Proc. Natl. Acad. Sci. USA
Vol. 76, No. 1, pp. 200–203, January 1979
Biochemistry

Helical repeat of DNA in solution

(DNA structure/linking number/DNA topoisomerase/closed circular DNA)

JAMES C. WANG

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Communicated by Francis Crick, October 23, 1978

ABSTRACT **The helical repeat of DNA in solution has been measured directly by analyzing the gel electrophoretic patterns of pairs of covalently closed DNAs with length differences between 1 and 58 base pairs, out of a total length of about 4350 base pairs per DNA molecule. The method is based on the observation that for a covalently closed DNA of a fixed size of n base pairs (n of the order of several thousand), under appropriate conditions, two topological isomers (topoisomers) differing by 1 in their linking numbers are well resolved by gel electrophoresis. If the size of the DNA is increased to $n + x$ base pairs, unless x is an integral multiple of the helical repeat h, the bands of the topoisomers with $n + x$ base pairs per molecule are all shifted relative to the bands of the topoisomers with n base pairs per molecule. The magnitude of the shift is directly related to the nonintegral residual of x/n. Analysis of the set with x ranging from 1 to 58 gives the DNA helix repeat in solution as 10.4 base pairs per turn under physiological conditions, with an estimated probable error of ±0.1. This result strongly supports the double helix structure of DNA and rejects the side-by-side model of Rodley *et al.* [Rodley, G. A., Scobie, R. S., Bates, R. H. T. & Lewitt, R. M. (1976) *Proc. Natl. Acad. Sci. USA 73*, 2959–2963]. The helical repeat of DNA measured in solution is significantly different from the value 10.0 base pairs per turn for the B form fiber structure.**

It has been a quarter of a century since the bihelical structure of DNA was first proposed (1). Two recent developments prompted me to reexamine the helical repeat—i.e., the number of base pairs per helical turn—of the DNA double helix in solution. First, a model radically different from the Watson–Crick structure has been proposed by Rodley *et al.* (2). In this model, two antiparallel strands held together by Watson–Crick type base pairing assume a side-by-side meshing rather than a bihelical intertwining. It is asserted that the model is also consistent with the x-ray diffraction pattern of B form DNA (2). Second, values for the helical repeat of DNA—9.3, 10, and 11 base pairs per turn for the C, B, and A structures, respectively—were measured for fibers in different states of hydration (3). The B structure observed for fibers at high humidity is generally taken to be the structure in solution. There has been no reliable measurement of the DNA helical repeat in solution. Small angle x-ray scattering and circular dichroism measurements of DNA in solution have been carried out (4, 5), but the interpretations of these results are sufficiently complicated and the data available do not permit an unambiguous conclusion.

Recent studies on chromatin structure also have brought into focus the lack of reliable measurement of the DNA helical repeat in solution. According to the structural model of Finch *et al.* (6), deduced from x-ray diffraction and electron microscopic analyses of crystals of nucleosome cores, the DNA in each nucleosome is wound into about 1¾ or more turns of a superhelix. The model requires that, if a closed circular DNA duplex consisting of a certain number of nucleosomes is relaxed by cycles of breaking and rejoining of the DNA backbone bonds, the linking number of the resulting DNA should be lower than that of the same DNA relaxed in the absence of the histones, by 1¾ turns or more per nucleosome. The experimentally observed difference was less than 1¼ turns per nucleosome. This discrepancy can be reconciled if one realizes that the helical repeat of DNA free in solution and around the nucleosome core might not be the same (7). Finch *et al.* argued that, if the helical repeat of DNA around the nucleosome is exactly 10 base pairs per turn as suggested by pancreatic DNase I digestion experiments with nucleosomes, then the helical repeat of free DNA in solution might be in the range 10.4 to 10.7 base pairs per turn. Theoretical calculations by Levitt (8) indicated that the helical repeat of a straight DNA segment is likely to be different from that of the same segment bent smoothly around a cylindrical core.

In a preliminary communication, measurements of the helical repeat of DNA in solution by the Gaussian center method have been presented (9). The method is based on the earlier work of Depew and Wang (10) showing that, if a nicked circular DNA is converted to the covalently closed form by ligase, the resulting topological isomers (topoisomers) differing only in their linking numbers (topological winding numbers) can be resolved by gel electrophoresis. The relative amounts of the topoisomers, when plotted as a function of their linking numbers, give a Gaussian distribution. In general the center of the Gaussian curve does not occur at an integral linking number and is separated from the linking number α_m of the most abundant topoisomer by a fractional number, ω. Depew and Wang (10) interpreted this separation in the following way. For a circular DNA with a single-chain scission, the most stable configuration of the DNA chain is such that, for the formation of the most stable closed circular species, which has an integral linking number α_m, the twist number Tw is altered by a certain fraction of a turn, ω. It immediately follows that, if a segment x base pairs long is inserted into the original DNA, the separation between the Gaussian center and the most intense band will not be affected if x is an integral multiple of the helical repeat h. If x is not an integral multiple of h, the nonintegral residual of x/h will modulate the position of the Gaussian center relative to the most intense topoisomer band.

Three pairs of DNAs have been examined by this method, with length differences of 99, 114, and 414 base pairs, respectively (9). Because the method measures only the nonintegral residual of x/h, to calculate h from these measurements it was necessary to have an approximate value of h. By assuming that h lies between 10 and 11, the data obtained gave $h = 10.4$ base pairs per turn, with an estimated probable error of ±0.1.

In the present communication, results obtained by a second method, termed "the band shift method," are presented. The electrophoretic patterns of a set of relaxed covalently closed

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

DNAs in agarose gel were examined. The set of DNAs can be considered as constructed from inserting segments ranging from 1 to 58 base pairs into a starting DNA.* The DNA helical repeat can be deduced directly from the set of data, without any prior assumption as to the value of h. It is concluded that, for a DNA with a sequence represented by the sequences of the inserts used in this work, its helical repeat in solution is 10.4 ± 0.1 base pairs per turn under physiological conditions. The band shift method as well as the Gaussian center method can also be applied to the determination of the sequence dependence of h.

PRINCIPLE OF THE METHOD

If a covalently closed circular DNA duplex n base pairs long is relaxed by breaking and rejoining of backbone bonds, a family of topoisomers differing only in their linking numbers is obtained. These topoisomers can be resolved into discrete bands by gel electrophoresis under appropriate conditions. The method to be described as well as the Gaussian center method is based on the postulates that these bands are indeed species differing only in linking numbers and that the difference between the linking numbers of adjacent bands is unity. Arguments supporting these postulates will be presented elsewhere (F. H. C. Crick, J. C. Wang, and W. Bauer, unpublished data).

The band shift method can be deduced from a consideration of the dependence of the electrophoretic pattern of the covalently closed topoisomers on the length of the DNA. Suppose, by the now well-developed methods of molecular genetics, a short segment x base pairs long is inserted into the original DNA, increasing its length to $n + x$ base pairs. For the time being it will be assumed that x is sufficiently small, compared with n, so that the length increment *per se* has no effect on the electrophoretic mobility. It is then straightforward to arrive at the following conclusions:

(*i*) If x is an integral multiple of the helical repeat h, the electrophoretic pattern is unchanged.

(*ii*) If x is not an integral multiple of the helical repeat h, each of the covalently closed DNA bands in the electrophoretogram will be shifted by an amount equal to the nonintegral residual of x/h times the spacing between two adjacent bands.

The first conclusion above is easy to understand because, if x is an integral multiple of the helical repeat, the twist number Tw and the linking number α will be increased by the same integral number, and therefore the writhing number Wr, which determines the electrophoretic mobility and is related to α and Tw by the equation $\alpha = Tw + Wr$ (11–14), remains the same.† In arriving at the second conclusion, it is assumed that, under electrophoresis conditions, because Wr of all topoisomers is relatively low, Tw is not significantly different from that of a relaxed DNA under the same conditions. Thus, the insertion of x base pairs is expected to increase Tw by x/h. Because α can only be changed by an integral value, if Tw is increased by a nonintegral value, Wr must also be changed by a nonintegral value because of the topological constraint.

For most of the measurements to be described in this work, the conditions chosen for the relaxation of the DNAs and for gel electrophoresis were such that during electrophoresis the covalently closed circular DNAs were all positively supercoiled—i.e., with positive Wr (10). Thus, for a topoisomer with a fixed linking number, an increase in Tw decreases Wr during electrophoresis, and an upward shift (a reduction in mobility) is expected. Depew and Wang (10) pointed out that, if the helical repeat is 10 base pairs per turn, the addition of 1 base pair, which increases Tw by 0.1 turn, will cause an upward shift of 0.1 times the spacing between adjacent topoisomers, or 0.1 turn.

MATERIALS AND METHODS

Materials. The plasmid DNAs used—pTR161, 182, 183, 188, 190, 193, and 199—were obtained from strains kindly provided by T. Roberts. The plasmids were constructed by inserting a fragment of phage λ DNA containing the *cro* gene into a plasmid derived from pBR322 (T. M. Roberts, R. Kacich, and M. Ptashne, personal communication). One end of the inserted λ fragment had been resected by digesting with *Escherichia coli* exonuclease III and *Aspergillus orizae* endonuclease S1 prior to insertion, and therefore the set of plasmids differed in lengths within a well-defined sequence. For each of the plasmid DNAs used here with the exception of pTR193, a *Hae* III restriction fragment containing the region of sequence variation was sequenced. For pTR193, sequencing of the fragment was not carried out but the fragment was sized on a sequencing gel. For the purpose of this paper, all DNAs can be considered as derived from pTR161 by deleting, from the left end, part or all of the 58-base-pair sequence 5′-G-A-T-C-C-G-G-G-A-C-T-A-T-T-T-T-A-C-C-T-A-T-G-G-C-G-G-T-G-A-T-A-A-T-G-G-T-T-G-C-A-T-G-T-A-C-T-A-A-G-G-A-G-G-T-T-G-T-3′. The numbers of base pairs deleted for the set were: pTR161, 0; pTR182, 53; pTR183, 26; pTR188, 35; pTR190, 58; pTR193, 47; pTR199, 25. The enzyme calf thymus DNA topoisomerase used for the relaxation of the DNAs was the preparation of K. Javaherian and L. Liu of this laboratory. Relaxation of the DNAs was carried out in 10 mM Tris·HCl, pH 8/0.2 M NaCl/0.1 mM Na_3EDTA at 0°C overnight. The reaction was terminated by the addition of buffer-saturated phenol to the reaction mixtures at 0°C. The phenol-extracted DNA samples were then dialyzed against the gel electrophoresis buffer to remove excess salt.

Gel Electrophoresis. The procedures described by Depew and Wang (10) were used. The 0.7% agarose slab was 13 cm wide × 22 cm long × 0.3 cm thick, and electrophoresis was usually done at 60 V for 16 hr. After electrophoresis the gel slab was stained with ethidium bromide and photographed on Polaroid type 55 film, and the negative of the film was traced with either a Joyce-Loebl or a Helena Laboratories microdensitometer.

RESULTS

Fig. 1 *top* illustrates the gel electrophoretic patterns of several DNA samples that had been relaxed under identical conditions. Comparison of the patterns shown in lanes 1 and 2 clearly indicates that the addition of 10 base pairs did not change the gel electrophoretic pattern. The same conclusion can be drawn more precisely by comparing the patterns of lanes 3 and 4. The shifts of the bands of pTR199 relative to the bands of pTR161 can be measured from a microdensitometer tracing of the negative of the photograph, and the shifts can be compared with those of pTR188 relative to the bands of pTR161. The patterns shown in the center lanes show that the addition of 25, 35, and 26 base pairs shifted the bands upward by 0.5, 0.5, and 0.6 turn, respectively. The effect of the addition of a single base pair can also be seen by comparing the patterns of lanes 4 and 5. By using the bands of pTR161 in these lanes as the common references, it is clear that the bands of pTR188 shifted upward from the corresponding bands of pTR183 by 0.1 turn.

Fig. 1 *middle* depicts the electrophoretic patterns of four additional pairs of DNAs, with length differences of 53, 27, 5, and 6, respectively. A microdensitometer tracing of one of the

* The actual construction of the set involved deletion of sequences from a DNA rather than insertion of sequences into a DNA.

† For an introduction to the definitions of α, Tw, and Wr and the relationship among them, see ref. 14.

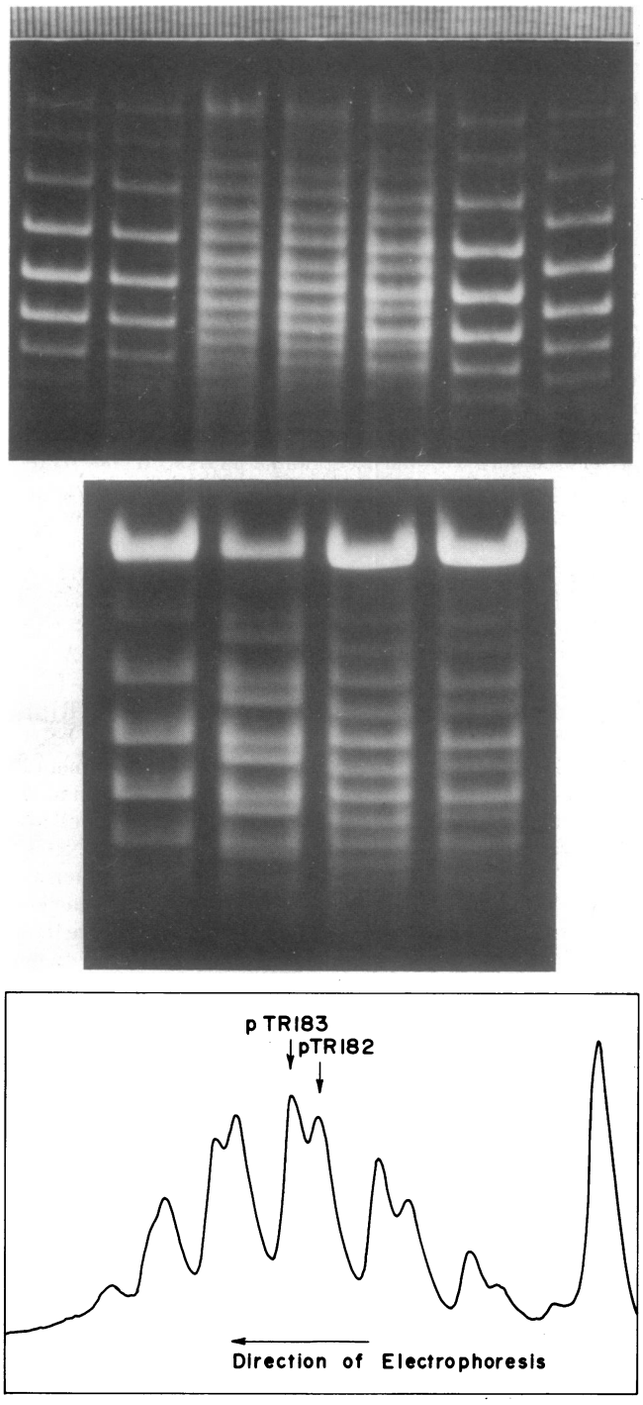

FIG. 1. (*Top*) Electrophoretic patterns of DNAs. From left to right: pTR188, pTR199, a mixture of pTR161 and pTR199, a mixture of pTR161 and pTR188, a mixture of pTR161 and pTR183, pTR183, and pTR161. The length increments of the DNAs pTR188, pTR199, and pTR183 over pTR161 are, in base pairs, −35, −25, and −26, respectively. Only a section of the gel was photographed. The scale on top is in millimeter units. (*Middle*) Electrophoretic patterns of pairs of DNAs. All samples contained DNA pTR182. The second DNA mixed in, from left to right, was pTR161, pTR183, pTR190, and pTR193. (*Bottom*) Microdensitometer tracing of the negative for the electrophoretic pattern of the pair of DNAs pTR183/pTR182.

patterns, that of the DNA pair pTR183/pTR182 with a length difference of 27 base pairs, is illustrated in Fig. 1 *bottom*. Results obtained for pairs of the DNA samples with length differences up to 10 base pairs are shown in Table 1.

The data in Table 1 show that very close to 10 base pairs make one DNA helix turn under electrophoresis conditions (in 40 mM Tris·HCl, pH 8/5 mM Na acetate/0.5 mM Na_3EDTA at ≈23°C). This preliminary result in turn allows the calculation of the helical repeat from shifts measured with pairs of DNAs of larger length differences. For example, the shift observed for the pair pTR161/pTR199, with a length difference of 25 base pairs, is approximately 0.5 turn. Knowing that the helical repeat is close to 10 base pairs per turn, however, allows one to deduce that the actual change in Tw must be 2.5 turns. Therefore by performing measurements with pairs of DNAs of increasing length differences, the accuracy of the measurements is improved.

Table 1. Helical repeat (h) of DNA from measurements of DNA pairs with length differences <10 base pairs (bp)

DNA pair*	Δ length, bp	Shift observed†	h calc.
pTR199/pTR183	1	0.10	10
pTR182/pTR190	5	0.51	9.8
pTR193/pTR182	6	0.59	10
pTR199/pTR188	10	−0.02	9.8

* For each DNA pair, the left one is the larger of the pair.

† The shift is the upward (opposite to the direction of electrophoresis) shift of a band of the larger DNA from the band of the shorter DNA immediately below it and is expressed as the fraction of spacing between two topoisomers differing by 1 in their linking numbers. Because, for each pair, two families of bands are present, the value given is the mean of all measured shifts of the various bands.

For DNA pairs with larger length differences, however, in addition to considerations of differences in their writhing numbers the dependence of the electrophoretic mobility on the molecular weight itself is no longer negligible. Fortunately, because the length correction applies only to the measurement of the fractional turns and not to the integral part, even for the longest length difference, 58 base pairs between pBR161 and pBR190, the correction amounts to only a few percent. Therefore, a rough correction suffices. It will be assumed that the length dependence for a closed circular DNA is the same as that for a nicked circular DNA. For the samples shown in Fig. 1 *middle*, the actual distance travelled by nicked circular pTR161 was measured to be about 13.9 cm. In the same run, dimeric and trimeric nicked circular pTR161 migrated about 7.4 and 4.7 cm, respectively. Therefore, the mobility is roughly linearly related to size. Under identical conditions, when nicked circular pTR161 and pTR190 DNA were run in adjacent lanes, with linear pBR322 DNA mixed in each as an internal marker, the difference in distances migrated was measured to be 0.12 cm. Linear interpolation gives a decrease of 0.12/58 or 0.0021 cm per base pair increase in length. The average spacing between two adjacent topoisomers shown in Fig. 1 *middle* is about 0.57 cm. Thus, the length effect is of the order of 0.0021/0.57 or 3.7×10^{-3} turn per base pair. This value will be used to correct for the length effect for DNAs with length differences greater than 10 base pairs. The results are shown in Table 2.

A second way of correcting for the dependence of mobility on length is as follows. A mixture of a pair of DNAs, A/B, is analyzed by gel electrophoresis under two different sets of conditions. In one set, the conditions are so selected that the DNAs are positively supercoiled. If A is longer than B by a nonintegral multiple of the helical repeat, A will be shifted upward from B. Because the length effect slows down the bands of A more than the bands of B, the measured upward shift of the A bands is an overestimate. In the other set, the electrophoresis conditions are so selected that the DNAs are negatively supercoiled. The increase in Tw of A over B because of the length increment will cause a downward shift, and the mea-

Table 2. Helical repeat h of DNA calculated from measurements of DNA pairs with length differences >10 base pairs (bp)

DNA pair	Δ length, bp	Shift Observed	Shift Corrected*	Δ Tw deduced	h calc.
pTR193/pTR190	11	0.10	0.06	1.06	10.4
pTR161/pTR199	25	0.49	0.40	2.40	10.4
pTR161/pTR183	26	0.58	0.48	2.48	10.5
pTR183/pTR182	27	0.69	0.59	2.59	10.4
pTR161/pTR188	35	0.47	0.34	3.34	10.5
pTR161/pTR193	47	0.65	0.48	4.48	10.5
pTR161/pTR182	53	0.24	0.05	5.05	10.5
pTR161/pTR190	58	0.78	0.57	5.57	10.4

* Corrected for length effect.

sured downward shift is an underestimate because of the length effect on mobility. By averaging the measured shifts under these two sets of electrophoresis conditions, the length effect can be cancelled out, and the average shift is then used to calculate h. The helical repeat h so calculated is the average of the h values under the two sets of electrophoresis conditions (see *Discussion*). The pair of DNAs pTR161/pTR182, with a length difference of 53 base pairs, was analyzed in this manner and the value of h so obtained was not significantly different from the value given in Table 2.

DISCUSSION

For a given pair of DNAs n and $n + x$ base pairs long, the helical repeat h obtained by the band shift method is that of a DNA with the same sequence as the inserted x base pairs under the electrophoresis conditions. The dependence of the helical repeat on temperature and ionic environment has been known for some time (13), and knowing the helical repeat h under one set of conditions usually permits the calculation of this quantity under a different set of conditions. The results of Depew and Wang (10), for example, indicate that the helical repeat h of DNA at 37°C in a dilute aqueous buffer containing Mg(II) is not significantly different from that under the electrophoresis conditions used: DNA covalently closed under the former set of conditions by ligase migrated close to nicked DNA under the electrophoresis conditions.

The Gaussian center method measures the helical repeat h of the x base pairs under the conditions used for covalent closure or relaxation of the DNAs. In the work of Wang (9), covalent closure by ligase was carried out at 25°C in a dilute aqueous buffer containing Mg(II), and the helical repeat h under these conditions is not expected to be significantly different from that under the electrophoresis conditions. Experimentally, both sets gave the same value, 10.4 base pairs per turn, for the helical repeat h. Because the base sequences of the inserts used in these measurements are different, the agreement between the measurements indicates that, at least for these particular inserts, the sequence dependence of the helical repeat h is not strong.

The measured DNA helical repeat h in solution under physiological conditions, 10.4, is close to the value calculated for the lowest energy configuration of unbent DNA, 10.6 (8). The agreement is probably fortuitous because the theoretical calculation neglects electrostatic and solvent effects, which are known to affect the helical repeat. For examples, from the previously measured temperature and ionic effects (10, 15), it can be calculated that, for h = 10.4 base pairs per helical turn at 37°C in a dilute aqueous buffer containing Mg(II), at 0°C in the same buffer h becomes 10.3 base pairs per helical turn and at 0°C in 3 M CsCl h becomes 10.2.

The directly measured DNA helical repeat h in solution lends additional strong support for the classical Watson–Crick double-helix and rejects the side-by-side double-stranded model of Rodley *et al.* (2). The measured value of 10.4 base pairs per helical turn in solution under physiological conditions, with an estimated probable error of ±0.1, is significantly different from the value 10.0 for the classical B structure, however.

The band shift method or the Gaussian center method, in conjunction with *in vitro* and *in vivo* methods for constructing DNA recombinants, can be applied to the determination of the effects of certain base sequences on the helical repeat. Such measurements should complement fiber and single crystal x-ray diffraction studies of DNAs of unique sequences.

I am grateful to Dr. Thomas Roberts who kindly provided me with strains carrying the plasmid DNAs used in this work and also made available to me the sequencing results of these plasmids prior to their publication. This work was supported in part by National Science Foundation Grant PCM 78-05892.

1. Watson, J. D. & Crick, F. H. C. (1953) *Nature (London)* **171,** 737–738.
2. Rodley, G. A., Scobie, R. S., Bates, R. H. T. & Levitt, R. M. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2959–2963.
3. Wilkins, M. H. F. (1961) *J. Chem. Phys.* **58,** 891–898.
4. Bram, S. (1971) *J. Mol. Biol.* **58,** 277–288.
5. Tunis-Schneider, M. J. B. & Maestre, M. F. (1970) *J. Mol. Biol.* **52,** 521–541.
6. Finch, J. T., Lutter, L. C., Rhodes, D., Brown, R. S., Rushton, B., Levitt, M. & Klug, A. (1977) *Nature (London)* **269,** 29–36.
7. Crick, F. H. C. (1977) *Cold Spring Harbor Symp. Quant. Biol.* **42,** 243.
8. Levitt, M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 640–644.
9. Wang, J. C. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43,** in press.
10. Depew, R. E. & Wang, J. C. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4275–4279.
11. Vinograd, J. & Lebowitz, J. (1966) *J. Gen. Physiol.* **49,** 103–125.
12. White, J. H. (1969) *Am. J. Math.* **91,** 693–728.
13. Fuller, F. B. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 815–819.
14. Crick, F. H. C. (1976) *Proc. Natl. Acad. Sci. USA* **73,** 2639–2643.
15. Wang, J. C. (1969) *J. Mol. Biol.* **43,** 25–39.