Proc. Natl. Acad. Sci. USA
Vol. 73, No. 8, pp. 2639–2643, August 1976
Mathematics and Biochemistry

Linking numbers and nucleosomes

(DNA double helix/twist/writhing number/chromosome structure/simian virus 40)

F. H. C. CRICK

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Contributed by F. H. C. Crick, May 21, 1976

ABSTRACT **In considering supercoils formed by closed double-stranded molecules of DNA certain mathematical concepts, such as the linking number and the twist, are needed. The meaning of these for a closed ribbon is explained and also that of the writhing number of a closed curve. Some simple examples are given, some of which may be relevant to the structure of chromatin.**

It is not easy to think clearly about the way in which double-stranded DNA twists into various coils and supercoils. The subject has been greatly clarified by the mathematician F. Brock Fuller in a paper entitled "The writhing number of a space curve" (1). This paper is written in a clear, concise manner but its very compactness makes it difficult to grasp for the average molecular biologist. This note is an expansion and clarification of part of his paper. See also an earlier paper by White (2) and a further paper by Fuller (3).

The reader should recall two elementary facts about helices and handedness. The first is that a right-handed helix is right-handed from whatever position one looks at it. If it is turned end to end it stays right-handed. The second is that if a right-handed helix is viewed in a mirror, or inverted through a center of symmetry, it becomes left-handed, and vice versa.

The basic ideas

The essential concept we use is that of a ribbon. This ribbon can be thought of as a pair of lines—its two edges. Mathematically these are considered to be a minute distance apart. In reality the ribbons we will be considering will have finite width but we shall have the physical restriction that our ribbons cannot interpenetrate. We shall mainly be considering *closed* ribbons, which join back on themselves. It is assumed that *each edge joins only with itself*—and not with the other edge as in a Möbius strip. To underline this and to relate our ideas to the physical structure of the double helix of DNA, whose two chains run antiparallel rather than parallel, we put arrows, all pointing the same way, on *one* edge of the ribbon (in an arbitrary chosen direction) and label the other edge with arrows in the opposite direction. If we have occasion to break lines and join them we can only join lines pointing in the same direction. That is, we assume that we cannot join, by chemical bonds, a DNA backbone of one polarity to one of opposite polarity.

The line running along the center of the ribbon, which we shall call its axis, is also important. For a closed ribbon it joins back on itself. It does not have a direction.

Now we have to grasp three distinct but related concepts. These are: (*a*) the Linking Number, L; (*b*) the Twist, T; (*c*) the Writhing Number, W. The first important thing to realize is that the first two, L and T, are *properties of a ribbon*. They have, in general, no meaning for a single curve such as the *axis* of the ribbon. The Writhing Number, on the other hand, is the property of a closed curve, such as the ribbon axis. Its value depends on the exact shape of the curve in space, but not where the curve is in space ("is invariant under rigid motions") nor on the scale ("invariant under dilatations"). The mirror image of any curve has a writhing number of the same magnitude but of opposite sign. Thus, the writhing number of any curve which is its own mirror image (such as a circle) is necessarily zero. A curve which has a center of symmetry also has a zero writhing number.

The essence of Fuller's definition of the writhing number is the equation:

$$W = L - T$$

In short, although both L and T are properties of a ribbon, their difference (where they are suitably defined) is a property only of the ribbon's axis and not at all of the way in which the ribbon is twisted about that axis.

The meaning of L, T, and W

We must now state more precisely what is meant by L, T, and W.

The linking number, L, is roughly speaking the number of times the closed line along one edge of the ribbon is linked, in space, with the closed line along the other edge. For example, a ribbon forming a simple (untwisted) circle has linking number zero, since the two distinct circles formed by the edges are not linked in space. The linking number for a closed ribbon is necessarily an integer but as we shall see it can be positive or negative. It is unaltered under *all* deformations of the ribbon which do not tear it ("which deform it smoothly") and is therefore a *topological* property. We shall not define it more precisely here but later in this paper we give an algorithm for calculating it.

In order to give a sign to the linking number we must, in effect, put arrows on the two edges of the ribbon. We have already chosen to have these arrows run in opposite directions on the two edges because of the structure of DNA. [Mathematically this is not essential. In Fuller (1) the arrows are defined to run in the same direction.] Then a strip which is twisted in a right-handed manner will be given a positive linking number. To make the sign convention quite clear we illustrate two strips, one right-handed and one left-handed, in Fig. 1a and b. We can deform these figures to give the arrowed lines shown in Fig. 1c and d, which also illustrate the convention. Fig. 1e illustrates a (deformed) ribbon with a linking number of +2. The mirror image of a ribbon, or the ribbon inverted through a center, has L of the same magnitude as the original ribbon but of opposite sign.

We must now tackle the twist T. The exact definition, following Fuller (1), is given in the *Appendix* but at this stage all the reader needs is an intuitive idea of the twist.

Note first that a simple bend (Fig. 2a) does not introduce a twist, nor does the bend shown in Fig. 2b, in which the deformed ribbon lies in one plane. On the other hand the twisted ribbon shown in Fig. 3c clearly has a twist under any definition. The *units* of twist are chosen so that the stretch of ribbon shown in this figure, which goes round 360° (that is, 2π radians) is defined to have a twist of 1. Since the twist is right-handed it

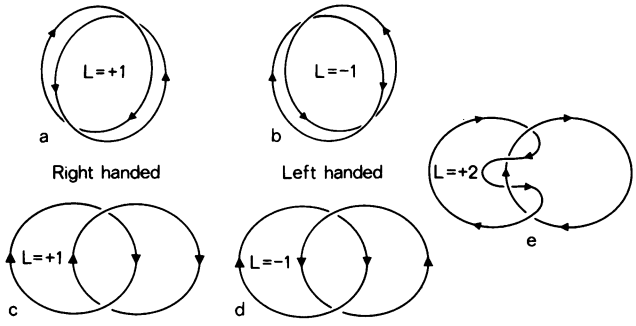

FIG. 1. Line drawings to show the sign convention for L. The pair of lines in (a) can be regarded as the two edges of a ribbon on which the arrows have been marked in *opposite* directions. (b) The mirror image of (a). Drawing (c) is topologically the same as (a), while (d) is the same as (b). If we reverse the direction of an arrow on *one* of the chains of a pair we alter the sign of L. (e) A pair of lines for which $L = +2$. Compare (c).

is called +1. The mirror image of a ribbon has a twist of the same magnitude but opposite sign, so that the (left-handed) reflected image of Fig. 2c has the twist −1.

The twist, T, of a ribbon is the integrated angle of twist (divided by 2π) along its length. For a closed curve it need not be an integer and in general it will not be. T is invariant under rigid motions and dilatations but its value depends on the exact shape of the ribbon in space and it is, in general, altered by arbitrary deformations. It is thus not a topological property but a metrical one. To illustrate the value of T for a slightly more complicated but still simple case we consider the twist of a ribbon wound flat on a cylinder and going round N times in a right-handed helix (see Fig. 3a, where $N = 2$). Then the mathematical definition of twist leads to the value

$$T = N \sin \alpha$$

where α is the pitch angle of the helix.

Care must be taken over the units of twist. One can easily but incorrectly assume that if $N = 1$ and α is small then the *angle* of twist is approximately equal to α. This is incorrect. The angle of twist is $2\pi \sin \alpha \doteq 2\pi\, \alpha$. However the twist T, which is the integrated angle of twist divided by 2π, does approximate to α in this case, α being measured in radians.

The twist of a ribbon wound flat on a cylinder tends to zero if the axis of the ribbon tends to a circle (each turn approximates to a simple untwisted circle) while the twist of a ribbon wound flat on a cylinder of vanishingly small radius (which gives a ribbon like that in Fig. 2c) tends to $T = N$. The twist for more complicated shapes must be calculated from the definition in the *Appendix*.

From the way that T is there defined it follows that in calculating it for a closed ribbon, by travelling along it, one can do so in sections. Thus, if there are "points" A, B, and C on the ribbon then one can calculate the value of T for the section AB, that for BC, and that for CA. The total T will be the algebraic

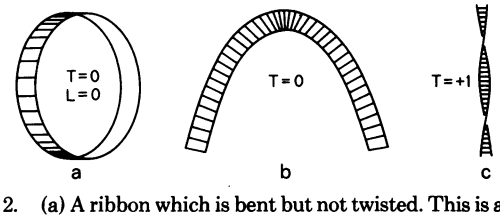

FIG. 2. (a) A ribbon which is bent but not twisted. This is also true for (b). In the latter case the ribbon lies entirely in the plane of the paper. (c) A ribbon which is twisted but not bent.

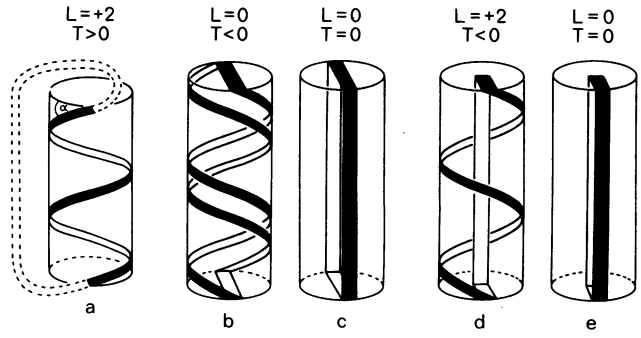

FIG. 3. (a) A short length of ribbon in the shape of a right-handed helix having two complete turns. To aid visualization a cylinder has been drawn in. The pitch angle of the helix is α. The broken lines show how the two ends of the ribbon can be joined to make it a closed ribbon. (b) A closed ribbon wound as a double helix, having two turns up and two turns down. Again the cylinder is shown to help visualization. (c) An untwisted ribbon. Note that by twisting (c) it can be deformed topologically into (b). (d) A closed ribbon forming a double helix with two left-handed turns joined along the helical axis. (e) An untwisted ribbon. Note that twisting (e) does not immediately give (d) (see *text*). Although in (d) the helix is left-hand, L is positive for this case.

sum of these three partial values. Notice that the definition of T is such that it has exactly the same value, both for magnitude and sign, if for the section AB one proceeds during the integration from A to B, or from B to A. Thus, not only can we calculate T in sections but we can do this in either direction for each section, as we please.

By contrast it is impossible to calculate the value of L in this way. L is a property of the entire ribbon and cannot usefully be calculated by travelling along it. One must view the ribbon as a whole. There is, however, a very neat way to obtain L from *any* outside view of the ribbon or of any topological deformation of the ribbon (F. B. Fuller, personal communication). The algorithm consists of marking first each edge of the ribbon distinctively (say, one with a red line and one with a black one), putting the necessary arrows on both of them. One then views the configuration (or any convenient topological deformation of it) from *any* chosen point outside the configuration. If this point is at infinity the view will be a projection but this is not essential. One now inspects each case where the red and black lines cross, arbitrarily choosing either the places where red is in front of black, or black is in front of red, ***but not both***. Each such cross-over can be assigned a value +1 or −1, according to the local direction of the arrows on the two lines at the cross-over and using the conventions of Fig. 1. Then L is the algebraic sum of all these assignments. In a loose way one can see that this will work because each cross-over represents a place where a cut must be made in one of the colored lines to allow it to be removed to infinity (either towards one, or away from one, as the case may be) while leaving the other colored lines in place.

The convention in Fig. 1 differs in sign from that of Fuller (1) because we have orientated the edges of a ribbon to run in opposite directions while Fuller orientated them to run in the same direction. Thus, the linking numbers for closed curves differ in sign from Fuller (1) but the linking number for ribbons agree.

What Fuller did was first to define L and T for a ribbon, on the lines sketched above, and then to define $W = L - T$. He showed that W was a property of the axis of the ribbon rather than of the ribbon itself. The properties of W, L, and T are summarized in Table 1.

One special result is of interest. It has been shown (1, 2) that

Table 1. Properties of L, T, and W

	L (linking)	T (twisting)	W (writhing)
Is a property of:	A closed ribbon	A ribbon	A closed line
It is:	Topological	Metrical	Metrical
Its value is:	An integer	Any number	Any number

the value of W for any closed curve lying on a sphere is zero. Obviously this is also true if it lies on a plane.

Although in general T can have any value there is a special case for which its values are restricted. Imagine that we are given a particular closed curve and told that our closed ribbon must have this curve as its axis but that we can put our ribbon on this axis in any way we please, provided that it has no discontinuity anywhere (i.e., the ends are joined up smoothly). Now because the axis of our ribbon is defined, the value of W is fixed. No matter how we choose to put the ribbon on the curve the value of L (which we are free to select) must necessarily be an integer. Thus, the possible values of T will be quantized. For example, if we are presented with a closed curve for which $W = +1.7$ then T must have a value chosen from the infinite set $\ldots -2.7, -1.7, -0.7, +0.3, +1.3, +2.3, \ldots$ since $L = W + T$.

Although L for a single closed ribbon must always be an integer there are special cases where the effective value of L need not be integral. Consider the case of a circular piece of DNA which is relaxed in solution under defined conditions so that it is the "unconstrained" state. Two sets of authors (4, 5) have shown that in practice one does not get a single value but a Boltzmann distribution of L, due to thermal effects. Thus, the effective value for the unconstrained molecule is the weighted average of these integral values of L, and this average will, in general, not be an integer. In the same way a solution of superhelical DNA molecules will have an effective nonintegral value of L. Note that while at any moment a particular molecule in solution will have definite values of L, T, and W, the thermal motion constantly changes T and W. Only L remains constant with time.

Some examples

In spite of all obvious differences between the way T and L are calculated it is not easy to obtain an easy familiarity with them unless one works through a few examples. We will consider two types of example: those which clarify the concepts and those which illustrate the pitfalls.

For clarification the reader might reasonably ask for an example of a ribbon for which $L = 0$ but T is not zero and the converse example where $T = 0$ and L is not zero. These are fairly easy to display.

Consider a regular left-handed *double* helix made of a ribbon wound flat on a cylinder and having any desired number of turns (see Fig. 3b). The double helix is made into a closed ribbon by joining the two ends at the top and the two ends at the bottom, as shown in Fig. 3b. Then it is easy to show that for such a structure [which, incidentally, is illustrated in Fuller's paper (Fig. 1 of ref. 1)] the value of L is zero. The proof is easy. Consider a closed ribbon wound on a cylinder in the simple way shown in Fig. 3c. Then by inspection we see that $L = 0$ since the two edges of the ribbon are not linked. Now imagine that the bottom of the cylinder is held still but that the top of the cylinder is rotated so that the whole structure is twisted. Clearly we can generate the double helix of Fig. 3b without tearing the ribbon so it, too, must have $L = 0$. Equally clearly T is *not* equal to zero since the contribution to the twist is negative as we go up one helix and also negative as we come down the other, while it can be made zero for the two loops at top and bottom by constructing them using twistless bends. In fact (again neglecting end effects) it is $N \sin \alpha$ where N is the sum of the turns up plus the turns down and α is the angle of inclination of the helix (1).

It is not quite so easy to grasp a ribbon for which $T = 0$ and $L \neq 0$. In principle one can do this for *any* closed ribbon whose axis has an integral (non-zero) value of W by simply laying the ribbon on the axis in such a way that $L = W$. Then $T = 0$. Probably the best example is one of the simplest: choose a helical ribbon like that in Fig. 3a having α such that $N \sin \alpha$ is an integer. Adjust the ribbon so that it is not laid flat on the generating cylinder but is given a counter-twist to make T zero. Then connect the two ends using twistless bends.

However, a better way to grasp the distinct nature of the L and T is to consider once again a ribbon wound flat on the generating cylinder as a regular right-handed helix with exactly N turns and inclination α, as in Fig. 3a. How do L and T change as α is varied and N is kept constant? The answer is that L *always* equals N, no matter what the value of α (provided it is between 0 and $\pi/2$) but that the twist, T, is equal to $N \sin \alpha$ (and is negative if the helix is left-handed). Notice also that the value of L jumps from N to $-N$ as α changes from positive to negative; when $\alpha = 0$ the ribbon degenerates and L is not defined. Such discontinuities in L *only* occur when a change cannot be brought about by a smooth topological deformation from one curve to the other, as in the cases above (3).

There is one other example that may be helpful. Can we define a closed ribbon for which $L = 0$ and for which the twist is everywhere zero? To do this in the most general way draw *any* closed line (which does not intersect itself) on a plane. Imagine a ribbon whose axis lies on this line and whose breadth is everywhere perpendicular to the plane. Notice that the ribbon is bent, possibly in a very elaborate way, but its twist is zero everywhere. This example makes quite transparent the difference between bending and twisting.

We now consider a few surprising examples. As we have just emphasized a regular helix with, say, two left-handed turns and which has its ends joined together outside the helix as shown in Fig. 3a has $L = -2$. But what is the value of L if the ends are joined together by passing the ribbon, untwisted, along the axis of the helix, as shown in Fig. 3d? (To avoid end effects the joins at top and bottom should be constructed using twistless bends and not exactly as illustrated here.) The main features of this new structure seem so similar to that of the old one that one feels at first that L cannot be too different. Bolder spirits often guess that $L = 0$. However, the surprising result is that $L = +2$ instead of -2. Joining the ends of the ribbon up the center has changed the sign of L. This can easily be proved by the algorithm described earlier and indeed is a good example to try one's hand on, but a neater proof is as follows. Consider once again a ribbon for which L is obviously zero but this time imagine one arm of the ribbon on the surface of a cylinder and the other up the axis, as shown in Fig. 3e. Now, as before, imagine the bottom of the cylinder to be stationary and the top to be rotated so that we generate a structure with two left-handed turns of flat ribbon on the outside and a twisted branch of the ribbon up the axis also having two left-handed turns. We now calculate the writhing number. The outside section of the ribbon contributes $-2 \sin \alpha$ to the twist while the contribution of the central section is -2. Thus, $T = -2(1 + \sin \alpha)$. Since $L = 0$ we have $W = +2(1 + \sin \alpha)$ for this structure. However the *axis* of the ribbon fol-

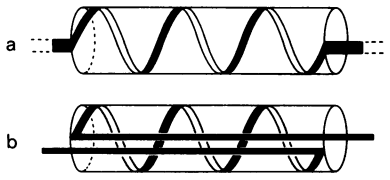

FIG. 4. (a) A ribbon with three left-handed helical turns, the ends uncrossed. (b) A similar ribbon having a little less than three turns but with crossed ends.

lows exactly the same configuration as the structure we are interested in, illustrated in Fig. 3d. Thus, for this structure, too, $W = +2(+\sin\alpha)$. However, for Fig. 3d the value of T is clearly $-2\sin\alpha$ since the central part of the ribbon is untwisted. Thus, since $L = W + T$ we obtain:

$$L = +2(1 + \sin\alpha) - 2\sin\alpha = +2$$

This calculation illustrates in a neat way one of the uses of the writhing number, W. It is sometimes easier to calculate L and T for a structure having the same *axis* as the one in which we are interested. Then if for our structure of interest, the calculation of T is easy, we can immediately obtain L for it, or vice versa.

The reader may be disturbed by the neglect of "end effects." This is usually justified though in real cases *the whole ribbon must be completely specified*. It may come as a surprise that any piece of ribbon can always have its ends joined together by an extension of the ribbon such that the extension has zero twist. This is done by using configurations which bend but are not twisted, such as those shown in Fig. 2a and b. Naturally in considering end effects one must be satisfied that this can be done without causing difficulties with L but in the cases quoted above this is not a problem.

As a second example, consider again a ribbon wound flat on a cylinder for a number of left-handed turns, not in this case necessarily an integer. However, instead of leading off the ribbon in the obvious way, as shown in Fig. 4a, we make a twistless bend at each end of the helix and lead off the ribbon in the opposite directions, as shown in Fig. 4b. Again the result is surprising. If the number of left-handed turns is just *under* 3, as for the structure shown in Fig. 4b, then L is not near -3 nor, as one might perhaps guess, near -2 but is close to -1. Again this can easily be checked by the cautious use of the algorithm given earlier. In fact, the value of L for structures of this type is plotted against N in Fig. 5. It will be seen that at each integer value of N there is a discontinuity, as might be expected since, for example, a structure with N slightly less than 3 cannot be smoothly deformed into one with N slightly greater than 3. To convert one to the other the ribbon must pass through itself, as one can easily verify with a wire model.

The reader may be puzzled by the fact that L, as displayed in Fig. 5, is not always an integer. What does it mean to say that when $N = -2\frac{1}{2}$, $L = -1\frac{1}{2}$? What is implied is that if a very large number of structures are joined end to end then the total value of L divided by the number of such structures is equal to $-1\frac{1}{2}$, again neglecting end effects due to making the ribbon into a continuous one. If the axes of the ribbon as it emerges on the two sides of each substructure are parallel (as in our example) then the value of L per substructure will only be an integer if the stretches of ribbon at these two points are untwisted relative to each other. For the case quoted above, if $N = -2\frac{1}{2}$, it will be found to be twisted by just half a turn. Thus, in using the algorithm to calculate L one should view a whole series of such substructures from one viewpoint (see *Appendix*) and find the *average* value of L per substructure. Strictly, L *has only meaning for the whole of a closed curve and attempts to factorize it into sections are only justified in special cases.*

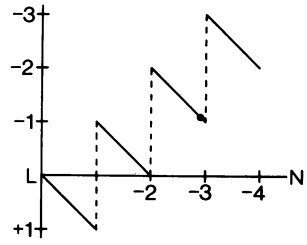

FIG. 5. The value of the Linking Number, L, is plotted against the number of turns, N, for structures like that shown in Fig. 4b. The marked point corresponds roughly to the values chosen for Fig. 4b itself. Both L and N are shown negative here because the helix has been taken as left-handed. For the definition of L, see the *Appendix*.

The structure of chromatin

The above example may perhaps be relevant to the structure of chromatin. The number of nucleosomes (or beads, or ν-bodies) on the simian virus 40 mini-chromosome was first estimated by Griffith (6). From measurements on gels, using the closed, circular, supercoiled DNA of simian virus 40 or polyoma virus, several workers (7–9) have estimated that the number of "supercoils" is about -1 or $-1\frac{1}{4}$ per nucleosome. This is done by using a "relaxing enzyme" and counting the number of different supercoiled species produced by spreading them out on an agarose or similar gel. The bands seen are clearly discontinuous, reflecting the integral nature of L. What is being measured is, in our terminology, the difference between the mean L for the mini-chromosome and the mean L for the completely relaxed form of DNA (in that medium) without supercoils. This estimate also fits with earlier, less direct, methods (10–12). Note that at any temperature there is, for unnicked, closed, circular DNA molecules, a Gaussian distribution of L about its mean value (4, 5).

The above example (Fig. 5) shows that a small modification to a structure with, in the loose sense, about three "supercoils" can produce a value of L, per nucleosome, not far from the observed value of -1. Whether this is the true explanation of the rather low observed value of L remains to be seen. If a model were required for which L was near zero a rather neat solution is to have the DNA folded as in Fig. 4b but with N a little less than 2 rather than a little less than 3.

What emerges clearly from the above examples is that, without some experience, it is not always obvious how to estimate the value of L for a structure and that one can easily make considerable errors. Thus, L should always be calculated carefully. It is fortunately easy to obtain an approximate value for a structure by constructing it from a piece of flexible ribbon of, say, some dressmaking material. One then simply pulls the ribbon "straight" and counts the number of twists. This will show if a gross error has been made and can also be used to astonish one's colleagues. However, for more exact work a careful estimate should be made from a model of the structure itself. A "kink" in DNA, for example although it mainly produces a bending, will often impart a small amount of twist and this may have to be allowed for. Moreover it is not obvious that the DNA in the nucleosome will be exactly in the normal B configuration unless other evidence suggests this.

In this paper we have been considering only configurations but mechanical properties not at all. An approach to these more difficult problems is sketched in the latter part of Fuller's paper (1).

APPENDIX

The definition of twist

For an exact definition see ref. 1, where it is called the *total twist number*.

At any point on the axis of the ribbon we define a vector, X, pointing along the axis and tangential to it at that point. We now define another vector, U, perpendicular to X and passing through the two edges of the ribbon at that point (the ribbon is strictly considered to have an extremely small width). As our point moves along the axis of the ribbon the vector U rotates about the vector X. Then the twist is the integrated angle of this rotation in radians, divided by 2π, integrated right around the closed ribbon. It need not be an integer. If the rotation is right-handed the twist is positive.

The calculation of L for Fig. 5

The special viewpoint required to calculate L correctly for Fig. 5 is defined as the view from infinity, perpendicular to a particular plane. Let there be M identical substructures, of the type shown in Fig. 4b, joined end-to-end without any deformation. Consider the axis of the ribbon itself where it enters the string of M substructures and also the axis of the ribbon where it leaves the string. Then the required plane is the plane containing these two short straight lines.

Calculate, using the cross-over algorithm, the total contribution to L from the entire string of M substructures. Then the average L per substructure is defined as the limit L/M as M tends to infinity. For this example the above viewpoint always allows one to construct a joining ribbon (of the general type shown by broken lines in Fig. 3a) which contributes nothing additional to the algorithm for L.

I wish especially to thank Prof. Fuller, for many useful points made in correspondence and for allowing me to quote unpublished work, and Dr. Graeme Mitchison of this laboratory for helpful explanations and discussion, and in particular for the neat proof shown in Fig. 3b and c. I also thank Dr. Aaron Klug and Prof. J. Vinograd for valuable comments.

1. Fuller, F. Brock (1971) *Proc. Natl. Acad. Sci. USA* **68**, 815–819.
2. White, J. H. (1969) *Am. J. Math.* **91** (3), 693.
3. Fuller, F. Brock. (1972) *Rev. Roum. Math. Pures Appl.* **17**, 1329–1334.
4. Depew, R. E. & Wang, J. C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4275–4279.
5. Pulleybank, D. E., Shure, M., Tang, D., Vinograd, J. & Vosberg, H.-P. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4280–4284.
6. Griffith, J. (1975) *Science* **187**, 1202–1203.
7. Keller, W. & Wendel, I. (1975) *Cold Spring Harbor Symp. Quant. Biol.* **39**, 198–208.
8. Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1843–1847.
9. Keller, W. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4876–4880.
10. Wang, J. C. (1974) *J. Mol. Biol.* **89**, 783–801.
11. Liu, L. F. & Wang, J. C. (1975) *Biochim. Biophys. Acta* **395**, 405–412.
12. Pulleybank, D. E. & Morgan, A. R. (1975) *J. Mol. Biol.* **91**, 1–13.