* Specific linking difference was measured by band counting method (17).

\*\* Values of  $h$  predicted from the relationship  $h = \frac{h_0}{1 + \sigma}$ .

insertion derivative of the smaller plasmid. Although in the multiple cloning site there are short range nucleotide sequence symmetries, no obvious long-range symmetries exist in the sequences examined. We therefore consider the inserts to be representative of random sequence DNA.

Fig. 2 shows typical mixed topoisomer distributions resolved by gel electrophoresis for the three insert lengths studied, at three different levels of underwinding. The insert lengths were chosen so that the bands in the overlapping topoisomer distributions interdigitated evenly over the range of  $\Delta Lk$  tested; in most cases fluorescence intensity fell to the background level between adjacent bands. A densitometer scan of one lane is shown in Fig. 2B.

The results of these measurements are summarized in Table 1 and presented graphically in Figs. 3A and 3B. The measured value for the helical repeat of relaxed DNA of  $10.49 \pm 0.06$  (S.D.) base pairs per turn is in excellent agreement with those previously obtained using the same method (1) and well within the variation of other methods (2-5). As the ethidium bromide concentration increased, the measured helical repeat also increased for all the insert sizes examined. Experimental scatter also increased with unwinding; this is probably due to the broadening of topoisomer distributions with increasing concentrations of ethidium bromide. Since the intensity differences between adjacent bands are smaller, slight errors in measurement of relative intensities are magnified. Multiple measurements were made to minimize this effect.

To compare the data with theory,  $\frac{1}{h}$  is plotted in Fig. 3B because this is linearly related to  $\sigma$ , whereas the plot of  $h$  vs.  $\sigma$  in Fig. 3A describes a barely discernible hyperbola. The formula for the theoretical line is  $\frac{1}{h} = 0.0952\sigma - 0.0952$ . A least squares regression fit to the experimentally measured values gives  $\frac{1}{h} = 0.0951\sigma - 0.0952$  ( $R = 0.97$ ) which almost exactly superimposes on the theoretical line.



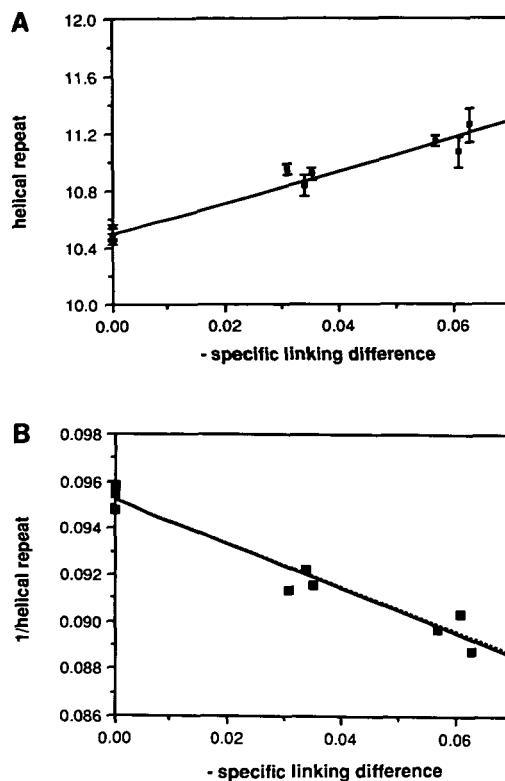


Figure 3. A. Helical repeat vs. specific linking difference,  $\sigma$ . The predicted variation of  $h$  with  $-\sigma$  calculated from the relationship  $h = \frac{h_0}{\sigma + 1}$  is plotted as a solid line following a shallow hyperbola. Experimentally obtained mean values of  $h$  are plotted  $\pm$  standard error of the mean. B. Reciprocal of helical repeat vs. specific linking difference. This shows the linear relationship between  $\frac{1}{h}$  and  $-\sigma$ . The formula for the line predicted from Eq. 7 (solid line) is  $\frac{1}{h} = -0.0952(-\sigma) + 0.0952$ . A least squares regression fit (dashed line) to the experimental values of  $\frac{1}{h}$  (squares) almost superimposes on the theoretical line; in this case,  $\frac{1}{h} = -0.0951(-\sigma) + .0952$  ( $R = 0.97$ ). The slope indicates that  $h$  increases by about 0.11 of a base pair per turn for every  $-0.01$  increment of  $\sigma$  over the range observed.

### DISCUSSION

We have demonstrated that the simple predicted relationship between  $h$ ,  $h_0$ , and  $\sigma$  in Eq. 7 is valid for underwound DNA with minimum writhe by independently measuring all three parameters. We have applied the Gaussian center method (1) and shown that this method is a

convenient way of measuring the change in helical repeat induced by a DNA-binding ligand, in this case *ethidium bromide*. This method can be applied to other ligands, providing that they do not induce writhe in the DNA sequence to which they bind. The Gaussian center method measures  $h$  only for the inserted DNA. In our experiments, the insert varied in size from 26 to 45 base pairs, well within the range of chemical synthesis. Moreover, the average number of intercalator molecules bound to this small region was at most 1 per 12 base pairs. Thus, the method is an extremely sensitive one.

The Gaussian center method is particularly well-suited to this application because the analysis of the DNA after relaxation may be done in any gel electrophoresis system. The information about the helical repeat is held in the "molecular memory" of the topoisomer distribution about its Gaussian center. Since the data are intensities, rather than electrophoretic mobilities, length corrections for different-sized molecules are unnecessary, and mobility anomalies due to ligands are irrelevant.

#### Implications for supercoiling

The Gaussian center method measures the helical repeat for the topoisomerase-relaxed DNA still bound with ethidium bromide, where  $SLk$  is zero. On removal of the dye, the writhe absorbed by the intercalator is released and the molecule deforms to the shape of negatively supercoiled DNA. Surface linking theory states that during this process the helical repeat remains unchanged because  $SLk$  remains zero (9). The rationale is as follows. The solution structure of supercoiled DNA is plectonemic, i.e., intertwined (18-20), and may therefore be modelled with the duplex axis winding helically on the surface of an imaginary cylinder which has hemispherical caps.  $A_E$  will be wound similarly except the surface will differ in size. The important point is that  $A$  and  $A_E$  will never wind around each other, so  $SLk$  is zero. Within limits, the exact shape of the virtual surface is not important because  $SLk$  is zero for all such sphere-like surfaces (9).

This extrapolation of our data to supercoiled DNA in solution depends on the validity of the surface linking theory. Support for the extrapolation comes from two other experiments. Analysis of multiply interlinked DNA catenanes produced *in vitro* by phage  $\lambda$  integrase shows that catenation induces supercoiling in the component rings (21). The surface linking theory predicts that  $h$  will increase with  $\sigma$  for this right-handed solenoidal supercoiling, and this was the result obtained. The sensitivity of  $h$  to supercoiling is also illustrated in the left-handed solenoidal winding of DNA around histones in the viral minichromosome of SV40 (22). Because the geometry of the nucleosome core particle is known, the value of  $h$  can be calculated from the theory using  $h_0$  and  $\sigma$ . The predicted  $h$  was 10.1 base pairs per turn (9), close to the experimental value obtained (23).

Our results demonstrate that when writhe is minimal,  $h$  varies with  $\sigma$  and  $h_0$  exactly as predicted. If future work demonstrates that the same relationship holds at another value of writhe then the surface linking theory will have strong experimental support.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health. T.C.B. was supported by an NIH postdoctoral fellowship.

\*Present address: British Biotechnology Ltd., Brook House, Watlington Road, Cowley, Oxford OX4 5LY, UK

REFERENCES

1. Wang, J.C. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 29-34.
2. Wang, J.C. (1979) *Proc. Nat. Acad. Sci. USA* **76**, 200-203.
3. Rhodes, D. and Klug, A. (1980) *Nature (London)* **286**, 573-578.
4. Shore, D. and Baldwin, R.L. (1983) *J. Mol. Biol.* **170**, 983-1007.
5. Goulet, I., Zivanovic, Y. and Prunell, A. (1987) *Nucleic Acids Res.* **15**, 2803-2821.
6. Peck, L.J. and Wang, J.C. (1981) *Nature (London)* **292**, 375-378.
7. Rhodes, D. and Klug, A. (1981) *Nature (London)* **292**, 378-380.
8. Bauer, W.R. (1978) *Ann. Rev. Biophys. Bioeng.* **7**, 287-313.
9. White, J.H., Cozzarelli, N.R. and Bauer, W.R., submitted.
10. White, J. (1969) *Am. J. Math.* **91**, 693.
11. Richmond, T.J., Finch, J.T., Rushton, B., Rhodes, D. and Klug, A. (1984) *Nature (London)* **311**, 532-537.
12. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
13. Maxam, A. and Gilbert, W. (1977) *Proc. Nat. Acad. Sci. USA* **74**, 560-564.
14. Messing, J. (1983) *Methods Enzymol.* **101**, 20-89.
15. Dynan, W.S., Jendrisak, J.J., Hager, D.A., and Burgess, R.R. (1981) *J. Biol. Chem.* **256**, 5860-5865.
16. Kolb, A. and Buc, H. (1982) *Nucleic Acids Res.* **10**, 473-485.
17. Keller, W. (1975) *Proc. Nat. Acad. Sci. USA* **72**, 4876-4880.
18. Vinograd, J., Lebowitz, J., Radloff, R., Watson, R. and Laipis, P. (1965) *Proc. Nat. Acad. Sci. USA* **53**, 1104-1111.
19. Mizuuchi, K., Gellert, M., Weisberg, R.A. and Nash, H.A. (1980) *J. Mol. Biol.* **141**, 485-494.
20. Spengler, S.J., Stasiak, A. and Cozzarelli, N.R. (1985) *Cell* **42**, 325-334.
21. Wasserman, S.A., Solomon, M.J., Varshavsky, A., White, J.H. and Cozzarelli, N.R. submitted.
22. Germond, J.E., Hirt, B., Oudet, P., Gross-Bellard, M. and Chambon, P. (1975) *Proc. Nat. Acad. Sci. USA* **72**, 1843-1847.
23. Drew, H.R. and Calladine, C.R. (1987) *J. Mol. Biol.* **195**, 143-173.