Chapter 3

KNOT THEORY, DNA TOPOLOGY, AND MOLECULAR SYMMETRY BREAKING

W.-Y. Qiu
The Center for Nonlinear Science, University of Science and Technology of China, Hefei 230026, P. R. China

3.1 Introduction 175
3.2 Chirality and Achirality 181
3.3 Genus, Point Group, and Seifert Construction 184
3.4 Closed Duplex DNA 187
3.5 Single-Stranded DNA Knots 201
3.6 Duplex DNA Knots 207
3.7 Duplex DNA Catenanes 218
3.8 Summary 224
3.9 Literature References 226

3.1 Introduction

The intimate relation between mathematics and chemistry may seem surprising to the layman, but to someone well acquainted with either filed, it appears as a natural evolutionary development. A theory invented by mathematicians to settle mathematical questions often turns out to be exactly what chemists and biologists need to advance their analyses and predictions of the molecular structures and chemical reactions. Mathematical chemistry should either introduce new mathematical methods or techniques for

the solution of chemical and biological problems, or develop new mathematical approaches or insights pertinent to any area of chemistry and biochemistry. DNA molecules have stimulated mathematical thinking at least since the discovery of the DNA double helix in 1953 [1,2]. The model of the DNA molecule proposed by Watson and Crick, with all its biological implications, has been one of the major scientific accomplishments of this century. In 1962 Watson, Crick and Wilkins shared the Nobel Prize in Medicine or Physiology, for their discovery of the DNA structure [1-4]. Understanding the mechanism of double helix and the consequences of this structural feature of DNA may be viewed as the foundation of our ultimate understanding of life. Thus chemistry, along with biology, can be regarded with mathematics as valid scientific approaches that are justifiable because they contribute to human knowledge in areas of universal philosophical significance.

DNA, the fundamental molecule of life, is both variable and flexible as the genetic material's double helix. In many forms of DNA the double helix itself is transformed into a new helix of a higher order, and even the new helix can tie itself in knots and links. Twisted, tangled, supercoiled, knotted or linked DNA, which can have important biological implications, is best described and analyzed by means of a simple mathematical model. Fortunately, the branch of mathematics known as topology which studies the properties of structures that remain unchanged when the structures are deformed, is able to offer substantial help in this effort. Topology is a branch of geometry and cannot actually help us solve equations. What it provides rather is a mathematical vocabulary (adjectives and nouns) and some algorithms (invariant) that allow a set of solutions to be discussed in a general way without actually being specified. The link between DNA and topology soon became obvious and this resulting branch of research, on the borderline between two fields, has attracted considerable attention for the last forty years. This is certainly due in part to the introduction of mathematical concepts into the description of molecules but also to the development of new synthetic methodologies, as well as powerful instrumental, and analytical techniques.

The study of DNA topology [5-7] involves the following five major endeavors:

(1) Elucidation of the mathematical implications of the DNA structures;

(2) Exploration of stereochemical theory from the topological perspective;

(3) Understanding of topological conversions and reaction pathways for the DNA replication and recombination;

(4) Use of topology to probe the hidden action of enzymes; and
(5) The design and synthesis of DNA stereoisomers with novel topology.

In summary, the study of DNA topology seeks to explain the molecular mechanisms underlying biological complexity. It has given rise to new ways of thinking about life. The potential gains from these exciting areas are considerable, and the life science disciplines are being challenged in a great variety of ways. How important such an approach will be remains to be seen; but we have explored only a minuscule fraction of this new domain!

3.1.1 The Ribbon Model

The first mathematical model, which is used to describe circular duplex DNA, is a closed ribbon in which the two edges of the ribbon represent the single strands of DNA.

In 1961 Călugăreanu [8] found the basic relationship between the geometrical and topological properties of a closed ribbon and in 1968 Pohl [9,10] gave much simplified proofs of Călugăreanu's results. In 1969 White, as a Ph.D. student of Pohl, proved [11] that a closed ribbon has both a linking number L_k and a twist T_w. These results were originally proved as a piece of pure mathematics, without any reference to DNA. In 1971 Fuller independently suggested [12] that the difference of the linking and twisting number is called the writhing number W_r so that

$$L_k = T_w + W_r .$$

This White-Fuller formula follows from rigorous examination of the results of Vingorad and co-workers and is also the starting point for mathematical methods in DNA research. In 1976 Crick [13] pointed out important applications of this formula, and in 1978 Fuller [14] deduced some conclusions from this formula, e.g., the linking-number difference. Since then the White-Fuller formulas, as basic relations, have become the fundamental tools of DNA topology [15-26]. This line of research is of fundamental importance in understanding the mechanism of supercoiling and the structural features of DNA molecules. Detailed background information on several newer results can be found in Refs. [7,21,25,27], whereas some of the earlier developments are reviewed in Refs. [15-17,20]. It is now known that each of the three quantities, L_k, T_w, and W_r, defined for a closed ribbon, has a desirable property not shared by the other two: L_k is topological, T_w and W_r are geometric. It is clear, however, that the present knowledge of supercoiling, both theoretical and experimental, provides a sound basis for the investigation of the surprising manifestation of

the double helix. Despite these successes, many questions arising from this line of research have remained unanswered or unanswerable.

3.1.2 Knot Theory

By a DNA knot we mean a curve of single-strand or double-strand DNA in three-dimensional spaces which begins and ends at the same point and does not intersect itself. A DNA link is a class of intertwined supermolecules and interlocked superstructures in which two circular duplex DNA molecules are mechanically interlocked. It is now clear that knots and links are new forms of the molecular structures and the stable knot- and link-like structures [28,29] represent solutions to the nonlinear field equations.

Links and knots have traditionally been considered as chemical curiosities, though researchers finally stopped looking at them as nothing but exotic chemical objects when DNA was clearly shown by biologists to form links and knots. A great variety of single- and double-stranded DNA knots have been observed in diverse biological systems [30-32], and by now have become a commonplace in biochemistry topology [6]. DNA links (or DNA catenanes) *in vivo* exist in plasmid DNA, in viruses, as well as in the replication of circular DNA [6,7,31]. Thus, chemistry and biology need knot theory to manipulate these substances. The development of analytical structural, synthetic, and theoretical techniques appropriate to this area of DNA knots and links represents an enormous challenge and opportunity for scientists in many different discipleines.

Knot theory [33-36], which is a part of topology, studies the placement problem of a one-dimensional curve traced in three-dimensional space. Before describing some of the methods that have been developed for knots that are different, however, it is necessary to decide when they can be regarded as being the same. To prove that two knots are equivalent it will suffice to deform one until it matches the other. On the other hand, in order to prove that the knots are different it is necessary to find some property that distinguishes them. Such a property of a knot is called an invariant. An invariant is a characteristic property of a knot that is unchanged by deformations of the knot. Such an invariant will always be the same if two knots are the same and will always be different if two knots are different. But the problem is that the invariants can be the same for two knots and even when the knots are not the same [37]. What mathematicians really want therefore are properties of knots that they can calculate that are genuinely invariant, that uniquely characterize knots. Thus, to completely classify and characterize knots or links, using a well-defined

invariant, is the fundamental task of knot theory.

To the best of our knowledge there are two areas of knot theory with potential for major impact on new knowledge of DNA topology:

3.1.2.1 The Polynomials of Knots and Links One of the most important invariant of a knot or link takes the form of a polynomial. There are many methods available for polynomial calculation [38]. Generally, such polynomials can be calculated by a simple iterative procedure starting with standard drawings of knots and links. More generally, the polynomials can encode data about knots and links in some very complicated way [37]. Just what the encoding pattern is poses an interesting challenge for mathematicians.

In 1928, John Alexander, an American mathematician, discovered a simple polynomial that is associated with knots and can determine whether knots are distinct [39]. But, the Alexander polynomial may be the same for two knots even when the knots are not the same. In 1985 Jones [40] discovered a one-variable polynomial. Almost immediately, six of mathematicians [41] independently and by different methods found a two-variable polynomial. By specifying one variable, we get the Alexander polynomial. By specifying another, we get the Jones polynomial [37]. This finding is the basis of the Lickorish-Millett polynomial [42-44]. It appears that the first applications of the new polynomials with the two-variable to DNA catenanes (links) and knots were made by White, Millett and Cozzarelli [45] in 1987. They utilized the polynomials to classify the topological state of DNA knots and links as well as to determine enzyme mechanism. At that time, Cozzarelli's group described three invariants of knots and links: the node number (or crossing point number) [46], the Schubert invariant [47], and the polynomials invariant [45]. The node number provides a ready first-level codification and is also critical for the structure of DNA in solution. The simple, integral Schubert invariant, β and α, have the advantage of classification uniqueness. Many different knots and links have the same number of nodes and, extremely rarely, have the same polynomials. Nonetheless, it was the development of the polynomial invariants which basically solved the classification problem for DNA knots and links [45].

3.1.2.2 The Tangle Model The mathematical notion of tangles, which was first introduced into knot theory by J. H. Conway, plays an important role in DNA topology. In 1967, Conway [48] defined a polynomial with integral coefficients for links or knots, which can be inductively computed from a regular projection of a link or knot. The Conway polynomial is a refined variant

of the classical Alexender polynomial and a more powerful invariant of oriented knots and links [35]. Conway's fraction theorem [35,36] allows us to quickly calculate polynomials of the numerators and denominators of rational tangles. Tangles are an encoding for knots and usually represented by their projections, called tangle diagrams. Rational tangles, of course, are important in studying DNA topology. In recent years, exciting advances have been made in the study and control of the tangle model. A rigorous mathematical treatment of tangles is found in Ernst and Sumners [49]. Mathematical results on the tangle equations can be found in Ernst [50,51]. More intuitive treatments of the mathematics are found in Sumners [52-55]. Using tangles, we can rigorously deduce enzyme mechanism from the topology of the DNA substrate and product (Summers, Ernst, Spengler and Cozzarelli, 1995 [56]; Ernst and Summers, 1998 [57]). Thus, the tangle model provides mathematical proof of structure and mechanism. More importantly, tangles can be used to prove that the unique mechanism for enzyme action has been found. The mathematically tractable type of tangles, rational tangles, can also describe DNA recombination. In short, a tangle provides an accessible, mathematically general, description of site-specific recombination and associated topological changes that permit definition of enzyme mechanism and prediction of substrate and product structures.

Indeed knot theory, an esoteric branch of pure mathematics [58], evolved from the interaction between chemistry and biology that appears here in new forms invigorated by new ideas. Although there is much to learn, the stakes are high and the odds for success are excellent.

The advance of molecular biology has undergone revolutionary acceleration during the past two decades, and chemists and biologists have exponentially increased the number of reaction types and classes of DNA and protein molecules. DNA can be cut apart, modified and reassembled; it can be amplified to form many copies; and perhaps most telling, with DNA one can generate RNA and then protein molecules of desired size and constitution [59]. The central experimental maneuver in these manipulations is the cloning of genes and the polymerase chain reaction [60]. Polymerase chain reactions (PCRs) involve a simple technique for copying a piece of DNA in the laboratory with readily available reagents. Because the number of copies increases exponentially, more than 100 billion can be made in only a few hours. It is the polymerase chain reaction, more than any other single factor, which has changed the face of biology. In 1993, K. B. Mullis won the Nobel Prize in

Chemistry, for his discovery of the polymerase chain reaction. Recently, PCR has been considered as the experimental basis of DNA computing [61].

In summary, molecular biology has become sufficiently detailed that many of its most important current questions must be phrased and answered using mathematical methods. Unfortunately, the theory of mathematical chemistry and DNA topology presently cannot provide fully adequate methods. This is an area ripe for development and deserves special attention. To understand the symmetry properties of large, very flexible molecules with novel topology, such as DNA links and knots, new methods are required. We introduce here point groups to the Seifert construction in knot theory for the first time and discuss the implications of our model for real DNA. Chirality or achirality, an interesting problem, has been characterized by applying a symmetry criterion to DNA links and knots. Our offering in this chapter represents only an initiateon to the relevant questions, but clearly such questions are well worth considering and eventually result in a better understanding of the life process and genetic control.

3.2 Chirality and Achirality

Most objects found in nature are not identical to their mirror image and therefore are said to possess chirality, or handedness [62]. To distinguish the two chiral forms, they are often designated right-handed or left-handed. Objects that are identical to their mirror image are said to possess achirality. Chemists refer to mirror-image molecules as L-enantiomers and D-enantiomers; L and D stand for *levo* (left) and *dextro* (right). Modern science has revealed that nature is asymmetric with respect to chirality and is symmetric with respect to achirality. Chirality is so much a part of daily life that we feel we understand it thoroughly. Although the left and right human hands are obviously different, left and right rubber gloves are essential the same. This is because the former is extremely rigid whereas the latter is completely flexible. A rigid hand cannot be deformed into its mirror image and therefore is chiral, whereas a flexible rubber glove can be deformed into its mirror image by inside-out motion or deformation and therefore is achiral. The rigidity is a geometrical property and the flexibility is a topological property. This fact tells us that a helix being left-handed or the right-handed helix is not a really criterion for chirality. Strictly speaking, chirality or achirality is an intrinsic property of the object. This requires proof!

Chirality can be classified into geometrical chirality and topological chirality according to object's rigidity or non-rigidity. A rigid object is geometrically chiral [63] if it cannot be superimposed on its mirror image by simple rotations and translations. A non-rigid object is geometrically chiral if its topology has the rigid presentation and two distinct forms. A non-rigid object is topologically chiral if its topology has the non-rigid presentation and the left-handed and right-handed forms are mirror images of each other.

Achirality can also be divided into two groups, geometrical achirality and topological achirality. In order to completely characterize achirality it is useful to understand the deformation of turning something "inside-out". As an example of this we can look at the two-handled torus with two holes. Stewart [64] sketched a series of pictures of the process, which is shown in Figure 3.1. First turn the whole thing inside out through its holes. Then the two handles on the outside became two tunnels on the inside. But then you can place your finger into each in turn and pull it out to create a handle on the outside. And then all you have to do is twist each handle around and you end up with the original ensemble, but turned inside out [64]. Intuitively it is not difficult to see that the two-handled torus with two holes can be converted into its own mirror image by a continuous deformation. Therefore, it is topologically achiral.

The abstract mathematical objects discussed above may be used as a model of the molecular structures. A molecule shaped like a rubber glove does not in fact exist in chemistry and biology. An approximate rubber glove can sometimes be used as a model of the inside-out deformation. Thus, a rubber glove is defined as a structure which can be turned inside-out (or reversed up and down) and in which the palm and thumb are differentiated from the back of the hand. A rubber glove is said to be Euclidean rubber glove if its topology possesses a rigid presentation and is said to be topological rubber glove if its topology possesses a non-rigid presentation. The Euclidean rubber glove is geometrically chiral and the topological rubber glove is topologically achiral. A torus with a hole (Figure 3.2) is a topological rubber glove, but the two-handled torus with two holes (Figure 3.1) is not a topological rubber glove. However, assuming the top edge and left frame are differentiated from the bottom edge and right frame, then our Figure 3.1(a) can be seem as an approximate model of the topological rubber glove. But, as noted above, a rubber glove has no presentations with proper and improper symmetry.

Flapan ever proved [65] that the union of figure-eight knot and circle ($G = K \cup C$) [66] is a topological rubber glove. The history of the particular problem

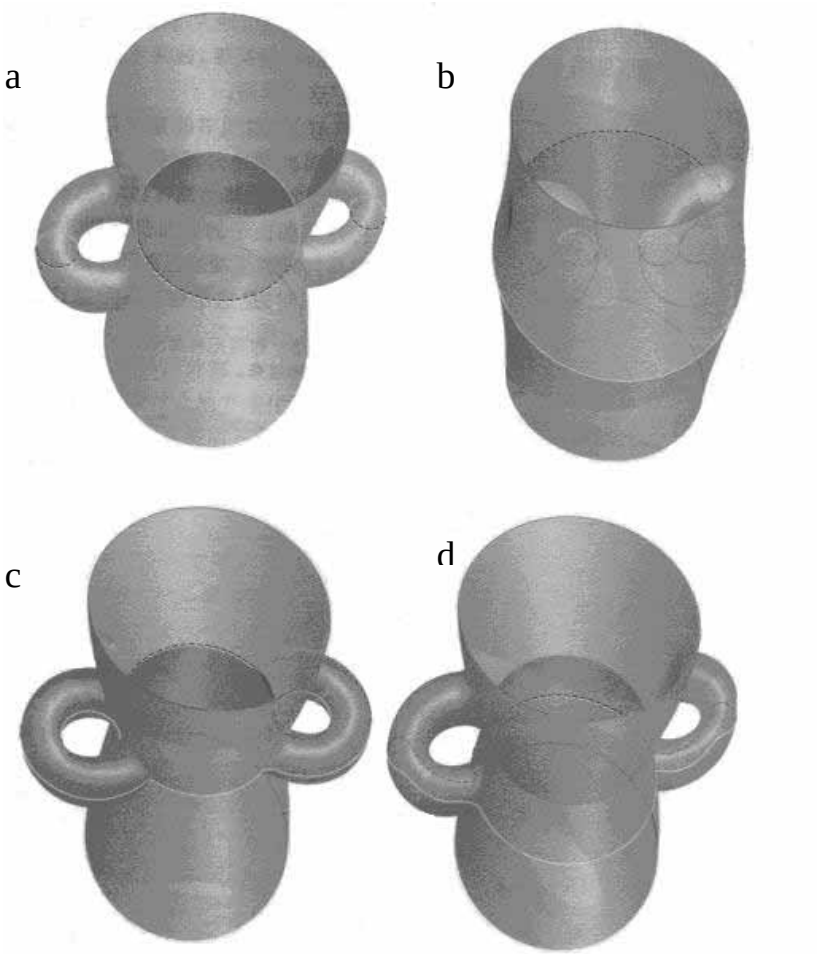

Figure 3.1: How to turn a two-handle torus inside out through two holes. Reproduced from Ref. [64] by permission of M. Goodman and Scientific American, Inc.

began with Mislow [67], who gave examples of disubstituted biphenyls that are achiral yet have no chemically accessible symmetry presentations. This led Walba [68,69] to find the knot 4_1 with a single colored point as a rubber glove. A molecular version of the rubber glove has recently been realized in single-stranded DNA tied into a figure-eight knot [70]. It is possible to rank the classes of molecular graphs by their "degree of chirality": from most chiral to least chiral [69]. A good insight into this domain of science may be found in Mislow excellent review [63] (also see Refs. [71-74]).

It is now clear that chirality reserves left and right [75] and achirality reverses up and down (Figure 3.1). Most molecular handedness is a geometrically

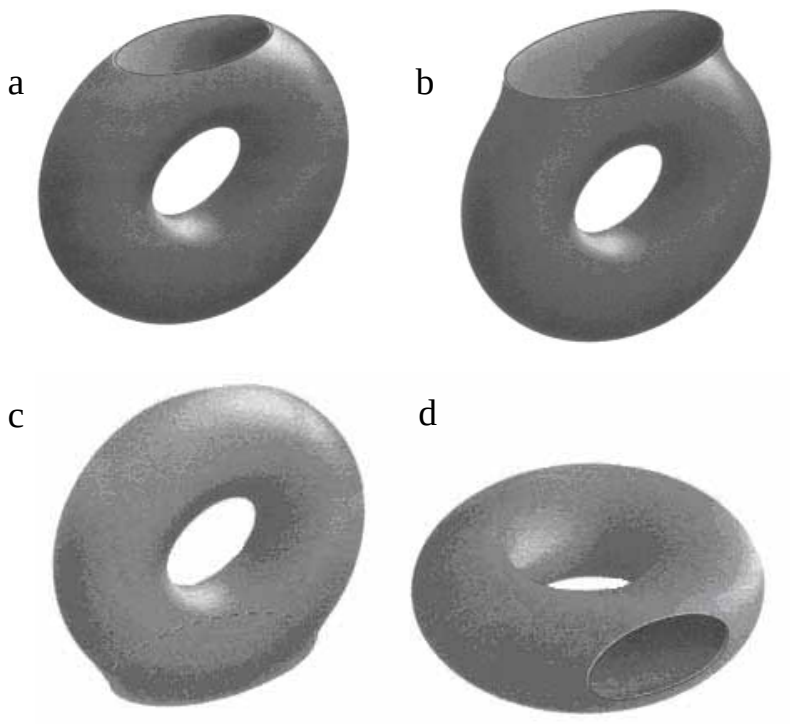

Figure 3.2: How to turn a torus inside out through one hole. Reproduced from Ref. [64] by permission of M. Goodman and Scientific American, Inc.

chiral [63], in essence, but topologically chiral molecules are very rare. Topological chirality is therefore only a special case of molecular chirality. Tsung Dao Lee and Chen-Ning Yang, who were awarded a Nobel Prize in 1957, proved in 1956 that the weak interactions do define a screw direction [76]: nature discriminates between left- and right-handedness. It is now evident that the world is chirally asymmetric at all scales, from the scale of elementary particles upward; mirror symmetry [75] is often absent in nature [62]: symmetry was broken and one kind of handedness became dominant, although the dominance of another was equally likely.

3.3 Genus, Point Group, and Seifert Construction

The Seifert construction [77] is a minimal surface with the smallest genus. The genus is a geometric invariant of a knot. The first general method for constructing such a surface was developed by the German mathematician

Seifert [78] in 1935. The genus and the related minimal surfaces hold a central position in knot theory and have proved to be valuable tools for treating a variety of knot problems [33-36].

The Seifert construction begins by considering a two-dimensional surface with a single edge that is embedded in three-space [77]. This type of surface may be either two-sided such as a disk or one-sided as an Möbius band. The two-sided surface is side to be orientable and the one-sided surface is said to be nonorientable. Orientable surfaces can be constructed for any knot so that the knot is the only edge of the surface. Two surfaces bounded by the trefoil knot (knot 3_1) are shown in Figure 3.3(a) and three surfaces bounded by the eight-figure knot (knot 4_1) are illustrated in Figure 3.3(b) [77]. The knot 3_1 and 4_1 were constructed by using two or three disks to fill in loops in the knots and using three or four ribbons with one half-twist to fill in the crossing-point areas, then connecting the disks with ribbons in 3-space, respectively.

The Seifert construction led to the definition of a geometrical invariant called the genus [36,37]. It can be proved [33] that the Seifert construction has the genus:

$$Genus = 1 - \frac{D + C - R}{2}$$

Where, D is the number of disjoint disks, C is the number of components and R is the number of ribbons with one half-twist. In particular, C is 1 for the knots and $C \geq 2$ for the links. It is easy to compute that the knot 3_1 and 4_1 are the knot of genus one. But the problem is that the knot 3_1 and 4_1 are different; to be specific, the knot 3_1 is invertible (chiral), but the knot 4_1 is both invertible and

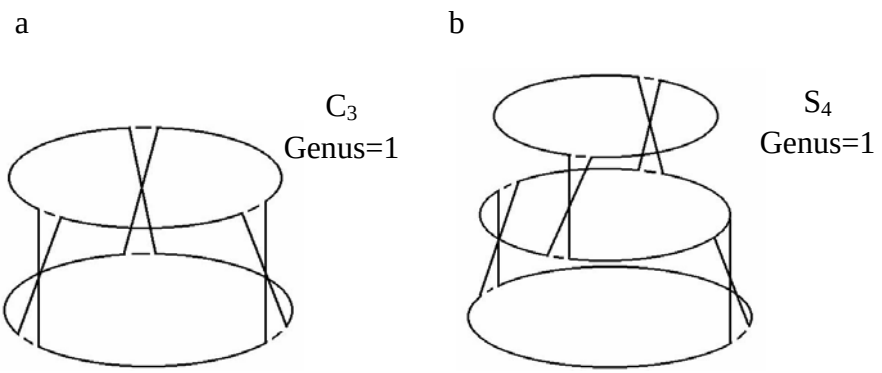

Figure 3.3: The Seifert construction of the knot 3_1 (a) and 4_1 (b)

amphicheiral [34] (achiral) though their genus is the same. How can you tell whether the knot 3_1 and 4_1 are different?

Seifert construction involves a structure that possesses a rigid presentation and led to the definition of another geometrical invariant called the point group [79-81]. This minimal surface of a Seifert construction is almost completely characterized by its point group. The point group is more powerful than the genus, and it is more likely to distinguish knots that are in fact different. Intuitively it is not difficult to see, from Figure 3.3, that the knot 3_1 has the point group C_3 whereas the knot 4_1 has the point group S_4.

The point group is very important in studying molecular structures [82]. In chemistry, a molecule is rigorously and precisely determined by the symmetry of the molecule or of the environment of the atom. Thus, from symmetry considerations alone, we can always tell what the qualitative features of a problem must be. We shall know without any quantitative calculations whatever, how many states there are and what interactions and transitions between them may occur. To put it another way, symmetry considerations alone can give us a completely and rigorous answer to the question “what is possible and what is completely impossible?”. Symmetry can tell us, in principle, that two states of the system must differ in their energy, but only by computation or measurement can we determine how great the different will be.

In general, molecules are chiral if they lack a center, plane, or axis of symmetry. If they have any one of these, they will be superimposable on their mirror image. In particular, we define a completely flexible structure to be geometrically chiral if and only if its Seifert construction has the point groups C_n (n = 1, 2,...). Otherwise, a Seifert construction with S_n symmetry is geometrically achiral. Thus, the knot 3_1 is geometrically chiral and the knot 4_1 is geometrically achiral. If Figure 3.3(b) has at least a single colored point, the knot 4_1 will have no symmetry (S_1’). In this case, the knot 4_1 is said to have a property of a Euclidean rubber glove. Where $S_1' \equiv \sigma'$, we have an operation of reversing through the plane in the Seifert construction, where the top frame is differentiated from the bottom frame and the effect of applying the same reverse operation twice is to bring all vertices into their original positions.

It is certainly intuitively obvious what we mean when we say that some molecules are more symmetric than others are, or that some molecules have high symmetry while others have low symmetry or no symmetry. But in order to make the idea of molecular symmetry as useful as possible, we must develop some rigid mathematical criteria of symmetry. It is also possible to rank the

classes of Seifert constructions by “degree of chirality [69] and achirality”: the higher the symmetry the greater the degree of achirality; the lower the symmetry the greater the degree of chirality. In other words, the point groups C_n ($n \geq 2$) are normally chiral, the point groups S_n (n = 1, 2,…) are normally achiral; whereas the no symmetry C_1 ($C_1 \equiv E$) [82] is the most chiral and the Euclidean rubber glove (S_1’) is the least chiral, and the topological rubber glove is the least achiral.

3.4 Closed Duplex DNA

The study of the closed duplex DNA began in 1963 with its discovery by Dulbecco and Vogt [83]. Vinograd’s group initiated the study of the physical properties of circular DNA in 1963 [84], and by 1965, had demonstrated that circular DNA extracted from cell is negatively supercoiled [85]. Lebowitz [86] described the history of this discovery and several reviews has considered the widespread influence of supercoiling on biological functions (Bates and Maxwell, 1993 [7]; Vologodskii and Cozzarelli [25], 1994; Stasiak,1996 [27]; Drlica, 1992 [87]; Kanaar and Cozzarelli, 1992 [88]; Wang,1992 [89]).

3.4.1 Circular Duplex DNA in the Crystal

There are three basis types of DNA double helix [90]: A-, B-, and Z-DNA. The B- and A-DNA are right-handed whereas the Z-DNA is left-handed. In the Watson-Crick model [1,4] of DNA, the B-DNA has an average of ten base pairs per turn of helix; the A-DNA has closer to 11 base pairs per turn. X-ray diffraction of single crystals [90-92] indicates that B-DNA and A-DNA, respectively, have 10.1 and 11 base pairs per turn (Wing et al, 1980; Conner et al, 1982). In 1979 and 1980, Wang [93] and Drew [94] and their co-workers, respectively, found that the Z-DNA and Z’-DNA have an average of 12 base pairs per turn. Z-DNA was crystallized from low salt and Z’-DNA from high salt concentrations [94]. In 1993, Dickerson’s group [95] reported the crystal structures of the B-DNA helix as having 10.48 and 10.6 base pairs per turn.

Recently, the study of mirror image DNA [96] has attracted attention. Urata et al. reported [97,98] in 1991 that the conformation of L-d (CGCGCG) in solution is the exact mirror-image of the corresponding natural D-hexamer under both low and high salt concentrations, and then in 1993 demonstrated the three-dimensional structure [99] of a racemic duplex containing both the natural left-handed and the unnatural right-handed Z-forms.

In order to completely characterize circular duplex DNA we define a DNA link [100-101] and denote the number of twists and rungs by T_{2n} and R (r), as shown in Figure 3.4 (where n represents the linking number (n = 1, 2,…)). A DNA link with 2n half-twists is a molecule shaped like a ladder which is made to join itself end-to-end, with r-rungs. Here the sides of the DNA link are chains of alternating deoxyribose sugar rings and phosphate groups, and the rungs are purinepyrimidine base pairs which are held together by hydrogen bonds. Attached to the sugar ring of each nucleotide is one of four bases: Adenine (A), guanine (G), thymine (T) or cytosine (C). An A normally on one strand is paired with a T on the other by two hydrogen bonds (A = T) and a G is paired with a C by three hydrogen bonds (G ≡ C).

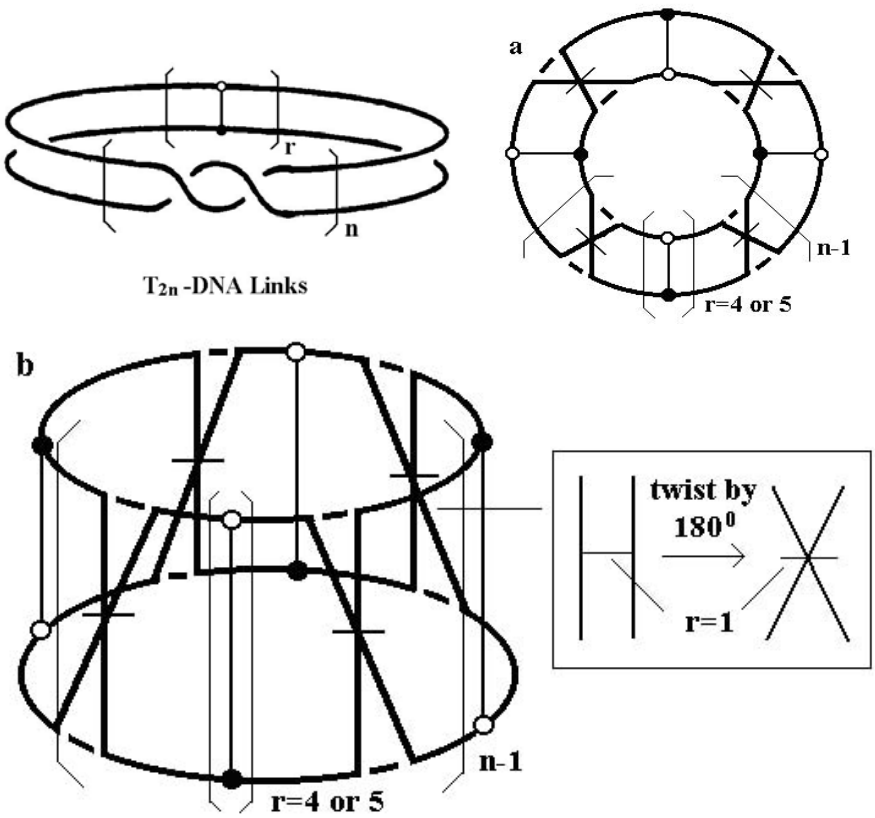

Figure 3.4: The Seifert construction for DNA links

The DNA links can be viewed geometrically as a surface of two components in two-space (Figure 3.4(a)). This surface is almost completely characterized by a Seifert construction and its point group. It is made by using two disks to fill in the loop in the DNA link and using 2*n* ribbons (*n* = 1, 2,…) with one half-twists to fill in the crossing-point areas, and finally connecting the disk with the ribbons in three-space (Figure 3.4(b)). In Seifert construction, the rungs are arranged in the optimally symmetrical position. From the symmetry property of the Seifert construction, we wish to prove that the B-DNA links have point groups C_n the 10-base-pairs whenever occur in a sequence which repeats itself exactly *n* times (*n* = 1, 2,…). Otherwise, they have point group C_1 (no symmetry).

Proof: If *n* = 1 and *R* (*r*) = 10, then the T_2-DNA link has the point group C_1, as shown in Figure 3.5(a). Here $C_1 \equiv E$ and E represents any combination of operations which takes the molecule to a configuration identical with the original one.

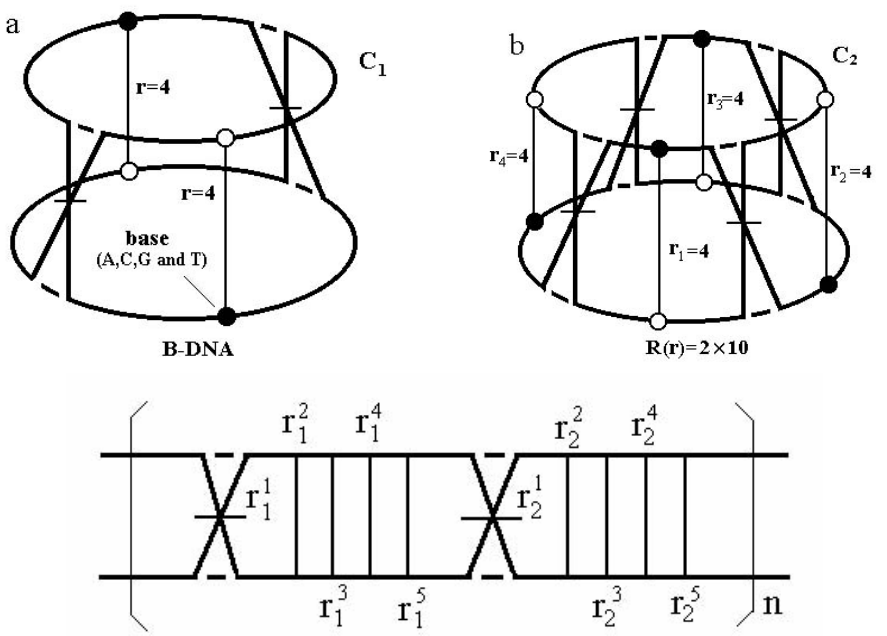

Figure 3.5: The Seifert construction for the T_2- and T_4-DNA links

If $n = 2$ and $R(r) = 2 \times 10$, then the T_4-DNA link (Figure 3.5(b)) has the point group C_2 when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) = (r_3^1, r_3^2, r_3^3, r_3^4, r_3^5)\}$$

$$\text{and } \{ (r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) = (r_4^1, r_4^2, r_4^3, r_4^4, r_4^5)\}\},$$

and has the point group C_1 when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) \neq (r_3^1, r_3^2, r_3^3, r_3^4, r_3^5)\}$$

$$\text{and } \{ (r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) \neq (r_4^1, r_4^2, r_4^3, r_4^4, r_4^5)\}\}.$$

If $n = k$ and $R(r) = 10k$, then the T_{2k}-DNA link (Figure 3.6) has the point group C_k when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) = \ldots = (r^1_{2j+1}, r^2_{2j+1}, r^3_{2j+1}, r^4_{2j+1}, r^5_{2j+1})\}$$

$$\text{and } \{(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) = \ldots = (r^1_{2j'}, r^2_{2j'}, r^3_{2j'}, r^4_{2j'}, r^5_{2j'})\}\}$$

$$(j = 0, 1, 2, \ldots, k\text{-}1; \quad j' = 1, 2, \ldots, k),$$

and has the point group C_1 when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) \neq \ldots \neq (r^1_{2j+1}, r^2_{2j+1}, r^3_{2j+1}, r^4_{2j+1}, r^5_{2j+1})\}$$

$$\text{and } \{(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) \neq \ldots \neq (r^1_{2j'}, r^2_{2j'}, r^3_{2j'}, r^4_{2j'}, r^5_{2j'})\} \}.$$

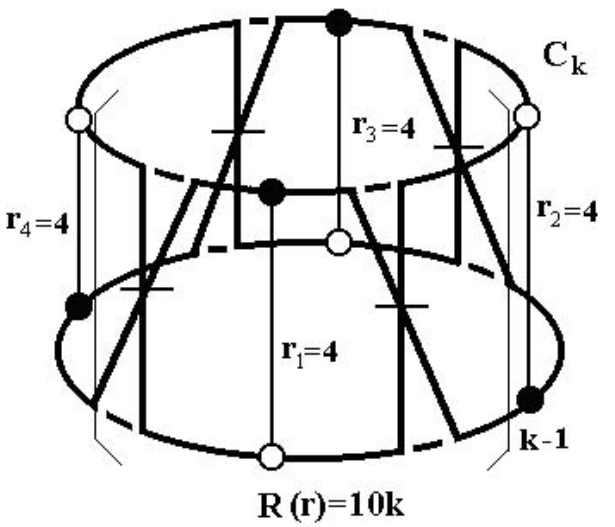

Figure 3.6: The Seifert construction for the T_{2k}-DNA link

Without loss of generality, we can conclude that the T_{2n}-DNA links (R (r) = 10n) have the point groups C_n when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) = \ldots = (r^1_{2j+1}, r^2_{2j+1}, r^3_{2j+1}, r^4_{2j+1}, r^5_{2j+1})\}$$

$$\text{and } \{(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) = \ldots = (r^1_{2j'}, r^2_{2j'}, r^3_{2j'}, r^4_{2j'}, r^5_{2j'})\}\}$$

$$(j = 0, 1, 2, \ldots, n\text{-}1;\ j' = 1, 2, \ldots, n).$$

Otherwise, they have only the point group C_1, and the linking number, n, increases by one (n = 1, 2,…).

We conclude from this proof that the following corollaries are true.

1. A-DNA links (Figure 3.7) have the point groups C_n when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) = \ldots = (r^1_{2j+1}, r^2_{2j+1}, r^3_{2j+1}, r^4_{2j+1}, r^5_{2j+1})\}$$

$$\text{and } \{(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5, r_2^6) = \ldots = (r^1_{2j'}, r^2_{2j'}, r^3_{2j'}, r^4_{2j'}, r^5_{2j'}, r^6_{2j'})\}\}$$

$$(j = 0, 1, 2, \ldots, n\text{-}1;\ j' = 1, 2, \ldots, n).$$

Otherwise, they have only the point group C_1.

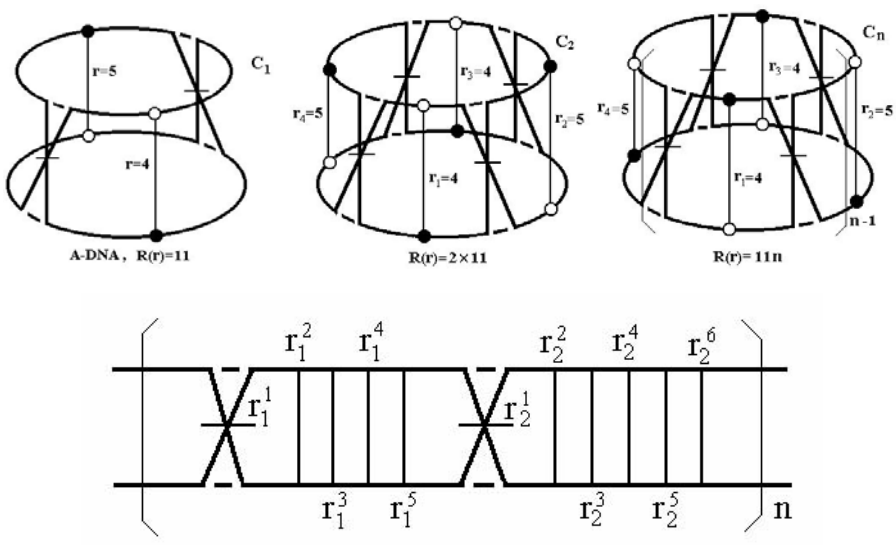

Figure 3.7: The Seifert construction for A-DNA

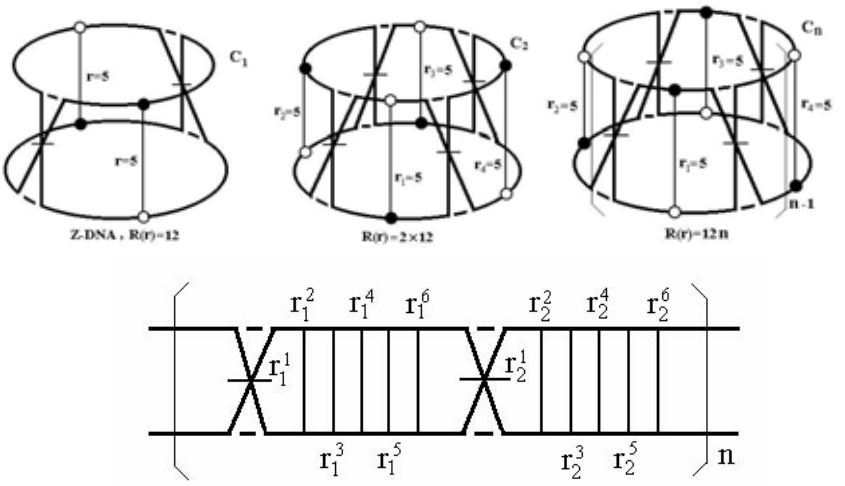

Figure 3.8: The Seifert construction for Z-DNA

2. Z-DNA links (Figure 3.8) have the point groups C_n when

$$\{\{(r_1^1, r_1^2, r_1^3, r_1^4, r_1^5, r_1^6) = \ldots = (r^1_{2j+1}, r^2_{2j+1}, r^3_{2j+1}, r^4_{2j+1}, r^5_{2j+1}, r^6_{2j+1})\}$$

$$\text{and } \{(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5, r_2^6) = \ldots = (r^1_{2j'}, r^2_{2j'}, r^3_{2j'}, r^4_{2j'}, r^5_{2j'}, r^6_{2j'})\}\}$$

$$(j = 0, 1, 2, \ldots, n\text{-}1;\ j' = 1, 2, \ldots, n).$$

Otherwise, they have only the point group C_1.

3. B'-DNA links (Figure 3.9) have the point groups $C_{n/2}$ when

$$\{\{(r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) = \ldots = (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r^1_3, r^2_3, r^3_3, r^4_3, r^5_3, r^6_3) = \ldots = (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3}, r^6_{4j+3})\},$$

$$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2) = \ldots = (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\}$$

$$\text{and } \{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}\ \}$$

$$(j = 0, 1, 2, \ldots, m - 1;\ j' = 1, 2, \ldots, m).$$

Otherwise, they have only the point group C_1, and the linking number increases by two ($n = 2m$, $m = 1, 2, \ldots$). In 1978 the mathematician Fuller [14] first noted the occurrence of causing even-number changes in the linking number.

In general, it is possible to synthesize artificial DNAs [90] with the point groups C_n. In particular, natural DNA sequences do not satisfy a condition of

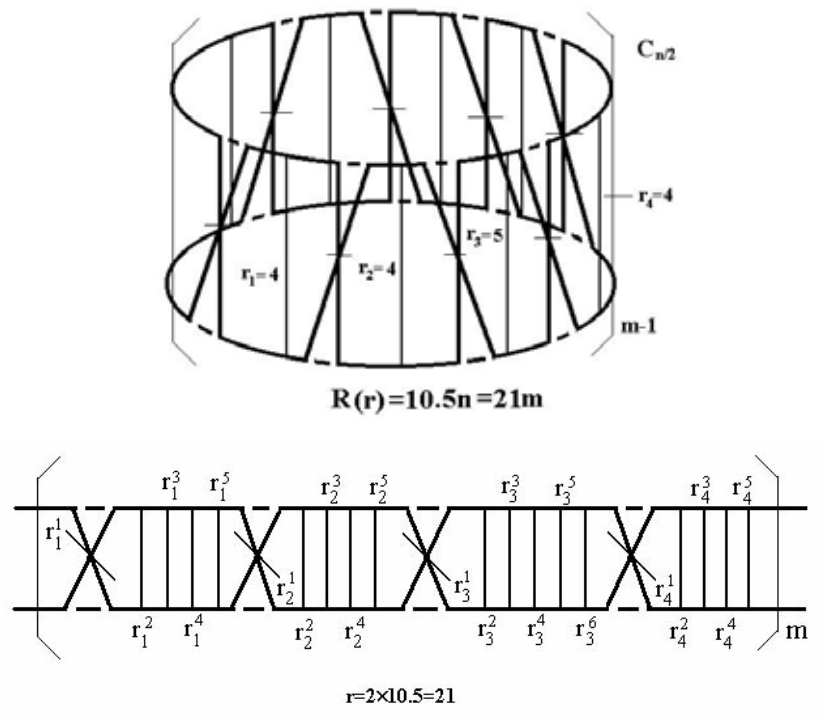

Figure3. 9: The Seifert construction for B'-DNA (*R*(*r*) = 10.5*n* = 21*m*)

regular repetition and have only a point group C_1. Thus, DNA links have some features which are regular, and some which are irregular. It is obvious that the C_n ($n \geq 2$) symmetries are very regular and the C_1 symmetry is completely irregular. The fusion of regular and irregular features is achieved admittedly only at the expense of the symmetry conformation. Such phenomena ($C_{n/2}$ or $C_n \rightarrow C_1$) are said to have the properties of intramolecular symmetry breaking, the intramolecular symmetry breaking being a mechanism in which a system by itself goes from a highly symmetric conformation to a nonsymmetric one.

On the other hand, the point groups $C_{n/2}$ or C_n tell us that the T_{4n}-DNA links are chiral and the twin forms are often distinguished by calling one "right" and other "left". No amount of inspection or measurement of one will disclose a property not possessed by the other, yet the two are clearly quite different. The numbers of right- and left-handed forms will be equal, in principle, and the state will be chirally symmetric. However, these processes are not observed in the natural DNA. Ordinarily, the natural left-handed B and A helix as well as the right-handed Z helix do not appear in the crystal. We cannot move the right-handed B and A helix within three-space to create the left-handed B and A

helix, and also cannot move the left-handed Z helix so as to create the right-handed Z helix. The symmetry between left- and right-handed DNA is broken spontaneously, the numbers of left- and right-handed forms are unequal, and the state is therefore chirally asymmetric. Such phenomena are said to have the property of intermolecular symmetry breaking [62]. The significance of intermolecular symmetry breaking is that mirror symmetry is always absent in natural B-DNA, A-DNA and Z-DNA, etc. The natural B-, A-, and Z-DNAs usually do display a preference for one kind of chirality over another (Figure 3.10). Furthermore, right- and lift-handed DNA must differ with regard to physical properties, such as energy, that depend on their handedness.

The situation is now clear about circular duplex DNAs being geometrically chiral. Yet we cannot move the right-handed DNA within three-space to create the left-handed DNA, and vice versa. How then does the chirality arise? How might intermolecular symmetry breaking have arisen in living systems? We shall attempt to answer this question, insofar as it is possible to answer it, beginning at the level of particles in solution.

3.4.2 The Hydration Structure of B-DNA

Water is an important factor in helix structure (for reviews, see Dickerson *et al.*, 1982 [102]; and Dickerson, 1983 [90]; Berman, 1991 [103]; 1994 [104]). Solution measurements [105-106] and theoretical calculation [107] show that

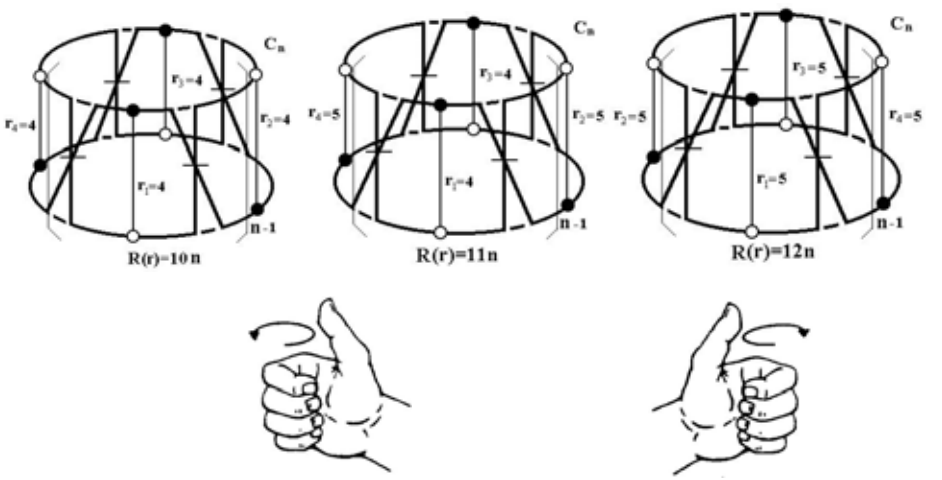

Figure 3.10: The natural B-, A-, and Z-DNAs usually display a preference for one kind of chirality over another.

for DNA molecules in solution the strands of the double helix make a full turn every 10.5 base pairs (Wang, 1979; Rhodes and Klug, 1980; Levitt, 1978).

The water structure has been examined in molecules of all three forms of the DNA double helix [90]. Detailed analyses of the water structure in crystalline DNA have been reported mainly from high-resolution studies of B-DNA helices [108] (Drew and Dickerson, 1981). The hydrating molecules are invisible to X-ray analysis [102], which can see only the average structure over all molecules, and hence only the ordered solvent positions. But cooling the crystals to 16 K [109], or transferring CGCGAATTCGCG crystals to 60 percent MPD (methylpentanediol), lowers the temperature factors and produced localized hydration along the backbone [102]. It is clear that the helix of B-DNA in solution is coated with a layer of water molecules [90], as showed in Figure 3.11.

In recent years the interest in hydration of DNA has increased and the role of water in DNA conformational transitions was well understood. Detailed studies on A-DNA hydration have been reported (Kennard *et al*., 1986 [110]; Eisenstein *et al*., 1990 [111]; 1995 [112]). Numerous experimental and theoretical attempts have been made to quantify hydration for different DNA conformations, and some of the several newer results can be found in Refs. [113-117] (e.g., Schneider *et al*., 1993[113]; Lipscomb *et al*., 1994 [114]; Schneider and Berman, 1995 [115]; Hummer *et al*., 1995 [116]; Denisov *et al*.,

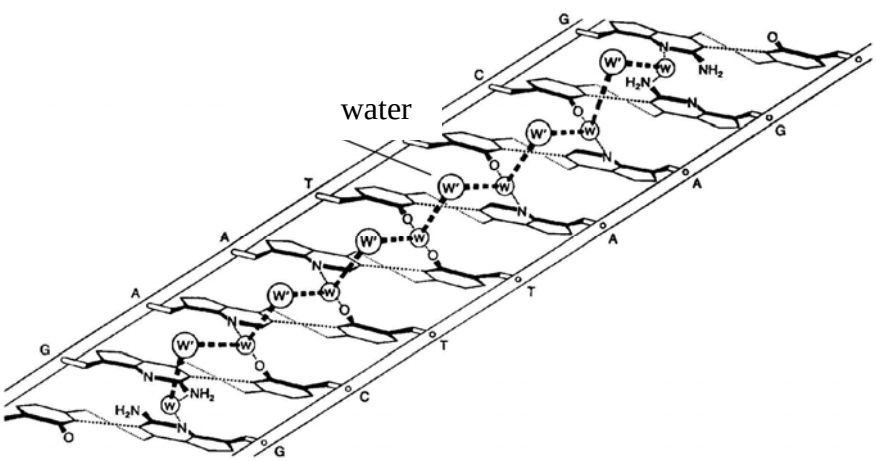

Figure 3.11: The hydration structure of B-DNA double helix coated with a layer of water molecules. Figure adapted from Ref. [90].

1997 [117]). No comparable systematic water structure has been found around either the A or the Z helix. Although water molecules are distributed liberally around atoms on the DNA that could take part in hydrogen bonds, there is nothing that is comparable to the minor groove spine of hydration in B-DNA and that could be expected to have structural integrity [90].

The Seifert constructions for DNA hydration structures are made by using the third disk in the middle of the previous Seifert constructions in the crystal to fill in the loop which represents a layer of water molecules [100-101]. Pictorially we represent this as in Figure 3.12. We wish to prove that the B-DNA hydration structure has the properties of a Euclidean rubber glove if the 21-base pairs occur in a sequence which repeats itself exactly m (or $n/2$) times ($n = 2m$, $m = 1, 2,\ldots$).

We can prove that the structures in Figure 3.13 and 3.14 have the point groups $S_{n/2}$ (or S_m) if they satisfy the following conditions:

1. There are a $C_{n/2}$ symmetric axis for the T_{2n} hydration structures when

$$\{\{(r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) =\ldots= (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r^1_3, r^2_3, r^3_3, r^4_3, r^5_3, r^6_3) =\ldots= (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3}, r^6_{4j+3})\},$$

$$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2) =\ldots= (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\}$$

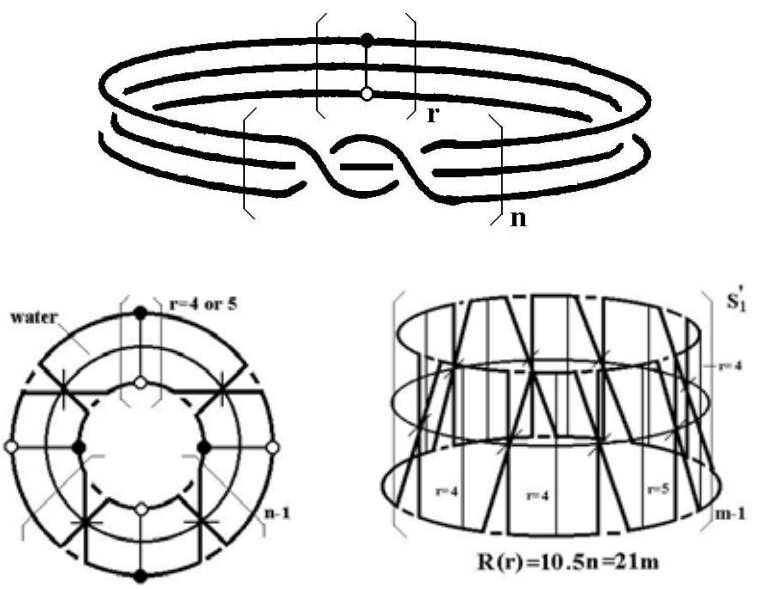

Figure 3.12: The Seifert constructions for the B-DNA hydration structure

and $\{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}$ }

($j = 0, 1, 2, \ldots, m - 1; j' = 1, 2, \ldots, m$).

Here the linking number increases by two ($n = 2m$, $m = 1, 2, \ldots$).

2. There is a symmetry plane when the vertices are the non-Watson-Crick base pairs. This means that the bases A, C, G and T on one strand must be in one-to-one correspondence with A, C, G and T on the other.

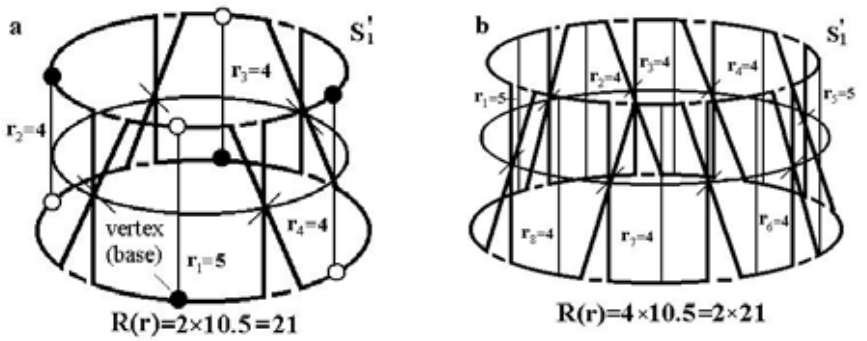

Figure 3.13: The Seifert construction for the T_4- and T_8-hydration structures

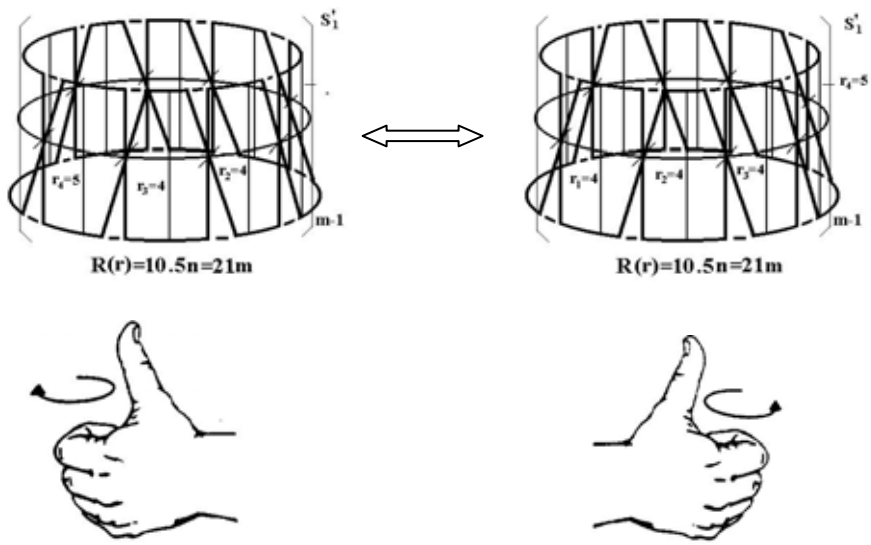

Figure 3.14: The Seifert construction and its symmetric state for the T_{2n}-hydration structures

However, the second condition above is not satisfied because two strands of DNA are self-complementary: Adenine (A) normally is paired with thymine (T) and guanine (G) is paired with cytosine (C). Here we show that objects with non-Watson-Crick base pairs have the point groups $S_{n/2}$, while the others with Watson-Crick base pairs have no symmetry but possesses a reverse operation ($S_1' = \sigma'$) which makes the top frame in the Seifert construction into a bottom frame. We have thus shown that the B-DNA hydration structure is a Euclidean rubber glove.

The Euclidean rubber glove tells us that the B-DNA hydration structure is the least chiral and the state will be achiarlly symmetric. We can now move the right-handed B helix within three-space to create the left-handed B helix in solution. One would expect equal numbers of right-handed and left-handed B helix to inhabit the living systems. Yet DNA usually displays a preference for chirality over achirality [118]. For some time it has seemed possible that the enzymes, designed to act on left-handed molecules, could not copy right-handed ones although the least chiral object and its mirror image are the nearly same.

According to this view, the hydration structure of B-DNA developed as a singular event, and did not possess the strongly chiral chemistry characteristic of modern life and so incorporated only a slight chiral asymmetry. It must be pointed out that the loop constructed by H_2O in the Seifert construction is an interface for the exchange of information and reaction with the environment [119]. The interface is unstable because it may not only disappear but can also be reconstructed by changed conditions.

3.4.3 Structural Transition in DNA

The observed solvent positions in three crystal structures have confirmed earlier fiber and solution measurements, and have led to proposals that have explained the transitions from B to A and from B to Z [120] helices. Two characteristic X-ray fiber patterns are observed when stretched fibers are dried [102]: the B pattern around 92 percent relative humidity (RH) and the A pattern when the fiber is dried to 75 percent RH in the absence of salt. Further dying below 55 percent RH leads to increasing disorder and deterioration of the quality of the diffraction pattern. The B form is stabilized by salt: a 10 percent salt content by weight is sufficient to prevent appearance of the A pattern no matter how low the RH, but with a salt content of 0.4 percent or less the A form can be observed to persist all the way up to 98 percent RH. One now has a clear molecular picture of the reasons for the B form's prevalence under most

general aqueous conditions [90] and for the transition to the A form as a result of dehydration [121].

The B-to-A helix conversion [122] takes place in a narrow humidity range, with a midpoint between 75 and 83 percent RH. Beyond about 80 percent RH, the grooves are completely filled with liquid water so that further hydration causes the fibers to swell [102]. In any case, high salt favors the transition of B to Z [123,124] but not of B to A. Under special conditions of stress, right-handed DNA with the purine-pyrimidine alternating sequence can be flipped over into the left-handed Z state [90]. Behe and Felsenfeld's results have indicated [123] that the spectral transition from the B to the Z form occurs at much lower salt concentrations, close to the usual physiological conditions. Peck and Wang's results shown [124] that negative supercoiling of the DNA changes the alternating C-G sequence from a right-handed helix with 10.5 base pairs (bp) per turn to a left-handed helix with 11.6 bp per turn under physiological conditions.

It is also interesting to note that the B-, A-, and Z-DNA differ in the number of rungs when the linking number is fixed. Accordingly, the difference between the two states is called the rung difference, i.e.

$$\Delta r = R_2(r) - R_1(r).$$

Where $R_1(r)$ and $R_2(r)$ are the numbers of rungs (base pairs) in the first and the second state, respectively. Supercoiling [16,17] is a precondition for DNA transition, replication and recombination. A ring for which $R(r) = 10.5n$ is said to be relaxed. Increasing or decreasing the ratio strains the double helix, which responds by supercoiling [16]. Negatively supercoiled (left-handed) [125] DNA rings are known to be common in intact cells, but positively supercoiled (right-handed) rings have been made only in the laboratory [17]. It should be noted that if one stranded of a supercoiled ring of either kind is cut, the molecule returns to the relaxed state.

In order to understand how the number of rungs influences the configuration and physical properties, it is necessary to appreciate the mechanism of enzymes and ethidium molecules (for reviews, see Bauer *et al*., 1980 [16]; and Wang: 1982 [17]; 1985 [126]; 1996 [127]). The enzymes can convert rings of DNA from one topological form to another and a single topoisomerase molecule can carry out the complete operation of breaking and resealing. There are two kinds of topoisomerase: one kind cuts a single strand of DNA; the other cuts two strands simultaneously. Ethidium, a planar molecule, slides between two base

pairs when it binds to the double helix. Each molecule of bound ethidium unwinds the double helix by about 26 degree [17], and so one or two ethidium molecules move the number of rungs by one. This is because the mean helical twist angle [90] from one base pair to next for A- and B-DNA is 33.1 and 35.9 degrees, and for Z-DNA is –51.3 (G-C) and –8.5 degrees (C-G), respectively.

We now suggest two mechanisms, shown in Figure 3.15 and 3.16, to describe the transitions first from B to A and secondd from B to Z helices, respectively. There are some clear instances in which the least chiral symmetry causes the most chiral asymmetry. These results lead us to infer that the chirality of DNA is an artifact of the life process. Further, it also provides a sound basis for the investigation of key questions about the origin of chiral

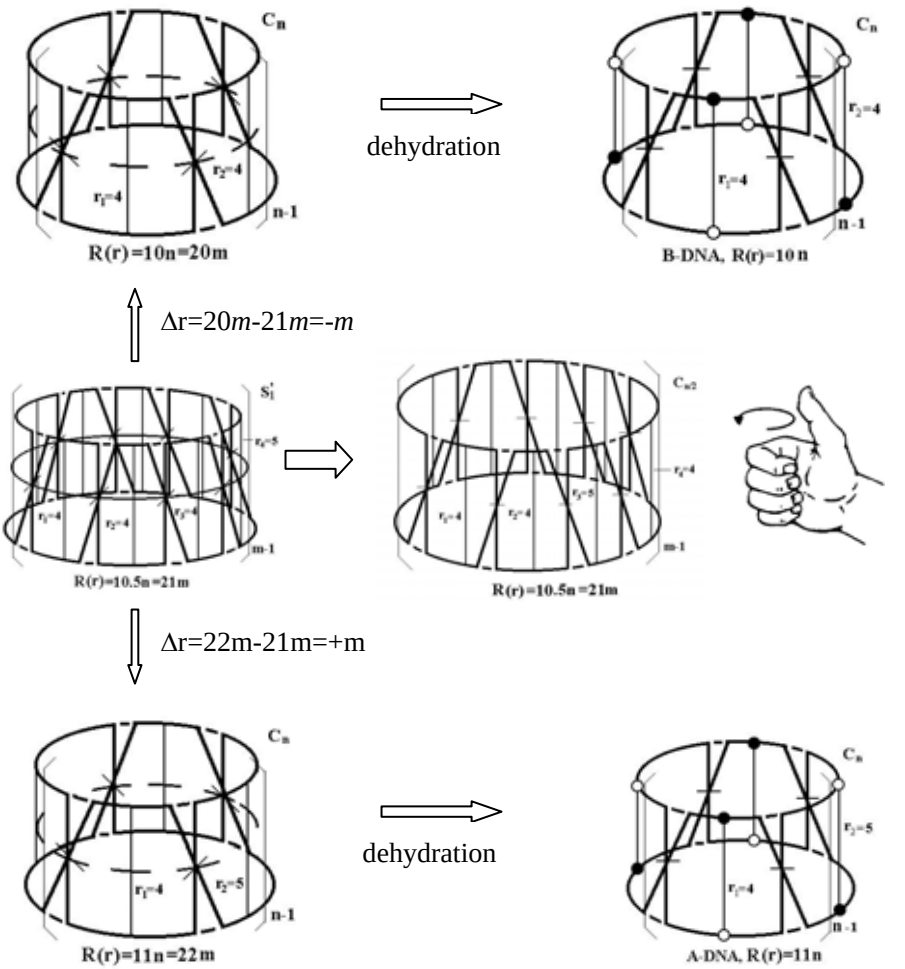

Figure 3.15: Topological conversions for the right-handed helix

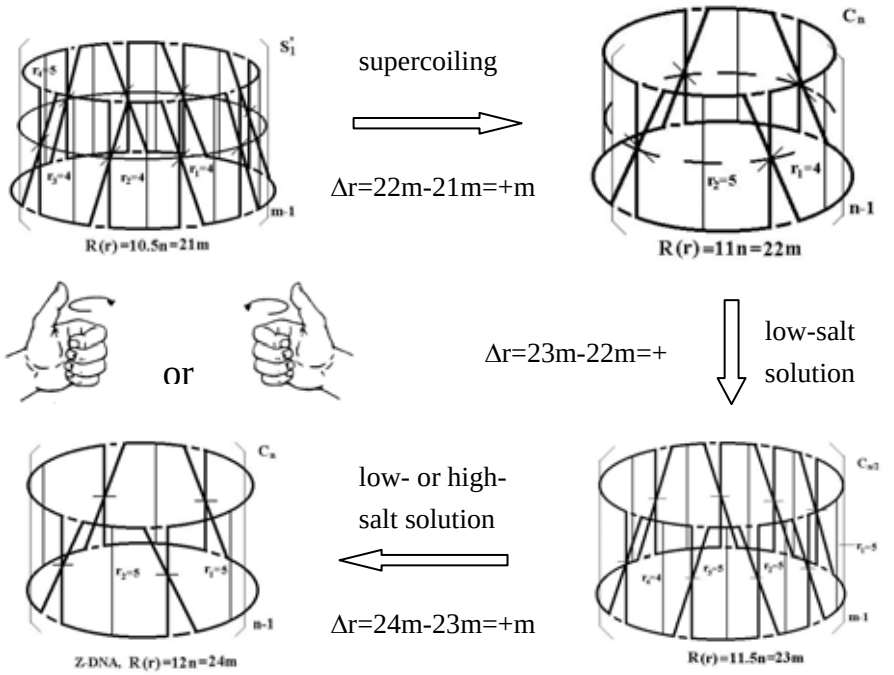

Figure 3.16: Topological conversions for the left-handed and right-handed Z helix A: left-handed B-DNA→left-handed Z-DNA; B: right-handed B-DNA→right-handed Z-DNA

asymmetry in life. We now have a clear molecular picture of the reasons for B-DNA stabilization by hydration and for the transition to the B- and A-DNA crystal forms as a result of dehydration with DNA supercoiling (Figure 3.15). It is worthwhile to note that the transitions from B to Z helices go through a three-step process (Figure 3.16). Our results also reveal the important fact that the structural phase transition takes place in DNA links when symmetries are changed from S_1' to C_1 or C_n and that geometrical chirality remains unchanged when rungs are cut down (Figure 3.17) by chemical reaction in DNA links. Intermolecular symmetry breaking is a mechanism by which a DNA link, acted upon by enzymes and ethidium molecules, goes from a symmetric state (Figure 3.14) to an asymmetric one (Figure 3.15, 3.16 and 3.17). The symmetric state is one with equal numbers of left- and right-handed forms; the asymmetric state is one in which one form dominates.

3.5 Single-Stranded DNA Knots

The observation of a knotted single-stranded DNA ring was first made by

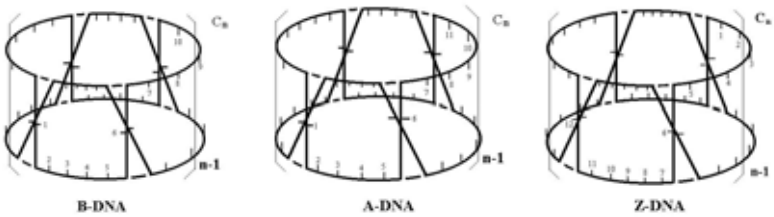

Figure 3.17: The Seifert constructions for the B-, A- and Z-DNA in which rungs are cut down by chemical reaction

Liu, Depew and Wang in 1976 [128], when circular fd DNA was treated with *E.coil* ω. In recent years, Seeman's group [129-135] has demonstrated in a most impressive manner how DNA can be used for the construction of topologically complex non-natural structures. The design and realization of a synthetic knot comprised of a single-stranded DNA molecule was described in 1991 [129]. Exciting advances have recently been made in the study and control of single-stranded DNA knots [134,135], e.g., knot 3_1 and knot 4_1. It was found that the trefoil knots formed by single-stranded DNA have two forms, containing either a left-handed form or a right-handed form. Seeman and his co-workers[132] have synthesized molecules containing 104,96,88,80,74,70 and 66 nucleotides. Helix repeats of DNA double helices are very well understood, e.g., S-, D-, C-, B-, A-, and Z-DNA have about 6, 8, 9, 10,11 and 12 base pairs for each turn (360°) of the double helix, respectively (for a review, see Leslie *et al*., 1980 [136]). But how many bases per half-twist (180°) do DNA knots have in solution? We shall try to answer this question from the symmetry perspective.

Doubled knots with q half-twists were introduced by Whitehead [137] in 1937 and formed an interesting class of knots with respect to certain invariant (Figure 3.18). It is interesting to note that the single-stranded DNA knots can be described as double knots [138]. Single-stranded DNA knots can be divided into two types depending on whether the twisting number is $2n + 1$ or $4n$. For convenience, we denote the number of twists and vertices by T_{2n+1}, T_{4n} (n = 1, 2,…) and V_p (nucleotide subunites), respectively. Our goal here is to construct some point groups pertaining to the single-stranded DNA knots on the basis of the Seifert construction in knot theory [77].

3.5.1 T_{2n+1} DNA Knots

The Seifert constructions for the T_{2n+1} single-stranded DNA knots can be made by using two disks to fill in two loops in the knot and using $2n$ +1 ribbons

left-handed right-handed

Figure 3.18: Doubled knots with *q* half-twist

with one half-twist to fill in the crossing-point areas, and then connecting the disks with the ribbons in three-space. It is enough to prove that the T_{2n+1} single-stranded DNA knots have point groups C_{2n+1} when

$$V_p = \sum_{j=0}^{2n} E_{4j+1}(p_1) + \sum_{j=0}^{2n} E_{4j+2}(p_2) + \sum_{j=0}^{2n} E_{4j+3}(p_3) + \sum_{j'=1}^{2n+1} E_{4j'}(p_4)$$

$$= \begin{cases} (2n+1)4I & \text{if } p_1 = p_2 = p_3 = p_4 = 1, 2, \ldots, 10;\ I = 1, 2, \ldots, 10. \\ (2n+1)(4I+1) & \text{if } p_1 = p_2 = p_4 = 1, 2, \ldots, 9; \\ & \quad p_3 = (p_1 + 1) = 2, 3, \ldots, 10;\ I = 1, 2, \ldots, 9. \\ (2n+1)(4I+2) & \text{if } p_1 = p_3 = 1, 2, \ldots, 9; \\ & \quad p_2 = p_4 = (p_1 + 1) = 2,3,\ldots, 10;\ I = 1,2,\ldots, 9. \\ (2n+1)(4I+3) & \text{if } p_1 = p_2 = p_4 = 2, 3,\ldots, 10; \\ & \quad p_3 = (p_1 - 1) = 1, 2,\ldots, 9;\ I = 1, 2,\ldots, 9. \end{cases}$$

Otherwise, they have the point group C_1. Where E represents the segment edge on the Seifert construction and p represents the number of vertices (bases) on the segment edge. Some examples are given in Figure 3.19.

In the case of Figure 3.19(a), if

$E_1(p_1) = E_5(p_1) = E_9(p_1)$, $E_2(p_2) = E_6(p_2) = E_{10}(p_2)$,

$E_3(p_3) = E_7(p_3) = E_{11}(p_3)$ and $E_4(p_4) = E_8(p_4) = E_{12}(p_4)$,

then the T_3 DNA knot has the point group C_3. Hence, the knot 3_1 satisfies the following conditions:

1. $p_1 = p_2 = p_3 = p_4$ and $p_1 = 1, 2, \ldots, 10$, and $V_p = 12, 24, 36, 48, 60, 72, 84, 96, 108, 120$.

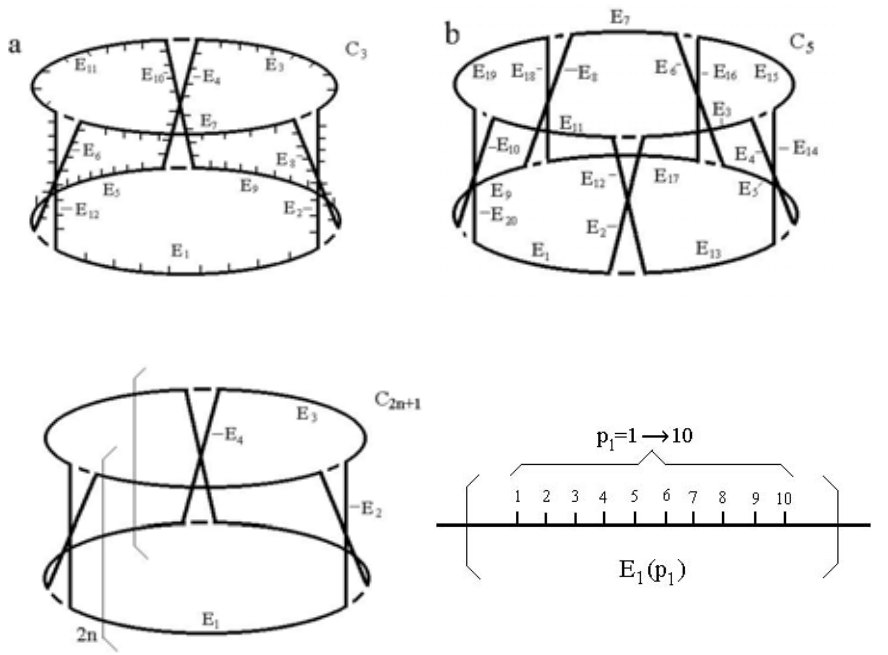

Figure 3.19: The Seifert construction for the T_{2n+1} DNA knots that have the point groups C_{2n+1}

2. $p_1 = p_2 = p_4$, p_1=1,2,…,9 and $p_3 = (p_1 + 1) = 2, 3, \ldots, 10$, and $V_p = 15, 27, 39, 51, 63, 75, 87, 99, 111$.
3. $p_1 = p_3 = 1, 2, \ldots, 9$ and $p_2 = p_4 = (p_1 + 1) = 2, 3, \ldots, 10$, and $V_p = 18, 30, 42, 54, 66, 78, 90, 102, 114$.
4. $p_1 = p_2 = p_4$, $p_1 = 2, 3, \ldots, 10$ and $p_3 = p_4 = (p_1 - 1) = 1, 2, ..., 9$, and $V_p = 21, 33, 45, 57, 69, 81, 93, 105, 117$.

Otherwise, the T_3 DNA knot has the point group C_1.

These results tell us that the trefoil knots containing 66 and 96 nucleotides have C_3 symmetry and that the trefoil knots containing 70,74,80,88 and 104 nucleotides have C_1 symmetry. This is the reason the 96 nucleotides ($p_1 = p_2 = p_3 = p_4 = 8$ and $3 \times 32 = 96$) generates the best yield for trefoil knot (87%) [132]. Thus, there are about 11 and 16 bases per half-turn for the 66- and 96-mers, respectively. However the trefoil knots, which have the 10,12,13,14 and15 bases per half-turn ($p_1 + p_2$) for the 60-, 72-, 78-, 84- and 90-mers respectively, are a target for the future experimental syntheses.

3.5.2 T_{4n} DNA knots

The Seifert construction for T_{4n} knots can be made by using three disks to fill in two loops in the knot and using 4n ribbons with one half-twist to fill in the crossing-point areas, and then connecting the disks with the ribbons in three-space. It is safe to say that the T_{4n} single-stranded DNA knots have point groups S_{4n} when

$$V_p = \sum_{j=0}^{4n-1} E_{4j+1}(p_1) + \sum_{j=0}^{4n-1} E_{4j+2}(p_2) + \sum_{j=0}^{4n-1} E_{4j+3}(p_3) + \sum_{j'=1}^{4n} E_{4j'}(p_4)$$

$$= \begin{cases} 4n \times 4I = 16nI & \text{if } p_1 = p_2 = p_3 = p_4 = 1, 2, \ldots, 10;\ I = 1, 2, \ldots, 10. \\ (4n)(4I+1) & \text{if } p_1 = p_2 = p_4 = 1, 2, \ldots, 9; \\ & \quad p_3 = (p_1+1) = 2, 3, \ldots, 10;\ I = 1, 2, \ldots, 9. \\ (4n)(4I+2) & \text{if } p_1 = p_3 = 1, 2, \ldots, 9; \\ & \quad p_2 = p_4 = (p_1+1) = 2, 3, \ldots, 10;\ I = 1, 2, \ldots, 9. \\ (4n)(4I+3) & \text{if } p_1 = p_2 = p_4 = 2, 3, \ldots, 10; \\ & \quad p_3 = (p_1-1) = 1, 2, \ldots, 9;\ I = 1, 2, \ldots, 9. \end{cases}$$

Otherwise, they have no symmetry (S_1'). Some examples are given in Figure 3.20.

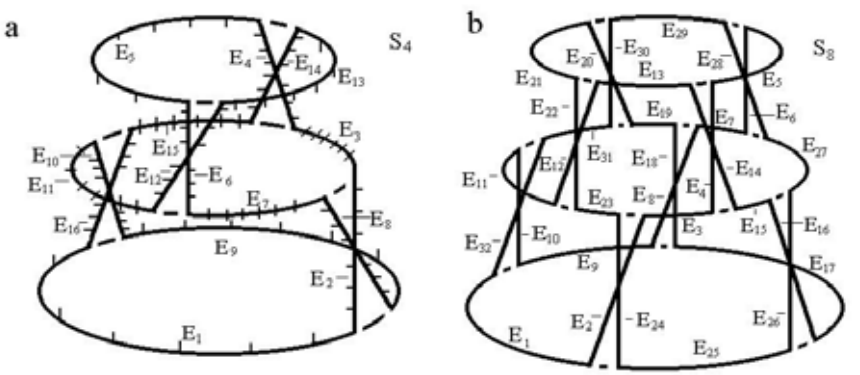

Figure 3.20: The Seifert construction for the T_4 and T_8 DNA knots

In Figure 3.20(a), if the sets

$\{ E_1 (p_1)=E_5 (p_1)=E_9 (p_1)=E_{13} (p_1), E_2 (p_2)$
$=E_6 (p_2)=E_{10} (p_2)=E_{14} (p_2), E_3 (p_3)=E_7 (p_3)=E_{11} (p_3)=E_{15} (p_3)$
and $E4 (p_4)=E_8 (p_4)=E_{12} (p_4)=E_{16} (p_4) \}$

are valid, then the T_4 DNA knot has the point group S_4. The S_4 symmetry must take the form:

1. $p_1 = p_2 = p_3 = p_4$, $p_1 = 1, 2, \ldots, 10$, and $V_p = 16, 32, 48, 64, 80, 96, 112, 128, 144$.
2. $p_1 = p_2 = p_4$, $p_1 = 1, 2, \ldots, 9$, $p_3 = (p_1 + 1) = 2, 3, \ldots, 10$, and $V_p = 20, 36, 52, 68, 84, 100, 116, 132, 148$.
3. $p_1 = p_3 = 1, 2, \ldots, 9$, $p_2 = p_4 = (p_1 + 1) = 2, 3, \ldots, 10$, and $V_p = 24, 40, 56, 72, 88, 104, 120, 136, 152$.
4. $p_1 = p_2 = p_4$, $p_1 = 2, 3, \ldots, 10$, $p_3 = (p_1 – 1) = 1, 2, \ldots, 9$, and $V_p = 28, 44, 60, 76, 92, 108, 124, 140, 156$.

Otherwise, the T_4 DNA knot has no symmetry (S_1').
The above results tell us that the figure-eight knots containing 80, 88, 96 and 104 nucleotides have S_4 symmetry and the figure-eight knots containing 66, 70 and 74 nucleotides have no symmetry (S_1'). This is the reason the 96 nucleotides ($p_1 = p_2 = p_3 = p_4 = 6$ and $4 \times 24 = 96$) generate the best yield for figure-eight knot (71%) [132]. Thus, there are about 10, 11, 12 and 13 bases per half-turn ($p_1 + p_2$) for the 80-, 88-, 96-, and 104-mers, respectively.

As has been said above, the T_{2n+1} DNA knots have the point groups C_{2n+1} and are chiral, while the T_{4n} DNA knots have the point groups S_{4n} and are achiral. The others, C_1 or S_1', are the most chiral or the least chiral.

On the other hand, if V_p is the same but the bases are different in equivalent positions, e.g., if

$\{p_1^1,p_1^2,p_1^3,p_1^4,\ldots \neq p_2^1,p_2^2,p_2^3,p_2^4,\ldots \neq p_3^1,p_3^2,p_3^3,p_3^4,\ldots \neq p_4^1,p_4^2,p_4^3,p_4^4,\ldots\}$,

then the T_{2n+1} and T_{4n} DNA knots have no symmetry and are C_1 and S_1', respectively.

3.6 Duplex DNA Knots

Duplex DNA knots are well known in molecular biology [7,139-143] and many excellent discussions have appeared concerning this subject from Cozzarelli's group [144-151]. In 1983 Krasnow *et al*. found [145] that the trefoil knots formed by duplex DNA can have two forms, containing either left-handed supercoils or right-handed supercoils. In 1985 Spengler *et al*. [149] determined the structure of the duplex DNA knots 3_1, 5_1, 7_1, 9_1, 11_1, 13_1, 15_1, 17_1, 19_1, 21_1, 23_1, etc. More recently, Rybenkov *et al*. [151] shown that duplex DNA knots are formed below equilibrium values by type Ⅱ topoisomerases. Shaw and Wang found [152] that supercoiling of a DNA trefoil perturbs differently the spatial writhe of its two chiral forms. Also, Stasiak *et al*. [153-155] reported the result of computer simulations which show that there is a linear relationship between the speeds of migration of different types of duplex DNA knots and the average crossing numbers of their ideal geometrical representations.

Topological structures of duplex DNA knots [156] have been investigated as an interesting class of superstructures [31], with the aim of understanding at a fundamental level the ways in which simple components come together to form larger and more complex assemblies and arrays. To solve the problem, we shall first divide duplex DNA knots into two sections depending on whether the twisting number is T_{2n+1} or T_{4n} (n = 1, 2,…), then introduce Seifert constructions for the duplex DNA knots, and finally discuss the implications of the model for real DNA. The molecular Seifert construction and its molecular symmetry may help us better understand and predict sequence/structure relationships, and open the door to controlling the supermolecular design and assembly of duplex DNA knots.

3.6.1 T_{2n+1} Duplex DNA Knots

The Seifert construction is a good mathematical model of the concrete molecular knot and is defined as a minimal surface with the lowest genus. The duplex DNA knots are considered as embedded in the Seifert construction in the most natural way, and so they have extremely simple, regular point groups that characterize them almost completely. Now is surely a time of special opportunity to deepen our fundamental knowledge of why and how knotting phenomena take place in living organisms.

In order to explain our results we first define some terms as follows: the total linking number of duplex DNA is defined as the product of the twisting

number of the duplex DNA knots and the local linking number of the duplex DNA which is distributed in steps of two. The local twisting number in the double helix is changed by four or by multiples of four, this being because the linking number is changed by two or by multiples of two [14,17] in topological reactions. We know that the linking number multiplied by 2 gives the twisting number [7], and that the local linking number of duplex DNA is controlled by the twisting number of supermolecular knots. This can be expressed as follows: $L_{\text{total}} = 2(2n + 1)m$, $T_{\text{w}} = (2 \times 2m)(2n + 1) = 4(2n + 1)m$ (n = 1, 2,…; m = 1, 2,..., ∞), and also visualized as in Figure 3.21. As described below, the total linking number (L_{total}) satisfies the topological requirement for Seifert construction and the total twisting number is defined as the number of generators on a minimal surface. Further, the total base pairs number in supermolecular knots is defined as $R\ (r) = 10.5L_{\text{total}} = 21(2n + 1)m$ or $R\ (r) = 10.0L_{\text{total}} = 20(2n + 1)m$, considering that there are 10.5 or 10.0 base pairs for each turn of the double helix.

The T_{2n+1} duplex DNA knots can be viewed geometrically as the double edges, or double boundaries, of a two-dimensional surface. The molecular Seifert constructions are assembled by using two disks to fill in two loops in the supermolecular knots (e.g., 3_1, 5_1, 7_1, 9_1,…) [33] and then using $2n + 1$ ribbons (n = 1, 2,...) with one half-twist to fill in the crossing-point areas, and finally connecting the disks with ribbons in three-space. Double edges represent two strands of DNA double helix coated with a layer of water molecules, and are made by using $4(n + 1)m$ (n = 1, 2,…; m = 1, 2,...) ribbons with one half-twist to fill in the crossing-point areas in the double helix.

We shall use mathematical induction to prove that the T_{2n+1} duplex DNA knots have either C_{2n+1} symmetry or C_1 symmetry for all positive integer n when

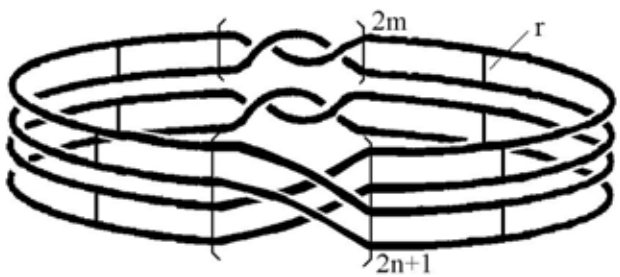

Figure 3.21: Cartoon representation of the T_{2n+1} duplex DNA knots.Here $2n+1$ represents the twisting number of a supermolecular knot, $2m$ represents the local linking number of a duplex DNA (m=1,2,…, ∞), and "r" is a collection of the base pairs which are held together by hydrogen bonds..

the 21- and 20-base-pairs occur in a sequence which repeats itself exactly $2n+1$ times.

For example, if $n = 1$, $L_{total} = 6m$, $T_w = 12m$ and $R(r) = 63m$ (Figure 3.22), then the T_{2n+1} duplex DNA knots have C_3 symmetry because they can be rotated three times when

$\{\{ (r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) = (r^1_5, r^2_5, r^3_5, r^4_5, r^5_5) = (r^1_9, r^2_9, r^3_9, r^4_9, r^5_9)\}$,

$\{(r^1_3,r^2_3,r^3_3,r^4_3,r^5_3,r^6_3) = (r^1_7,r^2_7,r^3_7,r^4_7,r^5_7,r^6_7) = (r^1_{11},r^2_{11},r^3_{11},r^4_{11},r^5_{11},r^6_{11})\}$,

$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2) = (r^1_6, r^2_6, r^3_6, r^4_6, r^5_6) = (r^1_{10}, r^2_{10}, r^3_{10}, r^4_{10}, r^5_{10})\}$

and $\{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4)=(r^1_8, r^2_8, r^3_8, r^4_8, r^5_8)=(r^1_{12}, r^2_{12}, r^3_{12}, r^4_{12}, r^5_{12})\}\}$.

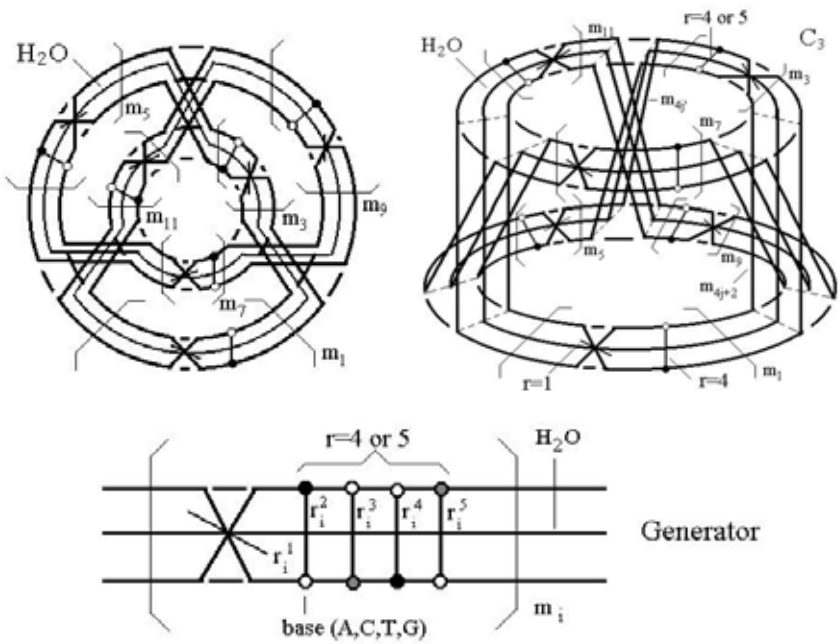

Figure 3.22: The Seifert construction for the T_3 duplex DNA knot that has C_3 symmetry. Here "*r*" represents T = A or G ≡ C ("=" or "≡" represents the number of hydrogen bonds). Note that

$$T_w = 12m = \sum_{i=1}^{12} m_i \equiv \sum_{j=0}^{2} m_{4j+1} + \sum_{j=0}^{2} m_{4j+3} + \sum_{j=0}^{2} m_{4j+2} + \sum_{j'=1}^{3} m_{4j'},$$

and that m_{4j+1} and m_{4j+3} are, respectively, arranged in the boundaries of the bottom and top disks, and that the others, m_{4j+2} and $m_{4j'}$, are unmarked in the crossing-point areas of the knot 3_1.

They have C_1 symmetry when

$$\{\{(r^1_1,\ldots, r^5_1) \neq (r^1_5,\ldots, r^5_5) \neq (r^1_9,\ldots, r^5_9)\},$$

$$\{(r^1_3,\ldots, r^6_3) \neq (r^1_7,\ldots, r^6_7) \neq (r^1_{11},\ldots, r^6_{11})\},$$

$$\{(r^1_2,\ldots, r^5_2) \neq (r^1_6,\ldots, r^5_6) \neq (r^1_{10},\ldots, r^5_{10})\}$$

$$\text{and } \{(r^1_4,\ldots, r^5_4) \neq (r^1_8,\ldots, r^5_8) \neq (r^1_{12},\ldots, r^5_{12})\}\}.$$

If $n = 2$, $L_{\text{total}} = 10m$, $T_w = 20m$ and $R(r) = 105m$ (Figure 3.23), then the T_5 duplex DNA knot has C_5 symmetry when

$$\{\{(r^1_1,\ldots, r^5_1) = (r^1_5,\ldots, r^5_5) = (r^1_9,\ldots, r^5_9) = (r^1_{13},\ldots, r^5_{13}) = (r^1_{17},\ldots, r^5_{17})\},$$

$$\{(r^1_3,\ldots,r^6_3)=(r^1_7,\ldots,r^6_7)= (r^1_{11},\ldots, r^6_{11})=(r^1_{15},\ldots, r^6_{15}) = (r^1_{19},\ldots, r^6_{19})\},$$

$$\{(r^1_2,\ldots, r^5_2)= (r^1_6,\ldots, r^5_6) = (r^1_{10},\ldots, r^5_{10}) = (r^1_{14},\ldots,r^5_{14}) = (r^1_{18},\ldots, r^5_{18})\}$$

$$\text{and } \{(r^1_4,\ldots, r^5_4) = (r^1_8,\ldots, r^5_8)=(r^1_{12},\ldots, r^5_{12}) = (r^1_{16},\ldots, r^5_{16})$$

$$= (r^1_{20},\ldots, r^5_{20})\}\},$$

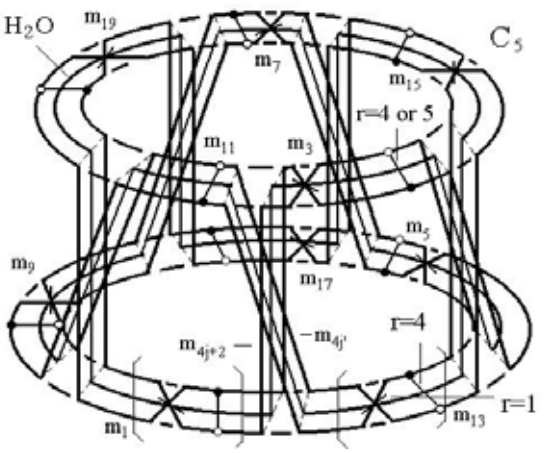

Figure 3.23: The Seifert construction for the T_5 duplex DNA knot that has C_5 symmetry.

Note that

$$T_w = 20m = \sum_{i=1}^{20} m_i \equiv \sum_{j=0}^{4} m_{4j+1} + \sum_{j=0}^{4} m_{4j+3} + \sum_{j=0}^{4} m_{4j+2} + \sum_{j=1}^{5} m_{4j}\ .$$

and that m_{4j+1} and m_{4j+3} are, respectively, arranged in the boundaries of the bottom and top disks, and that the others, m_{4j+2} and $m_{4j'}$, are unmarked in the crossing-point areas of the knot 5_1.

and has C_1 symmetry when

$$\{\{m_1 \neq m_5 \neq m_9 \neq m_{13} \neq m_{17}\}, \{m_3 \neq m_7 \neq m_{11} \neq m_{15} \neq m_{19}\},$$

$$\{m_2 \neq m_6 \neq m_{10} \neq m_{14} \neq m_{18}\} \quad \text{and} \quad \{m_4 \neq m_8 \neq m_{12} \neq m_{16} \neq m_{20}\}\}.$$

If $n = k$, $L_{\text{total}} = 2(2k + 1)m$, $T_w = 4(2k + 1)m$ and $R(r) = 21(2k + 1)m$ (Figure 3.24), then the T_{2k+1} duplex DNA knot has C_{2k+1} symmetry because it can be rotated by $2k+1$ times when

$$\{\{(r^1_1, r^2_1, r^3_1, r^4_1, r^5_1)=\ldots=(r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r^1_3, r^2_3, r^3_3, r^4_3, r^5_3, r^6_3)=\ldots=(r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3'} r^6_{4j+3})\},$$

$$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2)=\ldots=(r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\}$$

and $\{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4)=\ldots=(r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}\},$

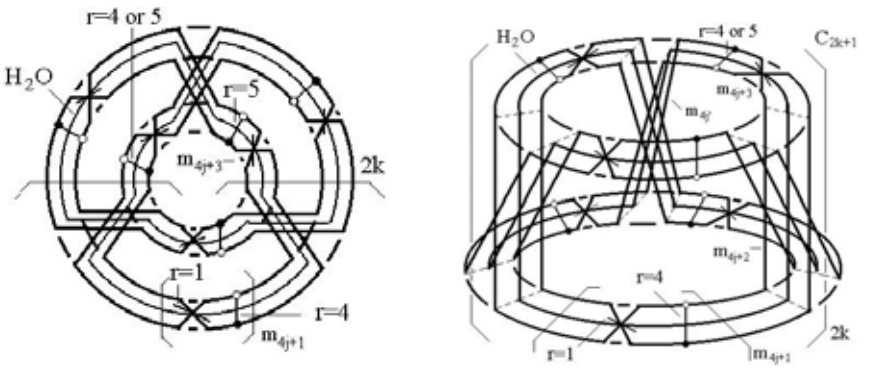

Figure 3.24: The Seifert construction for the T_{2k+1} duplex DNA knot that has C_{2k+1} symmetry. Note that

$$T_w = 4(2k+1)m = 4\sum_{i=1}^{2k+1} m_i \equiv \sum_{j=0}^{2k} m_{4j+1} + \sum_{j=0}^{2k} m_{4j+3} + \sum_{j=0}^{2k} m_{4j+2} + \sum_{j=1}^{2k+1} m_{4j}\,.$$

and that m_{4j+1} and m_{4j+3} are, respectively, arranged in the boundaries of the bottom and top disks, and the others, m_{4j+2} and $m_{4j'}$, are unmarked in the crossing-point areas of the T_{2k+1}-knot.

and has only a C_1 symmetry when

$$\{\{m_1 \neq m_5 \neq m_9 \neq \ldots \neq m_{4j+1}\},\{m_3 \neq m_7 \neq m_{11} \neq \ldots \neq m_{4j+3}\},$$

$$\{m_2 \neq m_6 \neq m_{10} \neq \ldots \neq m_{4j+2}\} \text{ and } \{m_4 \neq m_8 \neq m_{12} \neq \ldots \neq m_{4j'}\}\}$$

$$(j = 0, 1, \ldots, 2k\,;\, j' = 1, 2, \ldots, 2k+1\ ).$$

Without loss of generality, we can conclude that the T_{2n+1} duplex DNA knots have C_{2n+1} symmetry (n = 1, 2,…) if and only if

$$\{m_1 = m_5 = \ldots = m_{4j+1}\},\ \{m_3 = m_7 = \ldots = m_{4j+3}\},$$

$$\{m_2 = m_6 = \ldots = m_{4j+2}\} \text{ and } \{m_4 = m_8 = \ldots = m_{4j'}\},$$

and have C_1 symmetry if and only if

$$\{m_1 \neq m_5 \neq \ldots \neq m_{4j+1}\},\ \{m_3 \neq m_7 \neq \ldots \neq m_{4j+3}\},$$

$$\{m_2 \neq m_6 \neq \ldots \neq m_{4j+2}\} \text{ and } \{m_4 \neq m_8 \neq \ldots \neq m_{4j'}\}$$

$$(j = 0, 1, \ldots, 2n;\ j' = 1, 2, \ldots, 2n+1).$$

We conclude from the proof above that the following corollary is true.

If $L_{\text{total}} = 2(2n + 1)m$, $T_w = 4(2n + 1)m$ and $R\ (r) = 2 \times 10(2n + 1)m = 20(2n + 1)m$, then the T_{2n+1} duplex DNA knots have C_{2n+1} symmetry when

$$\{\{\ (r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) = \ldots = (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r^1_3, r^2_3, r^3_3, r^4_3, r^5_3\ ) = \ldots = (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3}\ )\},$$

$$\{(\ r^1_2, r^2_2, r^3_2, r^4_2, r^5_2\ ) = \ldots = (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\}$$

$$\text{and } \{(\ r^1_4, r^2_4, r^3_4, r^4_4, r^5_4) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}\ \}.$$

Conversely, if we suppose that

$$\{\{m_1 \neq m_5 \neq \ldots \neq m_{4j+1}\},\ \{m_3 \neq m_7 \neq \ldots \neq m_{4j+3}\},$$

$$\{m_2 \neq m_6 \neq \ldots \neq m_{4j+2}\} \text{ and } \{m_4 \neq m_8 \neq \ldots \neq m_{4j'}\}\}$$

$$(j = 0, 1, \ldots, 2n;\ j' = 1, 2, \ldots, 2n+1\ ),$$

then the T_{2n+1} duplex DNA knots have only C_1 symmetry.

The point groups C_1 or C_{2n+1} tell us that the duplex DNA knots are chiral and their states will be chirally symmetric. Therefore, right-handed and left-handed DNA knots are different in the sense that we cannot move the right-handed DNA knots within three-space to create the left-handed DNA knots. However, the symmetric balance between the two types of the natural DNA molecules is unstable and spontaneously evolves into an asymmetric state in which one type dominates. Mirror symmetry is, thus, often absent in natural duplex DNA knots: right-handed T_{2n+1} duplex DNA knots far outnumber left-handed ones. Intramolecular and intermolecular symmetry breaking always takes place in T_{2n+1} duplex DNA knots.

3.6.2 T_{4n} Duplex DNA Knots

In contrast with T_{2n+1} duplex DNA knots, the T_{4n} duplex DNA knots can be described as double knots (Figure 3.18) with double edges and have a different Seifert construction which uses three disks to fill in three loops in the supermolecular knots (e.g., 4_1, 8_{18},…) [33] and $4n$ ribbons (n = 1, 2,...) with one half-twists to fill in the crossing-point areas, and then connecting the disks with the ribbons in three-space. For convenience, we denote the local linking and twisting number by $2m$ and $4m$ (m = 1, 2,…), respectively. Thus the total linking and twisting numbers of duplex DNA in supermolecular knots are defined as $L_{\text{total}} = 2m \times 4n = 8nm$ and $T_w = 2 \times 2m \times 4n = 16nm$, respectively. Also, the total base pairs number is defined as $R(r) = 10.5L_{\text{total}} = 10.5 \times 8nm = 84nm$ or $R(r) = 10.0L_{\text{total}} = 10 \times 8nm = 80nm$ (n = 1, 2, …; m = 1, 2, …, ∞).

We shall use mathematical induction to prove that T_{4n} duplex DNA knots have either S_{4n} symmetry or no symmetry (S_1') for all positive integers n if and only if the 21- and 20-base-pairs occur in a sequence which repeats itself exactly $4n$ times.

For example, if n = 1, $L_{\text{total}} = 8m$, $T_w = 16m$ and $R(r) = 84m$ (Figure 3.25), then the T_4 duplex DNA knot has S_4 symmetry because it can be reversed up and down (i.e., a rotation by 90° followed by a reflection takes this operation to itself). In this case, we wish to show that the sets

$$\{\{(r^1_1,\ldots, r^5_1) = (r^1_5,\ldots, r^5_5) = (r^1_9,\ldots, r^5_9) = (r^1_{13},\ldots, r^5_{13})\},$$

$$\{(r^1_3,\ldots, r^6_3) = (r^1_7,\ldots, r^6_7) = (r^1_{11},\ldots, r^6_{11}) = (r^1_{15},\ldots, r^6_{15})\},$$

$$\{(r^1_2,\ldots, r^5_2) = (r^1_6,\ldots, r^5_6) = (r^1_{10},\ldots, r^5_{10}) = (r^1_{14},\ldots, r^5_{14})\}$$

$$\text{and } \{(r^1_4,\ldots, r^5_4) = (r^1_8,\ldots, r^5_8) = (r^1_{12},\ldots, r^5_{12}) = (r^1_{16},\ldots, r^5_{16})\}\}$$

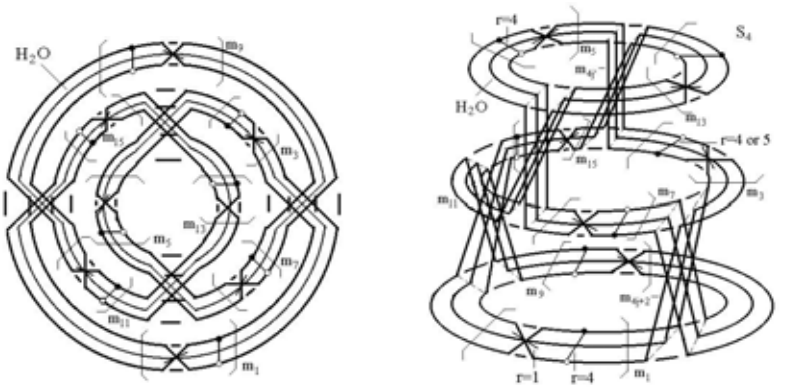

Figure 3.25: The Seifert construction for the T_4 duplex DNA knot that has S_4 symmetry. Note that

$$T_w = 16m = 4\sum_{i=1}^{4} m_i \equiv \sum_{j=0}^{3} m_{4j+1} + \sum_{j=0}^{3} m_{4j+3} + \sum_{j=0}^{3} m_{4j+2} + \sum_{j=1}^{4} m_{4j}\ .$$

and that m_{4j+1} is arranged in the boundaries of the bottom and top disks and m_{4j+3} is in the middle of the disk, and that the others, m_{4j+2} and m_{4j}, are unmarked in the crossing-point areas of the knot 4_1(m_i = 1, 2,..., ∞).

are valid. Conversely, if we suppose that

$$\{\{(r^1_1,\ldots,r^5_1)\neq(r^1_5,\ldots,r^5_5)\neq(r^1_9,\ldots,r^5_9)\neq(r^1_{13},\ldots,r^5_{13})\},$$

$$\{(r^1_3,\ldots,r^6_3)\neq(r^1_7,\ldots,r^6_7)\neq(r^1_{11},\ldots,r^6_{11})\neq(r^1_{15},\ldots,r^6_{15})\},$$

$$\{(r^1_2,\ldots,r^5_2)\neq(r^1_6,\ldots,r^5_6)\neq(r^1_{10},\ldots,r^5_{10})\neq(r^1_{14},\ldots,r^5_{14})\}$$

$$\text{and } \{(r^1_4,\ldots,r^5_4)\neq(r^1_8,\ldots,r^5_8)\neq(r^1_{12},\ldots,r^5_{12})\neq(r^1_{16},\ldots,r^5_{16})\}\},$$

then the T_4 duplex DNA knot has no symmetry (S_1').

If n = 2, L_{total} = 16m, T_w = 32m and R (r) = 168m, the resulting T_8 duplex DNA knot is shown in Figure 3.26. In this figure, if it has S_8 symmetry, this improper rotation must take the form:

$$\{\{(r^1_1,\ldots,r^5_1)=(r^1_5,\ldots,r^5_5)=(r^1_9,\ldots,r^5_9)=(r^1_{13},\ldots,r^5_{13})=(r^1_{17},\ldots,r^5_{17})$$

$$=(r^1_{21},\ldots,r^5_{21})=(r^1_{25},\ldots,r^5_{25})=(r^1_{29},\ldots,r^5_{29})\},\{(r^1_3,\ldots,r^6_3)$$

$$=(r^1_7,\ldots,r^6_7)=(r^1_{11},\ldots,r^6_{11})=(r^1_{15},\ldots,r^6_{15})=(r^1_{19},\ldots,r^6_{19})$$

$$=(r^1_{23},\ldots,r^6_{23})=(r^1_{27},\ldots,r^6_{27})=(r^1_{31},\ldots,r^6_{31})\},\ \{(r^1_2,\ldots,r^5_2)=(r^1_6,\ldots,r^5_6)$$

$= (r^1_{10},\ldots, r^5_{10}) = (r^1_{14},\ldots, r^5_{14}) = (r^1_{18},\ldots, r^5_{18})$

$= (r^1_{22},\ldots, r^5_{22}) = (r^1_{26},\ldots, r^5_{26}) = (r^1_{30},\ldots, r^5_{30})\}$

and $\{(r^1_4,\ldots, r^5_4) = (r^1_8,\ldots, r^5_8) = (r^1_{12},\ldots, r^5_{12}) = (r^1_{16},\ldots, r^5_{16})$

$= (r^1_{20},\ldots, r^5_{20}) = (r^1_{24},\ldots, r^5_{24}) = (r^1_{28},\ldots, r^5_{28}) = (r^1_{32},\ldots, r^5_{32})\}\};$

and, after performing this rotation eight times, every generator on the Seifert construction must return to its original position. On the other hand, if

$\{m_1 \neq m_5 \neq m_9 \neq m_{13} \neq m_{17} \neq m_{21} \neq m_{25} \neq m_{29}\},$

$\{m_3 \neq m_7 \neq m_{11} \neq m_{15} \neq m_{19} \neq m_{23} \neq m_{27} \neq m_{31}\},$

$\{m_2 \neq m_6 \neq m_{10} \neq m_{14} \neq m_{18} \neq m_{22} \neq m_{26} \neq m_{30}\}$ and

$\{m_4 \neq m_8 \neq m_{12} \neq m_{16} \neq m_{20} \neq m_{24} \neq m_{28} \neq m_{32}\},$

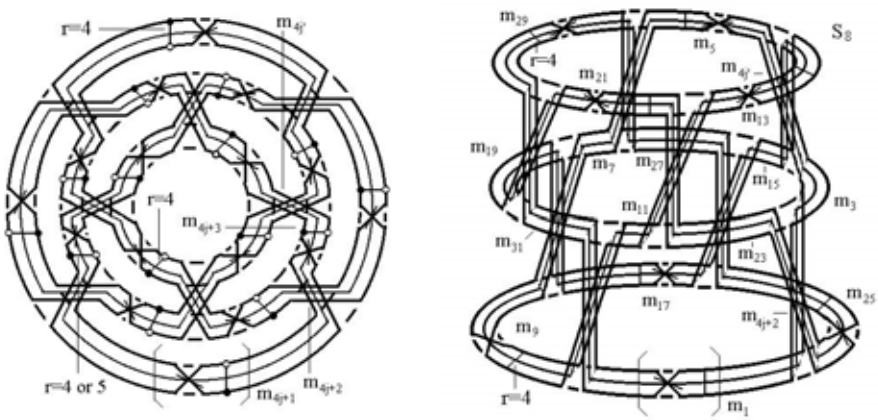

Figure 3.26: The Seifert construction for T_8 duplex DNA knot that has S_8 symmetry. Note that

$$T_w = 32m = 4\sum_{i=1}^{8} m_i \equiv \sum_{j=0}^{7} m_{4j+1} + \sum_{j=0}^{7} m_{4j+3} + \sum_{j=0}^{7} m_{4j+2} + \sum_{j=1}^{8} m_{4j}\ .$$

and that m_{4j+1} is positioned in the boundaries of the bottom and top disks and m_{4j+3} is in the middle of the disk, and the others, m_{4j+2} and m_{4j}, are unmarked in the crossing-point areas of the knot 8_{18}.

then the generator m_1 does not occur in a sequence that repeats itself exactly eight times. So the T_8 duplex DNA knot has no symmetry (S_1').

If $n = k$, $L_{\text{total}} = 8km$, $T_w = 16km$ and $R(r) = 84km$, then the T_{4k} duplex DNA knot (Figure 3.27) has S_{4k} symmetry when the generators are regularly equivalent, e.g.,

$$\{m_1 = m_5 = \ldots = m_{4j+1}\},\ \{m_3 = m_7 = \ldots = m_{4j+3}\},$$

$$\{m_2 = m_6 = \ldots = m_{4j+2}\} \text{ and } \{m_4 = m_8 = \ldots = m_{4j'}\}$$

$$(j = 0, 1, 2, \ldots, 4k\text{-}1;\ j' = 1, 2, \ldots, 4k).$$

In this case, the middle disk (which is a mirror plane) must reverse up and down and repeats itself exactly $4k$ times with a rotation. On the other hand, if the sequence of bases on a given chain is likely to be irregular

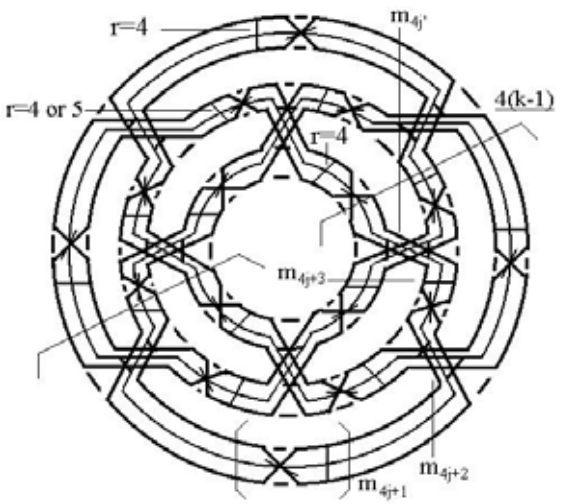

Figure 3.27: The Seifert construction for the T_{4k} duplex DNA knot is displayed in two-space. Note that

$$T_w = 16km = 4\sum_{i=1}^{4k} m_i \equiv \sum_{j=0}^{4k-1} m_{4j+1} + \sum_{j=0}^{4k-1} m_{4j+3} + \sum_{j=0}^{4k-1} m_{4j+2} + \sum_{j=1}^{4k} m_{4j}\ .$$

and that m_{4j+1} is positioned in the boundaries of the bottom and top disks and m_{4j+3} is in the middle of the disk, and the others, m_{4j+2} and m_{4j}', are unmarked in the crossing-point areas of the T_{4k}-knot.

and does not satisfy a condition of regular repetition, e.g.,

$$\{(r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) \neq \ldots \neq (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r^1_3, r^2_3, r^3_3, r^4_3, r^5_3, r^6_3) \neq \ldots \neq (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3'}\, r^6_{4j+3})\},$$

$$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2) \neq \ldots \neq (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\} \text{ and}$$

$$\{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4) \neq \ldots \neq (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}$$

$$(j = 0, 1, 2, \ldots, 4k\text{-}1;\ j' = 1, 2, \ldots, 4k),$$

this object possesses no symmetry (S_1').

By the induction property, we can conclude that T_{4n} duplex DNA knots have S_{4n} symmetry if and only if

$$\{(r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) = \ldots = (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r13, r^2_3, r^3_3, r^4_3, r^5_3, r^6_3) = \ldots = (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3'}\, r^6_{4j+3})\},$$

$$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2) = \ldots = (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\} \text{ and}$$

$$\{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\},$$

and have no symmetry (S_1') if and only if

$$\{m_1 \neq m_5 \neq \ldots \neq m_{4j+1}\},\ \{m_3 \neq m_7 \neq \ldots \neq m_{4j+3}\},$$

$$\{m_2 \neq m_6 \neq \ldots \neq m_{4j+2}\} \text{ and } \{m_4 \neq m_8 \neq \ldots \neq m_{4j'}\}$$

$$(j = 0, 1, 2, \ldots, 4n\text{-}1;\ j' = 1, 2, \ldots, 4n;\ n = 1, 2, \ldots).$$

We conclude from the proof above that the following corollary is true.

If $L_{\text{total}} = 2m \times 4n = 8nm$, $T_w = 4m \times 4n = 16nm$ and $R(r) = 10(2m \times 4n) = 80nm$, then the T_{4n} duplex DNA knots have S_{4n} symmetry when

$$\{\{(r^1_1, r^2_1, r^3_1, r^4_1, r^5_1) = \ldots = (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1})\},$$

$$\{(r^1_3, r^2_3, r^3_3, r^4_3, r^5_3) = \ldots = (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3})\},$$

$$\{(r^1_2, r^2_2, r^3_2, r^4_2, r^5_2) = \ldots = (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})\}$$

$$\text{and } \{(r^1_4, r^2_4, r^3_4, r^4_4, r^5_4) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}\},$$

and have no symmetry (S_1') when

$\{\{m_1 \neq m_5 \neq \ldots \neq m_{4j+1}\}, \{m_3 \neq m_7 \neq \ldots \neq m_{4j+3}\},$

$\{m_2 \neq m_6 \neq \ldots \neq m_{4j+2}\}$ and $\{m_4 \neq m_8 \neq \ldots \neq m_{4j'}\}\}$

$(j = 0, 1, \ldots, 4n\text{-}1; \quad j' = 1, 2, \ldots, 4n).$

The point groups S_{4n} tell us that the T_{4n} duplex DNA knots are achiral and their state will be achirally symmetric. In other words, right-handed and left-handed DNA knots are the same in the sense that we can move the right-handed DNA knots within three-space to create left-handed DNA knots. Thus, we need a mirror to get from one to the other. A mirror, in fact, makes a right-handed frame in three-space into a left-handed frame; of course, the mirror reverses up and down, and not left and right. On the other hand, the S_1' is shown to have the properties of a Euclidean rubber glove.

3.7 Duplex DNA Catenanes

Catenanes, from the Latin catena meaning chain, are molecules that contain two or more interlocked rings, which are inseparable without the breaking of covalent bond [31]. DNA duplex catenanes, called duplex DNA links, were first discovered in 1967 in Vinograd's laboratory [157] as naturally occurring linked dimers in the mitochondria of malignant cells. The first study of equilibrium catenation of DNA in solution was carried out in 1976 [158] by Wang and Schwartz, who measured the fraction of catenanes formed between phage 186 and λ DNA molecules upon their cyclization (for a review, see Wasserman and Cozzarelli, 1986 [6]). In 1985 Spengler *et al*. [149] determined the structures of the duplex DNA links 2_1^2, 4_1^2, 6_1^2, 8_1^2, 10_1^2, 12_1^2, 14_1^2, 16_1^2, 18_1^2, 20_1^2, etc. In 1992 Adams *et al*. [159] obtained the duplex DNA links having from two to more than 32 crossing points.

At thermodynamic equilibrium, a distribution of isoforms of closed circular DNA (superhelical coils, knots, and catenanes) is formed from linear DNA. DNA topoisomerases [126,127] can catalyze the strand movement between two DNA segments, thereby affecting DNA topology. DNA links and knots, formed by the type I DNA topoisomerase approach, are observed at equilibrium. More recently, however, Cozzarelli's group [151] have shown that with topoisomerase II the distribution of links and knots lies below the equilibrium values. These findings have implications in DNA replication and chromosome segregation [160,161].

Duplex DNA links are a fascinating phenomenon of the life science, and characterize many of the medically and biologically most interesting DNAs [162-164]. A molecular Seifert construction [165] based on knot theory [77] has been presented to describe duplex DNA links. Applying the point symmetry concept to Seifert construction has permitted solution of the determination of geometrical chirality. The novel topology of the Seifert construction for duplex DNA links can be used to help understand life processes, such as replication and recombination [6], although much remains to be learned. Here we describe some chemical implications of knot theory, illustrating an important application of topology in the fields of supermolecular design and assembly.

3.7.1 Definition of Intertwined Structures

Clearly the duplex DNA links have very similar properties to normal duplex DNA. In order to completely characterize duplex DNA links with 2*n* half-twists, we denote the twisting number of superstructures by T_{2n} (where *n* represents the linking number, $n = 1, 2,...$), as shown in Figure 3.28. The twin linking number of duplex DNA is defined as $L = n\times 2m = 2nm$ or $L' = n\times 2m' = 2nm'$; the twin twisting number is $T_w = 2n\times 2m = 4nm$ ($m = 1, 2,\ldots,\infty$) or $T_{w'} = 2n\times 2m' = 4nm'$ ($m' = 1, 2,\ldots,\infty$); the number of twin base pairs is $R(r) = 10.5L = 21nm$ or $R(r') = 10.5L' = 21nm'$ and $R(r) = 10.0L = 20nm$ or $R(r') = 10.0L' = 20nm'$, respectively. Thus $L_{total} = L + L'$, $T_{total} = T_w + T_{w'}$ and $R_{total} = R(r) + R(r')$, respectively, represent the total linking-number, total twisting-number and total base pairs in supermolecular links. The total twisting-number is equivalent to a certain number of generators on a minimal surface with the fewest genus.

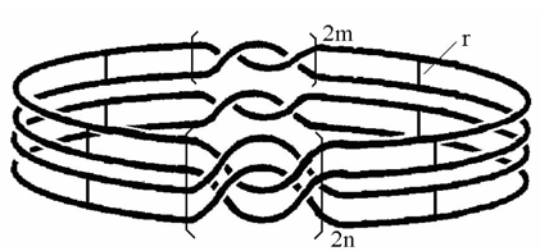

Figure 3.28: Cartoon representation of the T_{2n} duplex DNA links. *n* and 2*n* represent, respectively, the linking and twisting numbers of duplex DNA links; *m* and 2*m*, respectively, represent the local linking and twisting numbers of duplex DNA and "*r*" is a collection of the base pairs held together by hydrogen bonds.

3.7.2 Topological Features of Intertwined Structures

In general, the molecular knot and link can be viewed geometrically as the single edge, or boundary, of a two-dimensional surface [77]. In particular, the duplex DNA links can be viewed geometrically as the double edges, or double boundaries, of a two-dimensional surface. The Seifert construction with double edges is assembled by using two disks to fill in two loops in the supermolecular links (e.g., 2_1^2, 4_1^2, 6_1^2, 8_1^2, 10_1^2,...) [33] and $2n$ ribbons (n = 1, 2,...) with one half-twist to fill in the crossing-point areas in the supermolecular links, then connecting the disks with ribbons in three-space. The double edges represent two strands of the DNA double helix coated with a layer of water molecules, which are made by using $4nm$ and $4nm'$ ($m = m'$ = 1, 2,...) ribbons with one half-twist to fill in the crossing-point areas in the double helix. In molecular Seifert constructions, the total twist is arranged in the optimum symmetry position.

For variety, let us use mathematical induction to prove that the Seifert constructions for the T_{2n}-DNA links have either C_n symmetry or C_1 symmetry for all positive integers n when the 21- or 20-base-pairs occur in a sequence that repeats itself exactly n times.

For example, if n = 1, $L = L' = 2m$, $T_w = T_{w'} = 4m$ and $R(r) = R(r') = 2\times10.5m = 21m$ ($m = m'$ = 1, 2,..., ∞), then the T_2 duplex DNA link (Figure 3.29) has C_1 symmetry whether the generators are regularly equivalent or not.

If n= 2, $L = L'$ = $4m$, $T_w = T_{w'} = 8m$ and $R(r) = R(r') = 42m$ ($m = m'$ = 1, 2,..., ∞), then the T_4 DNA link (Figure 3.30) has C_2 symmetry when

$$\{\{ (r_1^1,\ldots,r_1^5) = (r_5^1,\ldots, r_5^5), (r_3^1,\ldots, r_3^6) = (r_7^1,\ldots,r_7^6),$$
$$(r_2^1,\ldots,r_2^5) = (r_6^1,\ldots, r_6^5) \text{ and } (r_4^1,\ldots,r_4^5) = (r_8^1,\ldots,r_8^5) \} \text{ and}$$
$$\{ (r'_1{}^1, \ldots, r'_1{}^5) = (r'_5{}^1,\ldots, r'_5{}^5), (r'_3{}^1, \ldots, r'_3{}^6) = (r'_7{}^1,\ldots, r'_7{}^6),$$
$$(r'_2{}^1, \ldots, r'_2{}^5) = (r'_6{}^1, \ldots, r'_6{}^5) \text{ and } (r'_4{}^1,\ldots, r'_4{}^5) = (r'_8{}^1, \ldots, r'_8{}^5) \}\},$$

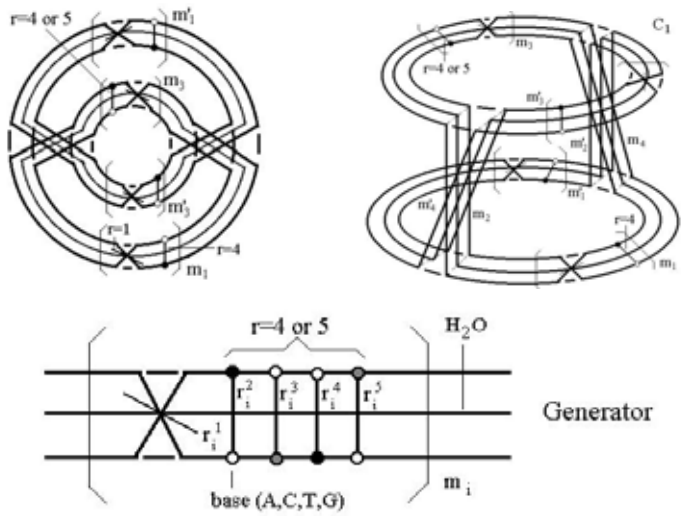

Figure3. 29: The Seifert construction for the T_2-DNA link that has C_1 symmetry. Here "r" represents T=A or G≡C ("=" or "≡" represents the number of hydrogen bonds). Note that $\{m_1, m'_1\}$ and $\{m_3, m'_3\}$ are, respectively, arranged in the boundaries of the bottom and top disks, and that $\{m_2, m_4\}$ and $\{m'_2, m'_4\}$ are unmarked in the crossing-point areas of the link 2_1^2 ($m_i = m'_i$ =1, 2,. ..., ∞).

and has C_1 symmetry when

$$\{\{m_1 \neq m_5,\ m_3 \neq m_7,\ m_2 \neq m_6 \text{ and } m_4 \neq m_8\}$$

$$\text{and } \{m'_1 \neq m'_5,\ m'_3 \neq m'_7,\ m'_2 \neq m'_6 \text{ and } m'_4 \neq m'_8\}\}.$$

Similarly, if $n = k$, $L = L' = 2km$, $T_w = T_{w'} = 4km$ and $R(r) = R(r') = 21km$ ($m = m' = 1, 2, ..., \infty$), then the T_{2k}-DNA link (Figure 3.31) has C_k symmetry. In this case, proper rotation must take

$$\{\{ (r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) = \ldots = (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1}),$$

$$(r_3^1, r_3^2, r_3^3, r_3^4, r_3^5, r_3^6) = \ldots = (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3}, r^6_{4j+3}),$$

$$(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) = \ldots = (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})$$

$$\text{and } (r_4^1, r_4^2, r_4^3, r_4^4, r_4^5) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}$$

$$\text{and } \{(r'^1_1, r'^2_1, r'^3_1, r'^4_1, r'^5_1) = \ldots = (r'^1_{4j+1}, r'^2_{4j+1}, r'^3_{4j+1}, r'^4_{4j+1}, r'^5_{4j+1}),$$

$$(r'^1_3, r'^2_3, r'^3_3, r'^4_3, r'^5_3, r'^6_3) = \ldots = (r'^1_{4j+3}, r'^2_{4j+3}, r'^3_{4j+3}, r'^4_{4j+3}, r'^5_{4j+3}, r'^6_{4j+3}),$$

$$(r'^1_2, r'^2_2, r'^3_2, r'^4_2, r'^5_2) = \ldots = (r'^1_{4j+2}, r'^2_{4j+2}, r'^3_{4j+2}, r'^4_{4j+2}, r'^5_{4j+2})$$

$$\text{and } (r'^1_4, r'^2_4, r'^3_4, r'^4_4, r'^5_4) = \ldots = (r'^1_{4j'}, r'^2_{4j'}, r'^3_{4j'}, r'^4_{4j'}, r'^5_{4j'})\}\}.$$

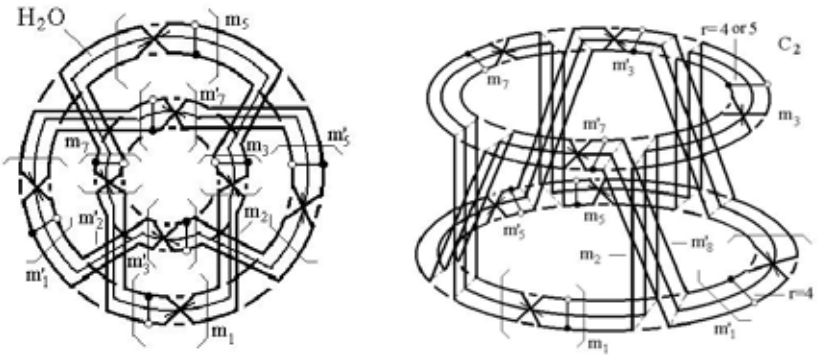

Figure 3.30: The Seifert construction for the T_4-DNA link that has C_2 symmetry. Note that

$$T_{w'} = 8m' = \sum_{i=1}^{8} m'_i = \sum_{j=0,1} m'_{4j+1} + \sum_{j=0,1} m'_{4j+3} + \sum_{j=0,1} m'_{4j+2} + \sum_{j'=0,1} m'_{4j'},$$

$$T_w = 8m = \sum_{i=1}^{8} m_i = \sum_{j=0,1} m_{4j+1} + \sum_{j=0,1} m_{4j+3} + \sum_{j=0,1} m_{4j+2} + \sum_{j'=1,2} m_{4j'},$$

and that $\{m_{4j+2}, m'_{4j+2}\}$ and $\{m_{4j'}, m'_{4j'}\}$ are unmarked in the crossing-point areas of the link 4_1^2.

However, if

$\{\{(m_1 \neq m_5 \neq \ldots \neq m_{4j+1}), (m_3 \neq m_7 \neq \ldots \neq m_{4j+3}),$

$(m_2 \neq m_6 \neq \ldots \neq m_{4j+2})$ and $(m_4 \neq m_8 \neq \ldots \neq m_{4j'})\}$

and $\{(m'_1 \neq m'_5 \neq \ldots = m'_{4j+1}), (m'_3 = m'_7 = \ldots = m'_{4j+3}),$

$\{m'_2 = m'_6 = \ldots = m'_{4j+2})$ and $(m'_4 = m'_8 = \ldots = m'_{4j'})\}\ \},$

then the T_{2k}-DNA link has C_1 symmetry (j = 0, 1, 2,…, k-1; j' = 1, 2,…, k).

It is quite surprising to realize, by the induction process, that the T_{2n} duplex DNA links have C_n symmetry when

$\{\{ (m_1 = \ldots = m_{4j+1}), (m_3 = \ldots = m_{4j+3}), (m_2 = \ldots = m_{4j+2})$

and $(m_4 = \ldots = m_{4j'})\ \}$ and $\{(m'_1 = \ldots = m'_{4j+1}),$

$(m'_3 = \ldots = m'_{4j+3}), (m'_2 = \ldots = m'_{4j+2})$ and $(m'_4 = \ldots = m'_{4j'})\}\ \},$

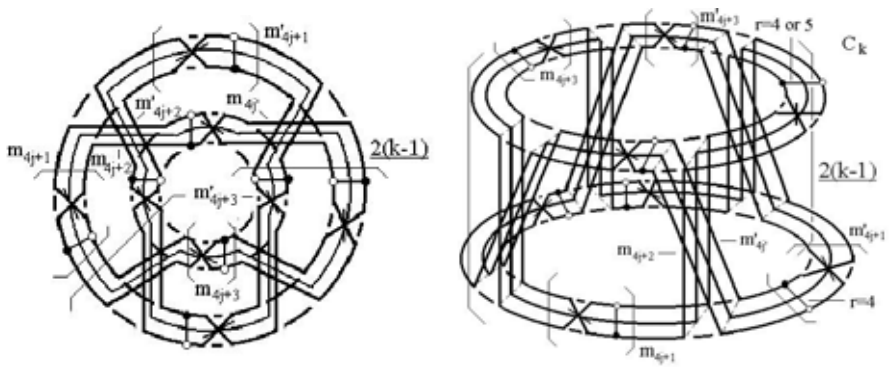

Figure 3.31: The Seifert construction for the T_{2k}-DNA link that has C_k symmetry. Note that

$$T_w = 4km = 4\sum_{i=0}^{k} m_i = \sum_{j=0}^{k-1} m_{4j+1} + \sum_{j=0}^{k-1} m_{4j+3} + \sum_{j=0}^{k-1} m_{4j+2} + \sum_{j'=1}^{k} m_{4j'},$$

$$T_{w'} = 4km' = 4\sum_{i=1}^{k} m'_i = \sum_{j=0}^{k-1} m'_{4j+1} + \sum_{j=0}^{k-1} m'_{4j+3} + \sum_{j=0}^{k-1} m'_{4j+2} + \sum_{j'=1}^{k} m'_{4j'},$$

and that $\{m_{4j+1}, m'_{4j+1}\}$and $\{m_{4j+3}, m'_{4j+3}\}$ are, respectively, positioned in the boundaries of the bottom and top disks; whereas the others, $\{m_{4j+2}, m'_{4j+2}\}$ and $\{m_{4j'}, m'_{4j'}\}$, are unmarked in the crossing-point areas of the T_{2k}-link.

and have only C_1 symmetry when

$$\{\{(m_1 \neq \ldots \neq m_{4j+1}), (m_3 \neq \ldots = m_{4j+3}),$$

$$(m_2 \neq \ldots \neq m_{4j+2}), (m_4 \neq \ldots \neq m_{4j'}), \ldots\}\}$$

for all positive integers n (j = 0, 1,…, n-1; j' = 1, 2,…, n; n = 1, 2,…).

In the same way, we can prove that the T_{2n} duplex DNA links ($L = L' = 2nm$, $T_w = T_{w'} = 4nm$ and $R(r) = R(r') = 20nm$, $m = m' = 1, 2, ..., \infty$) have C_n symmetry when

$$\{\{ (r_1^1, r_1^2, r_1^3, r_1^4, r_1^5) = \ldots = (r^1_{4j+1}, r^2_{4j+1}, r^3_{4j+1}, r^4_{4j+1}, r^5_{4j+1}),$$

$$(r_3^1, r_3^2, r_3^3, r_3^4, r_3^5) = \ldots = (r^1_{4j+3}, r^2_{4j+3}, r^3_{4j+3}, r^4_{4j+3}, r^5_{4j+3}),$$

$$(r_2^1, r_2^2, r_2^3, r_2^4, r_2^5) = \ldots = (r^1_{4j+2}, r^2_{4j+2}, r^3_{4j+2}, r^4_{4j+2}, r^5_{4j+2})$$

$$\text{and } (r_4^1, r_4^2, r_4^3, r_4^4, r_4^5) = \ldots = (r^1_{4j'}, r^2_{4j'}, r^3_{4j'}, r^4_{4j'}, r^5_{4j'})\}$$

$$\text{and } \{(r'^1_1, r'^2_1, r'^3_1, r'^4_1, r'^5_1) = \ldots = (r'^1_{4j+1}, r'^2_{4j+1}, r'^3_{4j+1}, r'^4_{4j+1}, r'^5_{4j+1}),$$

$$(r'^1_3, r'^2_3, r'^3_3, r'^4_3, r'^5_3) = \ldots = (r'^1_{4j+3}, r'^2_{4j+3}, r'^3_{4j+3}, r'^4_{4j+3}, r'^5_{4j+3}),$$

$(r'_2{}^1, r'_2{}^2, r'_2{}^3, r'_2{}^4, r'_2{}^5) = \ldots = (r'_{4j+2}{}^1, r'_{4j+2}{}^2, r'_{4j+2}{}^3, r'_{4j+2}{}^4, r'_{4j+2}{}^5)$

and $(r'_4{}^1, r'_4{}^2, r'_4{}^3, r'_4{}^4, r'_4{}^5) = \ldots = (r'_{4j'}{}^1, r'_{4j'}{}^2, r'_{4j'}{}^3, r'_{4j'}{}^4, r'_{4j'}{}^5)\}\ \},$

and have C_1 symmetry when

$\{\{(m_1 \neq m_5 \neq \ldots \neq m_{4j+1}), (m_3 \neq m_7 \neq \ldots \neq m_{4j+3}),$

$(m_2 \neq m_6 \neq \ldots \neq m_{4j+2})$ and $(m_4 \neq m_8 \neq \ldots \neq m_{4j'})\}$

and $\{(m'_1 \neq m'_5 \neq \ldots = m'_{4j+1}), (m'_3 \neq m'_7 \neq \ldots \neq m'_{4j+3}),$

$(m'_2 \neq m'_6 \neq \ldots \neq m'_{4j+2})$ and $(m'_4 \neq m'_8 \neq \ldots \neq m'_{4j'})\}\ \}$

$(j = 0, 1, 2, \ldots, n\text{-}1; j' = 1, 2, \ldots, n).$

We have shown that the duplex DNA links possess the point group C_n or C_1 when the double helix makes a full turn every 10.5 or 10.0 base pairs. This result demonstrates that the duplex DNA links are chiral and the right-handed and left-handed DNA links are distinct.

3.8 Summary

As mentioned above, circular duplex DNA, single-stranded DNA knots, duplex DNA knots and duplex DNA links have some features which are regular, and some which are irregular. Here are some of the phenomena we are now in a position to explore.

Seifert construction enlarged The molecular Seifert construction involves the language of geometry and is based on strict deterministic rules. In general, molecular Seifert construction is the result of a construction procedure or production rule that is often recursive (repeated over and over). The generators are a class of structures that look similar but are not exactly the same and reveal a growth process of order and synchronization, which are enlarged depending on either $2j + 1$ and $2j'$ or $4j + 1$, $4j + 3$, $4j + 2$ and $4j'$ ($j = 0, 1, 2,\ldots$; $j' = 1, 2,\ldots$). As the pattern is enlarged it reveals repetitive levels of detail, so that similar structure exists on all scales.

Geometrical chirality and achirality Chirality and achirality of the DNA links and knots are beginning to emerge that differentiate one from the other. The structures with C_1, C_n or the C_{2n+1}symmetries are geometrically chiral

while the structures with S_{4n} symmetries are geometrically achiral. The objects with geometrical chirality have two forms which are the left-handed and right handed; whereas those with geometrical achirality are not, since the left-handed and right-handed forms can be transformed in three-space from one to the other by symmetry operations. An object with no symmetry S_1' is shown to have the properties of a Euclidean rubber glove because it can reversed up and down and the palm and thumb are differentiated from the back of the hand. The no symmetry point group C_1 represents the most chirality and the no symmetry S_1' represents the least chirality.

Intramolecular symmetry breaking Most artificial DNAs may have the point groups C_n, C_{2n+1} or S_{4n}, whereas the natural DNAs have no symmetry and are C_1 or S_1'. The C_n ($n\geq2$), C_{2n+1} and S_{4n} symmetries are very regular whereas the C_1 and S_1' are completely irregular. The C_n, C_{2n+1} and S_{4n} symmetries possess characteristic sizes and are a geometrical invariant, whereas the C_1 and S_1' possesses no characteristic scales and are a scale invariant but infinitely repeated structures, such as $2n$, $2n+1$ or $4n$ (n=1,2,…). The artificial DNAs are either regular or irregular, whereas the natural DNAs must be irregular. The fusion of regular and irregular features is achieved admittedly only at the expense of the symmetry conformation. Thus, the transitions, C_n, $C_{2n+1}\rightarrow C_1$ or $S_{4n}\rightarrow S_1$', are referred to as intramolecular symmetry breaking. The intramolecular symmetry breaking is a mechanism by which a system of itself goes from a highly symmetric conformation to a no-symmetric one. The most chirality or the least chirality is the result of intramolecular symmetry breaking. It can take place in part because of the ubiquity of low-frequency $1/f^{\beta}$ noise and long-range fractal correlations as well as prominent short-range periodicities [166] in DNA base sequences [167]. The intramolecular symmetry breaking may be a basis of characterizing the complexity and variety of the medically and biologically most interesting DNAs.

Intermolecular symmetry breaking The origin of chiral asymmetry is a key question in life studies [168]. Artificial DNA is easily formed in thermodynamic equilibrium. At thermodynamic equilibrium, the numbers of left- and right-handed forms will be equal, and the state will be chirally symmetric. This explains why the right- and left-handed forms are found in the laboratory in essentially equal numbers [96]. A life system is open to the inflow of energy or matter; however, it is no longer in thermodynamic equilibrium [62]. Intermolecular symmetry breaking then can become operative and can throw

the natural DNA into a chirally asymmetric state, one that has unequal amounts of the left- and right-handed forms. A life system will therefore be open and far from equilibrium, ensuring that intermolecular symmetry breaking can take place. Intermolecular symmetry breaking is a mechanism by which a system spontaneously goes from a symmetric state to an asymmetric one.

Acknowledgments

The author wishes to thank his advisors, Professors Hou-Wen Xin, Bin-Hong Wang, and Guang-Can Guo, and Ke-Ling Wang, for their helpful guidance and encouragement during the preparation of this chapter. The author is much indebted to N. R. Cozzarelli, K. Mislow, N. C. Seeman, J. C. Wang, J. -P. Sauvage, A. Stasiak, C. Ernst, E. Flapan, D. M. Walba and H. M. Berman for communicating results prior to publication. The author also thanks Drs. D. Bonchev and D. H. Rouvray for valuable comments and linguistic editing on the manuscript. The author also thanks Drs. D. Bonchev, A. Stasiak, K. -C. Xu, and Prof. K. -H. Chen for useful help and discussion. This work was supported by Postdoctoral Science Foundation of China and the K. C. Wong Education Foundation of Hong Kong.

3.9 Literature References

1. J. D. Watson and F. R. C. Crick, A structure for deoxyribose nucleic acid, *Nature*, **171**,737-738 (1953).
2. J. D. Watson and F. R. C. Crick, Genetical implications of the structure of deoxyribonucleic acid. *Nature*, **171**, 964-967 (1953).
3. M. H. F. Wilkins, A. R. Stokes and H.R. Wilson, Molecular structure of deoxypentose nucleic acids, *Nature*, **171**, 738-740 (1953).
4. M. H. F. Wilkins, Molecular configuration of nucleic acids, *Science*, **140**,941-950(1963).
5. D. M. Walba, Topological stereochemistry, *Tetrahedron*, **41**, 3161-3212 (1985).
6. S. A. Wasserman and N. R. Cozzarelli, Biochemical topology: applications to DNA recombination and replication, *Science*, **232**, 951-960 (1986).
7. D. Bates and A. Maxwell, *DNA Topology*, Oxford University Press, New York, 1993.
8. G. Călugăreanu, Sur las classes disotopie des noeuds tridimensionnels et leurs invariants. *Czech. Math. J.*, **11**, 588-625 (1961).

9. W. F. Pohl, Some Integral formulas for space curves and their generalization, *Am. J. Math.*, **90**, 1321-1345 (1968).
10. W. F. Pohl, The Self-linking number of a closed space curve, *J. Math.and Mech.*, **17**, 975-986 (1968).
11. J. H. White, Self-linking and the Gauss integral in higher dimensions, *Am. J. Math.*, **91**, 693-728 (1969).
12. F. B. Fuller, The writhing number of a space curve, *Proc. Natl. Acad. Sci. USA*, **68**, 815-819 (1971).
13. F. R. C. Crick, Linking numbers and nucleosomes, *Proc. Natl. Acad. Sci. USA*, **73**, 2639-43(1976).
14. F.B. Fuller, Decomposition of the linking number of a closed ribbon: a problem from molecular biology, *Proc. Natl. Acad. Sci. USA*, **75**, 3557-3561 (1978).
15. W. F. Pohl, DNA and differential geometry. *Math. Intel.*, **3**, 20-27 (1980).
16. W. R. Bauer, F. H. C. Crick and J. H. White, Supercoiled DNA, *Sci. Am.*, **243**(1), 118-133 (1980).
17. J. C. Wang, DNA topoisomerases, *Sci. Am.*, **247**(1), 94-109 (1982).
18. J. H. White and W. R. Bauer, Applications of the twist difference to DNA structural analysis. *Proc. Natl. Acad. Sci. USA*, **85**, 772-776 (1988).
19. J. H. White, N. R. Cozzarelli and W. R. Bauer, Helical repeat and linking number of surface-wrapped DNA, *Science*, **241**, 323-327(1988).
20. J. H. White, An introduction to the geometry and topology of DNA structure, In *Mathematical Methods for DNA Sequences*, ed. MS Waterman, Boca Raton, FL: CRC Press, 1989, pp. 225-153.
21. N. R. Cozzarelli and J.C. Wang (Eds.), *DNA Topology and its Biological Effects*, Cold Spring Harbor Laboratory Press. New York, 1990.
22. A. V. Vologodskii, *Topology and Physics of Circular DNA*, Boca Raton, FL: CRC Press, 1992.
23. W. -Y. Qiu and L.-J. Zhang, Topological stereochemistry of DNA, *Huaxue Tongbao*, **8**, 31-36(1990).
24. W. R. Bauer, R. A. Lund and J. H. White, Twist and writhe of a DNA loop containing intrinsic bends, *Proc. Natl. Acad. Sci. USA*, **90**, 833-837 (1993).
25. V. Vologodskii and N. R. Cozzarelli, Conformational and thermodynamic properties of supercoiled DNA, *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 609-643 (1994).

26. M. D. Frank-Kamenestskii, DNA topology, *J. Mol.Struct.* (*Theochem*), **336**, 235-243 (1995).
27. A. Stasiak, Circular DNA, In *Large Ring Molecules*, ed. J. A. Semlyen, John Wiley & Sons Ltd, 1996, pp. 43-97.
28. A. Malevanets and R. Kapral, Link, knots, and knotted labyrinths in bistable systems, *Phys. Rev. Lett.*, **77**, 767-770 (1996).
29. L. Faddeev and A. J. Niemi, Stable knot-like structures in classical field theory, *Nature*, **387**, 58-61 (1997).
30. C. O. Dietrich-Buchecker and J.-P. Sauvage, Interlocked and Knotted Rings in Biology and Chemistry, in: H. Dugas (Ed.), *Bioorganic Chemistry Frontiers*, Springer-Verlag, Berlin, 1991, Vol. 2, pp. 195-248.
31. D. V. Amabilino and J. F. Stoddart, Interlocked and intertwined structures and superstructures, *Chem. Rev.*, **95**, 2725-2828 (1995).
32. K. Mislow and C. Liang, Knotted structures in chemistry, biochemistry, and molecular biology. *Croa. Chem. Acta*, **69**(4), 1385-1403 (1996).
33. D. Rolfsen, *Knots and Links*, Publish or Perish Inc, Berkely, 1976, pp.118-123.
34. G. Burde and H. Zieschang, *Knots*, Walter de Gruyter Inc., Berlin, New York, 1985, pp.15, 17-18.
35. L. H. Kauffman, *On Knots*, Princeton University Press, Princeton, NJ, 1987, pp. 19-69.
36. C. Adams, The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots, New York: W. H. Freeman & Co., 1994.
37. G. Kolata, Solving knotty problems in math and biology, *Science*, **231**, 1506-1508 (1986).
38. P. G. Mezey, Tying knots around chiral centers: Chirality polynomials and conformational invariants for molecules, *J. Am. Chem. Soc.*, **108**, 3765-3984 (1986).
39. J. W. Alexander, Topological invariants of knots and links, *Trans. Amer. Math. Soc.*, **30**, 275-306 (1928).
40. V. F. R. Jones, A polynomial invariant for knots, *Bull. A.M.S.*, **12**, 103-111 (1985).
41. P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K. Mikkett and A. Ocneanu, A new polynomial invariant of knots and links, *Bull. A.M.S.*, **12**, 239-246 (1985).
42. W. B. R. Lickorish, A relationship between link polynomials, *Math. Proc. Camb. Phil. Soc.*, **100**, 109-112 (1986).

43. W. B. R. Lickorish and K. C. Millett, The reversing result for the Jones polynomial, *Pac. J. Math.*, **124**, 173-176 (1986).
44. W. B. R. Lickorish and K. C. Millett, A polynomial invariant of oriented links, *Topology*, **26**, 107-141 (1987).
45. J. H. White, K. C. Millet and N. R. Cozzarelli, Description of the topological entanglement of DNA catenanes and knots by a powerful method involving strand passage and recombination. *J. Mol. Biol.*, **197**, 585-603 (1987).
46. N. R. Cozzarelli, M.A. Krasnow, S. P. Gerrard and J.H. White, A topological treatment of recombination and topoisomerases, *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 383-400 (1984).
47. J. H. White and N. R. Cozzarelli, A simple topological method for describing stereoisomers of DNA catenanes and knots, *Proc. Nat. Acad. Sci., USA*, **81**, 3332-3326 (1984).
48. J. H. Conway, An enumeration of knots and links and some of their related properties. *Computational Problems in Abstract Algebra* (Proc.Conf. Oxford, 1967), Pergamon Press, Oxford, 1970, pp.329-358.
49. C. Ernst and D. M. Sumners, A calculus for rational tangles: Applications to DNA recombination, *Math. Proc. Cambridge Phil. Soc.*, **108**, 489-515 (1990).
50. C. Ernst, Tangle equations, *J. Knot Theory and its Ramification*, **15**(2), 145-159(1996).
51. C. Ernst, Tangle equations II, *J. Knot Theory and its Ramification*, **16**(1), 1-11(1997).
52. D. M. Sumners, The Role of Knot Theory in DNA Research, in: C. McCrory and T. Schifrin (Eds.), *Geometry and Topology*, Marcel Dekker, New York, 1987, pp.297-318.
53. D. W. Sumners, Untangling DNA, *Math. Intel.*, **12**, 71-80 (1990).
54. D. W. Sumners, Knot theory and DNA, *New Scientific Applications of Geometry and Topology* (D. W. Sumners, ed.), *Proc. Sympos. Appl. Math.*, Vol. 45, Amer. Math. Soc., Providence RI, 1992, pp. 39-72.
55. D. W. Sumners, Lifting the curtain: Using topology to probe the hidden action of enzymes, *Notices of the AMS*, **42**(5), 528-537(1995).
56. D. W. Sumners, C. Ernst, S. J. Spengler and N. R. Cozzarelli, Analysis of the mechanism of DNA recombination, *Quart. Rev. Biophys.*, **28**(3), 253-313(1995).

57. C. Ernst and D. W. Sumners, Solving tangle equations arising in a DNA recombination model, *Math. Proc. Camb. Phil. Soc.*, in press (1998).
58. A. Watson, Twist, tangles and topology, *New Scientist*, **132**(1789), 42-46(1991).
59. R. A. Weinberg, The molecules of life, *Sci. Am.*, **253**(4), 34-43(1985).
60. K. B. Mullis, The unusual origin of the polymerase chain reaction, *Sci. Am.*, **262**(4), 56-65(1990).
61. L. Kari, DNA computing: arrival of biological mathematics, *Math. Intel.*, **19**(2), 9-22(1997).
62. R. Hegstrom and D. K. Kondepudi, The handedness of the universe, *Sci. Am.*, **262**(1), 98-105(1990).
63. K. Mislow, A commentary on the topological chirality and achirality of molecules, *Croa. Chem. Acta*, **69**(2), 485-511 (1996).
64. I. Stewart, The topological dressmaker, *Sci. Am.*, **269**(1), 93-95 (1993).
65. E. Flapan, Rigid and non-rigid achirality, *Pacific J. Math.*, **129**(1), 57-66 (1987).
66. E. Flapan and P.College, Topological techniques to detect chirality, In *New Developments in Molecular Chirality*, ed. P. G. Mezey, Kluwer Acad. Publ., Dordrecht, 1991, pp. 209-239.
67. K. Mislow, Limitations of the symmetry criteria for optical inactivity and resolvability, *Science*, **120**, 232-233(1954).
68. D. M. Walba, Topological stereochemistry: Knot theory of molecular graphs, In *Graph Theory and Topology in Chemistry*, ed. R. B. King and D. H. Rouvray, Elsevier, Amsterdam, 1987, pp.23-42.
69. D. M. Walba, A topological hierarchy of molecular chirality and other tidbits in topological stereochemistry, In *New Developments in Molecular Chirality*, ed. P. G. Mezey, Kluwer Acad. Publ., Dordrecht, 1991, pp.119-129.
70. E. Flapan and N. C. Seeman, A topological rubber glove obtained from a synthetic single-stranded DNA molecule, *J. Chem. Soc., Chem. Commun.*, **22**, 2249-2250 (1995).
71. E. Flapan, Knots and graphs in chemistry, *Knot Theory and Its Applications. Chaos, Solitons & Fractals*, **9**(4/5), 547-560 (1998).
72. C. Liang and K. Mislow, A left-right classification of topologically chiral knots, *J. Math. Chem.*, **15**, 35-62(1994).
73. C. Liang and K. Mislow, Classification of topologically chiral molecules, *J. Math. Chem.*, **15**,245-260(1994).

74. C. Liang, C. Cerf and K. Mislow, Specification of chirality for links and knots, *J. Math. Chem.*, **19**,241-263(1996).
75. M. Gardner, About left- and right-handedness, mirror images and kindred matters, *Sci. Am.*, **198**(3), 128-133(1958).
76. R. K. Adair, A flaw in a universal mirror, *Sci. Am.*, **258**(2), 50-56(1988).
77. L. Neuwirth, The theory of knot, *Sci. Am.*, **240**(6), 110-124 (1979).
78. H. Seifert, Über das geschlecht von knoten, *Math. Ann.*, 110, 571-592 (1935).
79. W. -Y. Qiu, *Knot theory of molecular structure and DNA topology*, Ph.D. Thesis, University of Science & Technology of China, Hefei, P. R. China, 1997.
80. W.-Y. Qiu and H.-W. Xin, Molecular design and topological chirality of the T_q-Möbius ladders, *J. Mol. Struct.* (*Theochem*), **401**, 151-156 (1997).
81. W.-Y. Qiu and H.-W. Xin, Molecular design and tailor of the doubled knots, *J. Mol. Struct.* (*Theochem*), **397**, 33-37 (1997).
82. F. A. Cotton, *Chemical Applications of Group Theory*, John Wiley & Sons, Inc., New York and London, 1964.
83. R. Dulbecco and M. Vogt, Evidence for a ring structure of polyoma virus DNA, *Proc. Natl. Acad. Sci. USA*, **50**, 236-243 (1963).
84. R. Weil and J. Vonograd, The cyclic helix and cyclic coil forms of polyoma viral DNA, *Proc. Natl. Acad. Sci. USA*, **50**, 730-739 (1963).
85. J. Vinograd, J. Lebowitz, R. Radloff, R. Watson and P. Laipis, The twisted circular from of polyoma viral DNA, *Proc. Natl. Acad. Sci. USA*, **53**, 1104-1111 (1965).
86. J. Lebowitz, Through the looking glass: the discovery of supercoiled DNA, *Trends Biochem. Sci.*, **15**, 202-207 (1990).
87. K. Drlica, Control of bacterial DNA supercoiling, *Mol. Microbiol.*, **6**, 425-433 (1992).
88. R. Kanaar and N. R. Cozzarelli, Roles of supercoiled DNA structure in DNA transactions, *Curr. Opin. Struct. Biol.*, **2**, 369-379 (1992).
89. J. C. Wang, Template topology and transcription, In *Transcriptional Regulation*, ed. J.C.Wang, Cold Spring Harbor, New York: Cold Spring Harbor Lab. Press, 1992, pp.1253-1269.
90. R. E. Dickerson, The DNA helix and how it is read, *Sci. Am.*, **249**(6), 87-102 (1983).

91. R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, Crystal structure analysis of a complete turn of B-DNA, *Nature*, **287**, 755-758(1980).
92. B. N. Conner, T. Takano, S. Tanaka, K. Itakura and R. E. Dickerson, The molecular structure of d (CpCpGpG), a fragment of right-handed double helical A-DNA, *Nature*, **295**, 294-299(1982).
93. A. H. –J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich, Molecular structure of a left-handed double helical DNA fragment at atomic resolution, *Nature*, **282**, 680-686(1979).
94. H. Drew, T. Takano, S. Tanaka, K. Itakura and R. E. Dickerson, High-salt d (CpGpCpG), a left-handed Z' DNA double helix, *Nature*, **286**, 567-573(1980).
95. (a). A. Lipanov, M. L. Kopka, M. Kaczor-Grzeskowiak and R. E. Dickerson, The structure of the B-DNA decamer C-C-A-A-C-I-T-T-G-G in two different space groups: conformational flexibility of B-DNA, Biochemistry, 32,1373-13889(1993); (b). I. Baikalov, K. Grzeskowiak, K. Yanagi, J. Quintana and R. E. Dickerson, The crystal structure of the trigonal decamer C-G-A-T-C-G-6meA-T-C-G: a B-DNA helix with 10.6 base-pairs per turn, *J. Mol. Biol.*, **231**,768-784(1993).
96. J. Kawakami, K. Tsujita and N. Sugimoto, Stability of DNA duplexes containing mirror-image DNA, *Nucleic Acids Symposium Series*, **39**, 55-(1998).
97. H. Urata, K. Shinohara, E. Ogura, Y. Ueda and M. Akagi, Mirror-image DNA, *J. Am. Chem. Soc.*, **113**(21), 8174-8177(1991).
98. H. Urate, E. Ogura, K. Shinohara, Y. Ueda and M. Akagi, Synthesis and properties of mirror-image DNA, *Nucleic Acids Research*, **20**(13), 3325-3332(1992).
99. M. Doi, M. Inoue, K. Tomoo, T. Ishida, Y. Ueda, M. Akagi and H. Urata, Structural characteristics of enantiomorphic DNA: crystal analysis of racemates of the d (CGCGCG) duplex, *J. Am. Chem. Soc.*, **115**(22), 10432-10433(1993).
100. W. -Y. Qiu and H. -W. Xin, Topological chirality and achirality of closed circular DNA, *Chinese Sci. Bull.*, **42**(20), 1705-1707(1997).
101. W.-Y. Qiu and H.-W. Xin, Topological structure of closed circular DNA, *J. Mol. Struct.* (*Theochem*), **428**, 35-39 (1998).

102. R. E. Dickerson H. R. Drew, B. N. Conner, R. M. Wing, A. V. Fratini and M. L. Kopka, The anatomy of A-, B-, and Z-DNA, *Science*, **216**, 475-485 (1982).
103. H. M. Berman, Hydration of DNA, *Curr. Opin. Struct. Biol.*, **1**, 423-427(1991).
104. H. M. Berman, Hydration of DNA: take 2, *Curr. Opin. Struct. Biol.*, **4**, 345-350(1994).
105. J. C. Wang, Helical repeat of DNA in solution, *Proc. Natl. Acad. Sci. USA*, **76**, 200-203(1979).
106. D. Rhodes and A. Klug, Helical periodicity of DNA determined by enzyme digestion, *Nature*, **286**, 573-578 (1980).
107. M. Levitt, How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations, *Proc. Natl. Acad. Sci. USA*, **75**, 640-644(1978).
108. H. R. Drew and R. E. Dickerson, Structure of a B-DNA dodecamer, Ⅲ. Geometry of hydration, *J. Mol. Biol.*, **151**,535-556(1981).
109. H. R. Drew, S. Samson and R. E. Dickerson, Structure of a B-DNA dodecamer at 16 K, *Proc. Natl. Acad. Sci. USA*, **79**, 4040-4044(1982).
110. O. Kennard, W. B. T. Cruse, J. Nachman, T. Prange, Z. Shakked and D. Rabinovich, Ordered water structure in an A-DNA octamer at 1.7 Å resolution, *J. Biomol. Struct. Dynam.*, **3**, 623-647(1986).
111. M. Eisenstein, F. Frolow, Z. Shakked and D. Rabinovich, The structure and hydration of the A-DNA fragment d (GGGTACCC) at room temperature and low temperature, *Nucl. Acids Res.*, **18**(11), 3185-3194(1990).
112. M. Eisenstein and Z. Schakked, Hydration patterns and intermolecular interactions in A-DNA crystal structures: implications for DNA recognition, *J. Mol. Biol.*, **248**, 662-678(1995).
113. B. Schneider, D. M. Cohen, L. Schleifer, A. R. Srinivasan, W. K. Olson and H. M. Berman, A systematic method for studying the spatial distribution of water molecules around uncleic acids bases, *Biophys. J.*, **65**, 2291-2303(1993).
114. L. A. Lipscomb, M. E. Peek, F. X. Zhou, J. A. Bertrand, D. VanDerveer and L.D. Williams, Water ring structure at DNA interfaces: hydration and dynamics of DNA-Anthracycline complexes, *Biochemistry*, **33**, 3649-3659(1994).

115. B. Schneider and H. M. Berman, Hydration of the DNA bases is local, *Biophys. J.*, **69**, 2661-2669(1995).
116. G. Humner, A. E. Garcia and D. M. Soumpasis, Hydration of nucleic acids fragments: comparison of theory and experiment for high-resolution crystal structures of RNA, DNA, and DNA-drug complexes, *Biophys. J.*, **68**, 1639-1652(1995).
117. V. P. Denisov, G. Carlström, K. Vence and B. Halle, Kinetics of DNA hydration, *J. Mol. Biol.*, **268**, 118-136(1997).
118. W. C. Johnson, Jr., Determination of the conformation of nucleic acids by electronic CD, *Circular Dichroism and the Conformational Analysis of Biomolecules*, edited by G. D. Fasman, Plenum Press, New York, 1996, pp. 433-468.
119. W. –Y. Qiu and H. -W. Xin, Topology of circular duplex DNAs, *Science in China B*, in press.
120. R. E. Dickerson, DNA structure from A to Z, In *Methods in Enzymology: DNA Structure*, Vol. 211, Part A, edited by D. M. J. Lilley and J. E. Dahlberg, Academic Press, Inc., 1992, pp. 67-111.
121. C. -H. Lee, H. Mizusawa and T. Kakefuda, Unwinding of double-stranded DNA helix by dehydration, *Proc. Natl. Acad. Sci. USA*, **78**, 2838-2842(1981).
122. C. R. Calladine and H. R. Drew, A base-centered explanation of the B-to-A transition in DNA, *J. Mol. Biol.*, **178**,773-783(1984).
123. M. Behe and G. Felsenfeld, Effects of methylation on a synthetic polynucleotide: the B-Z transition in ploy (dG-m^5dC) ploy (dG-m^5dC), *Proc. Natl. Acad. Sci. USA*, **78**, 1619-1623(1981).
124. L. J. Peck and J. C. Wang, Energetics of B-to-Z transition in DNA, *Proc. Natl. Acad. Sci. USA*, **80**, 6206-6210(1983).
125. P. O. Brown and N. R. Cozzarelli, A sign inversion mechanism for enzymatic supercoiling of DNA, *Science*, **206**, 1081-1083(1979).
126. J.C. Wang, DNA topoisomerases, *Annu. Rev. Biochem.*, **54**, 665-697 (1985).
127. J.C. Wang, DNA topoisomerases, *Annu. Rev. Biochem.*, **65**, 635-692 (1996).
128. L. F. Liu, R. E. Depew and J. C. Wang, Knotted single-stranded DNA rings: A novel topological isomer of circular single-stranded DNA formed by treatment with Escherichia coli ω protein, *J. Mol. Biol.*, **106**, 439-452 (1976).

129. J. E. Mueller, S. M. Du and N. C. Seeman, Design and synthesis of a knot from single-stranded DNA, *J. Am. Chem. Soc.*, **113**, 6306-6308 (1991).
130. N. C. Seeman, The design of single-stranded nucleic acid knots, *Mol. Eng.*, **2**, 297 (1992).
131. S. M. Du and N. C. Seeman, Synthesis of a DNA knots containing both positive and negative nodes, *J. Am. Chem. Soc.*, **114**, 9652-9655 (1992).
132. H. Wang, S. M. Du and N. C. Seeman, Tight single-stranded DNA knots, *J. Biomol. Str., Dyns.*, **10**, 853-863 (1993).
133. N. C. Seeman, J. Chen, S. M. Du, J. E. Mueller, Y. Zhang, T.-J. Fu, Y. Wang and S. Zhang, Synthetic DNA knots and catenanes, *New J. Chem.*, **17**, 739-955 (1993).
134. S. M. Du, B. D. Stollar and N. C. Seeman, A synthetic DNA molecule in three knotted topologies, *J. Am. Chem. Soc.*, **117**, 1194-1200 (1995).
135. S. M. Du, H. Wang, Y.-C. Tse-Dinh and N. C. Seeman, Topological transformations of synthetic DNA knots, *Biochem.*, **34**, 673-682 (1995).
136. A. G. W. Lesliet, S. Arnott, R. Chandrasekaran and R. L. Ratliff, Polymorphism of DNA double helices, *J. Mol. Biol.*, **143**, 49-72(1980).
137. J. H. C. Whitehead, On doubled knots, *J. London Math. Soc.*, **12**, 63-71(1937).
138. W.-Y. Qiu and H.-W. Xin, Topological chirality and achirality of DNA knots, *J. Mol. Struct.* (*Theochem*), **429**, 81-86 (1998).
139. K. Mizuuchi, L. M. Fisher, M. H. O'Dea and M. Gellert, DNA gyrase action involves the introduction of transient double-strand breaks into DNA, *Proc. Natl. acad. Sci. USA*, **77**, 1847-1851(1980).
140. T.-S. Hsieh, Knotting of the circular duplex DNA by type II DNA topoisomerase from drosophila melanogaster, *J. Bio. Chem.*, **258**, 8413-8420 (1983).
141. J. D. Griffith and H. A. Nash, Genetic rearrangement of DNA induces knots with a unique topology: Implications for the mechanism of synapsis and crossing-over, *Proc. Natl. acad. Sci. USA*, **82**, 3124-3128 (1985).
142. P. Dröge, Recombination of nicked DNA knots by γδ resolvase suggests a variant model for the mechanism of strand exchange, *Nucl. Acids Res.*, **20**(23), 6159-6166(1992).
143. S. Y. Shaw and J. C. Wang, Knotting of a DNA chain during ring closure, Science, 260, 533 (1993).

144. P. O. Brown and N. R. Cozzarelli, Catenation and knotting of duplex DNA by type 1 topoisomerases: a mechanistic parallel with type 2 topoisomerases, *Proc. Natl. Acad. Sci. USA*, **78**, 843-847 (1981).
145. M. A. Krasnow, A. Stasiak, S. J. Spengler, F. Dean, T. Koller and N.R. Cozzarelli, Determination of the absolute handedness of knots and catenanes of DNA, *Nature*, **304**, 559 (1983).
146. S. A. Wasserman and N. R. Cozzarelli, Determination of the stereostructure of the product of Tn3 resolvase by a general method, *Proc. Nat. Acad. Sci. USA*, **82**, 1079-1083 (1985).
147. F. B. Dean, A. Stasiak, T. Koller and N. R. Cozzarelli, Duplex DNA knots produced by Escherichia coil topoisomerase I, *J. Biol. Chem.*, **260**, 4795-4983 (1985).
148. S. A. Wasserman, J. M. Dungan and N. R. Cozzarelli, Discovery of a predicted DNA knot substantiates a model for site-specific recombination, *Science*, **229**, 171-174 (1985).
149. S. J. Spengler, A. Stasiak and N. R. Cozzarelli, The stereostructure of knots and catenanes produced by phage λ integrative recombnation: implications for mechanism and DNA structure, *Cell*, **42**, 325-334(1985).
150. P. Dröge and N. R. Cozzarelli, Topological structure of DNA knots and catenanes, *Meth. Enzym.*, **212**,120-130(1992).
151. V. V. Rybenkov, C. Ullsperger, A. V. Vologodskii and N. R. Cozzarelli, Simplification of DNA topology below equilibrium values by type II topoisomerases, *Science*, **277**, 690-693 (1997).
152. S. Y. Shaw and J. C. Wang, Chirality of DNA trefoils: implications in intramolecular synopsis of distant DNA segments, *Proc. Natl. Acad. Sci. USA*, **94**,1692-1697(1997).
153. A. Stasiak, V. Katritch, J. Bednar, D. Michoud and J. Dubochet, Electrophoretic mobility of DNA knots, *Nature*, **384**, 122(1996).
154. A. V. Vologodskii, N. J. Crisona, B. Laurie, P. Pieranski, V. Katrich, J. Dubochet and A. Stasiak, Sedimentation and electrophoretic migration of DNA knots and catenanes, *J. Mol. Biol.*, **278**, 1-3(1998).
155. B. Laurie, V.Katritch, J. Sogo, T. Koller, J. Dubochet and A. Stasiak, Geometry and physics of catenanes applied to the study of DNA replication, *Biophys. J.*, **74**,2815-2822(1998).
156. W. -Y. Qiu, Topological structure of duplex DNA knots, in preparation.

157. D. A. Clayton and J. Vinograd, Circular dimer and catenate gorms of mitonchondiral DNA in human leukaemic leucocytes, *Nature*, **216**, 652-656 (1967).
158. J. C. Wang and H. Schwartz, Noncomplementarity in base sequences between the cohesive ends of coliphages 186 and lambda and the formation of interlocked rings between the two DNAs, *Biopolymers*, **5**, 953-966 (1967).
159. D. E. Adams, E. M. Shekhtman, E. L. Zechiedrich, M. B. Schmid and N. R. Cozzarelli, The role of topoisomerase IV in partitioning bacterial replicons and the structure of catenated intermediates in DNA replication, *Cell*, **71**, 277-288(1992).
160. D. E. Pulleyblank, Molecular biology of topo and Maxwell's dream, *Science*, **277**, 648-649 (1997).
161. A. D. Bates and A. Maxwell, DNA topology: Topoisomerase keep it simple, *Current Biology*, **7**, 778-781(1997).
162. S. A. Wasserman and N. R. Cozzarelli, Determination of the stereostructure of the product of Tn3 resolvase by a general method, *Proc. Natl. Acad. Sci. USA*, **82**, 1079-1083 (1985).
163. L. Yang, M. S. Wold, J. J. Li, T. J. Kelly and L. F. Liu, Roles of DNA topoisomerases in simian virus 40 DNA replication in vitro, *Proc. Natl. Acad. Sci. USA*, **84**, 950-954 (1987).
164. W. M. Stark and M. R. Boocock, Topological selectivity in site-specific recombination, In *Mobile Genetic Elements* (D. J. Sherratt ed.), ARL Press, Oxford Univ. Press, Oxford, 1995, pp.101-129.
165. W. -Y. Qiu, Topological structure of duplex DNA catenations, in preparation.
166. R. F. Voss, Evolution of long-range fractal correlations and 1/f noise in DNA base sequences, *Phys. Rev. Lett.*, **68**(25), 3805-3808(1992).
167. D. R. Hofstadter, Is the genetic code an arbitrary one, or would another code work as well? *Sci. Amer.*, **246**(3), 18-29(1982).
168. V. Avetisov and V. Goldanskii, Mirror symmetry breaking at the molecular level, *Proc. Natl. Acad. Sci. USA*, **93**,11435-11442(1996).