*Critical Reviews in Biochemistry and Molecular Biology*, 26(5/6):475–559 (1991)

# Paranemic Structures of DNA and their Role in DNA Unwinding

*Gad Yagil*

Department of Cell Biology, The Weizmann Institute of Science, Rehovot, Israel



**ABSTRACT:** A DNA structure is defined as paranemic if the participating strands can be separated without mutual rotation of the opposite strands. The experimental methods employed to detect paranemic, unwound, DNA regions is described, including probing by single-strand specific nucleases (SNN), conformation-specific chemical probes, topoisomer analysis, NMR, and other physical methods. The available evidence for the following paranemic structures is surveyed: single-stranded DNA, slippage structures, cruciforms, alternating B-Z regions, triplexes (H-DNA), paranemic duplexes and RNA, protein-stabilized paranemic DNA. The problem of DNA unwinding during gene copying processes is analyzed; the possibility that extended paranemic DNA regions are transiently formed during replication, transcription, and recombination is considered, and the evidence supporting the participation of paranemic DNA forms in genes committed to or undergoing copying processes is summarized.

**KEY WORDS:** DNA, paranemic, topoisomers, superhelicity, duplex, triplex, transcription, replication, enhancers, recombination, unwinding, unwound.

## I. INTRODUCTION*

In their paper to the Cold Spring Harbor Symposium of 1953, Watson and Crick wrote: "The third difficulty to this scheme (of DNA expression) involves the necessity of the two strands to unwind in order to serve as a template for a new chain. This is a very fundamental difficulty when the two chains are interlaced as in our model. The difficulty is a topological one and cannot be surmounted by simple manipulation. Apart from breaking the chains there are only two sorts of ways to separate chains coiled plectonemically: . . . one . . . simply pulls in axial direction . . . ; in the second way the chains must be directly untwisted . . . once every 34 Å . . . in our structure".[519]

Today, 38 years thereafter, a vast amount of knowledge has accumulated on how DNA templates are regulated and copied; a detailed solution to the unwinding difficulty is nevertheless not yet available. Mechanisms for breaking and rejoining DNA strands have been discovered and the enzymes involved have been characterized, but the exact relationship between strand breakage and DNA expression is far from understood. It is clear that mutual strand rotation must take place at some stage during replication or transcription. It is not yet known, however, at what locations strands are broken and what the state of DNA between two breakage-rejoining events can be. For DNA replication, a rotation around a "swivel" has long been considered, but only most recently have some of the details of that swivel begun to emerge.[16,75] We do not know yet the length of the rotating DNA or even whether the rotating strands are the yet to be replicated ones ahead of the fork or the just copied strands behind it. We do not know yet whether, for transcription, for instance, uncoiling takes place once

* A list of abbreviations used throughout this article appears after the acknowledgment section and before the reference list.

1040-9238/91/$.50

every passage of a polymerase, or whether at high transcription rates DNA uncoils for a batch of polymerases to pass, or whether, in very active genes, DNA may not pass altogether into a temporarily unwound state.

At least two basic approaches can be distinguished. According to one approach, transcription can proceed with minimal strand separation, superhelical turns providing the amount of unwinding necessary for transcription to proceed. Experimental evidence shows that the size of the necessary transcription "bubble" may not exceed 15 to 20 bp.[149] Transcription can then proceed by mutual rotation of DNA and product RNA (Figure 20). It is not known whether the polymerase and the nascent RNA chain rotate around the DNA template or whether DNA itself rotates to permit the transcription "bubble" to proceed. More recent work indicates that at least some of the rotation may be by DNA around the polymerase-nascent RNA complex.[286] That work demonstrates that transcription in a closed circle is accompanied by the build up of superhelical turns, which have to be released by the action of DNA topoisomerases. This implies that at least some chain breakage-joining follows, rather than precedes, the transcription event.[481]*

According to the other approach, larger segments of DNA become unwound, either ahead of or behind the template-copying machinery. This is supported by the requirement for topoisomerases, for transcription as manifested by the inhibition of the processes by specific topoisomerase inhibitors, or following topoisomerase gene disruption (see Section IV.B.4). Replication processes also require the participation of the topoisomerases[548] and are impaired in bacterial and yeast strains lacking topoisomerase activity. The detailed evidence is brought in Section IV.A. This requirement for topoisomerases gives rise to the possibility that the activation of frequently copied genes is accompanied by their transition into an at least partially unwound state, permitting the rapid passage of the copying machinery.

This review will concentrate mainly on the question of whether DNA can exist in unwound forms, how these forms can be identified, what detailed structure they may assume, and what their role in replication, transcription, and other gene expression processes might be. A number of different structures have been proposed recently for topologically unwound DNA. We shall designate these structures, following the proposal of Watson and Crick,[519] as *paranemic DNA structures*.** A DNA structure is considered as paranemic when its two strands can separate without the introduction of strand breakages and mutual rotation of the two strands. The cruciform and H forms of DNA are at present the most intensively studied paranemic DNA structures. Both structures contain locally wound segments, but these present no barrier to strand separation. The evidence for the existence of these and other proposed structures is described in detail in the third section of the review. Most experimental studies concerning the unwound structures are currently performed on closed, supercoiled DNA forms, including plasmids and viruses. The experimental methods employed to study unwound DNA follow this introductory section. Before proceeding to the experimental methods used to study unwound DNA, the concepts of superhelical winding and our knowledge of the enzymes playing a role in unwinding (topoisomerases and helicases) are briefly summarized.

* A list of abbreviations used throughout this article appears after the acknowledgment section and before the reference list.

* At least two mechanisms by which the DNA duplex can be copied without prior strand separation can be envisaged. By one mechanism, the bases of one strand swing out of the duplex to serve as a template for an incoming base. Alternatively, the incoming base can recognize an intact duplex by Hoogsteen-type base pairing, forming a DNA triplex as an intermediate. With both mechanisms, the new strand needs yet to revolve for the products to disentangle. No evidence in favor of or against these possibilities could be found.

** The terms *paranemic* and *plectonemic coiling* were proposed by Watson and Crick[519] to describe the nonintertwined and intertwined states of DNA. These terms were previously used to describe similar states observed in chromosomes during cellular replication. *Para* in Greek means beside, *plexus* — a knot, and *nema* — a thread. *Paranemic*, therefore, means threads arranged in a parallel, nonintertwined fashion (also referred to in this review as *unwound*), in contrast to entangled, knot-forming threads. The term *paranemic joint* was adopted by Radding and co-workers[33] to describe the noninterwined intermediate formed during DNA recombination (see Section IV.C).

## A. Early Developments

Several eminent molecular biologists considered solutions to the unwinding problem soon after it was found to exist;[67,110,148] these early suggestions are reviewed in detail by Gorski.[168] Elements of the mechanisms considered today, including a partial transition into left-handed DNA,[519] intermediate cruciform formation, and anti-coiling[148] can be found in these early papers.

The problem was temporarily laid to rest when it was first calculated that no serious time barrier to the rotation of linear DNA around its axis is expected in a medium of common viscosity.[276,289] This was soon supported by measurements of the rate of DNA unwinding, then believed to have a molecular weight of just a few mDa (~100,000 bp). Separation times were found to be well below 1 s, i.e., many thousands of bases per second, sufficient to allow for strand separation of bacteriophage DNA.[97,105,460,529] Separation times were, however, found to be proportional to the square of the molecular weight, reflecting the effect of both increased DNA length and of increased viscous drag on the rotating molecule.

The unwinding problem became acute again when it was realized that

1. Intact, unnicked DNA can be much longer than the linear phage DNA previously analyzed; chromosomal DNA was found to exist mostly as one long, uninterrupted thread[226]
2. Viral and organelle DNA are often circularly closed so that strands cannot start to separate without prior strand breakage[191]
3. Nuclear DNA is not free to rotate as in aqueous solutions, but is tightly coiled around proteins, presenting a further barrier to strand separation[249,316,491]
4. Chromosomal DNA is further anchored to a scaffold or matrix, from which it emerges in closed loops, manifesting an ultimate topological barrier to unwinding[30,307,350]

Two discoveries of the same period pointed, nevertheless, toward a resolution of the unwinding dilemma: first, it was found that DNA untwisting, or "nicking-closing", activities, later termed DNA topoisomerases, are present in the cell nucleus.[77,509] These provide a machinery for temporary strand breakage and enable the unwinding of specific gene segments. The existence of topoisomerases thus makes the notion that defined DNA segments can unwind and rewind independently of neighboring gene regions reasonable.

Second, most DNA extracted from cells was discovered to have an appreciable degree of negative superhelical winding. The release of the negative superhelical turns of a circular or looped DNA region can cause transition of up to 5% of the DNA genome into an unwound form without opening the circle and at minimal energy cost.[23,494]

While the combined operation of superhelical strain and topoisomerase activity removes the difficulties in principle imposed by circularity and the high molecular weight of DNA, the detailed mechanism of strand separation during DNA expression remains to be elucidated. Studies of replication and transcription were performed, until recently, mostly on short linear DNA templates. When a circular template was used, it proved hard to ascertain that template linearization had not preceded transcription. Most older reviews dealing with transcription therefore avoided the topological aspect of transcription initiation (see Reference 290, for example). Only a few studies addressed questions such as at what stage during the initiation of transcription, or replication, DNA unwinding commences, or what the size of the unwound regions ("transcription bubble") can be.

Answers to these questions may prove essential to the understanding of the control of DNA activity, and we shall return to this problem in the last section of the review.

## B. Supercoiled DNA

### *1. Closed DNA*

It is well established today that most natural extrachromosomal DNA, including organelles, plasmids, and many viruses, exists as covalent closed circles. Chromosomal DNA is also found in the nucleus, mostly in topologically isolated loops, anchored on some base structures with no local free ends from which unwinding can com-

mence. When freed from proteins, chromosomal DNA does not appear as extended, flat circles, but rather in highly twisted, supercoiled shapes. We shall call all forms of DNA, circular, anchored or other, *closed* (or *clamped*) if their ends are not free to revolve around each other.*

The first DNA that was established to exist as a closed circle (based on resistance to exonucleases) was the single-stranded form of bacteriophage $\phi x174$.[134] It was soon found that the duplex, replicating form of this virus is also circularly closed, as are several other bacterial viruses (reviewed in Reference 191). Other bacteriophages, like phage lambda, are closed by the association of cohesive ends (terminal repeats) during their replication cycle. The DNA of the small mammalian viruses polyoma and SV40 were also found to be covalently closed, both during the replicative stage and in the virion form.[123,521] Adenovirus forms a closed structure by virtue of terminal-associated proteins.[405] DNA of organelles, including mitochondrial and chloroplasts, was similarly established to exist in a circularly closed state.[347,401] A closed circular structure of DNA is thus a very widespread property of nonchromosomal DNA.

Considering chromosomal DNA, Paulson and Laemmli[376] observed that under mild preparative procedures nuclear DNA from metaphase chromosomes does not appear in an opened linear form, but rather as loops that exit and reenter the "nucleoid substance". This loop, or domain, structure has been confirmed by isolating the nucleoid substance from interphase nuclei. A brief treatment with DNAse I, or intercalation with ethidium bromide (EtBr), causes the structure to relax into a more "open" form.[30] A similar domain structure has been found for bacterial nucleoid preparations.[469,538] The size of the eukaryotic loop has been estimated to vary from a few thousand to several hundred thousand bases, based on unwinding experiments by intercalating agents, on psoralene cross-linking,[453] on restriction endonuclease cutting,[213] and on quantitative electron micrography. The intercalation and irradiation studies have further shown that each of these domains can be independently relaxed.[453] Each can therefore be expected to be an independent topological domain.

The chromosomal DNA loops have been found to be attached, in metaphase chromosomes, to an insoluble protein "scaffold".[332] In interphase nuclei, the DNA loops are attached to protein elements of the nuclear matrix. Specific DNA tracts that attach to these protein elements have been identified in the 5′ upstream region of a number of genes. These include mouse immunoglobulin kappa (MARs),[84] and human beta interferon genes,[328] as well as *Drosophila* histone, alcohol dehydrogenase, sgs-4, fushi tarazu, and Hsp-70 genes (SARs).[2,151] It has been further found that DNA topoisomerase II is concentrated in the nuclear matrix fraction[125,126] and in particular at the DNA attachment sites.[2] This is a strong indication that the DNA matrix attachment site may serve as the much sought for swivel point for DNA unwinding. Whether the DNA copying machinery joins the attachment site with the DNA rotating through it, or the machinery moves along the loop rotating with the DNA around the fixed swivel point, or DNA is detached from the matrix site altogether during copying cannot be said at present. In any case, a closed DNA structure seems to be a very widespread feature, making unwinding without strand breakage a virtual impossibility.

**The linkage deficit ($\Delta$Lk)** — The existence of part of a double-stranded molecule in a wholly or partly underwound state implies a decrease in the number of times one strand is wound around the other. A closed double-stranded DNA molecule is subject to the principle of linkage conservation, namely, the principle that the number of times that one strand passes over the other when projected on a suitable surface. This number is termed the *linking number*, Lk, and is constant (integral when the molecule is *circularly* closed). In the simplest case of a non-superhelical, fully relaxed DNA duplex, Lk is equal to the number of primary turns, e.g., one turn per 10.5 bases in B-form DNA. If, however, a section of the DNA is completely unwound, then the linking number is appropriately reduced. Several other structural features lead to decreased (or

* This use of the term *closed* is not to be confused with the closed and open forms of *transcription* complexes, where the extent of strand separation is implied. We shall occasionally use the term *open linear DNA* as an antonym to closed or clamped DNA in the sense used here, i.e., to DNA that cannot change its degree of linking.

increased) linking numbers. These include superhelical winding (described below), wrapping around protein complexes, such as nucleosomes (discussed in Section I.B.3), or several forms of knotting (beyond the scope of this review). Unwound and superhelical forms are topologically equivalent, i.e., they can pass into each other without opening or closing covalent bonds, as long as the linking number remains unchanged.*

For the present discussion, the point of importance is that the presence of an unwound section in a closed DNA molecule is manifested as a reduced linking number. The change in linking number is expressed by the *linking deficit*, $\Delta Lk$, which is the difference in linking number between the unwound DNA and the same molecule, fully wound and relaxed (with no superhelical turns). Linking deficits can be experimentally measured, as described in detail in Section II.C. It is common to describe a linking deficit not only the absolute number of turns, but also relative to the full length of the molecule, $\sigma = \Delta Lk/Lk$. $\sigma$ is called the superhelical density, because in purified DNA the linking deficit is generally (but not always, as will be seen) manifested as an interwound superhelix.** Lk of the relaxed molecule is the full length of the molecule divided by the number of bases per turn, 10.5 for B-form DNA. $\sigma$ can thus represent either the ratio of primary to secondary turns in a superhelical molecule, or the fraction of unwound turns in a locally unwound, closed DNA molecule.

## 2. Superhelicity

The establishment of the circularity of polyoma virus DNA was soon followed by the finding that the circle is extracted from cells as a highly twisted, supercoiled form.[493] This was initially deduced from the observation that the viral DNA extracted from the cell sediments in alkaline CsCl much faster than after being nicked once or a few times. In nicked circles one of the strands has an opportunity to open, revolve around the second strand, and form a "relaxed" circle. The number of superhelical turns could be measured by titration either with alkali or with ethidium bromide.[22] The number of superhelical turns (or superturns, distinguished from the primary turns of the DNA duplex, 1 per 10 bp) was initially determined to be 13 turns for the 5.5-kbp long polyoma and SV40 viruses.

More exact determinations, by counting bands on an agarose gel,[111,227,398] established an average number of 25 turns per viral DNA molecule, i.e., 1 superturn every 220 bp. The 25 superturns imply a linking deficit of $-25$ of the otherwise B form DNA plasmid. The ratio of the superturns to primary turns, i.e., the superhelix density, is thus $\sigma = -25/(5500/10.5) = -0.048$, or 4.8% of the primary turns, for the average polyoma viral DNA molecule. Negative superhelical densities of 4 to 5% were soon found for most cellular DNA preparations studied — viral, organelle, and plasmid.[508] Superhelical densities of the same magnitude were also found for chromosomal loop DNA[30] and the DNA released from bacterial nucleoid preparations.[380]

The sense of all naturally supercoiled DNA is negative,* which means that in a closed circle or restrained loop of right-handed, B-form DNA, the opening of each superhelical turn is accompanied by the *opening* of 1 turn (10.5 bases) of the primary duplex. Thus, a superhelical density of 5% means that approximately 5% of the primary turns are unwound when the negative superhelicity is completely released. In SV40 (5243 bp), for instance, the opening of the 25 superturns[155,227] permits the opening of 25 pri-

* There are several ways for the quantitative expression of linkage conservation:[25,26] The expression $Lk = Tw + Wr$ is mathematically proven but not always easy to dissect into primary or secondary (superhelical) winding; Tw is the twist, related but not necessarily equal to the number of primary turns and Wr is the writhe, related but not equal to the number of secondary turns. The original expression $Lk(\alpha) = \tau + \beta$, has been criticized in that experimentally titrable superhelical turns do not necessarily match the number of secondary turns present.[24] More recently, a concept of surface linking, SLk, has been proposed[530] to represent the wrapping of the duplex axis around a surface, and maintains $Lk = SLk + \Phi$ where $\Phi$ is defined as the winding number. This expression is helpful in analyzing situations involving changes in the helical repeat, as during the wrapping of DNA around nucleosomes. More exact definitions of superhelical and primary winding numbers, linking number, and superhelix density, can be found in References 424, 515, and 530.

** The following rule of thumb is helpful: the sense of the supercoil is considered negative when the opening of one supercoil necessitates the opening, rather than the closing, of one turn of a right-handed (e.g., B-form) primary helix.

mary turns of the duplex, i.e., some 250 bases. The superhelicity of SV40 DNA can thus provide to a certain degree the topological demands of strand separation, sufficient for creating an initial open complex, but far from enough for unwinding a complete SV40 gene, which can extend over thousands of bases, certainly not for the complete unwinding required for the replication of the virus. This is generally true — superhelical density is seldom sufficient to enable the unwinding of more than a fraction of a topologically isolated domain.

### 3. Restrained DNA

The role of supercoiling may be even more limited because it turns out that in the intact cell superturns are not "free", but rather "restrained" (or "constrained"), i.e., unavailable for unwinding, because of association with proteins or other nuclear components. This became clear when SV40 DNA was examined in its prevalent cellular form, i.e., associated with histones and organized in nucleosomes ("minichromosomes"). Minichromosomes, in contrast to free viral DNA, do not revert to relaxed circles when treated with nicking-closing activity.[155] The explanation is that DNA is wound around the histone cores, some 1.75 turns per histone core. This "toroidal" winding of DNA around the histone core is topologically equivalent to the interwound superhelical winding* usually observed with plasmid DNA. This can be readily verified by winding a two-stranded wire or rubber tube around an object such as a rod and removing the object while keeping the ends together. The 24 nucleosomes observed on SV40 DNA[425] thus imply that some $24 \times 1.75 = 42$ superturns ought to be engaged in the interaction with the histone cores. This is more than the 25 superhelical turns actually found (see above) for the viral DNA and was considered a paradox when first noticed. The difference was soon explained as being due to a concomitant change in the average angle between adjacent bases from 34.1° in free B-form DNA to 35.4° is DNA wrapped around the nucleosomes ($1.3 \times 5243/360 = 19$).[239,240]

The 25 superturns observed when viral DNA is extracted from the cell thus reflect the nucleosomal organization of the DNA. The superturns are not immediately available for unwinding of DNA organized in the chromosomal state, but only after the associated proteins have been removed by some process. The 25 superturns are probably introduced when nucleosomes are formed prior to DNA circle closure, remain restrained as long as nucleosomes are present, and are manifested as free superhelical turns only after the histones have dissociated by some mechanism. Alternative explanations for the generation of superturns are conceivable; for example, that the 25 superturns are introduced when a circle is closed before the last 5% of the duplex had been wound.[494] The winding of these 5% into a B-type duplex can then be a driving force for nucleosome formation. DNA has been found to also wrap around other nuclear proteins, for example, DNA gyrase[234] or RNA polymerase,[15] and this can help to generate additional superturns. Under cellular conditions, superturns are thus expected to be mostly restrained and not immediately available for unwinding.

There have been a number of attempts to measure the fraction of nonrestrained superhelical turns, i.e., the fraction immediately available for unwinding in the intact nucleus or cell. Bliska and Cozarelli,[36] using a probe undergoing site-specific recombination, estimate that nearly 40% of the superhelical turns are unrestrained in the intact bacterial cell. Sinden and Kochel,[452] using a probe undergoing a transition into Z-form DNA, estimate that only 3 to 5% are unrestrained in bacteria. The extent of restraining seems to depend on the DNA probe used. In any case, a process releasing restraining proteins, such as side-chain acetylation, must take place before most superhelical turns can be converted to unwound DNA. This has been considered repeatedly as a way to control template utilization.[122,176,243]

In summary, superhelical turns of DNA may

* The terms supercoils, superhelical turns, or in brief, superturns, are used interchangeably throughout this review, mostly in a general sense, encompassing toroidal wrapping around a core, formation of an intertwined superhelix, combinations thereof, as well as other geometric forms leading to a linking deficit or excess. The adjective interwound is added to superhelix when this type of supercoiling is specifically meant.

satisfy the topological requirements for strand separation, without breakage, or at most 5% of DNA organized in loops or circles, and this only when all the turns are available for the process. Alternative processes invoking strand breakage must be available if a larger fraction of the DNA has to be unwound.

## C. DNA Topoisomerases

Strand unwinding in excess of the 5% provided by natural superhelicity requires a DNA nicking-closing activity. Such an activity was first discovered in *E. coli* in 1971[509] and 1 year later in mammalian cells.[77] Many DNA nicking-closing activities have been characterized since, and are now termed *DNA topoisomerases* (reviewed in References 152, 463, and 510). DNA topoisomerases are classified into topoisomerases I and II, according to whether only one strand or both strands are cleaved. In prokaryotes, the principal type II topoisomerase requires ATP for its activity and is termed *DNA gyrase*.[153] The participation of topoisomerases in both replication and transcription has been established using specific topoisomerase inhibitors as well as mutants lacking the enzymes and various gene disruption strategies. Thus, either topoisomerase I or topoisomerase II activity can be absent when *E. coli* DNA synthesis takes place, but both activities cannot be absent together.[55,117] Similar data are available for yeast[169,483,546] as well as in SV40.[547] Thus, topoisomerase I can be replaced by topoisomerase II and vice versa, but one of them must be active for template copying to take place. Topoisomerase II is required for the final segregation of the replicated chromosomes.[117,417,548] In yeast, some transcription can still proceed in mutants lacking both topoisomerase I and topoisomerase II activities, so that a yet-uncharacterized topoisomerase might exist.[56,504] An *E. coli* topoisomerase III has been reported recently.[116] The role of supercoiling in transcription will be discussed further in Section IV.

DNA topoisomerases have generally been characterized by their ability to remove or introduce superhelical turns, or otherwise change the linking number of closed DNA. The superturns are removed by first nicking either one strand of the enzyme remaining attached to the end of one cleaved strand. Nicking is followed by revolution of one of the opened strands around the other, and finally nicked DNA ends are rejoined. Alternatively, in class II topoisomerases, both strands are cleaved but remain associated via the topoisomerase, and a change in winding can take place, e.g., by passage of DNA through the enzyme molecule. These modes of action, without the rejoining steps, can serve to unwind primary turns of DNA, as necessary for transition into an unwound state. Topoisomerases are likely candidates for primary strand unwinding, not only because of their ability to rejoin ends after revolutions have taken place, but also because they are localized either in chromatin (topoisomerase I)[138,160,216] or in the nuclear matrix (topoisomerase II).[2,126]

In case topoisomerases do participate in strand unwinding, the question arises as to where and how often they cut. Do they cut upstream or downstream of a transcribed DNA region? Do they cut the transcribed region at one site or in several sections? Do they cut at defined locations along a gene? The last question has been studied intensively in SV40. As many as 265 major cutting points were found when topoisomerase I was applied to purified SV40 DNA.[28,175,388] A prominent cutting point was nevertheless reported at base 4955, a few hundred bases into the early transcribed region, when camptothecin was used to enhance the reaction.[217] This pattern is not always found.[236,475] With topoisomerase II, a prominent cutting site was found at base 270, just downstream of the late protein transcription start.[548] When doxorubicin was used to stabilize the topo II-cleaved DNA complex, the main cleavage sites shifted, interestingly, to a region associated with matrix attachment (4200 to 4300).[71]

These results were largely obtained with isolated DNA as the target. A multitude of DNA gyrase cutting sites were detected upon application of oxolinic or naldixic acid to pBR322 in bacterial cells, but a consensus cutting sequence appearing at 11 locations could be identified.[287] DNA topoisomerase cutting sites along the *Xenopus laevis* r-DNA gene have been examined using intact nuclei as topoisomerase I targets.[99]

Many cleavage sites, spaced by approximately 200 bases, were found. This is approximately the distance between nucleosomes, so that in this case topoisomerase may act preferentially in linker regions. Whether the same sites will also be preferentially cut during the transcription of the r-DNA genes in the intact cell cannot be said at present. Isolated r-DNA was cut by topoisomerase I along its entire length. More on this topic can be found in Section IV.B.4. Thus there is little conclusive evidence for or against specific unwinding start sites at the beginning or the end of transcribed regions. Detailed studies of cell-free transcription of circular closed DNA may be needed for a definite answer.

## D. SSB Proteins and Helicases

The best known form of unwound DNA is melted, single-stranded DNA. The melting of linear DNA is well known to be a highly cooperative process. This implies that at temperatures well below the melting transition, including room temperature or even 37°, strand separation is thermodynamically highly disfavored. Strand separation can also be a kinetically slow process. It is therefore not surprising, in retrospect, that proteins affecting the position of the ds-ss DNA equilibrium, as well as enzymatic activities catalyzing the rate of ds DNA strand separation, have been discovered.

The best-studied group of proteins shifting the equilibrium between wound, ds DNA and unwound, ssDNAs are the single-strand binding proteins ("SSB proteins"). The first SSB protein found was gp32, a gene product of bacteriophage T4 necessary for the replication of the phage.[4] SSB proteins with similar properties were soon discovered in *E. coli*,[451] yeast,[222] and mammalian tissues.[197,228] Evidence that proteins HMG1 and 2, components of active chromatin, also have SSB properties has been reported.[178,439] The SSB proteins are well reviewed, and we shall not discuss them in detail here.[79,288] Most of them are present in cells in amounts sufficient for stoichiometric combination with long stretches of ssDNA. SSB proteins can therefore serve to stabilize at least temporarily the unwound, single-stranded state.

More recently, a group of enzymes catalyzing the *rate* of DNA unwinding and strand separation were characterized and are communally named helicases.[306] Helicases utilize ATP to catalyze the separation of the strands of nicked or linear DNA, but not of circular closed DNA. Helicase activity is measured by the rate at which a labeled complementary DNA fragment hybridized to a closed ssDNA circle is displaced after addition of the enzyme, the analysis being performed on acrylamide gels.[263,493] The first helicase identified was helicase I of *E. coli*[1] later found to be coded by the F plasmid of *E. coli*.[258] At least seven helicase activities from *E. coli* are known,[310,536] They include the dnaB protein (helicase III?), which is essential for *E. coli* replication, acting along the template in the 5′ to 3′ direction,[17,18] as well as protein PriA (Y,n′) acting along the 3′ to 5′ direction.[261,269] Both helicase activities are associated with the primosome and migrate with it during replication. While dnaB serves to separate the parental strands, PriA has been proposed to prepare the lagging strand for priming and polymerization.[18] SSB and DNA gyrase are required for the unwinding to proceed.

The gene product UvrD of *E. coli* (helicase II) is active in repair and can separate both DNA · DNA and DNA · RNA hybrids.[310,311] The UvrABC complex has also been found to exhibit helicase activity.[211,357] Several bacteriophages possess their own helicases, like protein 41 of phage T4[74,199] and protein 4 of phage T7.[210,346] The best-characterized mammalian helicase so far is the T antigen of SV40.[118,166,237,532] Helicases from other viruses have also been identified.[98] All of them require at least one nick on the DNA substrate and cannot unwind circular closed or looped duplexes.

We have seen that strand separation is a rapid process, e.g., 3 s for the 40-kb T7 genome studied by Davison.[105] This is about 8300 bases per second, compared with, e.g., a rate of 730 bases per second unwound and replicated in an *in vitro* replication system at 30°.[334] This raises the question of why helicases are needed at all. The answer may be that the experiments of Davison were carried out by raising the pH first to 11.6, beyond the pK at guanine N1 or at thymine N3 (pK of the free nucleotides are 10.00 and 10.47, respectively, see page 108 of Reference 424); so

that at neutral pH the unwinding rate can be 1000-fold slower. Association with nuclear proteins can be an additional impediment to the separation process, making the catalytic action of helicase essential in the compact environment of the nucleus.

This raises a second question: how can a helicase, as an enzyme present in catalytic amounts, shift the ds-ss equilibrium toward ss, highly disfavored under cellular conditions (37°, 0.15 *M* osmolarity)? Part of the answer is probably that in the cell the separated strands are first stabilized by interaction with SSB and HMG proteins as mentioned, then soon combine with the newly synthesized complementary strands. The electrophoretic assay described can nevertheless yield single strands without SSB or equivalent being added.[334,346,532] Here again helicase may act under steady-state rather than under equilibrium conditions: helicases require ATP for their activity, and the assay system contains sufficient ATP to keep the helicase running along the strands, keeping them apart until physically separated on the gel. In the intact cell, the copying machinery, and then the newly formed complementary strand, may follow the helicase closely enough to prevent parental strand reannealing.

Be that as it may, a helicase without a nicking activity cannot separate the strands of a closed duplex, and cannot alone help in overcoming the topological limitation. All it seems to do is accelerate strand separation, once an initial nick has been made. This may be essential under the crowded cellular conditions where DNA is both supertwisted and tied to a multitude of nuclear proteins.

## E. The Unwound State

In summary, we have seen that the biochemical machinery for transforming the DNA of closed, helically wound regions into an unwound state exists in the cell. High rates of gene expression require that many polymerases pass continuously along the template. Electron micrographs showing transcripts emerging simultaneously along entire genes are well known; long unwound regions would permit the transcription machinery to progress along the template without any topological impediment. The transition of a gene region to an unwound, paranemic, state could therefore be a critical step in the activation of genes producing large amounts of cell-specific products.

The simplest, least-structured unwound form is obviously DNA separated into single strands. Complete separation into single strands nevertheless requires a considerable free-energy input at physiological temperatures.

During the last decade a number of additional DNA structures, which from the topological standpoint are all unwound, have been proposed. A DNA structure can be classified as unwound or paranemic when the complementary strands are not intertwined in a way creating a topological barrier for their separation into single strands, when their ends are clamped to prevent mutual rotation. The alternative unwound DNA structures that are considered at present, and the evidence for each, are described in detail in Section III (see Figure 8). These paranemic forms include partly or wholly double-stranded (duplex), or even triple-stranded (triplex) structures, considerably reducing the energetic requirements.

The transition of closed wound forms of DNA, like B-form DNA, into any of the unwound structures described in Section III requires (1) opening of a covalent bond in the DNA backbone, (2) rotation of one strand around the other, and (3) transition into the ultimate structure. The unwinding of open linear DNA requires steps 2 and 3 only. In a negatively supercoiled circle, a DNA region can undergo a transition to an unwound state to the extent provided by the superhelical density, as described in Section I.B.2, using the available torsional energy.

Before describing the different paranemic structures, the experimental techniques currently used to detect and characterize unwound segments of DNA are described. It will be seen that DNA segments that are all purine on one strand and all pyrimidine on the other have an increased tendency to assume an unwound state. In the last section we summarize evidence on the extent of strand unwinding obtained in studies of DNA replication, transcription, and recombination. It is concluded that the intermediary formation of unwound DNA regions during these processes is not an unlikely proposition, but has yet to be

generally demonstrated. Several reviews concerning unusual DNA structures, including unwound forms, have been published recently[395,426,525-527,557]

## II. EXPERIMENTAL APPROACHES

In this section the experimental techniques currently employed for the detection and characterization of unwound DNA regions are surveyed. In recent years a number of techniques for distinguishing between different conformational states of DNA have been developed. Enzymatic and chemical probing were found to be particularly useful for the detection of unconventional forms of DNA, in circularly closed plasmids and in viral DNA. The use of these techniques is described in the first two subsections.

### A. Sensitivity to Single-Strand Specific Nucleases

#### 1. Single-Strand Specific Nucleases (SSN)

The presence of single-stranded segments in DNA can be detected by the application of single-strand specific nucleases. A number of such endonucleases, which normally do not cut B-form DNA, are available. SSN may of course cleave not only DNA present as single strands, but also, at a reduced rate, any DNA readily separable into single strands, i.e., any nonintertwined, paranemic form of DNA. The most widely used SSN is endonuclease S1 from *Aspergillus orizae*.[497] S1 nuclease has been used extensively for the detection and removal of ssDNA and RNA protruding or looping out from DNA and RNA hybrid structures. S1 and other SSN recognize and cleave DNA also at hairpin loops,[279] B-Z junctions,[456,457] untranscribed DNA templates, and many other situations where unpaired bases are present. Other single-strand nucleases used are BAL-31,[231,271] mungbean nuclease,[253] *Neurospora crassa* nuclease,[144] P1 nuclease,[146] T7 gene 3 nuclease,[455] T4 endonuclease VII, and snake venom (*A. crotalus*) phosphodiesterase I.[129] The literature concerning single-strand nucleases is well summarized in Volume 65 of *Methods in Enzymology*.[327]

S1 is singular among endonucleases in that it is effective only at somewhat acid pH values (4 to 5), while mungbean and P1 nucleases are nearly as active at neutral pH as at pH 4 to 5. BAL-31 and *N. crassa* endonuclease act optimally around pH 7.4 and venom phosphodiesterase around pH 9.[129] There are also differences in cutting specificities: BAL-31 tends to cut at the base of cruciforms, while S1 cuts in the loop; BAL-31 can recognize a single mismatched base.[438] S1,P1 and mungbean nucleases require $Zn^{2+}$ for maximal activity at acid pH (4.5 to 5.0). The necessity to work at pH 4 to 5 is a serious drawback for the use of S1, because under these conditions DNA, especially when in a superhelical state, may assume a structure different from its physiological form. P1 and mungbean are still appreciably active at pH 7 without Zn or even Mg ions present. The absence of $Mg^{2+}$ is important for discerning cruciforms from other sensitive conformations.[441]

#### 2. SSN-Sensitive Sites

In 1982 Larsen and Weintraub[260] reported the existence of S1-sensitive DNA sites at several positions along the chicken globin gene complex. The technique used in that and in many subsequent studies was to digest nuclei with nuclease S1, extract the DNA, and recut it with a restriction enzyme of one or a few cutting sites. The mixture is separated on agarose or acrylamide gels and new bands are identified by hybridization to suitably labeled probes. From the size of the bands appearing only in S1 treated samples, the location of cutting sites and the intensity of cutting at each site can be deduced. An example gel and the deduced cutting sites are shown in Figure 1.

Specific DNA sites sensitive to SSN were first detected when S1 was applied to erythrocyte nuclei from 14-d-old chicken embryos, in which the α- and β-globin genes are active. The same sites were not detected in erythrocyte nuclei from 5-d embryos, when these globin genes are still inactive.[260] No sensitive sites were found in brain

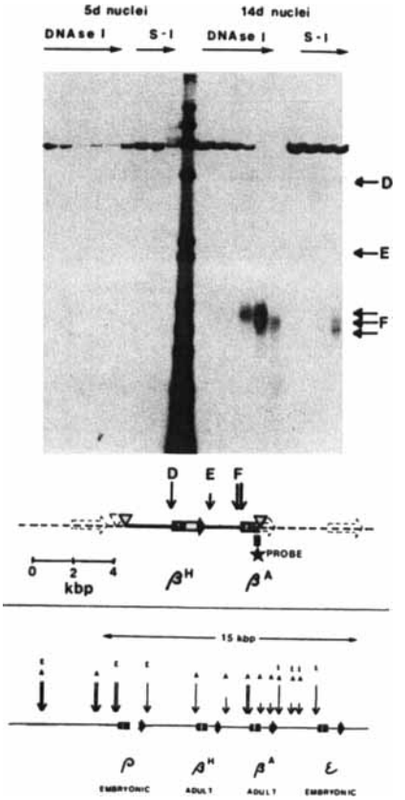


**FIGURE 1.** S1 and DNAse I cleavage site of globin DNA. Nuclei from 5- or 14-d chicken embryo red cells were digested with increasing concentrations of DNAse I and S1 nucleases. Nuclei were suspended in a Zn buffer at pH 4.5. DNA was purified, digested with *Hind* III, and probed with the fragment indicated by a star. D, E, F and arrows mark S1-specific cutting sites on the gel and on the diagram below it. The lower diagram summarizes studies of several embryonic and adult globin genes. A, cuts observed in 12- to 14-d nuclei; E, cuts observed in 5-d nuclei. (From Larsen, A. and Weintraub, H., *Cell*, 29, 609, 1982. With permission.)

nuclei, where no globin is made, from neither 5-d nor 14-d embryos. The gene for globin ρ, an embryonic form, was sensitive to S1 in 5-d nuclei, but not in 14-d nuclei. These observations indicate that the active globin gene contains a DNA region that is either in single-stranded form or in a form readily converted to single strands, i.e., with no topological barrier to being converted into single strands. The differential sensitivity of nuclei from different cells thus indicates a correlation between the appearance of an unwound structure and gene activation. Close proximity, but not identity with DNAse I cutting sites,[524] was noted. A similar sensitivity for S1 was observed on the β-globin gene in rabbit erythrocytes[305] and in human erythroleukemic cells.[70]

In a further step Larsen and Weintraub[260] applied S1 to supercoiled plasmids containing the β-globin region and showed that the sites sensitive to SSN in intact nuclei are also sensitive in a supercoiled plasmid. The location of the S1 cutting sites in plasmids corresponds in a general way to the sites observed in S1-digested nuclei, most but not all of the cuts being 5′ upstream from one of the globin genes (ρ, $\beta_A$, $\beta_H$, ε). This is a strong indication that the structural feature responsible for the sensitivity to SSN is similar in supercoiled plasmids and in intact nuclei. The sensitivity of the supercoiled plasmid DNA in the absence of cellular factors reflects an innate ability of the sensitive DNA sequence to assume an unwound state. The supercoiled DNA can thus serve as a model for the DNA unwound in the active nucleus. The linking deficit, possibly restrained in inactive nuclei, can be released upon activation, to permit the transition of a DNA segment into an unwound state, which can be recognized by the single-strand nucleases. Interestingly, the unwound structure is still detected when a plasmid containing the β-globin region is transfected into a mouse L cell.[522] The chromosomal globin gene remains inactive, however, but this does not dominate the expression of the globin gene on some transfected plasmids, as if local conditions on the plasmid prevail.

Many SSN-sensitive sites have been detected since, in a considerable number of genes cloned in plasmids. A partial list of these sites is given in Table 1. Unfortunately, only for a few of the examined genes has a corresponding sensitivity in intact nuclei or in chromatin been reported. The genes for which sensitivity to SSN has been demonstrated in intact nuclei include the genes for rabbit globin,[305] collagen,[318] the variant-specific glycoprotein of *Trypanosoma brucei*,[173] hsp70 genes of *D. melanogaster*,[180,181] hsp 26,[161] and SV40 minichromosomes.[21] Endogenic nucleases and the low pH required are factors that make the detection of SSN sites in intact nuclei[438] difficult.

The exact location of S1 cutting sites in the chicken β-globin gene (''fine structure'') has been

# TABLE 1
## Genes Susceptible to Single-Strand-Specific Nucleases

| Specs | Gene | Methods | Conditions | Region | Composition, sequence | Comments | Ref. |
|---|---|---|---|---|---|---|---|
| *E. coli* | oriC, 42bp | P1, MB | pH 7.5 | Repl. orig | 15:AAAAAGAAGA 26:AGAGATCTGTTCTATT | AT Rich (6G, 5C/42) | 485, 251 |
| pBR322 | — | S1, transfection | | 1.3;3;3.2 kb | | Sensitive in L cell nuclei | 522 |
| pBR322 | — | MB | Neutral,-Mg | 3206–3278 | $CA_5GCATCTTCACCTAGATCCT_4A_3TTA_5TGAACT_4AAATTTA_5TGAACT_4A_3TCAATCTA_3GTA$ | | 441 |
| pUC9 | lac | S1 | | Operator | CTTTACAC\|ACTTTATG | | 433 |
| pM2 | — | MB | Neutral | | | No fine structure | 251 |
| SC-2μ | ARS | MB | Neutral | | 3580–3278: 69%R; 3683-3722; 77% Y | AT rich | 486 |
| SC-2μ | leu2 | MB | Neutral | Terminator | 3570:12A,G | AT rich | 486 |
| SC | H4-ARS | MB, mutants | pH 7.5 | | | AT rich | 487 |
| *T. bruc.* | VSG genes | BAL31,S1 | Intact nuclei | | | No fine structure | 173 |
| *P. mil.* | Histone | S1 | No cuts in other repeats | Repeat 22 | (a): $(CA)_{10}(CT)_{22}$<br>(b): $AGAGGAAGG(GA)_{16}G_{11}AGGGAGAA$ | | 194 |
| *P. mil.* | Histone | S1, 2D | | Repeat 22 | $AGAGGAAGG(GA)_{16}G_{11}AGGGAGAA$ | | 293 |
| DM | Hsp70 | S1 | Intact nuclei | | | No fine structure | 180 |
| DM | Hsp70(22,26) | S1 | | −156−−100 | CTCTATCGTTTTGTGACTCTCCCTCTCTCTGTACTA\|TTGCTCTCTCACTCTGTC | | 300 |
| DM | Hsp23,26,28; | S1, BAL31, MB | | Intergenic (mainly) | | Many sites | 438 |
| DM | Hsp26 | S1, MB, P1, deltns. | pH 5.5,~6.5 | −99−−85 | −113TGCAAGAGAGAGAAGAGAAGAGAGAGAGATGAGAAAGGAAAAGACAGTG −64<br>********** | 1C,1T in 45R | 449 |
| DM | *Hsp26* | S1, DEPC | Mutations, nuclei | −99−−85 | same sequence | Affects HS inducibility | 161 |
| DM | Histone | S1 | With SSB | 5′ upstrm $(GA)_{10}$ | | | 162 |
| XL | 5S RNA | S1 | | 47–35 | 41 GATCTCAA\|GCGATACAGG 60 | Within transcribed region! | 406 |
| HS | U1 RNA | S1 | Low pH, NaCl favor[aa] | | ~1800: $CTCTT(CT)_{18}T$ −335: $CCTT(CT)_5T$ | >80% of sites cut[aa] | 208 |
| HS | U2 RNA | S1 | | 3′ end | 504–647: >90%$(TC)_n$ 5×(TGT), 1×(TCCGT) | Linears cut, on 5′ side | 209 |
| GG | Globin | S1 | Nuclei, plasmids | | | Beta,alfa,rho | 260 |
| GG | β-globin | BAA, S1 | Nuclei, plasmids | | | | 245 |
| GG | β-globin | S1 | Methylation | Promoter | −198−−178: GGGGGGGGGGGGGGGGCGGG | | 351 |
| GG | β-Globin | S1,MboII; SmaI; DMS(C3) | | Promoter | −64 TGCCCGGG\|GAAGAGGAGGG −46 | Cut when inverted | 443, 69 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| GG | β-Globin | S1,NC | | Intrn/Exn | 620:CGGAGGTGGGAGAGACAGGAGGGAAGGGACCAAACC | −198; 550–l590 NC only[bb] | 516 |
| HS | α-globin | S1 | | | −1145: CCCTTGTCC\|TCCTCTGTCCTTTCCCGACCTCCTCCCTC\|CTCCACCTCCTCCAC\| | α2<α1 | 442 |
| HS | β-glbn,adlt | S1 | | −211: | CCCGGTTC\|TCATATAGAATCTCCCTCCCGA | | 130 |
| | | | | −64: | AGGGAGGGCAGGAGCAGGGCTGGGCATAAAAGTCAGGGC\| AGAGCATCTATTGCTTACATTT\|GCTTCT | | |
| OC | β1-globin | S1, also in intact nuclei | | −400: | $(GA)_{12}$ | Bone marrow cells | 305 |
| GG | β-globin | DMS(N3C) | | | TGGGGGCCCCTCCT\|C\|TTCCCCGGGGGGCACCCCG | | 69 |
| Synt. | Gn.Cn | BAA,ClAA | | | | | 244 |
| Synt. | Gn.Cn | 2d | pH, salt dep. | | | Favored by high salt (rate effect) | 295, 368 |
| MM, GG | α2(I) collagen | S1 | pH4.6' NaCl | MM: −145–165 | TCCCTCCC $N_7$ TCCCTCCCC TGCTCCCCGCAG | [cc] | 318 |
| | | | | GG: −180–−200 | TCCCTCCC $N_4$ TCCTCCCC TCGGCCCCCCCTCC | | |
| GG | Pro-α1(I) collagen | S1 | pH 4.6; hi NaCl | | $(GGGAGA)_{15}$ | | 137 |
| GG | Pro-α2(I) collagen | S1 | pH 4 NaCl | 5′ upstream, | −210: CTTCCTCTTCCCTCCCT\|TCCTCCCTCCC\|$CGC_8TCCG$ −168 | | 136 |
| GG | Histone 5 | S1 | NaCl (favors) | −365, −287, +76; | −278: ATCACATCTCT\|CCCTGTCCC | Consensus: CCTC YTGTCY | 420 |
| GG | Myosin HC(emb). | S1, NC; | | | 220 $TTTTCTTTTTCTTCTTTTCCTTTTTCTCTG\|ATTTTCCTTTTCCTTTTCCT\|TACCT_6CT_5$ −367 | [dd] | 313 |
| HS | EGF receptor | S1, transfection | | Promoter | TTCTCCTCCCTCCTCCTCCTCGCATTCTCCTCCTCCTCTGCT ********** | | 218 |
| RN | Somatostatin | S1, 2d | | −660–−598 | CCTCCTCCTCCCTCCT\|GTGTGTGTGT$(GT)_{20}$\|$(CTGT)_5$CTATCTGTCTGT | [ii] | 196 |
| RN | Preproinsulin | S1 | | | CAAACAGCAAAGTCCAGGGG\|$TCAG_9$CCTGCTTTGGACTATAAA −25 | | 130 |
| HS | Thyroglobulin | S1 | | −512–−304: | 6 × (AAGGAAAGAA), within 209 R | S1 sensitive: ~475; ~360 | 82 |
| MM | IgH, const. | S1,P1,$OsO_4$, BAA,DEPC, DMS, 2d | | α switch | $(AGGAG)_6$; not $(CTGAG)_6$, not GTGTGTGC | P1 cuts, in neutral | 86 |
| ADV | Mj late | S1, 2d, BAA(neg) | Neutral | Promoter | AAGGGGGC\|TATAAAAGGGGGTGGGGGC (AD2) | Also T7 gene 3 protease[ee] | 230 |
| ADV | Mj late | S1, mutations | | Promoter, | −40–−11 | | 550, 165 |
| SV40 | Ori region | S1, MB | Acid, neutral | Enhancer | CTTCTGCCTGCTG\|GGGAGCCTGGGGACTTTCC ×2; TATA box: TAAATAAAAAAAA | | 130, 212 |
| HSV | Type 1 | S1, P1, BAA, ClAA | Neutral | DR2 | (CGCTCCTCCCCC), repeats 2–19 | Not mirror symmetric[ff] | 533 |
| c-Myc | | S1 | | −315–−263 | TCTCC.TC CCCACCTT.CCCCACCCTC.CCCACCCTCC.CCA\|TA\|AGCGC.CCCTCCCG | | 39 |
| | | | | −1443–1420: | GTA.AT\|ACTC\|C\|TCT.CCTCTTCTTT.G | RNA present | |

## TABLE 1 (continued)
## Genes Susceptible to Single-Strand-Specific Nucleases

| Specs | Gene | Methods | Conditions | Region | Composition, sequence | Comments | Ref. |
|---|---|---|---|---|---|---|---|
| Ki-ras | | S1 | | −311−−289 | CCTCCCTCCTTCCCTCCCTCCCGC | Bindg. protein identfd. | 201 |
| GG | — | S1 | | Repetitive | $(AGGAG)_{32}$ | | 124 |
| Crab, Bermuda | Land | S1, P1, MB | pH, σ, NaCl | Dom III: | $AC\|ATCC(CCT)_{15}CTT\|AACCCT.C_8\|C_{11}$ | Symmetries: local only[99] | 139 |
| RN,HS | r-RNA | S1 | | 3′ end of 28S | $(CT)_{23}.TTC.(TCCC)_{10}$ | Site conserved in$_2$ HS,RN | 135 |
| Synt. | — | S1,DMS | | | $(AG)_{45}$ | | 397 |
| Synt. | — | S1, VPD, DMS, DEPC, DNAseI, primer ext. | | | $(AG)_{38}$ | Linears are cut[hh] | 129 |
| Synt. | $d(A,G)_n$ | S1, P1, $OsO_4$, DEPC, BAA, DMS | | 6 components: | $G_{19};(TCC)_8;(AG)_{12};(AAGG)_6;(AAG)_8;(GAAA)_6;A_{20}$ | P1 OK at 4.6. salt eff. | 184 |
| Synt. | $d(A\text{-}G)_{22}$ | S1, $OsO_4$, DEPC + Zn | | | $(AG)_{22}$ | pH independent form | 31 |

*Note:* When not otherwise stated, the sensitive DNA tract was studied inserted into a plasmid and the nucleases and chemical probes were applied to the purified DNA of the supercoiled plasmid.

Abbreviations: DEPC, diethylpyrocarbonate; BAA, bromoacetaldehyde; ClAA, chloracetaldehyde; DMS, dimethyl sulfate; SC, *Saccharomyces cerevisae*; HS, *Homo sapiens*; *Gallus gallus* (chicken); OC, rabbit; ADV, adenovirus; Synt., synthetic; 2d, 2-dimensional topoisomer analysis; NC, *Neurospora crassa* nuclease; VPD, venom phosphodiesterase; MB, mungbean nuclease; |: strongest nuclease cutting sites; #, center of a mirror symmetry of ****** bases. Where no conditions are stated, application of S1 at pH 4.5–5, to the supercoiled plasmids, is implied.

[aa] Linearized plasmids are cut at pH 4.35 to 4.55.
[bb] Rate independent of NC conc, 7 to 10 h, pH dependent. Linears are cut.
[cc] Deletions suggest region is important for transcription.
[dd] pH independent (purines only, at 7.5). poly A site at 161.
[ee] Several point mutants loose S1 sensitivity, yet show enhanced expression.
[ff] P1 cuts at pH 6, 7, 7.5, 8.
[gg] Domain III is at the start of a 2089 × 16,000 repetitive region. Domain IV also sensitive?
[hh] VPD cuts nicely at pH 7 to 9 S1 cuts methylated N7, so no Hoogsteen pairing. Primers not extended on AG strand, so no ss., DEPC attacks, therefore A may be in syn. Favor a heteronomous structure.
[ii] Accent on alternating ("Z") seqs; these are flanked, however, by R.Y tracts.

established by Nickol and Felsenfeld[351] as well as by Wang and Hogan.[516] This was done by following the cleavage by S1 with one or two restriction cuts at nearby sites. The resulting fragments are denatured and analyzed by the procedures used in DNA sequencing, using appropriate labeling or probing techniques. The most prominent site is located between bases −197 and −178, 5′ upstream from the start site for β-globin transcription (Figure 2; note that the *N. crassa* nuclease cleavage sites are mapped at pH 7.5 and that lane 6 represents a linearized plasmid). The DNA sequence at this location is composed of 16 consecutive Gs followed, after a single C, by 3 more Gs. This motif, of a polypurine tract on one strand, complemented by a polypyrimidine on the opposite strand, characterizes the majority of S1-sensitive sites summarized in Table 1. Purine (R)- or pyrimidine (Y)-rich sequences are not the only structural motifs known to be sensitive to S1: inverted repeats are known to be SSN sensitive when in the cruciform state, and so are B-Z junctions. These DNA forms have, however, been studied mainly in synthetic inserts, and have not been reported in the many SSN-sensitive eukaryotic genes studied so far. Exceptions are several prominent cutting sites by mungbean nuclease in SV40[212] in the presence of $Mg^{2+}$, at neutral pH. In prokaryotic and in yeast genes, AT-rich regions have been the dominant sensitive theme.[485] These will be discussed further in Section III.A. The poly R · Y nature of some S1-sensitive regions has been established by deletion of the R · Y tract, which caused the sensitivity to disappear.[449] In other cases,[254,305,433] deletions caused the appearance of secondary cutting sites, revealing a hierarchy in the capacity of different DNA regions to undergo a transition into the SSN-sensitive conformation. The eukaryotic secondary sites again had a high R either on the coding or on the noncoding strand.

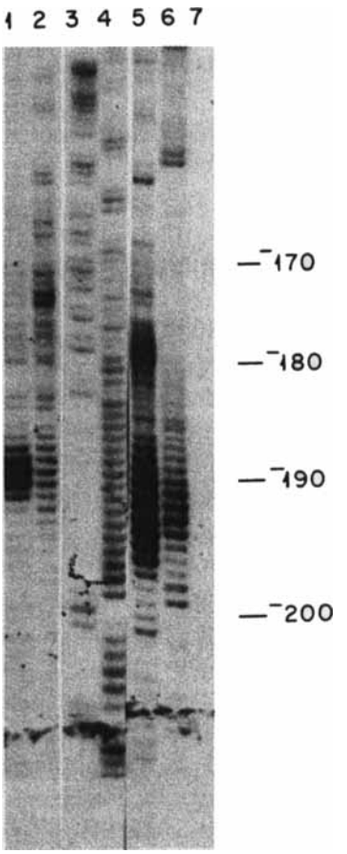


**FIGURE 2.** High-resolution mapping of SSN cleavage sites in the promoter region of the chicken β-globin gene (region F of Figure 1). A plasmid carrying the region −225 to −114 of the chicken $\beta_A$-globin gene was digested by S1 or NC nucleases, labeled, restricted with Pvu II, purified and analyzed on an agarose gel. Lanes 1,2, cut with S1 nuclease, at two conditions; 3,4, sequence ladders, as markers; 5,6, cut with *N. crassa* nuclease, at pH 7.5 (5: supercoiled plasmid; 6: linearized plasmid); 7, no nuclease, as control. (From Wang, J. N. C. and Hogan, M., *J. Biol. Chem.*, 260, 1985. With permission.)

### 3. Alternating d(AG · CT)$_n$

The most thoroughly studied SSN-sensitive polypurine · polypyrimidine sequence (R · Y sequence) is the alternating poly d(AG · CT) sequence, first reported to be S1 sensitive in a plasmid containing part of the histone gene complex (repeat 22) from *Psammechinus miliaris*.[194] Two S1 sites were mapped, one upstream and one downstream from the H1 region. The upstream site contained an alternating d(AG · CT)$_{16}$ tract, as part of a 60 consecutive purine · pyrimidine segment. The S1 cut in this region was "bimodal," i.e., cutting was most intense in the middle and toward the right end of the AG tract. This led to the suggestion that a slippage structure, discussed in Section III.B, is responsible. Ad-

ditional S1-sensitive poly d(AG · CT) tracts were soon discovered in *Drosophila* heat shock genes hsp70 and hsp22,[300] and in human U1[208] and U2 RNA genes (a 144-base region almost entirely CT[209]). It is noteworthy that a site cut with S1 in hsp26[300] is not cut by BAL-31,[438] which indicates some structure at the sensitive sites. The d(AG · CT) tract in these hsp genes was imbedded in a larger polypurine tract. Detailed studies of the sensitivity of artificial oligo $d(AG \cdot CT)_n$ tracts of various lengths, cloned into different vectors,[129,184,331,397] are discussed in Section III.E, in relation to the proposed triplex structure.

A distinct repetitive motif, as in $d(AG \cdot CT)_n$ or $d(G \cdot C)_n$, is not a requirement for SSN sensitivity. In most higher eukaryotic genes the sensitive regions are rich in purines in one or the other strands. An inspection of Table 1 shows, however, that most have no regular arrangement of A and G within the purine region. Some of the listed sequences contain even a few pyrimidines within a polypurine tract, like the sensitive region in the 5S RNA of *X. laevis*[406] or the 6 GGGCGG repeats near the SV40 origin region.[130] At least 10% Y in the poly R strand, and vice versa, can apparently be tolerated. In any case, the structural feature giving rise to SSN sensitivity is clearly not particular to poly dG · dC nor to strictly alternating d(AG · CT) tracts.

### 4. The Effect of Superhelicity

The sensitivity to SSN depends on the superhelicity of the plasmid studied, as first shown by Shishido.[445] The favorable effect of negative superhelicity is demonstrated, for example, by Kilpatrick et al.[230] Superhelicity can influence sensitivity to ss-specific enzymes and reagents in at least two ways: by providing the linking deficit necessary for local strand disentanglement and by providing the free energy needed for the process. The minimal number of superturns required varies from plasmid to plasmid. In several cases even *linearized* plasmids are cleaved by SSN.[194,208,516] Thus, a linearized plasmid with an $(AG \cdot CT)_n$ insert was sensitive to mungbean nuclease with n = 38 or n = 20, but not with an insert of n = 10.[129] The sensitivity $d(AG \cdot CT)_{38}$ insert in the linear form was found to depend on the nature of the flanking non-R · Y sequence: the entire plasmid, linearized by Ssp 1 restriction, was found to be sensitive, as was the 76-base AG · CT insert when flanked by just 3 and 4 bases. The same insert, with flanking regions of intermediate length, 37 and 151 bases, or 37 and 13 bases, was insensitive. Secondary structures influencing the state of the sensitive region were offered as possible explanations.[129] In any case, superhelicity, while an important factor in enhancing the sensitivity to ss-specific nucleases, is not an absolute requirement.

### 5. The Effect of Acidity

A second important parameter determining the sensitivity is the acidity of the digestion medium. This is an important parameter in view of the suggestions that cytosines are protonated in the sensitive structure, permitting the formation of Hoogsteen paired duplex[397] or triplex[294] structures, as discussed in Section III. The main problem with S1 nuclease, the main nuclease studied, and to some extent with P1 and mungbean nucleases, is that they exhibit their maximal activity toward separated single strands at pH 4.5 to 5.[251,497] This makes it difficult to determine whether increased activity at more acid pH is due to increased enzyme activity or to a special state of the DNA substrate. In most studies listed in Table 1, a range of pH values was examined; S1 sensitivity was highest between pH 4 and 5, diminishing considerably toward pH 7 (as far as can be judged from the gels shown).[209,396] This led to the conclusion that protonation of some sites in DNA is required, position N3 of cytosine being the most likely site (but alternative explanations are possible; see Section III.G.5). In some studies the effect of acid could be attributed to DNA protonation, by using cleavage by P1 or mungbean, whose dependence on pH is more limited,[396] or of *N. crassa* T7 nucleases or of venom phosphodiesterase.[129,184] The SSN sensitivity of a region in a repetitive chicken gene toward *N. crassa* nuclease was higher at pH 5.5, although this nuclease has its optimum pH at 8.0.[124] In no case was a systematic kinetic study to distinguish between the effect of acid on enzyme activity vs. effect on the state of substrate

reported. The slope of $d(AG \cdot CT)_n$ cleavage vs. pH was found to be 1.2,[397] (see Figure 15), which indicates that a single proton is involved. The implication of the pH dependence on the structure of the cleaved DNA is further discussed in Section III.G.5.

### 6. pH-Independent SSN-Sensitive Sites

There are nevertheless quite a number of cases where sensitivity to SSN at *neutral* pH is reported.[86,230,313,441,516] These include all bacterial and yeast sequences tested so far, exhibiting mungbean nuclease sensitivity in a pH-independent manner,[485] the exact cleavage pattern depending on the presence or absence of magnesium. The pH-independent behavior of bacterial and yeast sequences may be related to the different composition of the sensitive sequences, these sequences being A,T rich rather than R · Y "asymmetric". Different unwound structures may be responsible. The $d(AG \cdot CT)_{38}$-containing plasmid is sensitive to venom phosphodiesterase even at pH 9.0.[129] Hanvey et al.[184] report that a $d(AG \cdot CT)_{24}$ oligonucleotide and a $d(G \cdot C)_{19}$ oligonucleotide are cut equally well by P1 at pH 4.6 and 7.5. Sites in the *Drosophila* Hsp26 region were initially reported to be as sensitive to mungbean nuclease at pH 7.4 (but not to BAL-31) as to S1 at pH 4.5, indicating a structure insensitive to pH.[438] A later study[449] established, however, that the major site 5′ to hsp26 is much more sensitive at pH 5.5 than at pH 6.5, to S1, P1, and mungbean nucleases. Two different "unusual" DNA structures are proposed for the differently sensitive regions of the myosin heavy-chain gene.[313]

### 7. Ionic Effects

For some SSN-sensitive regions the extent of cleavage was found to diminish with increased NaCl concentrations.[129,208] The cleavage patterns were not affected. In other cases, high salt favored the sensitive conformations. These include sensitive regions in collagen[318] and oligo$(G \cdot C)_{19}$ (but not oligo $(AG \cdot CT)_{12}$), when cleaved by nuclease P1 at pH = 7.5 and by nuclease S1 at pH = 4.6.[184] Salt is well known to raise the melting temperature of DNA, i.e., to disfavor transition into single strands. The increased sensitivity of oligo $d(AG \cdot CT)_n$ inserts in low salt thus points toward a single-stranded state of the sensitive insert.

As for divalent ions, S1 nuclease requires $Zn^{2+}$ for optimal activity, while the neutral nucleases do not. An interesting effect of magnesium is encountered with bacterial plasmids; plasmid pBR322 is cut by mungbean nuclease at 37°, with no $Mg^{2+}$ present, mainly in a region extending from base 3200 to base 3290.[441] When $Mg^{2+}$ is present, at pH 7.0, the dominating cuts appear at the major and minor cruciforms of the plasmid and secondary S1 sensitive regions. Mungbean cleavage sites at potential cruciform-forming regions were also reported for SV40 DNA.[212]

Two questions arise: (1) The sensitive region studied by Sheflin and Kowalski[441] is located in a region present in most ColE1 derivatives, including the pUC series vectors and other vectors used. Why does this sensitive region not appear when these plasmids are used as vectors in cutting studies with inserts? The use of mungbean cannot be the explanation, as P1 was also tested. It must be remembered, however, that in a supercoiled plasmid a single cut by a SSN is sufficient to relax the circle, so only the most vulnerable sites are likely to be cleaved. Sites with polypurine inserts must be energetically favored over the vector sites. (2) The main sensitive region of pBR322, between 3200 and 3290, has been characterized by Umek and co-workers[485,486] as "A,T rich". A simple interpretation would be that in *Escherichia coli*, and possibly in yeast, the vulnerable structure is different from higher eukaryotes. A detailed inspection of the sensitive section of pBR322[441] nevertheless shows a distinct clustering into R and Y segments. Thus, of the 27 Y in the sensitive region, 8 are found in a TCTTCACCT tract, clearly Y rich (R in the complementary strand) and 7+3 in the TCCTTTTAAATTAAAAATGAA tract, where R · Y clustering is also evident; only 4 are found in a TATATAT tract and 3 in an ATTAT tract, both clearly of the alternating A,T type. A combination of R · Y asymmetry with an AT-rich composition may be responsible for the special

properties of the prokaryotic sequences. The sensitive region near the ARS of the 2-μm plasmid[486] can also be viewed as R · Y rich no less appropriately than A,T rich: between 3660 and 3700 there is first a tract with 18/23 R, then a tract of 20/27 Y. R · Y clustering is less evident in a mungbean nuclease-sensitive region of plasmid pM2.[440] There are additional reports that A,T regions can be S1 sensitive when in supercoiled plasmids.[314,471] Purine · pyrimidine is therefore not the only compositional motif of SSN-sensitive regions.

### 8. Artifacts

Two types of artifacts have been mentioned as possible causes for S1 sensitivity: plasmids can be denatured during the dehydration that accompanies alcohol precipitation,[474,520] causing the formation of single strands previously not present. The presence of contaminating RNA in the plasmids examined can be a factor determining SSN sensitivity.[39] We shall return to this effect of RNA in Section III.G. Contamination of SSN by exo- and other nucleases has also been considered[129] and can be controlled, e.g., by an internal ϕx174 probe.[124]

### 9. Genetic Location and Base Composition of Senstive Sites

Table 1 summarizes the various genes and gene subregions where SSN-sensitive sites have been located. The cleavage sites are in many cases located in 5′ promoter regions of genes.[39,351] In other cases they may be within the gene,[406] at the 3′ end,[420] or in introns.[516] Thus, it is not possible at this stage to assign a specific promoter function for the SSN-sensitive sites. An enhancer-like function, in the sense that the effect is not dependent on the exact site or orientation of the sensitive sequence is more likely (Section IV.B.5). The base sequence of the sensitive site may not be a sufficient signal for cleavage. For instance, in the chicken histone H5 gene a certain sequence, CTCCTTGTCT, is cut in 3 locations, but not in a fourth one within the same gene complex.[420]

In a recent analysis, we have shown that R · Y tracts are highly overrepresented in all higher eukaryotic genomes.[61] The long R · Y tracts are avoided in coding regions when compared with introns, but are abundant in promoter regions. This is documented for a cellular organelle genome in Table 2. The lack of local mirror or dyad symmetry in the long homotracts listed is evident (except for blocks of T, etc.). Only a minority of these long homotracts have been documented to be SSN sensitive, but a systematic study has not been carried out; therefore, it is premature to suggest that a correlation between all purine or pyrimidine composition and tendency to unwind exists.

Are the SSN-sensitive sites essential for gene function? To answer this question, a S1-sensitive site was either deleted or mutated. The expression of a CAT reporter gene in transfected animal cells decreased severalfold when the 5′ region of the collagen gene containing the S1 sensitive site was omitted.[318] A major pH-independent SSN region is found in the major late promoter of adenovirus. Several mutations in that region abolish S1 sensitivity, yet increase the transcription of the late major region in transiently infected Hela cells.[550] Transcription of the adenovirus genome in the infected cell may, however, take place on a template topologically restrained by cellular factors not necessarily present when a transcript of the artificial plasmid is tested *in vitro*.

## B. Chemical Reagents

Chemical reagents have long been used as tools for probing the structure of nucleic acids (for older reviews see References 58 and 454). This is because duplex B-type DNA reacts very slowly to a series of reagents that react well with ss DNA and with free bases.[58] The use of specific reagents received a booster when high-resolution gel analysis made it possible to obtain separate information on the reactivity of each nucleotide in a DNA molecule.[356] The development of structure-specific reagents makes it possible to determine the local conformation of a DNA segment in a plasmid or a synthetic oligonucleotide. In this section we shall survey the application of conformation-specific reagents to the study of the

**TABLE 2**
**Frequency of Homopurine and Homopyridine Tracts in the Tobacco Chlorplast Genome**

| | Sequences | | Bases | | | |
|---|---|---|---|---|---|---|
| Tract length l | PYRS Y | PURS R | Found F = l (R + Y) | Expctd. E* | Diff. F−E | Ratio F/E |
| 1 | 15122 | 15274 | 30396 | 38953 | −8557 | 0.78 |
| 2 | 8011 | 8159 | 32340 | 38945 | −6605 | 0.83 |
| 3 | 4582 | 4664 | 27738 | 29208 | −1470 | 0.95 |
| 4 | 2481 | 2484 | 19860 | 19476 | 383 | 1.02 |
| 5 | 1391 | 1282 | 13365 | 12177 | 1187 | 1.10 |
| 6 | 793 | 720 | 9078 | 7311 | 1767 | 1.24 |
| 7 | 489 | 424 | 6391 | 4268 | 2122 | 1.50 |
| 8 | 314 | 275 | 4712 | 2441 | 2270 | 1.93 |
| 9 | 198 | 159 | 3213 | 1374 | 1838 | 2.34 |
| 10 | 121 | 116 | 2370 | 764 | 1605 | 3.10 |
| 11 | 97 | 64 | 1771 | 421 | 1349 | 4.20 |
| 12 | 44 | 37 | 972 | 230 | 741 | 4.22 |
| 13 | 33 | 32 | 845 | 124 | 720 | 6.76 |
| 14 | 28 | 27 | 770 | 67 | 702 | 11.4 |
| 15 | 15 | 17 | 480 | 36 | 443 | 13.2 |
| 16 | 15 | 13 | 448 | 19 | 428 | 23.1 |
| 17 | 10 | 7 | 289 | 10 | 278 | 28.0 |
| 18 | 5 | 0 | 90 | 5.4 | 84 | 16.4 |
| 19 | 5 | 4 | 171 | 2.9 | 168 | 58.9 |
| 20 | 2 | 5 | 140 | 1.5 | 138 | 91.4 |
| 21 | 3 | 3 | 126 | 0.8 | 125 | 156 |
| 22 | 1 | 1 | 44 | 0.4 | 44 | 103 |
| 23 | 2 | 1 | 69 | 0.22 | 69 | 309 |
| 24 | 2 | 0 | 48 | 0.12 | 48 | 411 |
| 27 | 2 | 0 | 54 | 0.02 | 54 | 3248 |
| 29 | 1 | 0 | 29 | 0.00 | 29 | 6433 |
| 35 | 1 | 0 | 35 | 0.00 | 35 | 38189 |
| | | Sum: | 155844 | | | |

* On statistical basis; see Reference 61.

**Tracts of 22 Bases and Longer**

| Region | Length | Position | Sequence |
|---|---|---|---|
| Intervening | 24 | 5480 | TCTTTTTTTTTCTTCTTTCTTCCT |
| *Promoter* | 27 | 8773 | TCTTTCTTTCTCTCTTTTCTTCTTTCT |
| ORF | 24 | 31450 | TCTTTCCTTTCCCCTCCTTTCTTC |
| Termination | 23 | 36985 | TTCCTTTTTTCTTTTTTTTTTTC |
| Termination | 29 | 45119 | TTTTTCTTTTTTTTTTTTTTTTTTTCTTTCT |
| Intergenic | 23 | 50493 | TCTCCTCTTTTTTTTCTCCTCTC |
| *Intergenic* | 35 | 69750 | TCCCTCTCCCCCTCTTTTCTCTTTTTTCCCTTTTT |
| ORF | 22 | 95047 | GAAGAAGGGGAGGGAGAAGGAG |
| ORF | 27 | 129073 | TCCTTTTCCTTTTTCTCTTCTCTTTTT |
| ORF | 23 | 129189 | AAAAGAAAAAAGAAAGGGGGGGG |
| ORF | 22 | 147461 | CTCCTTCTCCCTCCCCTTCTTC |

ORF: Open reading frame. "Promoter" and "Terminator" regions are defined as the 100 bases preceding or following the coding region. The strand listed in the EMBL data bank was analyzed. For further details see Reference 61.

paranemic DNA structures. An example gel of reaction products is shown in Figure 3.

The application of most of the reagents currently employed presents certain problems of interpretation. Diethylpyrocarbonate (DEPC), for instance, has been proposed as a reagent that differentiates between B and Z DNA.[195,221] DEPC

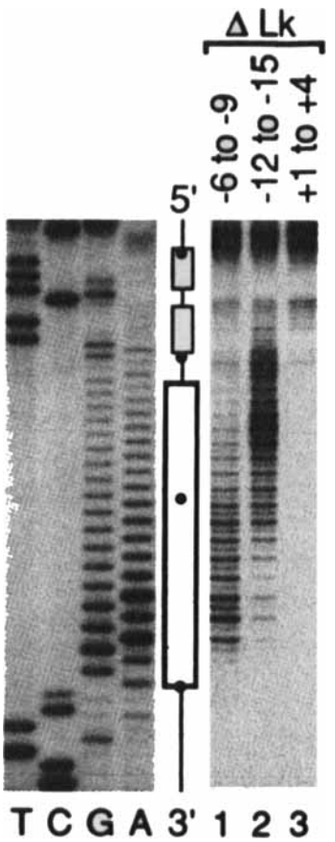


**FIGURE 3.** Modification of a d(AG · CT)$_{18}$ insert with DEPC. The position of the AG · CT tract is indicated by the box with a dot in the center of tract. The four lanes on the left are sequencing ladders of the same insert. At low linking deficit (lane 1) the 3′ half of the oligopurine strand shown is modified, while at high linking deficit (lane 2) the modification is shifted to the 5′ half and beyond, consistent with a single strand related state of that half. A plasmid with a linking excess is not modified (lane 3). Supercoiled plasmids containing the insert at different average ΔLk values were incubated with DEPC (1:40) in buffer of unspecified pH. The purified products were transcribed with SP6 RNA polymerase and NTPs, including labeled UTP. Transcription stops one base ahead of a DEPC-modified base. The products are purified and analyzed on an 8% acrylamide gel. (From Htun, H. and Dahlberg, J. E., *Science*, 243, 1571, 1989. With permission.)

is, however, a classic reagent for ss DNA, reacting at N7 of adenine and guanine. DEPC reacts with Z-form DNA by virtue of its more exposed N7, related to the syn conformation of guanines in Z DNA. Therefore, one should regard DEPC mainly as a reagent *not* recognizing DNA of the B (and A?) family, or, more exactly, not reacting with DNA forms where guanine residues adopt the anti rather than the syn conformation. A second problem is that DNA need not be predominantly in the reactive conformation in order to react with the chemical (or enzymatic) probes. It is sufficient that a small fraction is in the reactive conformation, as long as this fraction is in mobile equilibrium with the bulk of the DNA. In a closed non-supercoiled sequence of DNA a mobile equilibrium with unpaired strands is impossible as long as the two strands are helically interwound, but is possible when the DNA duplex is in the unwound form. A reasonably fast reaction with a single-strand-specific reagent therefore indicates a large degree of unwinding of the reactive DNA substrate. A detailed kinetic study, with proper positive controls, is necessary in order to determine what fraction of the DNA is actually in the reactive state under particular conditions. A kinetic analysis of this type has been carried out for a myc-carrying plasmid by Wang and Hogan[516] and Boles and Hogan;[39] see Section III.A. In the following, the main experimental findings with the various reagents are described. The discussion of the structural implications follows in Section III.

### 1. Bromoacetaldehyde (BAA) and Related Aldehydes

Aldehydes are well-known reagents for ss DNA, attacking positions N3 of cytosine and N1 of adenine, positions that are shielded by the hydrogen bonds in dsDNA. The simplest aldehyde, formaldehyde, is a classic reagent for keeping strands apart. A drawback of formaldehyde is that it reacts very slowly, and molar quantities are needed.[131,317] Most aldehydes, including formaldehyde and glutaraldehyde, tend also to polymerize and form hard-to-characterize products. Bifunctional derivatives like kethoxal ($CH_3$·$CH(OC_2H_5)$·CO·CHO)[352] and glyoxal

(OCH·CHO)[321,498] circumvent some of these problems.

More recently, bromoacetaldehyde (BAA; $BrCH_2O$), and its chloro analog (chloracetaldehyde; $ClCH_2O$) have been found to give reasonable reaction rates with ssDNA as well as a mostly unique product.[257] An etheno ring (-N6·CHOH·CH2·N1-) is formed between N6 and N1 of adenine or between N3 and N4 of cytosine.[245,246,319] BAA reacts well with loops of cruciforms, but was initially reported not to react with regions bridging B- and Z-form segments of DNA (B-Z junctions).[225] Subsequent studies[319] showed some reactivity in the junctions, supporting the notion that these junctions have a partly unwound character.

BAA was applied to a supercoiled plasmid containing the S1-sensitive oligo dG · dC sequences in the chicken β-globin gene, and was found to react efficiently at both pH 4.6 and at neutral pH.[245] Two high-resolution studies of a plasmid containing a 31-bp oligo dG · dC insert were performed. Initially, S1 was applied subsequent to BAA and chloroacetaldehyde, to cleave the DNA at the BAA reactive site,[246] revealing one intense reactive region in the middle of the oligo(dG · dC) tract and a second reactive region in a flanking segment. Later, when hydrazine-piperidine treatment was used for secondary cleavage, the main chloracetaldehyde reactive sites could be located along the 5′ half of the dC strand of the $(dG \cdot dC)_n$ tract (Figure 18), in full accordance with the triplex model to be discussed in Section III.[244] These results were obtained with chloroacetaldehyde as the reagent, and with at least 2 m*M* $Mg^{2+}$ added.

In the presence of Mg the 5′ half of the dG strand was the reactive one. Moreover, the reaction was almost pH independent between pH 5 and 8. The reactive form of poly dG · dC in the presence of magnesium ions may thus be of a structure different from that of other oligopurine tracts. Kohwi and Kohwi-Shigematsu[244] propose that a $Mg^{2+}$- and pH-dependent equilibrium between two different triplexes exists; one triplex is composed of (C · G · CH+) triplets and the other of (G · C · G) triplets, only the first triplex being protonated and therefore acid dependent. This is discussed further in Section III.G. The studies by the Kohwis demonstrate that bromo- and chloroacetaldehyde, being small molecules, have a high capacity to detect short, unpaired segments. They are therefore useful probes for the identification of ss DNA and other unwound segments.

### 2. Diethyl Pyrocarbonate

($C_2H^5$·CO·O·CO·$C_2H_5$, DEPC). This reagent reacts primarily with N7 of adenine and of guanine, leading to imidazole ring opening and formation of a carbethoxy derivative.[272,495] Position N6 of adenine can also be reactive. Herr,[195] as well as Johnston and Rich,[221] showed that diethyl pyrocarbonate (DEPC) attacks preferentially guanines in Z-form DNA over those in B-form DNA. The explanation is that guanines are reactive only in the more exposed syn conformation, adopted in Z DNA and not in the anti conformation, where neighboring groups interfere with the reaction. DEPC does also react with nonpaired adenines, as in the loop of some cruciforms.[145]

The reaction of DEPC with the S1-sensitive oligo (AG · CT) inserts was studied by Johnston[220] as well as by Htun and Dahlberg[205] and by Hanvey et al.[184] The three authors agree that DEPC attacks primarily the adenines in the 5′ half of the $d(AG)_n$ strand. This is consistent with a structure where the adenines along the 3′ half are in a shielded state, i.e., part of a triplex. Htun and Dahlberg[206] report that at somewhat higher linking deficit ($\Delta Lk = -12$ to $-15$; $\sigma = -.04$ to $-.05$), the reactivity switches to the 3′ half of the $d(AG)_n$ strand, indicating a second possible isomer of the triplex, in which the 5′ half of the purine strand participates in the triplex (H-y5, see Figures 3 and 14). Evans and Efstratiadis[129] also report that along the 76-bp AG insert, the 3′ half is reactive toward DEPC, at pH 4.5. (In earlier studies reactivity toward DEPC was in the 5′ half even at $\sigma = >0.1$;[129] see Reference 205.) The flanking AT-rich tract is even more reactive. The reaction of $(AG \cdot CT)_n$ with DEPC is most efficient at pH 6, at moderate superhelicity, and disappears toward pH 8.[205] Hanvey et al.[184] compared the reactivity of oligo d(AG · CT) with several other simple A,G-containing inserts, and found that the 5′ flank is the reactive one in all

these inserts, as well as in more complex polypurine inserts (see table). Oligo (dG · dC), having no As, is not reactive; Oligo (dA · dT) is only slightly reactive above the B-form background. In summary, DEPC detects, in addition to unpaired adenines, purines that are exposed in certain duplex DNA conformations.

### 3. Carbodiimide

(*N*-cyclohexyl *N'* beta. (methyl morpholinium)ethyl carbodiimide:

$$[O(CH_2CH_2)^2{\cdot}N(CH_3)^+]{\cdot}CH^2{\cdot}CH^2{\cdot}N{=}C{=}N{\cdot}C_6H_{11}$$

This electrophilic reagent has been shown to react primarily with the imino N3 of guanine and to some extent with N1 of thymine. It has been found to react with superhelical phage PM2 DNA, but not with the relaxed form.[262] Buoyant density studies of the products suggest that cruciforms are extruded from the superhelical form, with some transient ss formation. The bulky size requires that the bases be well separated. Carbodiimide can therefore serve as a valuable complement to BAA. The action of this reagent seems not to have been studied yet by high-resolution techniques.

### 4. Dimethyl Sulfate

($CH_3O{\cdot}SO^2{\cdot}OCH_3$, DMS). This alkylating agent is known to react well with B-form DNA, primarily at position N7 of guanine and to a certain extent with N3 of adenine. The product is a positively charged *N*-methyl derivative, readily cleaved in moderate basic medium.[312] Further treatment leads to cleavage of the deoxyribose phosphate backbone, i.e., to DNA cleavage. The reactivity of the alkylating agents toward position N7 of guanine is well known to increase significantly in the order: free nucleotide ≤ ss DNA ≤ dsDNA ≤ nucleosomal DNA.[389,549] The enhanced reactivity of nucleosomal DNA may be due to bending into the minor groove.[549] In ss DNA, position N3 of cytosine, which is hydrogen bonded in ds DNA, is also reactive, and its reactivity can be mapped by hydrazine cleavage.[233] Cantor and Efstratiadis[69] studied the action of DMS toward cytosine N3 of the S1-sensitive region of the β-globin gene at base −55, inserted into a supercoiled plasmid. No cytosine bands were seen, suggesting that the sensitive insert is base paired in the unwound state.

In a subsequent study, DMS was found to react with the SSN-sensitive $d(AG)_{38}$ insert to a similar extent in a supercoiled and in a linearized plasmid.[129] The reaction proceeded to a similar extent as in non-SSN regions, at pH 4.5 and 8, with or without $Mg^{2+}$. This speaks against the N7 position being shielded in the insert, as would be expected if Hoogsteen base pairing to a third strand took place.[129] A clear protection of guanine N7 was nevertheless found in a $d(AG)_{45}$ insert.[397] This is supported by Hanvey et al.[184] for $d(AG)_{12}$ and the related inserts described. Partly (but not fully) protected areas (at pH 5) are observed in the 3' half of several of the inserts. There was no protection of Gs in the inserts at pH 7.6 or in linear DNA. An oligo dG insert was not protected at pH 5, and an oligo dA insert was not reactive. In summary, DMS is a good reagent to detect bent B-form DNA sections, DMS also serves to establish the extent to which the major groove is protected from outside interactions. It should be noted that the guanine protection data indicate the state of *bulk* DNA, rather than that of a reactive minority.

### 5. Osmium Tetroxide

$OsO_4$ forms an addition compound with the 5-6 bond of thymine and, to a minor extent, of cytosine.[65,163,281] The 5-6 bond resides in the major groove of a B-form DNA, which is unreactive toward $OsO_4$. $OsO_4$ efficiently reacts in low salt and at pH 8 with ss DNA.[315]

McClellan et al.[315] used $OsO_4$ to characterize different states of alternating oligo dA-dT. This sequence, when inserted into a supercoiled plasmid, can adopt a cruciform structure. $OsO_4$ attacks only thymines in the nonpaired loop region[315] of the cruciform, but not along the stem. The linear form of the same plasmid is reactive toward $OsO_4$ *throughout* the oligo dA-dT insert, as if no cruciform stem is formed. This $OsO_4$ reaction is not observed with the linearized form

of the dyad symmetric $T_{15}A_{15} \cdot T_{15}A_{15}$ insert. This dyad symmetric insert apparently assumes, even in the linear state, the cruciform structure. The alternating (dA-dT) insert, when in the linear state, is suggested to assume an alternating structure of the type proposed by Klug et al.[238] In that structure, deoxyribose residues are in the 3′ endo conformation, exposing the T in the major groove toward the reagent. In a subsequent study, McClellan and Lilley[314] further explore the condition under which a $d(A\text{-}T)_{34}$ insert is modified uniformly throughout (''U'' or uniform pattern), i.e., is present in a non-B, noncruciform, unwound state. Low salt and high T conditions favored the transition into this unwound state in a closed supercoiled plasmid. Some ions, like $Al^{3+}$, also enhanced the U pattern of $OsO_4$ reactivity, while spermine and $Co_3(NH_3)^{6+}$ favored the ''C'' (central) or cruciform pattern. Two-dimensional electrophoresis indicates complete unwinding of the oligo dA-dT insert (cruciforms are topologically unwound structures; Section III.3). The facile equilibrium between U and C patterns strengthens the evidence that both U and C forms represent unwound states of DNA, while the structure proposed by Klug et al.[238] for dA-dT is not unwound.

Johnston and Rich[221] find that $OsO_4$, and hydroxylamine (Section II.B.7), attack very efficiently the thymines of a TCCCT or TCCTCT tract flanking a Z-forming alternating dG-dC insert. $OsO_4$ can thus serve as a reagent recognizing B-Z junctions.[147] In purine · pyrimidine regions, $OsO_4$ reacts mainly at the center of the insert.[205,220,498]* Thymines flanking the inserts are also reactive, on both strands. The reaction with an oligo d(AG · CT) is acid-dependent (see Figure 2 in Reference 498). This is in accord with the triplex model, in which the major groove side of the bases is highly shielded along both halves of the CT strand, but not at the center loop. $OsO_4$ reacts with thymines in the various oligo R · Y inserts studied by Hanvey et al.[184] in the same two subregions: with thymines at the center of the CT strand and with thymines flanking the insert (including the flanks of an oligo dG insert). In summary, $OsO_4$ can be regarded as a reagent that indicates the degree of exposure of thymines in a DNA structure.

### 6. Potassium Permanganate ($KMnO_4$)

This reagent, like $OsO_4$, oxidizes the 5-6 bonds of pyrimidines when in an exposed state.[418] $KMnO_4$ operates at neutral pH under mild conditions. $KMnO_4$ was used to localize a bend in a DNA loop formed by the simultaneous association of the lac repressor with two remote operator sites.[44] $KMnO_4$ has also been applied to show that T antigen unwinds supercoiled SV40 DNA when it binds to its major binding sites on the viral DNA.[43] $KMnO_4$ appears a promising reagent for detecting unwound DNA regions and has been applied recently to potentially triplex-forming dyadic sequences; guanines, in addition to thymines, were attacked.[164] $KMnO_4$ can recognize single unpaired C and T residues.[167]

### 7. Hydroxylamines

Hydroxylamine, $NH_2OH$ and its *O*-methyl ester, methoxylamine, $NH_2OCH_3$, react with carbons 4 and 6 cytosine to form an addition product.[58,62] Hydroxylamine becomes protonated at low pH, pK = 6.5, so that the reaction is expected to be slowed below pH 6. The reagent reacts well with ss DNA but not with linear B-form DNA.[221,419] $NH_2OH$ reacts well with cytosines in a B-Z junction (formed by either a $d(C\text{-}G)_{32}$ or a $d(T\text{-}G \cdot A\text{-}C)_{31}$ insert). Not every C in a junction is reactive, suggesting that junctions are not entirely single stranded, but rather in a mobile state.[221]

$d(AG \cdot CT)_n$ inserts in supercoiled plasmids are reactive mainly in the center of the CT region, although some activity through the CT region is evident. The reaction is strongest at pH 4.5, despite the reduced activity of the reagent.[220] Cytosines at the junction of the insert with the rest of the plasmid, particularly at the 3′ end of the CT strand, are also reactive. This indicates that the 3′ end is opposite the ss half of the d(AG) strand in the triplex, and has difficulty in rejoining the plasmid duplex. Htun and Dahlberg[205] obtain a similar result with methoxylamine (see Figure 3) in favor of the hinged version of the triplex proposed by them. In summary, hydroxylamines serve primarily to detect exposed cy-

* It should be noted that oligo R · Y inserts cannot form cruciforms.

tosines, complementing the information obtained by $OsO_4$ and $KMnO_4$ on thymidines.

### *8. Hydroxyl Radicals*

A number of transition metal complexes can generate OH radicals, which attack DNA. The primary target of the radicals is the deoxyribose ring, abstracting a H atom from deoxyribose carbon C4 and leading to ring opening.[387] The abstracted hydrogen is exposed toward the minor groove. Single-stranded DNA is unreactive toward $Cu^{2+}$ phenanthroline, a commonly used OH-generating complex; so is Z DNA. A-form DNA is much less reactive than B-form DNA, probably because of a narrower minor groove. Oligo dA-dT is particularly reactive to this reagent.[471] OH radicals formed by a second OH-generating system, iron-EDTA-$H_2O_2$, have been shown to avoid positions where the minor groove is narrowed because of DNA bending.[64] Moser and Dervan[341] attached the $Fe(EDTA)_2^{2-}$ complex to specific sites on an oligo $dT_{15}$ carrier and showed that an oligo d(A · T)-containing plasmid is cleaved in a manner consistent with the formation of a $d(T \cdot A \cdot T)_{15}$ triplex between the ss oligo $dT_{15}$ and the oligo d(A · T) duplex. The cleaved site is opposite the base to which the EDTA is attached, if the $T_{15}$ oligomer joins the duplex in a direction antiparallel, rather than parallel, to the $dT_n$ insert. This establishes the directionality of the triplex formed. A further example of the capacity of hydroxyl radical-generating reagents to differentiate between DNA conformations are chiral trisphenanthroline complexes of rhodium, which can attack specifically at the base of a cruciform[235] and to distinguish between A- and B-form DNA regions.[322]

All these studies were performed with isolated plasmid DNA. At least one study where reagents were applied to intact nuclei was reported.[173] No evidence for unwound regions was found. In summary, while most previous reagents attack the nucleotide bases, in the major groove, OH-generating reagents attack the deoxyribose moiety, and can serve to report the state of the minor groove side of a duplex.

## C. Two-Dimensional Topoisomer Electrophoresis

A direct method for establishing the extent of unwinding of a DNA segment in a closed circle is by electrophoretic analysis of the topoisomers of that circle. A change in the degree of primary winding of an internal segment must, buy the principle of linkage conservation, be compensated by an equivalent change in superhelical winding. In a negatively supercoiled plasmid (the common cellular state, see Section I.A.2) unwinding of any right-handed DNA segment must be accompanied by a reduction of negative superturns, one superturn opened per ~10 bases unwound. The electrophoretic mobility of circular DNA in agarose is determined primarily by the number of *superhelical turns* actually present. Unwinding of primary turns in a topoisomer, on account of superhelical turns, is therefore manifested as a reduced migration rate of that topoisomer. Unwinding in a closed circle requires a certain superhelical torsion energy, i.e., only topoisomers above a minimal linking deficit are unwound under a particular set of conditions. Upon electrophoresis on agarose, one observes a sudden discontinuity in the well-known ladder produced by a topoisomer mixture, and an irregular pattern results.[377] When a topoisomer, e.g., topoisomer −7 migrates, under a set of conditions favoring unwinding, only as far as topoisomer −4 under a previous set of conditions, we know that 3 superturns have been opened. This can be interpreted in a number of ways, for example, that a ~30-bp segment became completely unwound, or alternatively, that ~15 bp reverted to a *left*-handed helix, or else that ~60 bp have been just partly unwound.

To obtain better resolution, the pattern obtained upon electrophoresis in a single direction can be rerun in a perpendicular direction (second dimension), in the presence of chloroquine. Intercalation of chloroquine into partly unwound topoisomers causes them to snap back into their right-handed B form (slightly unwound by the intercalation of chloroquine), and the missing negative superturns are reintroduced into the topoisomers that ran retarded in the first direction.

This leads to a well-resolved discontinuity, or "jump", in the final pattern (see Figure 4). The two-dimensional chloroquine technique was first worked out for plasmids containing a $d(GC)_{16}$ insert, which converts at high superhelical density to left-handed Z DNA.[264,515] In the example shown in Figure 4, the band immediately following the discontinuity migrated in the first direction only as far as the fifth topoisomer below it (marked x in the figure). This means that under the conditions of the first run it had 5 superturns less than during the second run (now in B form), i.e., it was unwound by 5 turns. The $d(G\text{-}C)_n$ insert was 32 bases long. If we assume that it underwent a transition from a right-handed helix of 10.5 bases per turn to a left-handed helix of 12 bases per turn, we obtain an expected unwinding of $32*(1/10.5 + 1/12) = 5.7$ turns. This is slightly more than the observed change of 5 turns. Detailed quantitative treatments that integrate the information from all bands have, however, been applied and a helical repeat close to that found for crystalized Z oligonucleotides was obtained.[7,328,500] Comparison with the vector, i.e., with the same plasmid without the inset, can serve to delineate the location of the unwound region. Independent information on the length of the unwound region is needed to state the degree of unwinding of the structure formed, or vice versa.

Two-dimensional topoisomer analysis has been used to determine exactly the degree of unwinding of Z DNA — 11.5 bases per left-handed turn in solution.[183,378] A discontinuity was also observed in plasmid pUC7, in which the polylinker region is present twice, in a head-to-tail (dyad symmetric) arrangement. The transition was attributed to cruciform formation.[154] To form a cruciform, the strands must separate, and this can be achieved only at the expense of the negative superturns. The two-dimensional pattern of pUC7 showed a clear jump when topoisomer −9 was reached. Most of topoisomer −9 reverted to a

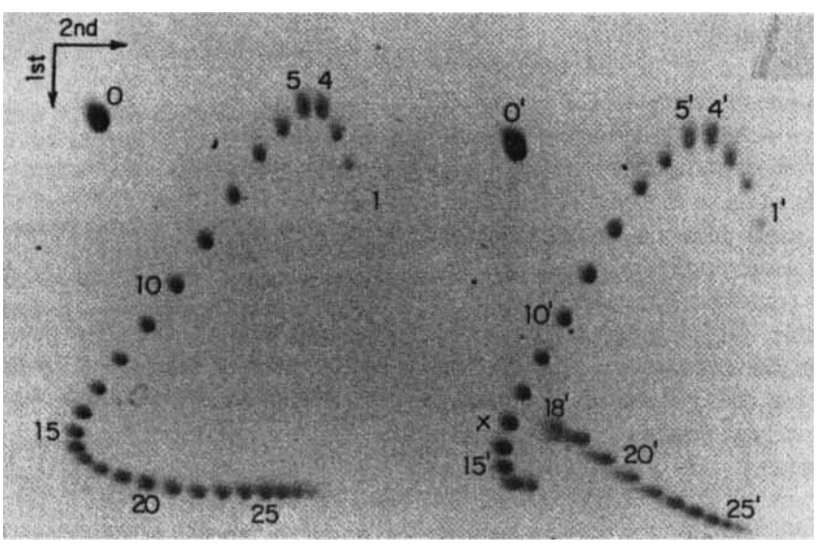


**FIGURE 4.** Two-dimensional topoisomer analysis of a plasmid without (gel on the left) and with a $d(G\text{-}C)_{16}$ insert (right). A family of topoisomers of each plasmid was prepared and loaded in two wells at the top left of two square 0.7% agarose gels. The gels were run first in Tris-borate-EDTA at room temperature for 20 h, soaked in 1.25 μ*M* chloroquine for 6 to 8 h, then run in perpendicular direction for 18 h in the presence of chloroquine. The gels are stained with ethidium bromide (negatives are shown). Topoisomer numbers are shown. 0 and 0′ are nicked, relaxed molecules. The topoisomer marked x migrates in the first direction closest to the first post-jump topoisomer, and is assumed therefore to have a similar number of superturns. (From Wang, J. C., Peck, L. J., and Becherer, K., *Cold Spring Harbor Symp.*, 47, 85, 1983. With permission.)

form migrating slightly less than topoisomer −5, compatible with the opening of 4 superturns and unwinding of 40 to 42 base pairs, as expected when a cruciform is formed by the palindromic polylinker. Unwinding upon cruciform extrusion was similarly shown with the minimal length plasmid pA03 (1683 bp).[298] Cruciform extrusion can be slow: $t_{1/2}$ for, e.g., topoisomer −11 of AP03 at 25° is 9 h.[366]

At higher temperatures, an additional set of transitions can be observed and was studied in detail by Lee and Bauer.[265] A 5420-bp plasmid derived from resistance factor R12 undergoes a series of unwinding transitions between 37° and 65°. The two-dimensional patterns show that during the first transition, at 37°, an average of 67 bp were unwound, i.e., 6 or 7 turns of the B helix (Figure 5). At least 4 additional transitions can be distinguished as temperature increases. The high extent of unwinding detected is due to the large number of topoisomers (+16 to −38) analyzed in this elegant study. The superhelical energy, ΔG, required for the first transition (deduced from the linking number of the topoisomers undergoing transition) decreased linearly, from 35.1 Kcal/mol at 37° to 6.6 Kcal/mol at 65°. ΔG extrapolates to zero at 71.4°, within 3° of the melting temperature of the linear plasmid. The linear free energy change from 37° up to complete melting suggests that complete unstacking and strand randomization, rather than formation of an unusual duplex or triplex, accompany the various transitions. The lack of sequence information for the R12 plasmid prevents a correlation between the composition of individual base segments and the different melting transitions. We cannot say whether special symmetries, or A,T content, or R · Y composition are responsible.

An alternative and no less elegant way to observe early stand separation is to load a topoisomer mixture on a gel with a gradient of denaturing agents (like formamide-urea) perpendicular to the electric field[299] (see Figure 6). A family of complete transition curves is visible directly on the electrophorogram. It is immediately evident that the amount of denaturant required for melting decreases as the superhelical strain goes up, transition free energies at different points along the gradient can be estimated directly from the pattern.

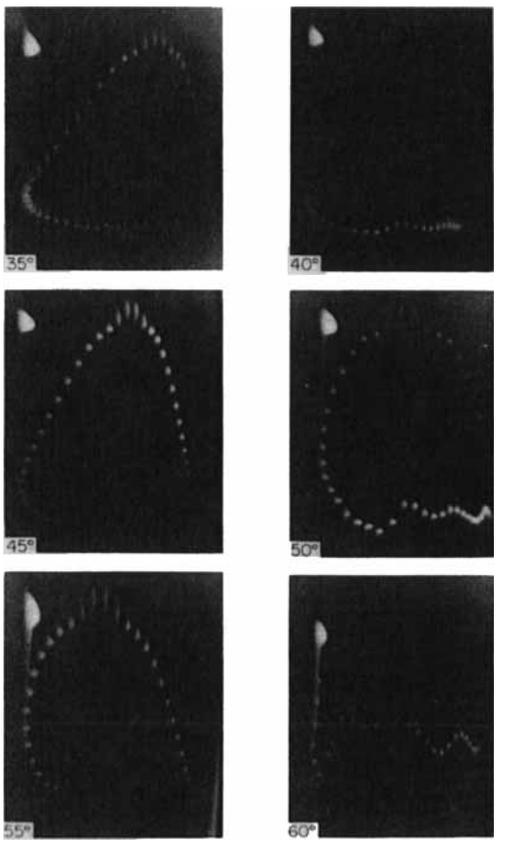


**FIGURE 5.** Two-dimensional topoisomer analysis of plasmid pSM1 at several temperatures. The experimental procedure is essentially similar to that used in Figure 4. The various temperatures at which the topoisomer mixture was analyzed are indicated on each gel. It can be seen that, for example, the first transition occurs at a topoisomer number that decreases with each temperature examined, reflecting the decreased torsional free energy required for the unwinding of the DNA region undergoing the transition. (From Lee, F. S. and Bauer, W. R., *Nucleic Acids Res.*, 13, 1665, 1985. With permission.)

Two-dimensional topoisomer analysis of a R · Y structure was first reported by Lyamichev et al.[293] and by Pulleyblank et al.[397] A transition was found by analysis of pUC plasmids containing a $d(AG \cdot CT)_{16}$ or a $d(AG \cdot CT)_{45}$ insert. The transition was favored by low pH (4.5 to 5.2) and by low salt (40 m*M*), conditions that also favor SSN sensitivity as well as reactivity toward

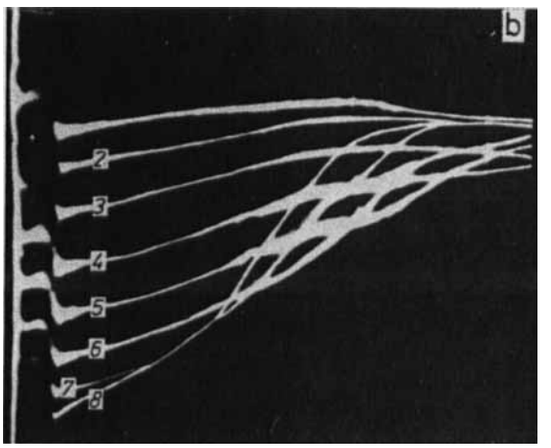


**FIGURE 6.** Topoisomer analysis on a formamide-urea gradient. A topoisomer mixture of plasmid pA03 was loaded along an extended slot on top of a polyacrylamide gel with urea-formamide gradient, the concentration of which increases from left to right. The numbers indicate the $-\Delta Lk$ of each topoisomer. It is evident that topoisomer 8 is the first melting topoisomer. (From Lyubchenko, Yu. L. and Shlyakhtenko, L. S., *Nucleic Acids Res.*, 16, 3269, 1988. With permission.)

BAA, DEPC, etc. When the topoisomer mixture was treated with S1 prior to electrophoresis, the "post jump" topoisomers were absent, confirming that unwinding caused by high superhelicity is responsible for SSN sensitivity[293] (Figure 16). The superhelix density required, deduced from the first topoisomer undergoing transition, increases linearly with pH down to pH 4.3,[293] indicating that the pK of the sensitive form is lower than that pH. The retarded migration of the topoisomers that underwent the transition corresponds to 4.5 turns unwound for the 45-base d(AG · CT) insert.[396] This means that complete unwinding, but no opposite rewinding, is involved in the formation of the acid homopurine structure. A similar two-dimensional pattern was found also for a $(dG \cdot dC)_{31}$ insert,[295] consistent with the unwinding of close to 30 bases.

The occurrence of a local topological transition in the mungbean nuclease-sensitive region of pBR322[441] was confirmed by two-dimensional electrophoresis.[254] Local topological transitions elsewhere in the vector can be an obstacle to the identification of the transition of a particular insert under neutral conditions. We have detected by two-dimensional electrophoresis a similar transition in the related plasmid pUC8. Cleavage with P1 nuclease confirmed the sensitive site in the major mungbean sensitive region of Kowalski and co-workers[556] (see Figure 11).

## D. Electrophoretic Analysis

Retarded electrophoretic migration of linear DNA fragments is currently associated mainly with bent DNA, assumed to retain its B-form state.[248,479,480] The retarded migration is explained as the result of the nonlinear DNA having to reptate through the network of the gelified medium. There are nevertheless several reports of anomalous migration related to DNA unwinding. Anomalous migration of the formally unwound cruciforms was shown quite early. Extrusion of a cruciform results in changed migration of pUC7 in agarose.[154] Htun and Dahlbert[205] have shown that a SSN-sensitive $d(AG \cdot CT)_{18}$ containing a restriction fragment exhibits retarded migration on a 5% acrylamide (1:60 bis) at pH 4. The retardation was maximal when the TC insert was in the middle rather than at the ends of an approximately 400-bp fragment. A shorter inset, $d(AG \cdot CT)_5$, showed normal migration, supporting the formation of a hinged triplex (H-DNA) structure by the longer insert.[206]

Electrophoresis of the circular closed DNA on agarose can also be diagnostic for unwound forms. Dean et al.[107,108] have detected a fast band that is formed when replication and prereplication complexes of SV40 are formed *in vitro*. This U (unwound) band moves far ahead of the normal supercoiled form of the virus under proper conditions. A similar fast-moving band has been detected by Stettler et al.[464] with form V DNA, a structure formed by association of two circular closed complementary strands, which must be formally unwound (Section III.F). The unwound paranemic joint, formed when two complementary DNA strands are incubated with rec A protein, is also detectable by its anomalous migration on gels[100,539] (see Section IV.C).

Bending, and therefore retarded migration, can probably be expected for most structures proposed for unwound DNA, certainly for melted DNA, B-Z functions, and the nonhelical duplex

proposed by us[545] (Section III.5). The utility of electrophoretic approaches may be in that the bulk of the tested material must assume the expected conformation in order to be detected at the appropriate experimental conditions. This is in contrast to sensitivity to nucleases or reagents that can be manifested only when a minority of the molecules is in reactive conformation.

## E. DNA Melting

DNA melting is a well-studied phenomenon and extensive theoretical treatments are available (see Reference 38). The advent of high-resolution melting techniques[496] demonstrated that many classic melting curves of cellular DNA are actually a superposition of a multitude of individual transitions, each representing the melting of a particular DNA segment. Readily unwound DNA regions should be among the earliest melting "thermalites" (see Reference 496). High-resolution melting curves are available now for a number of small viruses and plasmids of known base sequence.[502] This should make it possible, in principle, to ask what base compositions tend to be the early melting ones. The assignment of a particular thermalite to a particular DNA segment, however, has proven to be a difficult task. The procedure generally adopted was to assume that melting occurs in the order of decreasing AT content and to try to fit a curve calculated on that assumption to experimental data.[170,303,502,503] An unequivocal fit could, however, not be obtained until auxiliary techniques were introduced.

An important auxiliary technique is electron microscopy. DNA is heated to a certain temperature, melted regions are prevented from reannealing with formaldehyde or glyoxal, the molecules are linearized with a once cutting restriction enzyme, and the position of the observed "bubbles" is determined by EM. Using this technique, Kalambet et al.[223] were able to assign 3 to 4 individually melting regions of the 1683-base plasmid pA03 to definite DNA regions. Plasmid pBR322 exhibited up to 8 separate peaks in its high-resolution melting curve (Figure 7).[223,499] The v marks on the figure, and the associated numbers, indicate which section of pBR322 is melting, as identified by EM at the same temperature.

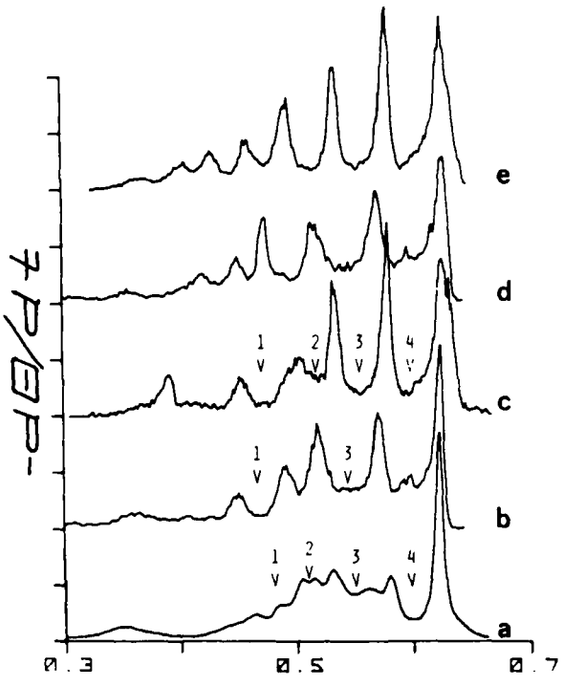


**FIGURE 7.** Fine-structured melting profile of pBR322. (a) Theoretically expected differential melting curve; (b) experimental melting curve in SSC, Eco RI linearized plasmid; (c) as (b), in 0.1SSC; (d) experimental curve of Pst I linearized plasmid in SSC; (e) as (d), in 0.1SSC. The v notch in (b) and (c) indicates temperatures at which the location of the melted region was established by EM. The notch in (a) indicates theoretically predicted minima. (From Kalambet, Yu. A., Borovik, A. S., Lyamichev, V. I., and Lyubchenko, L., *Biopolymers*, 24, 359, 1985. With permission.)

Interestingly, the lowest melting regions of pBR322, below v1, map by EM to bases 3050 to 3450 and around 4200. In these regions, Kowalski and co-workers[254,441] located the major SSN-sensitive sites of pBR322. The SSN-sensitive sequences have the expected high AT content but are not devoid of polypurine clusters. Cruciform-forming sequences are also present. The resolution of the EM mapping method is too low ( ± 100 bases at most) to determine definitely whether the A,T-rich region or the cruciforms, or both, are responsible for the early melting. The second region observed by EM to melt, bases 4000 to 4350, also includes a secondary S1-sensitive region. It would be of interest to compare the 4 to 5 thermal unwinding transitions detected by two-dimensional electrophoresis by Lee and

Bauer[265] with the high-resolution melting profile of the linearized plasmid. Unfortunately, a high-resolution study of that plasmid is not available.

A further potentially helpful technique is the gradient gel technique of Lerman and coworkers[273,274] and Myers et al.[345] Here the entire restriction fragment mixture of a plasmid is loaded along a slot of an acrylamide gel and then run perpendicular to a formamide-urea gradient. A complete denaturation profile of each fragment is obtained. Thus the melting profiles of the five longest Alu I fragments of pBR322 can be obtained on a single gel. Interestingly, the 521-bp fragment B, which contains the SSN-sensitive region of the plasmid, has been found to be the lowest melting fragment.

All these studies concern SSN-sensitive DNA regions characterized by a high A,T content[51,485] known to favor early strand separation. In higher eukaryotic genes, the readily unwound eukaryotic regions are, however, characterized by a high purine · pyrimidine asymmetry. DNA sequences containing eukaryotic SSN inserts, like $d(AG \cdot CT)_n$, have yet to be studied by high-resolution melting procedures. By their A,T content these inserts can be expected to be middle melting. The melting profile should be very different if the actually melting structure were non-B (e.g., a triplex). Inspection of Table 1 shows that there is no particular preference of A and G or of T over C in SSN-sensitive regions. Can we reconcile the high A,G content of many SSN-sensitive regions with the early melting of A,T-rich regions? One answer may be that the two compositions need not be exclusive. Alternatively, $\Delta H$ and $\Delta S$ might not be constant but vary with temperature, making the unwinding step a favorable one at physiological temperatures. The detailed measurements of Breslauer et al.[54] do not favor this; it should be noted, however, that the $\Delta H$ and $\Delta S$ values are low for ApG and CpT pairs, compared, for example, with ApC or GpT pairs. $\Delta H$ values are determined primarily by stacking interactions with adjacent bases, rather then by the H bonding of complementary bases. This supports a third possibility, namely, the polypurines in sensitive regions adopt an unwound but still stacked conformation, leading to a favorable $\Delta G$ at physiological temperatures. Melting studies of synthetic oligo R · Y inserts could serve to examine this possibility.

## F. Structural Methods

The ultimate method for establishing the structure of a particular unwound form of DNA is by molecular analysis employing X-ray crystallography, NMR, IR, Raman, and CD spectroscopy. A survey of these methods is clearly beyond the scope of this review. Circular dichroism has already been instrumental in delineating the pH and temperature ranges in which a DNA triplex is stable, as described in Section III.G.1. The recent determination of the melting profile of an artificial triplex is an example.[542]

Two-dimensional NMR spectroscopy is rapidly emerging as a most useful method for determining the structural details of DNA in its natural solution environment.[375,541] This technique has been valuable in establishing the range of existence of cruciform structures. The stem and loop part of the cruciform have been studied using partly self-complementary DNA fragments as model hairpin loops.[37,361,528] These studies established the B-type character of the stems, and the single-stranded, yet structured, character of the loop.

NMR analysis has most recently been applied to complexes formed by 2:1 mixtures of oligopyrimidine and their complementary oligopurine strands.[402,403] The complexes have been shown to adopt triplex structures at pH 5. Crosspeaks were observed in a $d(T\text{-}C)_4 \cdot d(A\text{-}G)_4 \cdot d(T\text{-}C)_4$ complex between H8 protons of guanines, and N3 protons of cytosines of the third strand.[402,403] This means that the two hydrogens are less than 4 Å apart, consistent with the formation of a Hoogsteen-type H bond between C-N3 and G-N7 opposite to it. The detailed assignment of interacting peaks indicates that the second dTC strand fills the major groove of a normally paired d(AG · TC) duplex, running antiparallel to the Watson-Crick H-bonded dTC strand. The proposed structure is essentially the same as proposed for the SSN-sensitive segments at acid conditions.[205] In a second study, new low-field resonances were identified in a $(Y \cdot R \cdot Y)_{11}$

complex and assigned to the imino protons of thymine (at position N3), again, at pH 5 and not at pH 6.8. Crosspeaks resulting from thymine N3 hydrogen participation in Hoogsteen H-bonds have been identified in all four T · A · T base triplets expected in the 11-base-long triplex studied, confirming the suggested structure of the triplex.

X-ray crystallography has also served to establish the existence of certain unwound structures. The first structure was a platinum complex ($[(bipyridyl)_2 \cdot Pt(ethylenediamine)_2]^{2+}$ intercalated into calf thymus DNA fibers[10] (see Figure 9). Crystals formed by intercalation of the quinoxaline peptide triostin A into the self-complementary hexanucleotide CGTACG duplex were analyzed. The two CpG pairs that accommodate the quinoxaline ring were found to be largely unwound.[506] Intercalation and unwinding are accompanied, in both structures, by stretching of the deoxyribose-phosphate backbone, increasing the distance between base planes to 5 to 7 Å. Unwound structures of this intercalative type can play a role in unwinding *in vivo*, when potential intercalating residues, for example, certain protein side chains, are available to intercalate and to stabilize the unwound state. The structural methods will certainly play a major role in the ultimate characterization of any unwound structure of demonstrated biological significance.

## III. THE PARANEMIC DNA STRUCTURES

We have seen in the previous chapter that many genes contain DNA regions that are sensitive to single-strand specific (SSN) nucleases when examined either in intact nuclei or after insertion into supercoiled plasmids. It is reasonable to assume that the DNA in these sensitive regions is in a topologically unwound state, otherwise ready separation to single strands would be hard to envisage. Unwinding does not necessarily mean complete separation into single strands. It means that DNA in the sensitive regions is in a conformational state topologically distinct from the conventional B state. A number of energetically favorable structures have been proposed, in all of which DNA is in a formally unwound state. In this section, the various structures proposed for the unwound, paranemic state are described, and the evidence in favor of or against the existence of each one is discussed. Some of these structures have particular symmetry and energy requirements. These are specified for each model and are summarized in the last subsection. The alternating all-purine sequence $d(AG \cdot CT)_n$ has been found to be the most SSN-sensitive one, and possesses both mirror symmetry and direct repeat features. It is the best-studied SSN-sensitive motif, and much of the discussion will necessarily focus on the structure of this R · Y sequence in the paranemic state.

### A. Single Strands

Separation of the DNA of specialized genomic regions into single strands has been considered a number of times over the years[95] and is a straightforward explanation for sensitivity toward SSNs. Unwinding of DNA into single strands has no symmetry requirements, and offers a direct explanation for SSN sensitivity. Under physiological conditions, a large input of free energy is nevertheless required to keep the strands from reannealing. This free energy can be provided either by the available superhelical (torsional) energy, or, in the intact cell, by interaction with proteins like SSB, by association with RNA, or with nuclear factors, as discussed in Sections III.I and III.J.

The experimental evidence in favor of separation of SSN-sensitive regions into single strands is not conclusive. In order to be cleaved by SSN, only a minor fraction of the sensitive DNA needs to be actually in the ss state, but can open just transiently by "breathing". One way to establish what fraction of the DNA is in the sensitive state would be to compare the rate of cleavage with a denatured single-stranded control. A proper single-stranded control should be circularly closed to exclude possible exonucleolytic cleavage, or other end effects. Extents of cleavage are, however, generally determined by gel analysis, which does not readily yield accurate rate data. In one study where a linear, denatured DNA was included as a control, this single-stranded control DNA yielded a S1 banding pattern quite similar to that of the supercoiled

plasmid studied; in the ss DNA the pattern was nevertheless evident already after 0.25 and 1 min of digestion, while it took 10 min for the same pattern to appear in the ds supercoiled plasmid.[432] This means that the S1-sensitive region in the closed plasmid was not actually single stranded.

Wang and Hogan[516] studied kinetically the cleavage of a sensitive site in the second intron of the chicken β-globin gene. A site near base +620 is sensitive to S1 at pH 4.5, and to the *N. crassa* nuclease at pH 7.5, even in linear form (see Table 1). The rate of cleavage of the linearized plasmid, as judged from gel analyses, was found to be practically independent of the amount of enzyme present. This suggests that a slow internal transition of the DNA must occur before it can be cleaved by the nuclease. In the linearized molecule, this slow transition can either be unwinding of the helix, or separation of the strands of an already unwound structure. In either case, at most a small minority of the attacked DNA can be present in the straight single-stranded form. An enzyme-independent rate of cleavage was also found in a supercoiled myc-containing plasmid,[39] but can be explained by the presence of a stabilizing RNA (see Section III.H).

Fine structural acrylamide gel analysis of the products of SSN cleavage (Sections II.A and III.G.5) indicates that only selected bases along a polypurine region are attacked. If the polypurine region was in a free single-stranded state, all bases would be equally cleaved (no sequence specificity of S1, P1, etc. has been reported).[327] This means, again, that the vulnerable unwound region may not be actually separated into single strands. Single-strand-specific reagents like DMS or BAA (Section II.B) also lead to localized rather than to general modifications along a SSN-sensitive region. Thus, N3 of cytosine in −55 chicken β-globin SSN-sensitive DNA was not reactive to DMS, as it should be if it was in single strands.[69] Restriction enzymes Mbo II and Sma I have recognition sites within the sensitive region, and cut the region efficiently within a supercoiled plasmid.[433] This would not be expected if the region was dissociated into single strands. The reactivity of oligo d(AG · CT) toward BAA is also not as expected for single strands.[184] Chloracetaldehyde is only partially active toward an oligo $d(G \cdot C)_n$ insert,[244] as discussed further in the subsection on triplexes (III.G). These findings are nevertheless not conclusive evidence against a single-stranded state, because supercoiled DNA relaxes after the first hit by an enzyme or reagent, preventing further cleavage, so that only the most readily attacked sites of the sensitive region may be detected.

Kowalski et al.[254] studied the cleavage by mungbean nuclease of two A,T-rich regions in plasmid pBR322 (bases 3180 to 3301 and 4133 to 4252). Complete melting rather than "breathing" is offered as an explanation, mainly because two-dimensional electrophoresis shows a clear structural transition of all topoisomers of ΔLk beyond −18. These, and not the topoisomers before −18, are also the mungbean nuclease-sensitive ones. This means that the bulk of the DNA present is in the sensitive state. The transition is observed at 37° and not at 23°. The sensitive regions contain symmetric subregions of quite limited extent (at most a 9;6 hairpin loop at 3223 and a 7;3 one at 4174), and therefore they are not likely to assume symmetric paranemic structures. These observations, and the high A,T content of the sensitive regions, led Kowalski et al.[254] to prefer complete separation into single strands. Lee and Bauer[265] used two-dimensional topoisomer analysis (see Section II.C) of a bacterial plasmid to show that the transition of successively unwound regions extrapolates directly to the melting point of the plasmid DNA, consistent with direct transition to single strands.

The prokaryotic SSN-sensitive regions differ from eukaryotic ones not only in that neutral conditions are optimal, but also in their A,T-rich rather than R · Y-asymmetric composition (Section II.A.7). Early melting of AT-rich regions is a well-established phenomenon in molecular biology.[142,215,336,337] For supercoiled plasmids, McClellan and co-workers have found that under some conditions oligo dA-dT can behave like R · Y segments[314,315] They find (see Section II.B.4) that oligo dA-dT inserts can undergo transition into two distinct forms. One form (the "C" form) is a cruciform, having the required dyad symmetry; the other form (the "U" form) is uniformly S1 and $OsO_4$ sensitive, favored by low salt and high T, and is therefore either in the completely melted, or in a related unwound state. The A,T-rich motif, rather than the all-purine

motif, may well be the prevalent one for prokaryotic and yeast SSN-sensitive DNA.[486,487] The low melting A,T-rich regions are thus the most likely ones to be separated into single strands when in the unwound state.

## B. Slippage Structure

This structure, depicted in Figure 8b, requires the presence of a direct repeat in the DNA sequence. One strand of the first repeat hybridizes with the complementary strand of the second repeat, with the opposite and in-between regions looping out of the structure. Two conformers are possible, according to which strand of the first repeat pairs with the complementary strand of the other repeat. The slippage structure is topologically at least partly unwound, because the two strands of B-form DNA must untwine and separate before the slippage structure can be formed. A slippage structure was first proposed by Hentschel[194] to explain the SSN sensitivity of the $d(AG \cdot CT)_{19}$ region in the sea urchin histone gene (Table 1). Subsequently, Mace et al.[300] considered this structure to explain the sensitivity of a T,C rich region in *D. melanogaster* heat shock genes. A slippage structure was also considered for the pH-independent SSN-sensitive sequence in the major late adenovirus promoter,[550] which contains a direct repeat. The slippage region is expected to be protected against SSN and chemical reagents only in the paired repeats, and to be reactive throughout the free loops. Such a "bimodal" S1 cleavage pattern was indeed observed for oligo d(AG · CT) inserts,[194,208] but was not substantiated by fine structural analysis of the cleavage products on gels.[205,220,501] The direct repeat in the major late adenovirus promoter was carefully tested with BAA, but only slight reactivity was found along the supposedly looped out segments.[230] The polypurine insert studied by Mirkin et al.,[331] discussed in Section III.G.6, shows the manifestations of an SSN-sensitive sequence, but possesses no direct repeat, excluding a slippage structure. Slippage structures certainly cannot explain many of the SSN-sensitive sequences in Table 1, for which no direct repeat of significant length can be found. Energetically, a large degree of strand separation is still required in order to form a slippage structure. The weight of present evidence can thus not be considered as favoring slippage structure models for unwound DNA regions.

## C. Cruciforms

The cruciform structure, also referred to as the double hairpin, stem and loop structure, was first proposed by Gierer[158] and has attracted much attention as a potential genetic recognition site (reviewed in References 280 and 282). The formation of a cruciform requires a dyad symmetrical arrangement of the bases (often referred to as inverted repeats, or palindrome, actually misnomers) to permit stem formation. The cruciform (Figure 8c) is a paranemic structure because the two strands of B-form DNA must be separated before they can form the stem and loop structures. In a closed circle cruciforms can be formed only by opening negative superhelical turns or equivalent source of linking deficit.[93,154,279] The bases along the stems of the crucifom are rewound to form an intrastrand B-form helix, so that the stem region becomes geometrically, but not topologically wound. The energy investment is limited to the unpairing of the bases in the loop and at the base regions. A transition into cruciform is therefore less energy costly than transition into single strands or into a slipped structure. The alternating $d[(T\text{-}G)_6\text{-}(C\text{-}A)_6]$ segment, which can form either a cruciform or left-handed Z DNA, has been found to prefer the cruciform state over the Z state.[320]

Cruciforms are well known to be sensitive to nuclease S1, by virtue of the free loop at the end of the stem.[278,365,457] Dyad symmetric regions have been noted in quite a number of genes, ranging from simple plasmids like pBR322 to well-studied eukaryotic genes.[344,525] The actual extent of cruciform extrusion in the intact genome is less clear. The well-known inverted repeat of the lac operator region, for instance, does not form a cruciform.[44,356,451] Neither poly d(AG · CT) nor most other R · Y SSN-sensitive regions in Table 1 possess dyad symmetry, so that cruciforms cannot explain SSN cleavage of most purine-rich regions. The only entries in Table 1 in which a cruciform may be implicated are the secondary

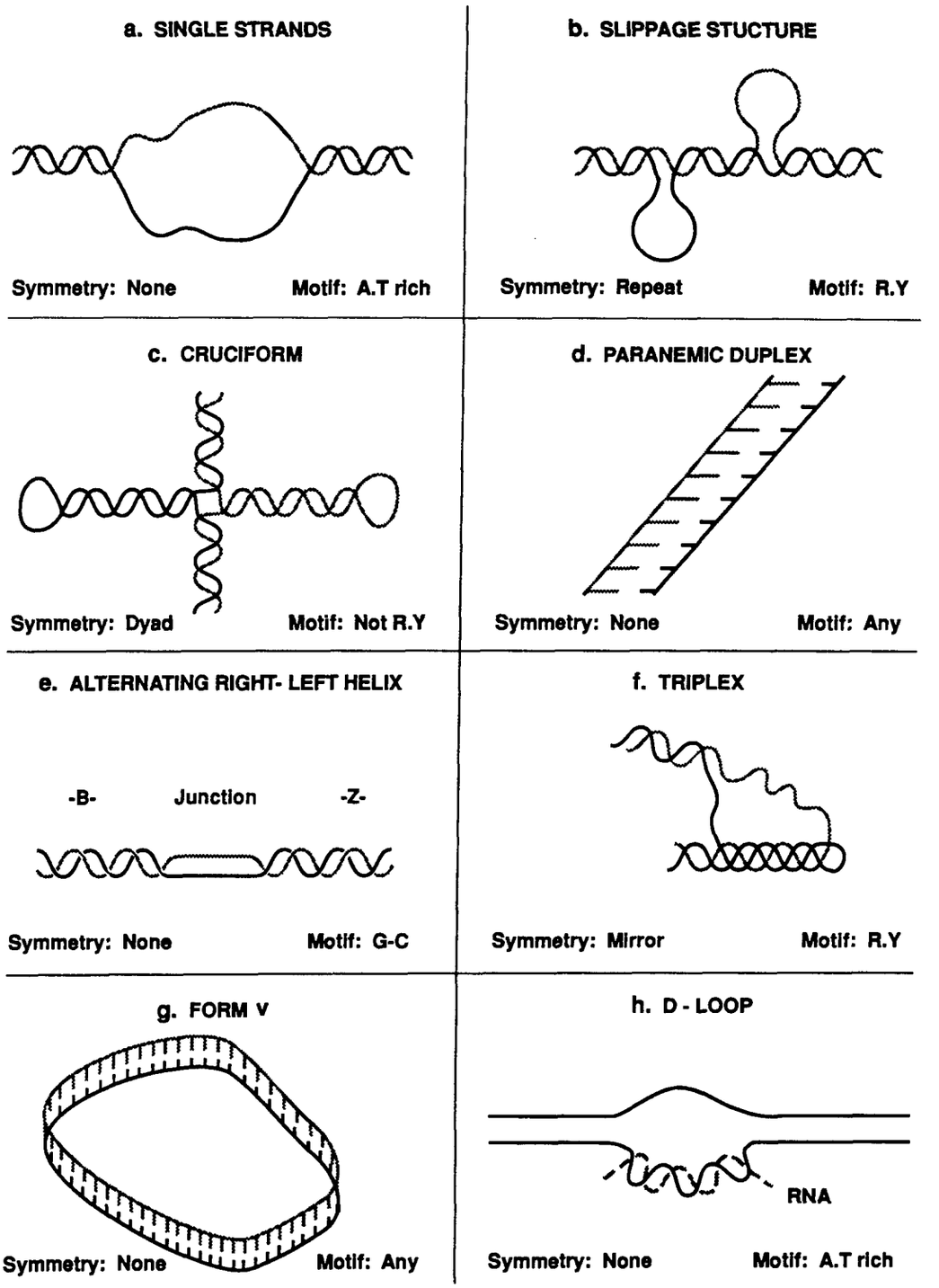


**FIGURE 8.** Eight structural forms considered for paranemic DNA. The symmetry requirements of each structure, and the principal sequence motifs reported, are indicated. Additional less regular forms can be expected to be formed by interaction with transcription factors, and other nuclear DNA-binding proteins.

S1-sensitive sites of pBR322, cleaved in the absence of $Mg^{2+}$,[441] and the oligo(A-T)$_n$ inserts studied by McClellan and Lilley[314] (see Section III.A). To conclude, dyad symmetric regions have the potential to form unwound regions by cruciform extrusion, but the extent to which this happens in the intact cell is not yet clear.

### D. Left-Handed DNA and B-Z Junctions

Left-handed forms of DNA were first detected in high-salt solutions.[385] The crystallographic determination of the Z-type structure for d(G-C)$_6$ served to break the premise that the right-handed helical form is the only stable form of DNA.[505] Left-handed DNA by itself cannot be considered as paranemic. A left-handed Z DNA segment of n bases can, however, combine with an adjacent right-handed B segment of an equivalent number of turns, to form a paranemic region U of 2n bases (Figure 8e):

$$B_n + Z_{n'} \rightleftharpoons U_{2n} \rightleftharpoons 2(SS)n$$

A mobile equilibrium of the unwound region and its two separated strands (SS) can be expected; there is no topological barrier to this process, and it can take place in a circular closed DNA. There are also neither symmetry requirements nor energy barriers to the process. An alternating B-Z structure of this type has been proposed for form V DNA, as discussed in Section III.F. In one study a left-handed model has been proposed for a SSN-sensitive region. The proposal was based on a square relation between the length of the plasmid studied and the superhelicity required for SSN sensitivity.[69] That relation fits the presence of a left-handed segment, but can also fit other paranemic conformations. There has been no further consideration of this proposal.

Z DNA, its structure, and evidence for *in vivo* roles are well reviewed[35,353,409] and will not be reviewed here. The potential to form left-handed DNA has been found mainly in alternating purine-pyrimidine tracts, particularly of the dG-dC class, rather than in homopurine · homopyrimidine (R · Y) sequences. This is a strong argument against a major contribution of left-handed segments to SSN sensitivity. Alternating oligo d(G-T · A-C) can also form left-handed DNA, but oligo d(A-T), which we have seen to be SSN sensitive under some conditions, can be induced to form Z-type structures only when imbedded in d(G-C) or d(G-T · A-C) segments.[128,229,320] Few if any of the eukaryotic SSN-sensitive sequences in Table 1 have alternating R-Y tracts of sufficient length to form a Z-type structure. An exception may be the long alternating d(G-T · A-C) sequences, 170 bp, in the rat prolactin promoter region.[348] Two-dimensional electrophoresis shows the opening of 5 superturns. This is interpreted as formation of a 28-base Z DNA region. The data can nevertheless also be interpreted as the unwinding of 56 bases. The S1 sensitivity of the same region is in fact in favor of the second alternative. More extensive evidence in favor of coexistence of alternating left- and right-handed regions has been found for the case of form V DNA discussed in detail in a subsequent section.

Systematic analysis of genomic sequences has shown that long alternating dG-dC segments are rather rare in genomic DNA.[247] This is in sharp contrast to purine- or pyrimidine-rich sequences, which are highly overrepresented in most genomes.[61] The relatively few occurrences of Z-form DNA segments in cellular DNA are in line with other indications that Z-forming sequences are underrepresented in actual genomes.[9,202,552] The low occurrence of Z-type DNA may still signify an important role for Z-type segments, including the controlled unwinding of very special DNA regions.

**B-Z junctions** — The junction of a Z-form insert in a plasmid to adjoining B-form segments has been recognized to have some features of an unwound structure; the junction region was found to be insensitive to B-recognizing restriction enzymes[14,349] and to be sensitive to nuclease S1.[456] The high reactivity of hydroxylamine and $OsO_4$ at the B-Z junction, rather than within the Z region, led Johnston and Rich[221] and Palecek et al.[364] to propose these reagents as suitable for SSN-sensitive regions. This does not necessarily imply that junction regions are actually separated into strands. The lack of reactivity of BAA in the junction region led Kang and Wells[225] to propose that the bases in the junction are in a paired

state. This conclusion was, however, modified when a better technique was worked out.[319] Thus it remains unclear to what extent B-Z junctions are not only unwound, but actually strand separated. Further study of B-Z junctions might provide important leads to the structure of unwound DNA segments.

## E. Paranemic Duplexes

The idea that the DNA duplex can exist in an unwound, nonhelical state (the "paranemic" vs. the "plectonemic" state) was raised by Watson and Crick as soon as the double helix was proposed.[519] It has been reconsidered a number of times since.[101,384] Detailed "side by side" models were proposed as alternative explanations of DNA fiber diffraction data leading to the classic B-form structure.[5,329,415,427] In these models the helix winds first five turns in the right direction, then five turns in the left direction, repetitively, so that an overall unwound structure results. Direct determinations of the helical winding number of B DNA from topoisomer shift data[96] and from DNAse I digestion patterns of adsorbed DNA[408] established definitely that the bulk of the DNA extracted from cells exists in a B form and not in an unwound form. These experiments do not rule out the possible participation of unwound duplex forms in selected gene regions, or during intermediate stages of gene expression processes.

In parallel, evidence was accumulating that the twist angle between adjacent bases can be significantly smaller that 36°. Arnott et al.[10] showed by X-ray diffraction that upon intercalation of a bipyridyl-ethylenediamine Pt complex ($[bipy)Pt(en)]^{2+}$) a ladder-type DNA structure is formed (Figure 9). Adjacent base pairs adopt a parallel arrangement and the distance between bases alternates between 3.4 Å and 6.8 Å to allow the intercalation of the complex. Low twist angles have been found in other complexes, like the crystals formed upon intercalation of the quinoxaline antibiotic triostin A into a 6-bp oligonucleotide.[506] In the Z-II form of DNA, the bases in the CpG step have a winding angle of $-8.5$ rather than $-36°$.[114] A highly unwound helix has been proposed by Olson.[359] To form this helix, the nucleotides must be in the high anti conformation, in which the glycosidic angle $\chi$ is close to 180°. This high anti angle has been shown to exist in nucleotides in which carbon C′2 is linked to purine C8 via a S or an O bridge.[484]

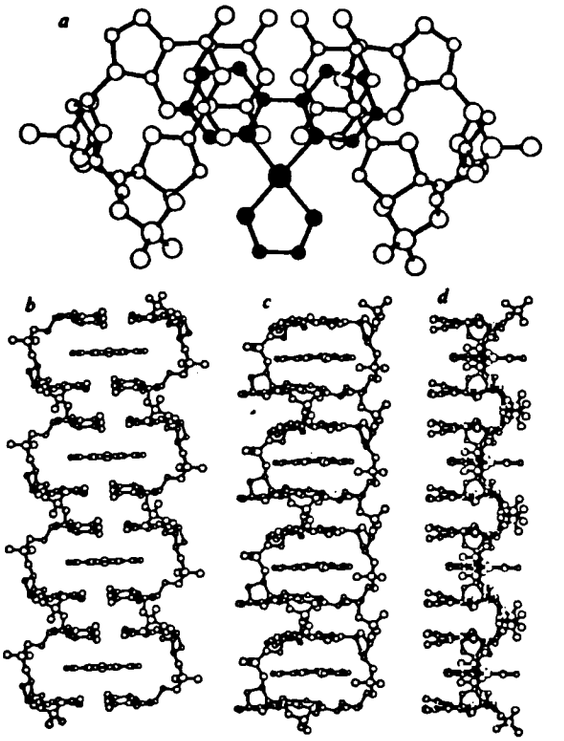


**FIGURE 9.** The crystal structure proposed for DNA unwound by intercalation with a platinum complex. (From Arnott, S., Bond, P. J., and Chandrasekharan, R., *Nature,* 287, 561, 1980. With permission.)

The structural and energy requirements of unwound DNA duplexes were reinvestigated by computerized energy minimization procedures by Yagil and Sussman.[545] When the energy minimization program EREF[277] was applied to a crude model, the paranemic structure shown in Figure 10 resulted. This structure is completely unwound, yet the base pairs are hydrogen bonded and stacked in the conventional fashion. The distance between adjacent base pairs ranges from 3.6 to 3.9 Å, with a propeller twist close to that found in B-form DNA. The structure arrived at differs from previous unwound models by having a very high tilt, of 41°, leading to a diagonal appearance of the structure (Figure 8d). The high tilt is a consequence of an unusual *trans* conformation of backbone angle $\beta$ (P-O5-C5-C4), lead-

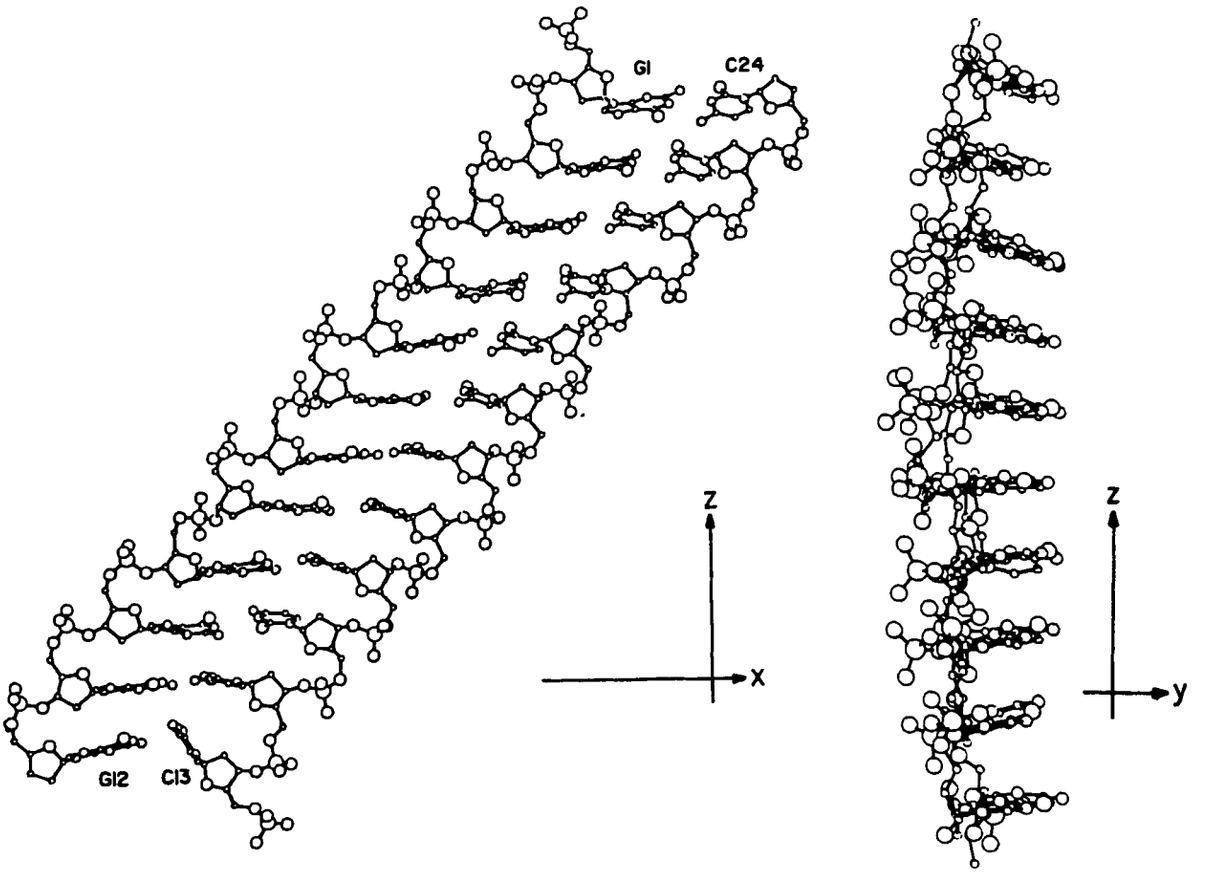


**FIGURE 10.** A computer-minimized model for paranemic DNA in the duplex state. (a) Front view; (b) side view. Note the 41° tilt and the alternating positions of backbone phosphates. Note also the exposed state of the bases in the side view, facilitating recognition by various nuclear components.[545]

ing to a high slide between adjacent bases. The unusual *trans* state of the β angle has been shown by two-dimensional NMR analysis to exist in certain autocomplementary DNA loops.[362] A further feature of the model is an alternation in the conformation of adjacent nucleotides, primarily on the purine-rich strand, as can be seen in Figure 10. This is in line with an alternating sensitivity of adjacent bases to S1 and DMS reported in a number of studies.[129,397]

The transition of $dG_{12}.dC_{12}$ from the B-form state to the proposed nonhelical state ("N-DNA") requires an input of 8.2 Kcal/mol bp, compared with 12.4 Kcal/mol for complete melting of 1 bp (Table 3). The ΔH associated with the melting of a GG · CC pair is, according to Breslauer et al,[54] 11.0 Kcal/mol bp. The reason for the difference can be that neither the effects of counterions nor those of solvation are considered by EREF, as by most current energy minimization programs. The entropy component of free energy is also not included. Nearly identical results are obtained when energy minimization is done by EREF, AMBER, or a dynamic modeling procedure based on EREF (see Table 3). The polypurine sequence found to require the least energy for the transition from B state to the paranemic form was indeed $d(AG)_n.d(CT)_n$ (Table 3); alternating dG-dC required even less and was exceptionally stable in the Z form.[545]

The two strands of the paranemic duplex can of course separate without prior rotation of one strand around the other, like in all other structures in this section. There is, in addition, no symmetry or nucleotide sequence requirement for the formation of the paranemic duplex; DNA of any

**TABLE 3**
**The Energy of Transition of B DNA to an Unwound Duplex**

| | $\Delta E^a$ (Kcal/mol bp) |
|---|---|
| $dG_{12}.dC_{12}$ ($N_2$) | 8.2 |
| $(dG\text{-}dA)_6.(dC\text{-}dT)_6$ | 7.4 |
| $dA_{12}.dT_{12}$ | 10.8 |
| $dG\text{-}dC_6.(dG\text{-}dC)_6$ | 7.0 |
| $dG_{12}.dC_{12}$ (dynamic) | 7.2 |
| Melting[b] | 3.1 ($\Delta G_0$) |
| | 11.0 ($\Delta H_0$) |

[a] Difference in energy of the oligonucletide between unwound duplex and its B form states.
[b] Data from Reference 54.

For additional information see Reference 545.

base sequence can assume this structure. Also, no protonation of cytosine or adenine is necessary, as needed for the formation of triplex structures discussed in the next section. Detailed examination of the proposed structure shows that all the bases in the structure are highly exposed to recognizing factors, both in the minor and in the major grooves. All this makes the paranemic duplex an attractive alternative model for SSN-sensitive regions, in particular for those gene regions where the transition is pH independent.

We are currently examining the two-dimensional topoisomer pattern of several plasmids containing SSN-sensitive regions under neutral conditions. The two-dimensional pattern of plasmids pUC8 and pBSTN1, containing the 19-G region of the β-globin promoter, is shown in Figure 11. A clear transition is visible with both plasmids. We are presently using a "small circle" technique[354] to determine whether the transition in pBSTN1 is due to the same SSN site or to the chicken globin tract inserted. As for pUC8, digestion with P1 demonstrated that the region most readily attacked is the region which was found by Sheflin and Kowalski[441] to be mungbean and P1 sensitive in pBR322. A similar result published earlier,[366,369] in a related plasmid, was assigned to a transition into a cruciform state. Under our conditions, the transition occurs when a linking number difference of 13 is reached. The first post-transition topoisomer migrates like topoisomer 10, which means that 3 turns are unwound. The patterns demonstrate that sequences without any obvious symmetry can, under neutral conditions, undergo a transition into a state of reduced winding, when sufficient torsional energy is supplied. Kowalski et al.[254] stress that the two-dimensional technique detects a stable unwound state and not one in a mobile equilibrium with bulk DNA. Transition to a completely melted state is the favored interpretation, as discussed in Section III.A; the formation of paranemic duplex can nevertheless not be excluded on the basis of the data brought.

Three additional duplex models have been considered as explanations of the properties of altered DNA in SSN-sensitive regions. Evans and Efstratiadis[129] proposed a heteronomous duplex, the distinctive feature of which is that the opposite strands have unequal backbone conformations. The term heteronomous was coined by Arnott et al.[12] to describe a poly dA · dT duplex in which opposite strands have different ribose-phosphate backbone conformations. The evidence in favor of a heteronomous state of SSN-sensitive regions is the alternate cutting pattern of both the $d(TC)_n$ and the $d(AG)_n$ strands of the plasmid studied, obtained by S1 nuclease as well as by mungbean nucleases and by venom phosphodiesterase at neutral pH. ApG and TpC steps are attacked by S1 in preference of GpA and CpT steps.[129] Pulleyblank et al.[397] find that, for $d(TC)_{45}$, the CpT step is the one preferentially cut. In both studies, the d(CT) strand is cut more frequently than the d(AG) strand, but, again, a single-strand control to normalize for possible base specificity is not included. No molecular details for the heteronomous state beyond the unequal conformation of the two strands and an alternating structure for consecutive base pairs are specified. It should be pointed out that in the Yagil-Sussman model, successive bases have the alternating feature, one oxygen pointing upward and the next one downward (arrows in Figure 10). This is the consequence of the alternating conformations that the deoxyribose rings assume in successive bases: C2′ exo-O1′ endo in the dG strand of $d(G \cdot C)_{12}$, complemented by C3′ endo-O1′ endo in the dC strand (the exact conformation is sensitive to details of procedure, and should not be taken as definite).

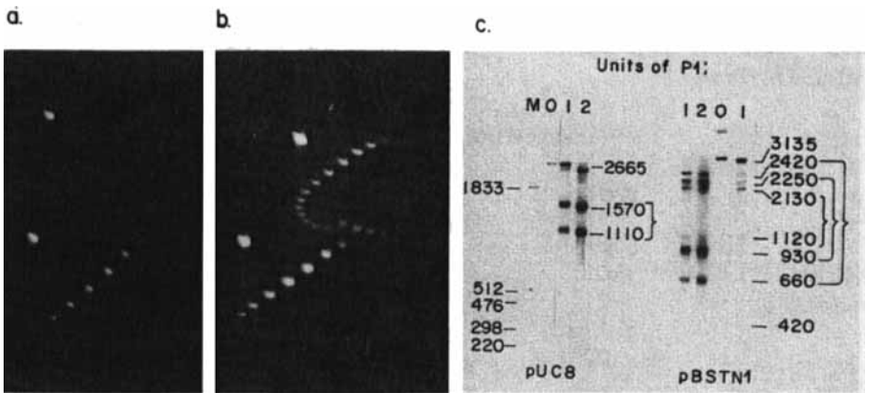


**FIGURE 11.** Two-dimensional topoisomer analysis of pUC8 and pBSTN1, at pH 8.3 (a) 20 m*M* spermine present; (b) no spermine present, bottom patterns; pUC8, top patterns, pBSTN1; (c) P1 digestion patterns of pUC8 and pBSTN1. The two-dimensional gels suggest a role for spermine in stabilizing the paranemic state (F. Shimron and G. Yagil, unpublished). pBSTN1 is pUC8 into which 470 bases from the chicken β-globin promoter (−303 to +16) have been inserted (Reference 351, kindly provided by Dr. Felsenfeld). The supercoiled plasmids were digested first with nuclease P1 at 37° for 1 h (P1 units/μl are shown on top of each lane), then restricted with Bam H1, end labeled, purified, and analyzed on agarose gel. It should be noted that the major bands group into pairs that add to the full length of the plasmids, 2665 and 3135 bases. This places the P1 cutting sites of the pUC8 vector in the mungbean nuclease cutting region established by Sheflin and Kowalsky,[441] as well as near the major and minor cruciform-forming regions described by Lilley.[278] Band lengths were established with a phage lambda digest as marker. A band mapping to the oligo G region of the globin promoter was detected with S1, at pH 4.5 only (not shown).

A quasi-duplex termed anisomorphous has been proposed by Wells and co-workers.[526] In that model, the distance between bases in one strand is somewhat larger than the interbase distances in the other strand.[525,533] This causes an occasional base to loop out, to become sensitive to the nuclease, and to nucleate an unwound region when superhelical strain increases. Wells and co-workers later adopted the triplex model to explain their results.

Finally, Pulleyblank and Haniford[396] proposed a duplex model, in which G and C bases are paired by Hoogsteen rather than by Watson-Crick H bonds, to explain their results with $d(AG \cdot CT)_n$. The formation of a Hoogsteen H-bonded structure requires the transition of the guanosines into the syn conformation as well as protonation of cytosines at position N3, explaining the acid conditions required for S1 sensitivity. DNA in the proposed duplex is, however, *not* unwound. Also, oligo $d(AG \cdot CT)$ was found to be cut by S1 equally well when position N7 in guanines is methylated,[129] so that guanines cannot participate in Hoogsteen base formation. We should also expect position N3 of cytosine to be more available to BAA and DMS than reported. The recent experimental establishment of Hoogsteen base pairing, both by crystallography[399] and by two-dimensional NMR (Section II.F) should encourage further exploration of structures containing Hoogsteen paired bases.

## F. Form V DNA

This intriguing form of DNA was discovered by Stettler et al.,[464] when the two complementary single strands of bacteriophage ϕx174, each separately closed, were cohybridized. A band migrating on acrylamide even faster than the supercoiled double-stranded phage DNA ("form I") appeared. The fast-migrating material was isolated by zonal centrifugation. Electron microscopy examination showed that it is more than

90% double stranded! This was unexpected, as the two circles were established to remain closed, so that no strand intertwining could take place (see Figure 8g). The fast-migrating form was termed form V (read five) DNA. In subsequent work (Brahms et al.[49]), CD and Raman spectra of this and similar duplexes were shown to be intermediate between Z-DNA and B-DNA; anti-Z antibodies were found to react with form V DNA (DiCapua et al.[113]). Therefore, it was concluded that form V duplex consists of alternating left- and right-handed segments, overall linking remaining zero.

A serious difficulty with this conclusion was pointed out: the frequency of alternating purine-pyrimidine sequences in all form V-forming plasmids (including pBR322 and derivatives, PM2, as well as the SV40 viral DNA) is far too low to explain the 40 to 45% Z DNA formation required.[49] For instance, not more than 0.8% alternating G-C tracts of more than 5 bases were found in the pBR322 derivative pβG. Only 7.7% of this genome consists of alternating R-Y tracts of more than 6 bases, A-T segments included.[50] Either yet unidentified left-handed helix-forming motifs exist, or we have to assume a different unwound structure to explain the existence of form V DNA. Lack of extensive symmetric regions on most form V-forming DNAs leaves the paranemic duplex proposed by us as the only reasonable structure for form V DNA (Figure 8).

This point was accentuated by the results of Brahmachari et al.,[47] and Shouche et al.,[447] who applied restriction endonucleases and methylases for form V pBR322. These enzymes recognize B-form DNA, but not Z-form or ssDNA. About half the sites cleaved or methylated in supercoiled form I plasmid were resistant when in form V of pBR322. This means that the resistant sites are not in B conformation. Only a fraction of the resistant sites reside in alternating R-Y tracts, and have no sequence feature known to form Z DNA.[447] A paranemic form that is not B-Z alternating is therefore not unlikely. Studies with structure-specific chemical and enzymatic probes are clearly called for. Whatever the structure, at present form V DNA seems to be an excellent experimental model for the unwound state. Its study may well yield important information on the conditions of existence and nature of paranemic DNA.

## G. Triplex Structures (H-DNA)

### 1. Synthetic Polynucleotides

It was observed early that certain ds polynucleotides disproportionate into a triple strand plus a single strand.[133] The early literature concerning triple helices is summarized by Felsenfeld and Miles.[132] Diffraction studies of r(A · U · U) established that the additional rU strand joins the double helix in the major groove, running parallel to the purine strand and antiparallel to the other pyrimidine strand.[11,13] The third strand pairs with sites N7 and N6 of adenine via Hoogsteen-type hydrogen bonds (Figure 12a). The conformation of the sugar ring is 3′ endo, as in A-form DNA with its deep major groove. A similar structure was found for the triple helices formed by dT and dA · dT, by dT and rA · rU, and by several other polynucleotides.[11,413] The existence of an rA · rA · rU triplex was more recently established.[57] The triplex nature of a (dT · dA · dT) helix has been confirmed by CD and NMR analyses.[382]

The situation in the G · C system is more complex, mainly because Hoogsteen-type base pairing is possible only when N3 of cytosine is protonated. A triplex structure composed of poly dG and poly dC has not been reported until recently (see Section III.G.7). The joining of an oligo $dC_n$ strand to a $d(G \cdot C)_n$ insert has been reported recently.[296] A triplex can be formed between poly dTC and poly dAG;[339] a 1:1 mixture of dTC and dAG disproportionates readily into a d(TC · AG · TC) triplex plus single-stranded dAG.[267] Triplex formation is supported by quantitative titration of dTC by dAG, by DNAse I or S1 nuclease digestion, and by melting studies. The triplex can be isolated on a $Cs_2SO_4$ gradient. A poly d(TCC · AGG · TCC) triplex can also be obtained.[172]

The CD spectrum of the $d(TC)_n + d(AG)_n$ system was investigated over a broad pH range by Antao et al.[8] At neutral pH a conventional B-form $d(TC \cdot AG)_n$ duplex predominates (Figure

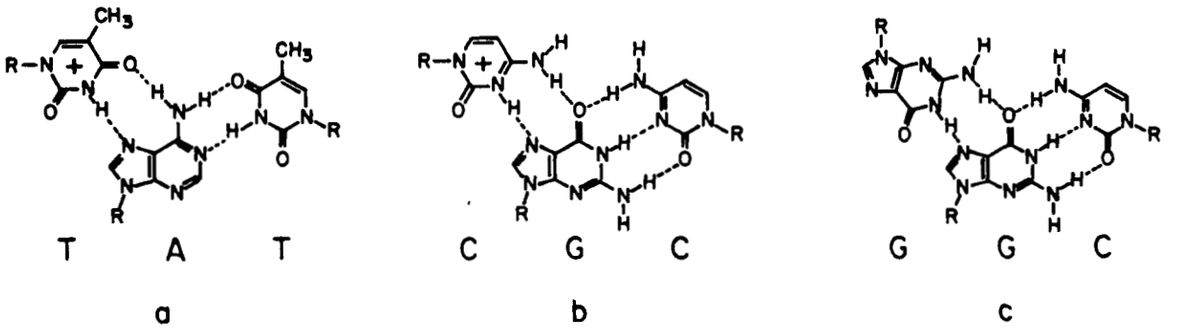


**FIGURE 12.** Base triplets formed by Hoogsteen base pairing.

13). Below pH 7, a d(TC · AG · TCH$^+$) triplex begins to be formed, in parallel to increased cytosine protonation. This triplex is the main species down to pH 3. The excessive dAG strands are liberated at first as free single strands; then, at pH 5 to 3, they form an apparently highly structured self-complex. When pH reaches 2.5, the triplex is disrupted and free single strands are formed. When pH is raised again, the initial product is a previously characterized protonated, self-structured form of poly dTC.[59,171,172] The structure d(TC)$_n$ does not tend to reform a triplex with poly dAG when pH is further raised, leading to an apparent hysteresis in the system, indicating slow interconversion processes. The conclusion that the triplex and other C-protonated forms predominate between pH 7 and 3 is in line with the acid titration curve of poly d(TC), which manifests an apparent pK of 6.3.[59] Poly C has a pK of 7.4[188,214] vs. 4.3 for deoxycytidine,[140] the negative charge of the phosphate backbone promoting cytidine N3 protonation. When 5-methyl cytosine replaces C, the triplex structure is already stable at neutral pH.[268] Immunochemical evidence for the *in vivo* formation of triplex DNA has been reported.[269] X-ray diffraction pattern of some poly d(purine · pyrimidine)fibers also favor a triplex structure.[68]

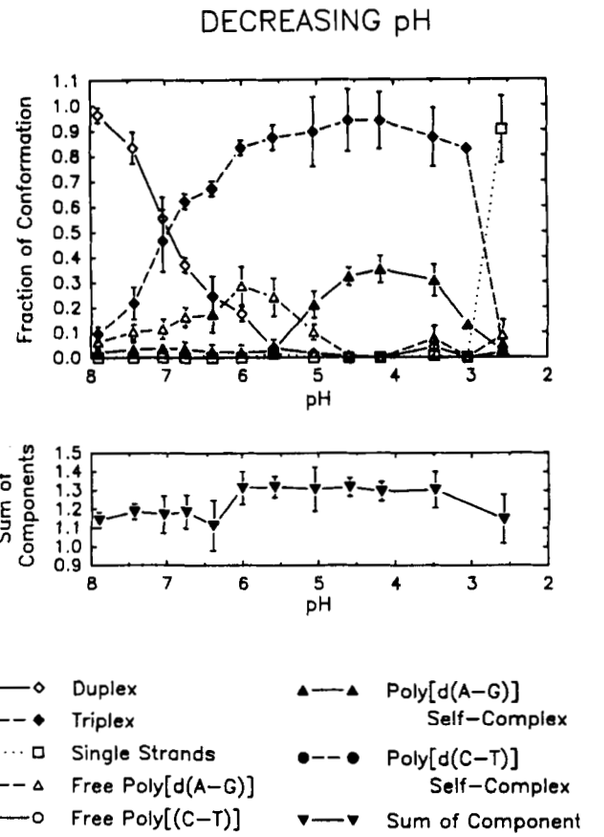


**FIGURE 13.** The composition of the poly d(AG · CT) system as deduced from the CD spectra at a range of pH conditions (From Antao, V. P., Gray, D. M., and Ratliff, R. L., *Nucleic Acids Res.*, 16, 719, 1988. With permission.)

### 2. Linear Plasmids and DNA Fragments

The capacity of a dT sequence to join a d(A · T)$_{15}$ segment within a plasmid to form a triple helix was demonstrated by Moser and Dervan[341] (Section II.B.8). A reactive Fe-EDTA complex was attached to the last base of a dT$_{15}$ oligonucleotide and subsequently incubated with the plasmid. Gel analysis reveals a specific cleavage of the dA$_{15}$ insert opposite to one end of the

associated $dT_{15}$, consistent with antiparallel rather than parallel alignment of the two pyrimidine strands. This was confirmed by attaching the reactive group to the 5′ end of the 5′ nTTTTTCTCTCTCTCT oligonucleotide and reacting it with the corresponding duplex. Cleavage at the 5′ end of the A5 tract was observed. Similar experiments with a Cu-phenanthroline reagent have been performed.[141,472] Cooney et al.[91] demonstrated the joining of a complementary purine-rich strand to a mostly (purine · pyrimidine) 27-base tract from the promoter region of the human c-*myc* gene (see Table 1). Cooney et al. point out the regulatory potential of the phenomenon. A triplex formed by self-folding of a single-stranded synthetic nucleotide of suitable composition was demonstrated recently.[182] Lyamichev et al.[296] demonstrated that a single-stranded $d(CT)_5$ oligomer did comigrate with a $d(AG \cdot CT)_{45}$-containing plasmid, forming a Y · R · Y triplex. A $d(AG)_5$ oligomer did not migrate with the $d(AG \cdot CT)_{45}$-containing plasmid, i.e., a R · Y · R triplex was not formed. The Y · R · Y triplex was formed with both superhelical and linear plasmids, but only at pH 4.5 to 6. The association with the third strand prevented formation of the internal triplex described in the next section.

Baran and co-workers[19] found evidence for the formation of a triplex structure when they studied the amplified replication initiated at the polyoma virus origin of replication in a polyoma-transformed rat cell line. Replication terminated at a cellular $d(AG \cdot CT)_{27}$ segment located downstream from the polyoma origin of replication. The replication stopped in the middle of that segment.[461] This observation was extended to SV40-origin initiated replication when cloned near the same sequence in plasmids. It is proposed that as soon as newly replicated DNA is extended to the middle of the oligopyrimidine template, the yet nonreplicated half of the oligopyrimidine segment folds back, to join the just replicated AG · CT duplex as a third strand, blocking further progress of the replication machinery.[259] The phenomenon is acid dependent, as expected for a triplex.

The application of two-dimensional NMR analysis to the determination of structural details of nucleic acids in solution was described in Section II.F. This technique confirmed triplex formation and the formation of Hoogsteen-type hydrogen bonds. Rajagopal and Feigon[402,403] studied the one-dimensional PMR of a 1:2.5 mixture $d(G-A)_4$ and $d(T-C)_4$ octamers at pH 5.5, in $H_2O$. Highly downshifted peaks assigned to N3 imino protons of thymine and cytosine were evident. The two-dimensional pattern shows crosspeaks between these resonances and the H8 resonances of the guanines, as expected upon Hoogsteen-type H bond formation (Figure 12). Sugar proton distances correspond to the 3′ endo conformation, as in A-form DNA structures. De los Santos et al.[106] studied a triplex formed from the 11-base d(CTCCCTCTTCC) and its purine complement, as well as from a shorter 7-base sequence. Downshifted resonances (14.5 to 16 ppm) of some of the imino protons are again observed. Crosspeaks corresponding to the shifted imino peaks spectrum permitted assignment of the resonances to 9 out of 11 GCG and TAT triplets potentially formed. Resonances for most cytosine and adenine amino protons (at N4 and N6), as well as for adenine H8, could also be identified. Adenine H8 protons interacted with the imino protons of the thymines of the third strand; only the protonated H3 of cytidines gave a resonance too broad to be assigned. The observed interactions confirm the Hoogsteen base pairing between the purines and the third-strand pyrimidines. A triplex model also explains the NMR spectrum obtained when a The d(TCTCTC-TTT-GAGAGA) molecule shows a B-type NMR spectrum at neutral pH, due to hairpin formation. At lower pH a spectrum consistent with intrastrand triplex is obtained,[335] formed by monomer-hairpin association. The NMR studies thus give strong support to the formation of triplexes by linear R · Y oligonucleotides in mild acid solutions.

### 3. Supercoiled Plasmids: H-Form DNA

The explanation that sensitivity to single-strand nucleases is due to transition to a triplex state was raised by Lyamichev et al.[294] (see also Reference 82), based on a previous proposal by Lee et al.[268] It is proposed that the sensitive $d(AG \cdot CT)_n$ insert opens and that one half of the pyrimidine strand loops back to join the major

groove of the other half, to form a triple-stranded structure, as shown in Figure 14a and b. The complementary purine half strand remains in the free single-stranded state. The direction of the joining pyrimidine strand is parallel to the purine strand and antiparallel to the other pyrimidine strand, as in the linear triple helices just described (and in contrast to the rec A catalyzed three-stranded array, where the two pyrimidine strands must run in parallel, Section IV.D). The bases of the third strand pair with the purine strand via two Hoogsteen-type hydrogen bonds, one bond between N7 of the purine and N3 of the pyrimidine and the other bond between N6 of adenine and O4 of thymine, in a TAT triplet, or between O6 of guanine and N4 of cytosine, in a C · GC · $CH^{++}$ triplet (Figure 12). Nitrogen 3 of cytosine must be protonated to form the H bond, i.e., each CGC triplet carries a positive charge. The participation of the proton can explain the acid dependence of SSN sensitivity (Figure 15). Examination of the partly unfolded triplex in Figure 13b demonstrated that H-DNA is a paranemic structure. Two isomers are possible, according to whether the 3′ half or the 5′ half of the purine strand remains unpaired. These isomers are termed H-y5 and H-y3 by Htun and Dahlberg[206] (see Figure 14). Two analogous structures in which purine-purine-pyrimidine (R · R · Y) triplets are formed are also possible[244] and are described further on.

Three conditions must be met for the triplex to form:

1. The sequence must possess a mirror symmetry, i.e., it must read, from the middle outward, the same in both directions. Poly d(AG · TC) is a mirror symmetric sequence; so is poly d(G · C). Mirkin et al.[331] showed that many of the then known SSN-sensitive genes have a region of mirror symmetry. The length of the symmetric region is limited; only in one case does the length of the mirror symmetry exceed one turn of a triplex helix (10 bp). An inspection of the sequences listed in Table 1 shows that extended mirror symmetry is far from being a general property of SSN-sensitive sequences. More rigorous tests of the mirror symmetry requirement are described later.
2. The strands must be unwound to the length of the $(AG \cdot CT)_n$ segment for the triplex to be formed. This was verified in the orig-

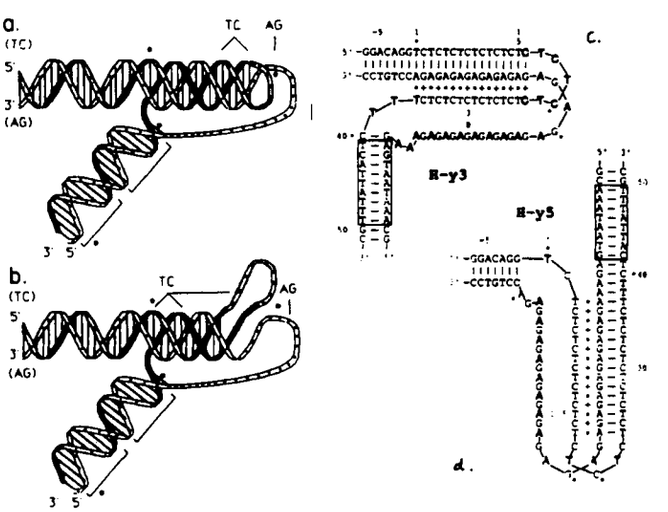


**FIGURE 14.** The structure of H-DNA. (a) The model. The third strand is drawn in black. The detailed structures of the single-stranded part and of the loop are unknown. (b) A partly unwound triplex, demonstrating the nonintertwined nature of the triplex. (c) The H-y3 isomer. (d) the H-y5 isomer. (From Htun, H. and Dahlberg, J. E., *Science*, 241, 1791, 1988. With permission.)

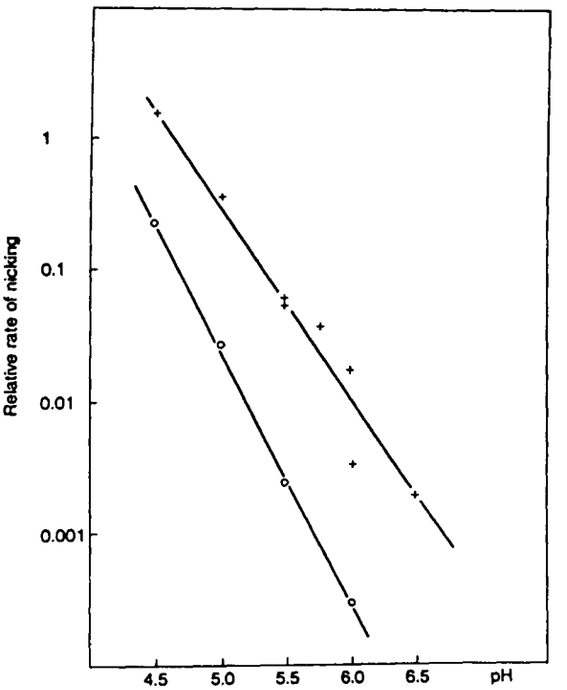


**FIGURE 15.** Effect of pH on rate of cleavage of an $(AG \cdot CT)_n$ insert by S1 nuclease. (0) Cleavage of the plasmid; (+) cleavage of the vector into which the 45-base A-G insert was introduced. Cleavage rates were measured at 23° in 0.1 *M* $Na^+$, acetate or cacodylate, and 0.2 *M* $ZnCl_2$. (From Pulleyblank, D. E., Haniford, D. B., and Morgan, A. R., *Cell*, 42, 271, 1985. With permission.)

inal paper by two-dimensional topoisomer analysis[293] (see Figure 16).

3. The cytosines must be protonated at position N3, to form an H bond with N7. The implied acid dependence was demonstrated by all studies of oligo dTC sequences and by a variety of methods[129,184,293,397,449] (see Table 1, Figures 15 and 16). A problem with assigning the proton dependence to cytosine protonation is raised in Section III.G.5 below.

The energy gain from triplex formation has been estimated recently from the superhelical energy required for the transition on the two-dimensional gel.[297] The transition free energy is pH dependent, and may reach −.42 Kcal/mol/bp at pH 4.2. It becomes positive, i.e., unfavorable, above pH 5. The studies of isolated triplexes described in the previous section[267] substantiate the role of protons in providing the required free energy. The energy of transition consists of a gain due to the formation of triplets and associated stacking, offset by the need to keep half a strand and the center loop single stranded. The measurements of Lyamichev et al.[297] lead to an energy of nucleation (i.e., the length-independent component of the energy) of 18 Kcal/mol; this value determines the minimum length of the insert that enables triplex formation.

An earlier analysis of dependence of σ for d(AG · CT) transition on pH leads Lyamichev et al.[293] to conclude that every second base is protonated in the unwound structure, consistent with protonation of all cytosines. The term H-form DNA was proposed for the triplex structure formed in the plasmids. This term was adopted later by Htun and Dahlberg[205] to describe the *hinge*-like shape of the internal triplex structure. Their suggestion is that triplex-forming segments serve as a hinge point where DNA can make a sharp turn, as part of chromatin packaging and unpackaging processes. Sharp turns, or hinges, have indeed been observed by EM.[468] This is in line with the notion that nuclear acidification plays a role in DNA packaging. It should, however, be stressed (see Figure 16b) that almost double the natural superhelicity is required to extrude the H form at physiological pH (in bare DNA!).

### 4. The Evidence: Topoisomer Analysis

The initial evidence in favor of triplex formation came from two-dimensional topoisomer analysis[293] (a temporary model of a $CH^+$-C self complex was abandoned in a subsequent paper[294] in favor of the triplex). The transition of B-form DNA into a triplex requires unwinding of the duplex (condition 2 above). In a closed plasmid, unwinding of primary turns of a DNA segment occurs at the expense of negative superhelical turns; this is manifested experimentally as a discontinuity of the topoisomer pattern, or "jump", in a two-dimensional gel, as described in Section II.C. Lyamichev et al.[293] demonstrated that a jump does occur when a 3160-base pUC19 plasmid containing a $d(AG \cdot CT)_{16}$ segment from a histone sea urchin gene is examined by two-di-

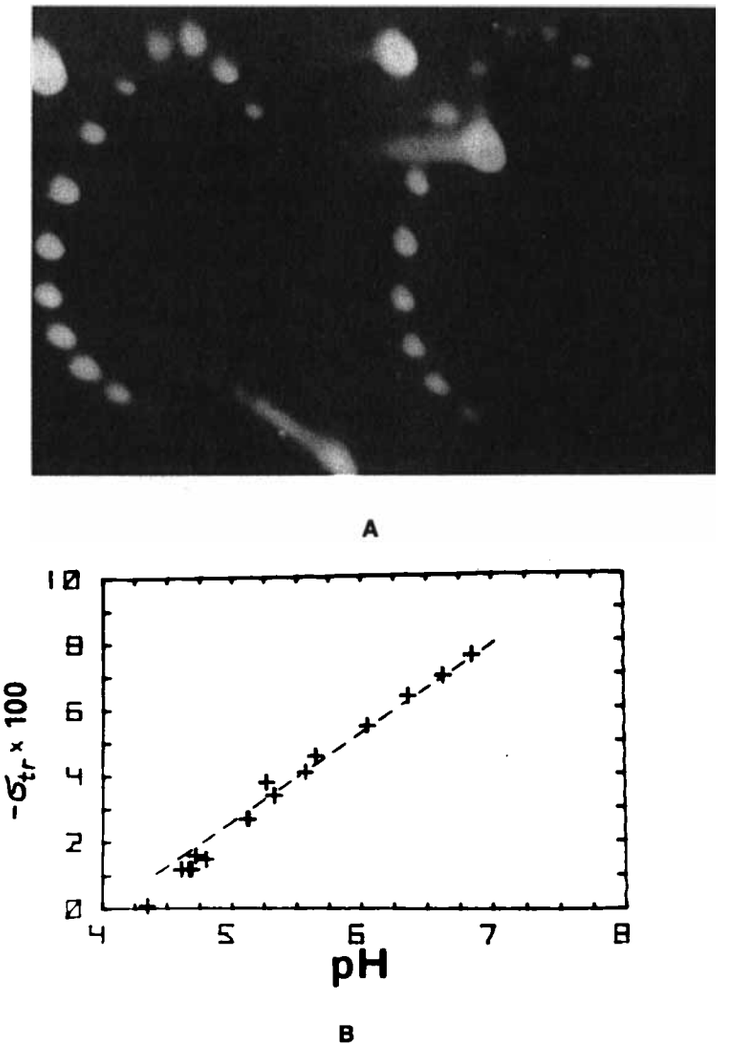


**FIGURE 16.** (A) Two-dimensional topoisomer analysis of a d(AG · CT)$_{16}$-containing plasmid, run at pH 5.2 (100 m*M* Na citrate). (B) pH dependence of transition of a series of patterns as in (A). The average superhelix density of the topoisomers at transition midpoint, $\sigma_{tr}$, is taken as a measure of transition into the proposed triplex state. For patterns at other pH values, see original paper. (From Lyamichev, V. I., Mirkin, S. M., and Frank-Kamenetskii, M. D., *J. Biomol. Struct. Dyn.*, 3, 327, 1985. With permission.)

mensional electrophoresis at pH 5.1 (Figure 16a). The discontinuity is observed between topoisomers 9 and 10, i.e., at a linking deficiency of 10 and at $\sigma = -10*10.5/3160 = 0.033$. Topoisomer 10 migrates in the first dimension like topoisomer 6.7, indicating that 3.3 turns are opened, corresponding nicely to a complete unwinding of the 32 base pairs of the d(AG · CT)$_{16}$ insert. Treatment of the topoisomer mixture with S1 prior to electrophoresis showed that topoisomers 10 and beyond are the only sensitive ones. (Figure 16a, pattern on right). This indicates that the bulk, and not a minority, of the higher topoisomers are in the sensitive state. The condition that

the strands be unwound for triplex formation is thus experimentally met by the oligo d(AT · CT) insert. It should be remembered, however, that all other unwound structures discussed in this article are topologically equivalent and will show a similar transition at a proper pH value. The only structure that $d(AG \cdot CT)_n$ cannot form is a cruciform, for lack of dyad symmetry.

Htun and Dahlberg[206] examined the two-dimensional pattern of plasmids with $d(AG \cdot TC)_n$ inserts of various lengths, run in the first direction in the presence of 1 m*M* magnesium acetate at pH 5. Several transitions were observed, depending on the length of the insert. One transition was assigned to a region present also in the vector; a second jump was attributed to triplex formation with the 5′ half of the purine strand unpaired (H-y3); a third jump, mainly with longer inserts at lower supercoiling, was assigned to a transition into the isomeric triplex, in which the 3′ half of the purine strand is the free one (H-y5). The identification of the free half was based on whether DEPC-modified bases (Section II.B.2; Figure 3) map to the 3′ or 5′ half of the purine strand. The non-even modification of successive pairs of bases indicates some structure for the single-stranded region.[207] Seven $d(AG \cdot CT)_n$ inserts, varying in length from 18 to 60 bases, were studied. A plot of the number of supercoils opened (deduced from the topoisomer patterns) vs. the length of the insert was linear (Figure 17). The slope of the straight line corresponded to one superturn lost per 11 base pairs opened, as expected for a topologically unwound, paranemic structure. Additional transitions attributed to partial triplex formation in the longer inserts were observed under some conditions. The sharpness of the spots in the presence of magnesium is interpreted by Htun and Dahlberg as indicating that the transition between the different isomers is slow. This indicates that acid, and Mg, may play a *kinetic* role in determining the pattern observed. A detailed mechanistic model of the transition processes is offered.[206,207]

The unwinding of several other SSN-sensitive DNA segments has been established by two-dimensional topoisomer analysis, including the murine C-α immunoglobulin region $(AGAGG)_6AG$[86] and the Herpes simplex DR2 SSN-sensitive region, of composition GATCCC CGCTCCTCCCCC CGCC CGCTCCTCCCCC CGATC;[533] notice the 16-bp direct repeat and the absence of a substantial mirror repeat.

### 5. The Evidence: Nuclease Sensitivity

The triplex structure, as shown in Figure 14, is expected to be digested by SSN primarily along the free half of the purine strand. The pyrimidine strand is expected to be protected, except for the very middle, where it loops back into the structure, and for one end, where the pyrimidine strand leaves the triplex and rejoins its complementary strand. This pattern of SSN sensitivity has been observed for the pUC 19 plasmid containing the $d(AG \cdot CT)_{16}$ insert.[293] The purine strand was shown to be S1 sensitive along its left, 5′ half, while 8 further purines adjacent to the alternating tract, with no mirror-symmetric counterparts, remained insensitive; also S1 sensitive were 19 purines beyond the 3′ end of the $(AG)_n$ region. Another example concerns the $d(CT)_n$ tract in the 5′ upstream region of a rabbit globin gene, which is S1 sensitive in the middle and at one end of the tract.[305] A further example is a homopyrimidine insert from the hsp 26 gene of *D. melanogaster*.[449] More examples of sensitive $d(AG \cdot CT)_n$ inserts are listed in Table 1. A clear demonstration of the cleavage pattern expected for a triplex was obtained with 7 synthetic R · Y inserts with mirror symmetry[184] (see next section).

The SSN sensitivity of oligo d(AG · CT) is expected to increase with decreasing pH, condition 3 above. The expected change with pH was indeed observed with a $d(AG \cdot TC)_{45}$ insert,[397] as shown in Figure 15. The linear dependence on acidity was confirmed with endonuclease P1, which is much less pH sensitive than nuclease S1.[184] We have also seen that the acid dependence of the topoisomer pattern is as expected for triplex formation (Figures 15 and 16b). However, there are two difficulties with assigning the effect of acid dependence to cytosine protonation at N3. While the pK of free cytosine is close to 4.3, the pK for protonation of N3 in poly d(CT) was found to be close to 6.3 (see Section III.F.2)[59] The midtransition for the formation of linear triplexes was indeed at pH 7.0 (Figure 13). The pH dependence in supercoiled plasmids

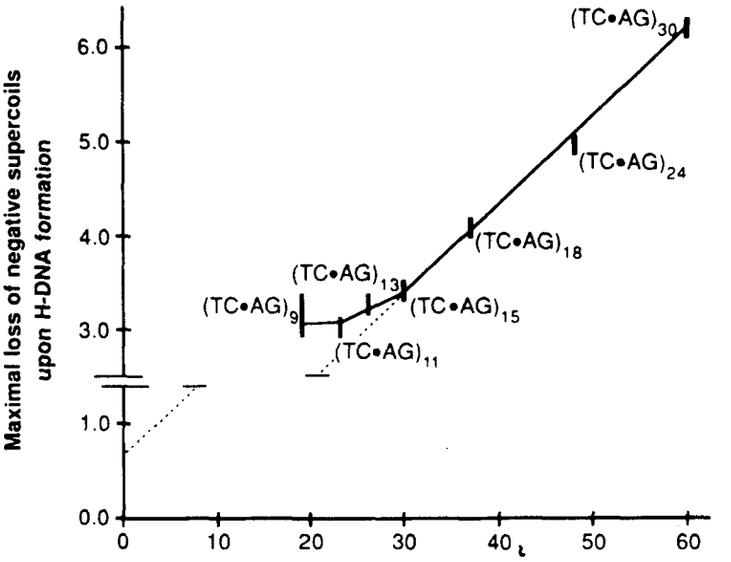


**FIGURE 17.** Number of supercoils opened (and helix turns unwound) as a function of length of a d(AG · CT)$_n$ insert. The slope of the straight line for n >13 shows that one superturn is unwound for every 11 inserted base pairs, in line with complete unwinding but not with left-handed rewinding of the insert. (From Htun, H. and Dahlberg, J. E., *Science*, 243, 1571, 1989. With permission.)

should therefore level off below that value, i.e., below pH 6.3. There is no sign of such a leveling off in Figures 15 and 16a or in most other pH dependence studies. One explanation can be that the pK of C-N3 is very different when the cytosine is within a triplex. An alternative explanation could be adenine or phosphate group protonation that is far from saturated above pH 4.5 (see Reference 424). Neutralization of phosphate $O^-$ can assist the rotations of the DNA backbone necessary for a transition into the paranemic conformation. The role of acid might thus be a kinetic rather than an equilibrium one: to enhance the rate of transition into the sensitive form, which is not necessarily a triplex.

A second difficulty is connected with the slope of the line correlating the rate of cleavage by S1 with pH shown in Figure 15.[397] The slope of this line for the insert is 1.45; this means that 1 or 2 protons at most are involved. This also favors a kinetic effect, because an equilibrium between a multiprotonated structure like the triplex and unprotonated B DNA should show a high order of protonation, i.e., a much sharper dependence on $H^+$ concentration. Unless the very first protonation (not necessarily on C) is the rate-limiting one (or, less likely, many independent foci of transition exist), we may assume that acid plays a kinetic rather than an equilibrium role. The alternative, that phosphate rather than cytidine protonation is involved in the transition to a nonprotonated paranemic structure cannot be ruled out from the pH dependence of the reactions studied and ought to be considered further.

## 6. The Evidence: Chemical Probes

Further evidence consistent with triplex formation by oligo (AG · CT)$_n$ inserts in supercoiled plasmids comes from experiments with conformation specific reagents (Section II.B). Both DMS and DEPC react strongly with the adenines in the 5′ half of an (AG · CT)$_n$ insert, but not in the 3′ half, which is presumably a part of the triple helix.[220] The reaction of a second insert with DEPC results in a similar pattern, most pronounced at pH 5 to 6 and diminishing toward pH 8.[205] At pH 4.5 a reaction on the 3′ half of the AG strand becomes evident, particularly at high

superhelicities, suggesting that a 5′ triplex (H-y3) can also be formed[206] (see Figure 3).

As for pyrimidine reagents, $OsO_4$, a reagent for thymine, reacts weakly with B-form DNA but strongly with thymines in B-Z junctions and ss DNA[221,498] (see Section II.C). $OsO_4$ reacts strongly at the center of the $(TC)_n$ insert, as expected for a triplex. So do methoxyl and hydroxylamine, reagents toward cytidine.[205,220] A reaction at the 3′ end, where a duplex begins to reform, is also observed. The extent of reaction diminishes with increasing pH, as expected for a protonated structure.

The SSN sensitivity and chemical reactivity of a series of homopurine · homopyrimidine (R · Y) inserts, all having both mirror and direct repeat symmetries, was studied by Hanvey et al.[184] The oligodeoxynucleotides studied were $d(G \cdot C)_{19}$, $d(AGG \cdot CCT)_8$, $d(AG \cdot CT)_{12}$, $d(AAGG \cdot CCTT)_6$, $d(AAG \cdot CTT)_8$, $d(AAAG \cdot CTTT)_6$, and $d(A \cdot T)_{20}$. In most cases the 5′ half of the A,G strand, rather than the 3′ half, was sensitive to S1 and P1 nucleases, and the pyrimidine strand was sensitive around the center, as expected for the triplex model. All inserts except oligo dA · dT were reactive toward BAA when negatively supercoiled, indicating that the reactive region is indeed unwound. Only sites in the middle of the pyrimidine strand reacted with BAA to a certain extent, as expected for the protonated triplex model. The reaction with BAA was pH dependent, but reactivity was still apparent at pH 7.6. Glyoxal also reacts with oligo d(AG · CT), mainly along one half of the purine strand.[498] Oligo dA · dT was not attacked by S1 or BAA, while the 75% A insert was sensitive like the rest. More recently, $(dA \cdot dT)_{69}$, but not $(dA \cdot dT)_{33}$, insert was found to react with DEPC and $OsO_4$ structure as expected for the H-DNA triplex.[139a] For $d(G \cdot C)_{19}$, see the next section. $OsO_4$ attacked the center of the pyrimidine strand, as expected, in all inserts except for $d(A \cdot T)_n$. DEPC attacked (less clearly) the 5′ half of the purine strands. DMS attacked the 3′ half of the purine strand. The 5′ half in which N7 of G ought to be buried in a triplex was at least partly protected against this reagent, again not in linearized plasmids, or at pH 7.6, and not in poly d(A · T) (see, however, DMS protection experiments in Section II.B.4). The unwinding of most inserts was also demonstrated by two-dimensional analysis. The extensive study of Hanvey et al.[184,185] strengthens considerably the case for the formation of H-DNA by mirror symmetric DNA sequences in acid conditions. More recently, triplex formation by two $d(AT)_7$ segments interrupted by a 14-base mainly d(A-T) was shown[372] to be stable at neutral pH.

It was mentioned that all of the 7 R · Y inserts possess mirror symmetry. What happens when mirror symmetry is distorted? Mirkin et al.[331] studied four inserts of composition AA-GAGAAXGGGGTA|TAGGGGYAAGAGAA by two-dimensional topoisomer analysis. When X and Y were equivalent, either both G or both A, i.e., when mirror symmetry was perfect, the expected two-dimensional transition was manifested. When symmetry was perturbed, then either there was no transition (X = G; Y = A), or much higher superhelical torsion was required (Y = G; X = A). This demonstrates the importance of mirror symmetry for the unwinding transition. For further details see Belotserkovskii et al.[29] In a similar study, Hanvey et al.[185] examined the reactivity of an $(AAG)_8$ insert, in which the A · T pair at position 21 was converted into a T · A pair. The center part of the pyrimidine strand remained active toward $OsO_4$, indicating that a triplex was still formed. Moreover, the T in position 4, which is mirror symmetric to 21, is now unable to make a TAT triplet (it has now to Hoogsteen pair with a T instead of an A). This T becomes indeed reactive toward $OsO_4$. The recognition by T-4 of the change in its mirror symmetric partner, A-21, can be regarded as a diagnostic for triple helix formation. In a further experiment by Voloshin et al.[501] an A · T pair in a mirror-symmetric insert of 32 bases was replaced by a G · C pair. The insert remained reactive toward DEPC and DMS as in a triplex, but no conclusion on a *trans*-insert interaction could be drawn because the pyrimidine strand was not tested. Glover and Pulleyblank[164] probed with $KMnO_4$ and DEPC two plasmids with two adjacent mirror-symmetric tracts arranged in a dyad symmetric manner: $d(TC)_8 \cdot N_3 \cdot d(AG)_8$ and $d(AG)_8\ N_3\ d(TC)_8$. Either two triplexes or one cruciform can be formed. Both constructs showed altered reactiv-

ity at pH 4.25, but the pattern of at least the second construct was not consistent with H-DNA formation. The potential cruciforms were not formed.

The myc gene contains a S1-sensitive region around base −125 that can form an 11-base-long triplex with one interruption.[232] The bases in that region are reactive toward DEPC, DMS, and $OsO_4$, as expected for the H-DNA structure, including acid dependence. This indicates again that at least one deviation from complete symmetry does not prevent triplex formation. The insert (GAA)9 · TCC · (GAA)4 can form either a perfect 12-base triplex or a 16-base triplex but with 3 interruptions. Reagent studies indicate that the shorter perfect triplex is formed.[186] The less perfect a triplex is, the lower its thermal stability, and the higher the superhelicity required for its formation[186] (see also Reference 87). The effect of loop size has also been investigated.[444]

### 7. Poly dG · dC

Triplex formation was detected also for oligo dG sequences, the sensitive element in the S1-cleaved β-globin region discovered by Larsen and Weintraub[260] (see Reference 351). The chemical reactivity of chloro- and bromoacetaldehyde toward oligo dG was studied thoroughly by Kohwi-Shigematsu and co-workers.[244-246] Several post-reaction cleavage procedures were employed. In all of them, the poly dC strand was the reactive one and this over the 5′ half only (Figure 18). The reaction was pH independent, and was observed only in the presence of $Mg^{2+}$ ions (2 m*M*) This means that in the presence of magnesium half of the *pyrimidine* strand is free, and indicates that a $d(C \cdot G \cdot G)_n$ triplex is formed when $Mg^{2+}$ is present (an H-r3 triplex by the terminology of Htun and Dahlberg). A Hoogsteen base-paired triplet, as depicted in Figure 12c, is proposed. Hydrogen bonds must be bent to a considerable extent to form the suggested C · G · G triplet (bending of H bonds is expected for other Hoogsteen base pairs as well, including those now established by NMR).[402,403]

Hanvey et al.[184] examined poly $d(G \cdot C)_{19}$

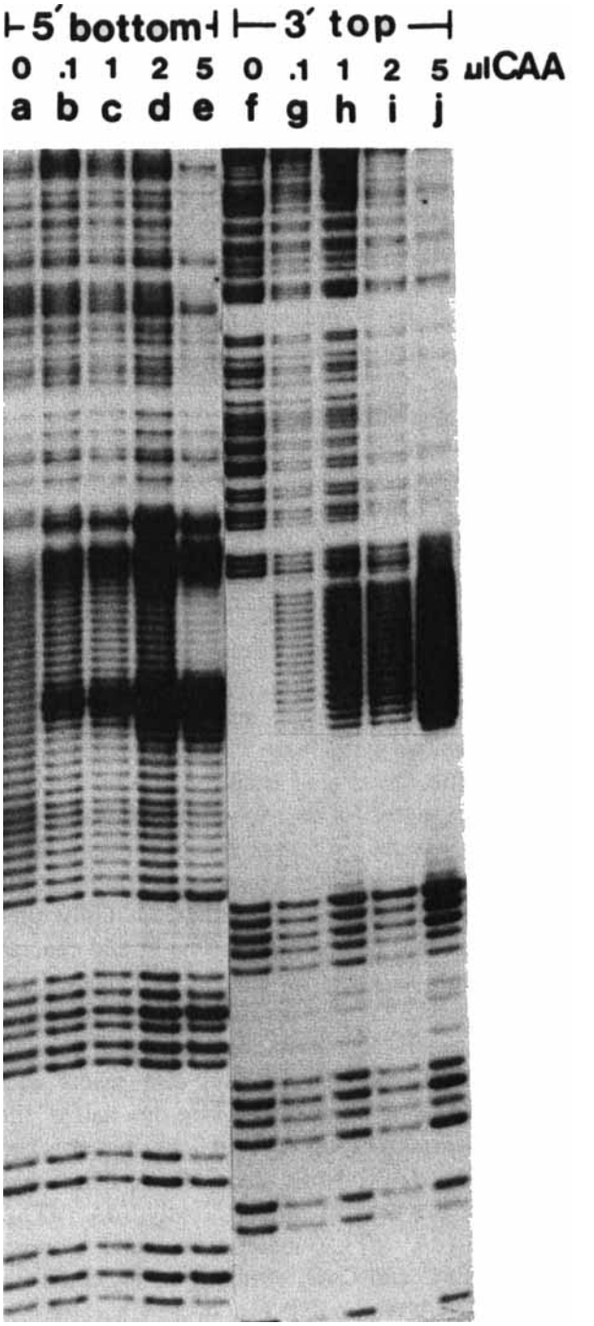


**FIGURE 18.** Modification of an oligo $d(G \cdot C)_{30}$ insert by chloroacetaldehyde. A plasmid containing a $d(G \cdot C)_{30}$ insert was reacted with chloroacetaldehyde, at the indicated concentrations, in 100 μl of Na acetate, pH 5 and 1 m*M* mg acetate. An end-labeled 242-bp Ban I/Hind III DNA fragment was subjected to the Maxam-Gilbert chemical reaction with formic acid, which cleaves chloroacetaldehyde-modified sites, and analyzed on an acrylamide gel. The 30 G can be clearly identified in the left 5 lanes. Enhanced labeling of the central Gs, proposed to reside in the ss loop of a G · C · G (H-r5) triplex, is clearly evident at the higher reagent concentrations. So is the modification of the 3′ half of the cytosine strand. (From Kohwi, Y. and Kohwi-Shigematsu, T., *Proc. Natl. Acad. Sci. U.S.A.*, 85, 3781. With permission.)

with their battery of reagents, and their results agree with the existence of special oligo($G \cdot C)_n$ structures at neutral pH. At pH 5.5 the "usual" H-y3 triplex is formed, cf. Lyamichev et al.[295] The two-dimensional spectrum of oligo $d(G \cdot C)_n$ inserts was recently examined by Panyutin et al.[368] A transition was found in neutral pH when the insert contained 46 GC pairs, but not when 37 or less consecutive GC pairs were present. The individual single strands, dC and dG, are suggested to form a self-associated structure. Following modification by DMS, this structure shows a periodic pattern (rather than protection), as if single-stranded $dG_{46}$ consists of 3 "fingers" rather than of one continuous triplex-forming strand. Panyutin et al.[368] hesitate therefore to suggest $C \cdot G \cdot G$-type triplex formation with the $dC_n$ strand and prefer the general term "G structure" for the observed structure. This is discussed further in the following section on quadruplexes (III.H).

In summary, the formation of a triplex structure (H-form DNA) by oligopurine inserts can be considered to be well supported in cases where all three conditions, mirror symmetry, topological unwinding, and acid dependence, are met. Slight deviations, such as the inversion of an $A \cdot T$, can be tolerated. Triplex formation cannot, however, be the explanation in cases where mirror symmetry is absent or limited. One clear example to SSN sensitivity without mirror symmetry is in the intron region of the chicken β-globin gene; no more than 5 mirror-symmetric bases can be discerned in the sensitive 26/29 R tract (see Table 1).[516] The case of the sensitive Herpes simplex region has been mentioned above.[533] In these cases the role of acid can be kinetic, possibly via phosphate rather than base protonation. Even when mirror symmetry is present, paranemic structures other than triplex are not necessarily excluded.

Bernues et al.[31] have reported that a $d(AG \cdot CT)_{22}$ insert in a supercoiled plasmid is SSN and reagent sensitive in a pH-independent manner, when Zn is present. The formation of a $d(C\text{-}T \cdot G\text{-}A \cdot G\text{-}A)_n$ triplex ("H* DNA") is proposed. Cd and Mn cations are also effective.[32] A recent study of the purine-rich, mainly AG, segment of the hsp 26 gene in *Drosophila*[161] showed that $(TC)_n$ segments that react with DEPC in supercoiled plasmids were unreactive in intact nuclei. This observation and the behavior of non-H-forming mutants indicates that H-DNA may not play a role in the control of heat shock gene expression *in vivo*. The physiological significance of H-DNA formation is thus unclear at present and must await further studies, including studies of $R \cdot Y$ inserts in their natural environment,[552] complemented by studies that include oligopurines that lack any particular sequence symmetries.

## H. Quadruplex DNA

An association between backfolded single stands of DNA has been proposed for special repetitive G-rich sequences identified at chromosome ends (telomers).[192] A detailed structural model, in which four oligo guanylate strands form a quadruplex structure, has been proposed recently[473] (see also References 143 and 297). In this structure, four guanosines, alternating between the syn and anti conformation, form a quadruplet linked by Hoogsteen-like H bonds, with a considerable deviation from linearity. The structure is supported by reactivity of oligo G inserts to chemical probes[367] and by melting behavior.[266] The quadruplex structure is topologically unwound, but is formed from four single-stranded G-rich tracts, without the participation of the complementary C-rich strands. Cytidine protonation is thus not involved.

## I. RNA and Proteins in DNA Unwinding

### 1. RNA

Evidence has been accumulating that in the intact cell, and possibly in plasmids, the unwound state of DNA can be stabilized by interaction with complementary RNA. This is energetically plausible, because no net formation of single strands takes place in the reaction:

$$\text{DNA.DNA} + \text{ss RNA} \rightleftharpoons \text{DNA.RNA} + \text{ss DNA}$$

except possibly at the junctions, so that the en-

ergy required at physiological conditions is, on first approximation, very small. There are also no symmetry requirements for association with a complementary RNA strand. The process requires, nevertheless, the separation of the two DNA strands. In a closed DNA circle, this also requires negative superhelicity. Evidence for RNA-stabilized strand separation was first found when RNA polymerase and other transcription proteins were denatured in active transcription complexes. A persistent hybrid between RNA and closed superhelical DNA (but not with relaxed DNA) was detected.[78,411] A specific RNAse H activity might be needed to initiate displacement of nascent RNA from the DNA template in a mammalian *in vitro* pol II transcription system, indicating a prefered stability of the RNA · DNA hybrid.[224]

An analogous DNA · DNA hybrid was found to be formed between superhelical bacteriophage DNA, and single-stranded complementary strand. This ssDNA — closed dsDNA — complex was studied in detail by Wiegand et al.[531] The complex was found to be sensitive to S1 nuclease and appears in the electron microscope as distinct displacement loops (D loops). In a closed circle, D loop formation was shown to require negative superhelicity.[400] A RNA · DNA D loop (R loop) is schematically shown in Figure 8h. D loop formation, whether of the DNA · DNA type or the RNA · DNA type, requires strand separation and therefore implies the presence of an open, unwound DNA region.

The formation of RNA · DNA D loops during transcription is well known, and was also established in two prokaryotic replication systems: Baker and co-workers[16] demonstrated that D loops are formed during the prepriming steps of *E. coli* replication at ori C (see Section IV.A). The formation of the prepriming complex is preceded by a transcription event that is inhibited by rifampycin. This is not the priming event (which involves a primase activity not inhibited by rifampycin), but rather an event that helps the dnaA protein to form an unwound DNA bubble (open complex). This unwound region serves as an entry point for dnaB protein, a helicase, for the next step of the initiation process. The open region, in a purified enzyme system, is up to 250 bp long, is nuclease P1 sensitive, and is also sensitive to RNAse H, which destroys RNA hybridized to DNA.[18,51] Baker et al. therefore suggest that DNA is separated into a single strand plus a DNA · RNA strand, i.e., a R loop is formed. The RNA must be complementary, and is most probably formed by the early transcription event. The capacity of the R loop-forming region to unwind was confirmed recently by two-dimensional topoisomer analysis.[252] Certain mutative changes that impair replication also abrogate the transition on the two-dimensional gel.

Participation of RNA was observed also during the replication of Col E1 and derived plasmids.[330] Synthesis of a replication essential RNA, RNA II, starts at position 3088 of pBR322, 555 bases from the ori, between the two cruciform-forming sequences of the plasmid. A SSB (Section I.D) is required, and is suggested to stabilize a R loop intermediately formed (see Reference 358). The open region is stabilized further by interaction between part of RNA I and complementary RNA II[478] (see Section IV.A.2). The major SSN-sensitive region of pBR322, located somewhat further upstream (3150 to 3330), is apparently not connected with R loop formation, because according to Kowalski et al.[254] this region can be deleted and replication of the plasmid can still proceed, indicating a nonessential role of the SSN-sensitive region. A secondary SSN-sensitive site may be involved.

The possible role of RNA in unwinding finds support in an intriguing observation with a mammalian gene. In a detailed study of the SSN sensitivity of a myc gene promoter-containing plasmid, it was observed that sensitivity to S1 nuclease diminishes when the CsCl-purified plasmid is further purified by passage through a Sepharose 6B column.[39] Sensitivity is restored when a lower-molecular-weight fraction from the column is re-added to the DNA. That fraction absorbs at 260 nm and is destroyed by RNAse and alkali, but not by proteases. The authors suggest that association with a low-molecular-weight RNA is responsible for conferring SSN sensitivity on the specific myc gene region. It is proposed that either a R loop or a triplex with RNA is formed.[39] No data are brought to indicate whether the low-molecular-weight RNA is complementary or not. More recently, a ri-

bonucleoprotein factor, which binds to a S1-sensitive element around base −125 upstream of the myc gene, was identified in nuclear extracts of c-*myc*-expressing cells. This factor was sensitive both to RNAse H and to proteinase K.[104,232] The possible stabilization of DNA in the unwound state by RNA is a possibility that ought to be examined in additional systems.

### 2. Protein Factors

Paranemic forms of DNA can also be stabilized by interaction with nuclear proteins. One way to achieve this is by interaction with a single-stranded part of the structure. While most known proteins, like histones and prokaryotic regulatory proteins, prefer dsDNA, the stabilization of single strands, e g., by SSB proteins, is well known. SSB proteins are a necessary component in most replication systems and support cellular unwinding by helicases, as discussed in Section I.D. A number of other proteins with roles in template processing, including DNA and RNA polymerases,[233] bind to ssDNA no less well than to dsDNA. HMG 1 and 2, which play an important role in active chromatin, have been shown to prefer single strands of DNA.[178,439] All these proteins are likely to pull an unwinding equilibrium in favor of the unwound state. In the intact cell, association with proteins may well be a major factor in maintaining important DNA regions in an open, unwound state. The dnaA protein appears to fill this role in the *E. coli* replication system.[51]

There are now several examples of specific proteins binding to SSN-sensitive DNA regions. The nuclear RNP factor found to associate with the myc gene SSN-sensitive region[104] has already been mentioned. Two protein factors binding specifically to the SSN-sensitive region of the human epidermal growth factor (EGF) receptor gene have been identified.[218] The promoter of this gene contains a SSN-sensitive region with a direct repeat of 11 pyrimidine bases (10-base mirror). The two protein factors were identified by gel retardation and DNAse I assays, when a DNA fragment containing the Y-rich region was incubated with nuclear extracts from human cells. One factor turned out to be transcription factor sp1. Most recently, a nuclear factor from HeLa cell nuclei, which binds to the sensitive region of the insulin receptor gene, a R · Y tract, was reported.[201] The extent of unwinding of the DNA fragment employed to identify the protein factors has not yet been established. Transient infections with a construct in which the promoter region is coupled to a reporter CAT gene showed that the R · Y tract must be intact for the CAT gene to be expressed. It remains to be established whether the nuclear factor actually converts the R · Y region into an unwound state.

Felsenfeld and collaborators have reported the purification, from adult chicken erythrocytes, of a protein BGP1, which binds preferentially to poly dG tracts.[83] At least 7 dG residues are required. The question of whether BGP1 prefers B DNA, H-DNA, or possibly ssDNA is being studied. Transcription factors like the GAGA binding protein[34] may also stabilize DNA in an unwound form by binding preferentially to ssDNA. The stabilized structure can be single stranded, but may also be of any of the paranemic structures mentioned, depending on the nature of the complex formed. Other regular paranemic DNA structures formed by interaction with specific proteins can be imagined, the structure formed with the RecA protein during nonspecific recombination being an example. Interaction with nuclear factors may play an important role in DNA unwinding during gene expression in the intact cell.

## J. Conclusions

Most of the recently obtained evidence concerning the state of DNA in R · Y rich, SSN-sensitive plasmids favors the formation of the triple-stranded H-DNA structure. The conditions that favor this structure in superhelical plasmids include:

1. A high purine · pyrimidine composition
2. Perfect or near-perfect mirror symmetry
3. Susceptibility to single-strand-specific enzymes and reagents, concentrated in one half of one strand and in the center region of the other

4. Maximal reaction of pH 4 to 5, diminishing toward neutral pH values
5. Two-dimensional topoisomer patterns that indicate a change in primary winding of the expected magnitude

The existence of linear triplexes of the T · A · T and C · G · C type, established by a range of physical and chemical techniques, supports the formation of H-DNA. The proposed Hoogsteen-type base association within the proposed triplets has now been validated by NMR analysis. Triplex formation is further supported by single-base substitutions, which interfere with the expected reactivity toward enzymatic and chemical probes.

There are nevertheless several considerations that point toward the existence of unwound structures other than H-DNA:

1. Many single-strand-specific regions do not possess the required mirror symmetric base sequence
2. The pK established for poly dC suggests that conversion to the protonated form should already by completed near pH 6, rather than increase beyond pH 4
3. Many sequences are SSN sensitive in a pH-independent manner
4. Most prokaryotic SSN-sensitive sequences are A,T rich rather than R · Y asymmetric
5. There is evidence that association with RNA or other cellular components can be important in stabilizing the unwound state

These considerations imply that unwound structures other than the triplex can be formed and may participate as intermediates during DNA unwinding processes. Systematic studies of the type performed by the groups of Wells and Lyamichev on near-perfect mirror-symmetric sequences ought, therefore, to be extended to less symmetric homopurine sequences.

Among the alternative paranemic structures considered, the dyad symmetric cruciforms (Figure 8c) cannot be formed by R · Y sequences; cruciforms are possible in A,T-rich sequences, the major SSN-sensitive motif in prokaryotic cells. Slippage structures (Figure 8b) are so far not favored by the SSN and reagent cleavage patterns obtained, even where a sufficiently long direct repeat is present. Alternation of left- and right-handed sections (Figure 8e) will be a realistic notion only in case the range of Z-forming motifs, presently mainly $d(G\text{-}C \cdot G\text{-}C)_n$ and $d(A\text{-}C \cdot G\text{-}T)_n$, is considerably broadened.

Two alternatives for paranemic DNA remain when a nonsymmetric DNA insert undergoes an unwinding transition. One alternative is that of the paranemic duplex structures, whether Watson-Crisk (Figure 8d) or Hoogsteen paired. This alternative is particularly attractive under neutral pH conditions. The other alternative is complete separation into single strands (Figure 8a), the original explanation of Weintraub.[522] This alternative is most likely, for the non-symmetric, low-melting A,T-rich sequences in prokaryotic SSN-sensitive regions. Separation into single strands may also be the preferred solution in the intact nucleus, where one or both strands can be stabilized by interaction with protein and RNA components formed (Figure 8h). Experiments that distinguish single strands from duplex and triplex structures, like temperature dependence of the transition,[265] as well as careful kinetic comparisons to single strands, have yet to be made with most SSN-sensitive plasmids studied so far.

All these considerations are relevant to gene expression on the premise that a considerable segment of DNA does indeed unwind during, or in preparation, for template copying processes, including replication and transcription. The available knowledge on the role of unwound DNA forms in these processes is the subject of the last section of this review.

## IV. DNA UNWINDING IN DNA-DIRECTED PROCESSES

This review cannot be completed without relating the information on unwound forms of DNA to the genetic processes in which they can be expected to participate. The techniques of molecular genetics have led in recent years to considerable advances in the understanding of the mechanistic details of replication and transcription processes. Cell-free systems, in which replication and transcription can be studied with purified components, are now available. Studies of

replication have been reviewed recently in a comprehensive volume published at the Cold Spring Harbor Laboratory;[85] see also Stillman.[467] Studies on transcription are reviewed by Reznikoff et al.,[407] Yager and von Hippel,[543] Wasylyk,[517] Hames and Glover,[179] Mitchell and Tjian,[333] Johnson and McKnight,[219] and Parker.[371] In most of these reviews, the problem of how DNA strands do unwind in transcription is barely mentioned. In this section we shall discuss those recent studies that relate specifically to the topological aspects of DNA template utilizing process. Previous discussions can be found in Gellert,[152] Wang,[510] Giaever et al.,[156] Wang and Giaever,[513] Pruss and Drlica,[392] Sternglantz,[463] and Borowiec et al.[40]

## A. Replication

### 1. *E. coli oriC*

The best-understood prokaryotic replication system is, at present, that of *E. coli oriC*, the principal origin of replication of *E. coli*. Most of the information on the role of unwinding in replication has been obtained in this system. Information from plasmid Col E1 (pBR322) and phage lambda systems is also available.

Six sequential stages have been identified for the initiation of DNA replication at *oriC*[16] (see Figure 19):

1. In the first step, 20 to 40 monomers of the DnaA protein bind to a DNA region that contains a 9-mer motif repeated four times, to form an initial complex. This association, detected by DNAse I footprinting, occurs at 20° and is assisted by HU protein and ATP. In bacteriophage lambda, phage protein O replaces dnaA to form an "O some" as an initial phage-replication complex.[120]
2. In the next step, at 38°, ATP is hydrolyzed and a region of about 45 bases of DNA is melted to form an "open" complex. The formation of an unwound region is manifested by the appearance of "D loops" on electron micrographs, and is substantiated by digestion with the SSN P1.[436,551] The melted DNA region is characterized by a 13-bp motif, repeated 3 times.
3. Proteins dnaB and dnaC join to form a prepriming complex. dnaB, which is a helicase, begins to unwind further stretches of the DNA double helix, counterclockwise on the *E. coli* map. In the presence of SSB and DNA gyrase, the unwinding can be very extensive.[17]
4. RNA polymerase is activated and an RNA primer segment is initiated. An analysis of the *oriC*-initiated RNA *in vivo* reveals not a single but a multitude of priming and switching sites in the origin region.[416]
5. DNA polymerase III is added to the initiation complex and the RNA primer is extended by deoxynucleotides to start the leading DNA strand.
6. The RNA primers are excised by DNA polymerase I and lagging DNA strand synthesis commences. Helicase n′, which moves along the template in the 3′ to 5′ direction (now called PriA),[270] joins and facilitates opposite strand synthesis, first as the leading strand of the opposite fork, then as the lagging strand of the initial fork.[261,269] Subsequently, the replication bubble, with the two forks on the opposing ends, grows until all the DNA is replicated.

Most of the components of the replication complex mentioned so far have been shown to be necessary participants in the replication of *oriC*-containing plasmids in a cell-free system.[16] The plasmids used are several thousand bases long, so that their natural superhelicity should provide for the opening of up to 250 to 500 bases. This can be sufficient for the formation of the prepriming complex, and probably for priming and initiation of DNA duplication as well. For further progress of the replication machinery along the DNA circle, participation of a nicking closing activity would eventually be needed, and a requirement for DNA gyrase has indeed been established.[17]

Whether nicking-unwinding-closing occur ahead of the moving replication machinery (replication fork), i.e., on the yet unreplicated DNA, or behind the machinery, i.e., on one or both arms of the newly replicated DNA, or behind the

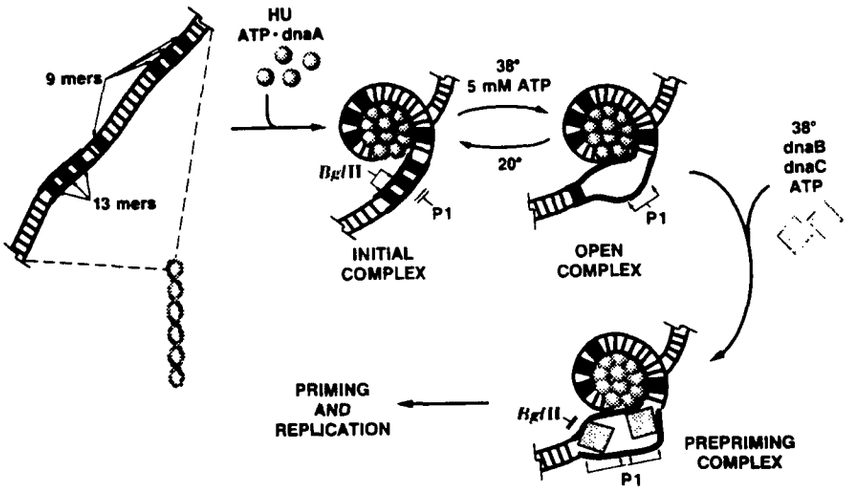


**FIGURE 19.** A scheme for initiation of DNA *E. coli* replication at oriC. Note the unwound, P1-sensitive regions in the open and prepriming complexes. (From Baker, T. A., Bertsch, L. L., Bramhill, D., Sekimizu, K., Wahle, E., Yung, B., and Kornberg, A., *Cancer Cells,* 6, 19, 1988. With permission.)

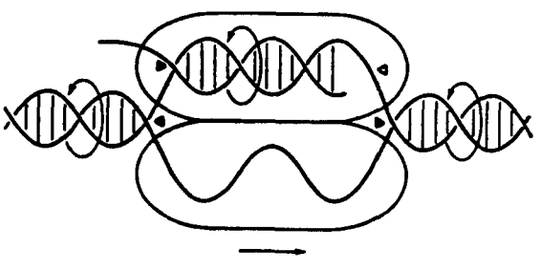

**FIGURE 20.** A topological model for DNA transcription. Swiveling centers are denoted by black arrowheads. The catalytic site is denoted by an empty arrowhead. The transcription bubble proceeds in the direction of the horizontal arrow. The mutual rotation of the two ovally enclosed strands, in the direction indicated by the anticlockwise arrows, causes the unwinding of the two DNA strands on the left and their rewinding on the right. (From Gamper, H. B. and Hearst, J. E., *Cell,* 29, 81, 1982. With permission.)

machinery, i.e., on one or both arms of the newly replicated bubble is yet unknown. The occurrence of intermediary nicked replication bubbles, expected if nicking occurs behind the replication machinery, has not been reported. Such a nick should be particularly evident if one single nick persisted throughout the replication cycle. Nicking-unwinding-closing ahead of the replication fork (with or without positive supercoil accumulation) seems so far a more likely possibility.

More knowledge has become available concerning the formation of the initially unwound complex (steps 1 and 2) from studies of an *oriC*-containing plasmid under defined cell-free conditions. Digestion of the dnaA protein-DNA preinitiation complex with the SSN P1, showed that the open region (initial bubble) is formed in a region of 4 to 5 helix turns, containing three 13-bp repeats of the motif GATCTNTTNTTTT.[51,52] Sensitivity to SSN of a DNA segment in a closed structure implies that this segment is in a topologically unwound state, either single-stranded or readily converted to single strands. The unwinding can be induced by high superhelical density without dnaA protein being present, as verified by mungbean nuclease digestion and by two-dimension topoisomer analysis.[252] An unrelated SSN-sensitive region can replace one of the three 13-bp tracts in conferring mungbean sensitivity.[252] The unwound region contains 47/55 AT residues,[555] but also has a certain degree of R · Y asymmetry (one $R_{10}$ tract). A SSN-sensitive region, 60 bps to the right of the origin, has also been reported for phage lambda.[432] Thus, the S1-sensitive region of lambda contains 18 consecutive R on the I strand and 18/21 R immediately adjacent.[432] Bramhill and Kornberg[52] pointed out the occurrence of 6-

16-base, A,T-rich tracts, repeated several times in the origin regions of 11 different bacterial and bacteriophage genomes. A detailed examination of these repetitive tracts shows that they are as pyrimidine (or purine) rich, as they are A,T rich. Homopurine · homopyrimidine rich is the motif that we have seen in Sections II and III to be the most readily convertible to the unwound state. It is therefore not unlikely that one of the paranemic structures discussed in Section III (Figure 8) will participate in the formation of the replication complex.

The formation of the open preinitiation complex was found to be inhibited by rifampicin, signifying that RNA synthesis is involved. This RNA transcript was found to be initiated several hundred bases to the left of the dnaA binding site.[16] This transcript is not required as primer RNA, because once a prepriming complex is formed, rifampicin (which does not inhibit primase) no longer inhibits subsequent priming; also, a transcript with a blocked 3′ end can initiate DNA replication. Activation takes place even when transcription proceeds away from the priming region.[446] It has been proposed[16] that the RNA stabilizes a melted region by forming a RNA-DNA hybrid with one of the DNA strands, which can be manifested as a R loop, as discussed in Section III.H. In support, RNAse H, which digests RNA in DNA · RNA hybrids, prevents the formation of the preinitiation bubble. The promoter for the RNA transcript need not be the promoter for the adjacent gene; several unrelated promoters can substitute. One finding that is not readily reconciled with a R loop stabilizing role of the RNA transcript is that the transcript need not reach the preinitiation bubble; it may even be directed away from the *oriC* region.[18,446] The RNA stabilized R loop need not, however, be started at the dnaA binding site; it can migrate there when the associating proteins arrive.

A transcription event toward the origin has also been found in phage lambda.[325] That transcription is, however, not required when replication is carried out *in vitro* with purified proteins, provided that excessive amounts of protein HU, a major component of the bacterial chromatin, are not present. It is proposed that cellular HU has a restraining effect on the negative supercoils necessary for unwinding, and the RNA transcript helps to overcome this restraint. Future studies will have to clarify the exact role of the RNA transcript.

In summary, the formation of an unwound region is an essential step in the initiation of bacterial replication. The first unwound region contains both A · T-rich and homopurine/pyrimidine sequence motifs. Of the many structures possible for unwound DNA, evidence for the RNA-stabilized D loop (Figure 8h) is most substantial. The other paranemic structures in Figure 8, including structure 8d may, however, participate when the unwound loop is extended along the bacterial genome.

### 2. Col E1-Related Plasmids (pBR322)

Detailed investigations have revealed that the replication of these plasmids within *E. coli* or *S. typhimurium* includes preinitiation steps similar in some respects to those of host replication (Masukata and Tomizawa,[309] Parada and Marians[370]). In common with *E. coli oriC*, the formation of a priming complex is preceded by the synthesis of an RNA molecule. This RNA, RNA II, starts 555 bases upstream from the origin, at base 2532.[340] In contrast to *oriC*, here the RNA does serve as a primer for leading strand synthesis. RNA II can displace the nontemplate strand, forming a persistent DNA · RNA hybrid.[103,308] This hybrid may be stabilized by specific RNA-DNA interactions.[309] In order to form the RNA · DNA hybrid, the two DNA strands must stay unwound. In pBR322 (4363 bases), 555 bases means that 12% of the bases are unwound, more than the native superhelicity (5%) can provide. Parada and Marians have shown[370] that to initiate replication *in vitro*, DNA gyrase must be active, and in its presence a highly underwound form appears. This form, termed form I*, is detected on agarose gels following RNAse treatment of the initiation complex and deproteinization of the DNA. Form I* migrates even faster than the naturally supercoiled plasmid, accentuating the role of the topological change in the initiation of plasmid DNA replication. In the absence of RNAse H, unwinding can be very extensive, extending over most of the plasmid.[370]

A second RNA molecule, RNA I, is copied

from the DNA strand complementary to the RNA II coding strand, starting 108 bases downstream (at base 2979) so that it is complementary to the first 108 bases of RNA II. These 108 bases contain three potential stem and loop structures, corresponding to minor cruciforms of pBR322 DNA, long known to be S1 sensitive.[278] RNA I serves as an inhibitory factor for plasmid DNA synthesis, by associating with RNA II, not only by hybridization, but also by interaction of the three loops that both RNA species can form.[127] This association, probably helped by a protein, may operate by preventing the necessary unwinding.[478] RNA I and II do not extend into the major mungbean-sensitive region of pBR322 (3210 to 3290).[441] Indeed mutants lacking this sensitive region seem to be viable,[254] so that the significance of this SSN-sensitive region for plasmid unwinding cannot be clearly stated at present.

### 3. Eukaryotes: The SV40 System

Detailed information on the role of unwinding in eukaryotic replication comes from the study of the replication of simian virus 40.[40,85] Today SV40 replication can be studied in cell-free systems containing the origin of replication region of SV40 cloned in a suitable plasmid, purified SV40 T antigen, several host cell proteins, ATP, and deoxynucleotides.[109,482,534] Soluble nuclear extracts from primate cells like HeLa, but not from rodent cells, can also support initiation and propagation of cell-free SV40 DNA replication.

Initial evidence on unwinding events preceding initiation came when purified T antigen was added to origin DNA (−151 to +150) in the presence of SSB and topoisomerase I, and in the absence of dNTP, so that no DNA synthesis could take place.[107] Upon electrophoresis of the deproteinated DNA in agarose gel, a new band, which migrates close to form I DNA, appears. When chloroquine is included in the gel the band is shifted to a much lesser extent then form I DNA, indicating that it is more negatively supercoiled than form I DNA. The DNA form represented by this band is termed U ("unwound")-form DNA.[107] The negative supercoils of the U form can be considered topologically equivalent to unwound DNA in the intact replication complex. Subsequent pulse and chase experiment showed that a subband of the U form is labeled in advance of DNA synthesis and is chased into higher molecular weight bands as replication proceeds.[63] This subband could represent the initial form of DNA in the replicative complex and is termed $U_R$ DNA. U and $U_R$ DNA are proposed to be steady-state precursors of leading strand initiation. A very good correlation was found between the percentage of unwound DNA and the rate of DNA duplication, in 43 point mutants in the origin region.[108]

The origin of replication of SV40 was found to consist of a "core region" of 64 bp (−33 to +31), which is both necessary and sufficient to support virus DNA replication.[421,476] The core region can be divided into 3 subregions or elements. In the center, there is a 27-base dyad symmetric sequence ("inverted repeat", bases −14 to +13), which serves as T antigen binding site II. Tight binding to site II of probably phosphorylated T antigen,[237,390] assisted by ATP or nonhydrolyzable analogs, is required for the initiation of DNA replication. To the left of the inverted repeat (toward the early region) resides the element termed "early palindrome" (−33 to −15). The actual continuous synthesis of leading strand DNA commences in both directions at base −33. To the right of site II, the 17-base A · T element is located (bases 15 to 31).

More detailed information on the degree of unwinding of core-region DNA was obtained by application of chemical probes to plasmids carrying this region in the presence of T antigen and ATP or analog.[43] When DMS was applied, all guanines within the "inverted repeat" (bases −14 to +13) were found to be protected from the reagent. This indicates mainly that T antigen sits firmly on site II. More interesting, flanking cytosines, in particular in the "early palindrome" (bases −33 to −15), became reactive toward DMS even without hydrazine treatment, indicating ready conversion to a single-stranded or topologically related state.[43] The analogous region in polyoma contains 18 successive purines.[193] The "early palindrome" region also reacts with $KMNO_4$ (see Section II.B.6) and can be unwound in the complete absence of either T antigen binding site II and/or the A,T element.[374] On the "late" side of the origin, the 17 thymines on two

strands of the A,T element become very reactive to $KMNO_4$ when T antigen is bound to site II, indicating a highly non-B structure.[43] An alternative explanation could be bending of the A,T element; a protein factor facilitating this bending has been described.[27] Some of the paranemic structures are nevertheless likely to be readily bendable, so that no contradiction is necessary. The formation of the unwound region has been confirmed by topoisomer analysis, using a version in which a minicircle (571 bp) closed by ligase in the presence of T antigen (and topo I) is examined.[414] A highly supertwisted topoisomer ($\Delta Lk = -5$, $\sigma = 0.087$) appears, indicating the unwinding of 50 bp by T antigen (or partial unwinding of a larger sequence). The unwinding thus seems to extend through most of the 64-base core region, the bound T antigen masking the unwinding of the central element from detection by the chemical probe. The appearance of the $-5$ topoisomer can in principle be explained as representing wrapping of 5 turns of DNA around the T antigen, but that would be hard to reconcile with the chemical modification data.

Two auxiliary regions, to the right and left of the ori core region, also contribute to the initiation of virus replication. Small-circle topoisomer analysis has been used to demonstrate the unwinding of these two regions.[177] The region to the left of the core region, bases $-50$ to $-33$, includes part of site I for T antigen binding, and part of a 14/16Y tract (on the noncoding strand). The auxiliary region to the right (bases $+31$ to $+72$) contains three GGCGGGA tracts and is thus highly purine enriched. It should be added that the ori region is SSN sensitive in negatively supercoiled plasmids in the absence of T antigen; in particular the A · T element, which contains $8+3$ consecutive adenines.[130,212] The correlation between high purine-pyrimidine asymmetry and propensity for unwinding seems to hold for both core and auxiliary SV40 origin of replication region.

The unwinding processes discussed so far do not require SSB[43,414] and are limited to the origin region. When SSB (and a topoisomerase) are present, extended unwinding of circular closed SV40 is initiated even in the absence of DNA polymerase or dNTP; EM studies show that a large part of the SV40 DNA circle can be unwound (manifested by the appearance of a large protein-coated D loop).[108,118,431] The extended unwinding is due to the helicase activity of the T antigen, discussed in Section I.D, and when SV40 DNA is in closed circular form, is strongly assisted by topoisomerase action, as well as by SSB.[228]

Three questions come to mind:

1. How often is a circle nicked and closed during unwinding? A nicking event once each turn of the helix (10 bases) would be energy costly, but not inconceivable. The other extreme is a single nick for the unwinding of the entire circle; this requires the presence of a nicked intermediate, not reported so far.
2. Do nicking-closing events occur ahead of or behind the replication fork? As with *E. coli*, no conclusive evidence is available.
3. Does unwinding occur prior to progress of the replication machinery or after it? In the latter case, positive supercoils can be expected to accumulate ahead of the replication forks, as proposed for transcription by the twin supercoiling domains model of Liu and Wang,[286] described in the following section. In that case, the topological problem is reduced to a mechanism for removing at intervals the accumulated positive supercoils. Energy for storing the linking excess in positive supercoils must, however, by supplied.

An interesting piece of information in this respect is that cellular SV40 has its most frequent cutting sites by topoiosomerase II between bases 4070 and 4376.[386] This 306-bp, A,T-rich region (68%) is also a matrix attachment region (MAR: see Section I.B.3), as established by competition with other MAR DNA sequences for nuclear matrix preparations. The DNA of the SV40 MAR region is proposed to interact with matrix-associated topoisomerase II[386] and can in principle serve as a major unwinding initiation site — the long-sought swivel. The unwound region at the origin is 100 helix turns away. Whether the turns unwound by topoisomerase II can migrate physically to the origin as a melted bubble, by axial rotation of the intermediate DNA (or nucleopro-

tein), or be stored first as negative superturns to be released at the origin, or by some other mechanism, will have to be established by future studies.

A second hint that attachment to a structure may be involved is the report by Sekimizu et al.[437] that the *E. coli* dnaA protein has a high affinity to phospholipids. This might anchor and stabilize the unwound complex, but an alternative model, in which a change in phospholipid composition releases the initial complex, has been proposed recently.[355]

### 4. Other Eukaryotic Systems

The relation between SSN-sensitive regions and replication has been studied in yeast autonomous replication sequences (ARS).[487] SSN-sensitive regions have been detected in negatively supercoiled ARS carrying plasmids when digested by mungbean nuclease at neutral pH (Table 1). The sensitive regions are located on the 3′ flank of the H4 ARS and are A,T rich (mainly A), as in *E. coli*. Deletion mutants that prevent replication of the H4 ARS-containing plasmids also eliminate their sensitivity to the nuclease. When an unrelated unwinding promoting sequence (the SSN-sensitive region of pBR322) is inserted into the mutant, both nuclease sensitivity and the ability to replicate are restored, which indicates a strong correlation between the ability of a sequence close to the origin to unwind, and the ability of the ARS to enter DNA replication.[485] The SSN-sensitive region is therefore termed a DNA unwinding element (DUE) by Umek and Kowalsky. The substitution of the DUE by two yeast SSN regions elsewhere on the replicating plasmid did not support the replication of the plasmid.[488]

A DNA region that serves as a genomic origin of replication in human cells has been identified 3′ of a highly amplified dihydrofolate reductase gene. One of these amplified ori regions has been cloned recently and sequenced.[66] The 6157-base sequence contains a number of distinct elements, including two Alu I homologous regions, several elements homologous two yeast ARS, a readily bent region, and relatively long tracts of repetitive sequences, including poly A, poly T, poly AC · TG, alternating (A-T) and (G-C)! tracts. In addition, the ori region includes a very long R · Y tract of 180 bases, interrupted once by a $(CAGA)_4$ tract, and is preceded by $(G-C)_5$ and $(A-C)_{18}$ tracts. The long R · Y tract is sensitive to mungbean nuclease at neutral pH, but less than the $(A-T)_{23}$ tract. An $A_{31}$ tract is also sensitive. An SSN sensitivity of the A-T and R · Y-rich regions makes them likely candidates for initial unwinding during the initiation steps of cellular replication.

In summary, there are presently available quite a number of replication systems in which the various aspects of DNA unwinding can be studied. Techniques to establish the degree of unwinding are also becoming rapidly available. Firm evidence on the participation of unwound regions in replication has already been obtained. More definite information on the structural aspects of the unwound complexes, as well as dynamic information on the sequence of events leading first to unwinding and then to rewinding can be expected in the near future. All this information will be needed before the long-proposed but still unclear concept of replication swivel becomes a molecular reality.

## B. Transcription

The knowledge so far accumulated on how unwinding and rewinding are accomplished during transcription comes from three main directions: (1) studies of the effect of the superhelicity on the transcription of closed plasmids; (2) studies of the topological changes circular templates undergo during transcription; (3) studies on the role of unwinding activities (topoisomerases). These studies were performed most intensively in *E. coli*, but considerable knowledge from model eukaryotic systems, mainly yeast and SV40, has also accumulated.

### 1. The Effect of Supercoiling

A strong indication that an unwound region is formed during transcription comes from the observation that negative supercoiling, which favors unwinding, also favors various steps in tran-

scription.* It was reported early that the presence of up to 2.5% superturns in phage lambda DNA accelerates cell-free RNA synthesis several-fold.[46] Electron micrographs showed that part of the additional transcription comes from promoters inactive in the relaxed (linear) state; the newly unwound segments were assigned to low melting, presumably A,T-rich regions.[45] Transcription of phage PM2 by *E. coli* polymerase *in vitro* can be accelerated up to sixfold when 30 to 65 superhelical turns are introduced into the template.[410] The acceleration is due to more RNA polymerases bound by each phage molecule.

The extent of stimulation by negative supercoiling depends on the promoter studied: Brahms et al.[48] studied the enhancing effect of supercoiling on all five identified promoters of plasmid pBR322. The promoters fall into 3 groups: the promoters of beta lactamase and tetracycline, which are maximally active at 1.1% superhelicity; the promoters for RNA I and RNA II (Section IV.A.2), which are maximally active at 5.5%, the natural superhelicity, while the short RNA5 promoter becomes most active only at 10% superhelix. The effect of superhelicity thus varies with each gene and promoter present. The decline in initiation beyond the optimal sigma is proposed by Brahms et al.[48] to be due to the transition of 8 to 10 alternating base sequences into the Z conformation.

A favorable effect of superhelicity is not always found. Very little stimulation was observed in the trp transcription system.[459] In some well-studied cases, negative supercoiling was observed to inhibit gene expression. For instance, transcription of the DNA gyrase gene or of the *recA* gene proceeds most efficiently on relaxed DNA.[326] These genes nevertheless have a function in the DNA unwinding processes, and specific mechanism associated with feedback regulatory circuits seem to be involved.[72]

Which step in the transcription is affected by the negative superhelicity? It is generally accepted that initiation frequency (for exact definition see Yagil[544]) determines the rate of RNA transcript production.[543] In the first step, DNA-dependent RNA polymerase, associated with a sigma subunit, recognizes a promoter DNA element to form a "closed" initiation complex.[76] The initially recognized promoter DNA is most probably in the classic right-handed duplex form.[543]

In the next step, the rate of which is no longer dependent on polymerase concentration, the "closed" complex is converted into an "open", presumably melted, state. Very large accelerating effects of negative superhelicity on the expression of the *E. coli* lac gene were observed, in particular in cell-free systems.[41] Detailed kinetic analysis has revealed that the rate constant of the conversion of closed to open complex is particularly affected.[301] The kinetic data suggest that there exists at least one additional intermediate,[60] which isomerizes into the ultimate open complex. This intermediate is partly unwound, but bases are still stacked, as deduced from the large activation energy needed for transformation into the ultimate open complex. It should be noted, however, that the stability of the intermediate complex is much lower when a supercoiled rather than a linear template is examined so that its significance in the cell may be limited.

Comparisons of wild-type and several mutant lac promoters[6,42,323] demonstrate that formation of the open complex is maximal when 2 to 10 supercoils are present. These data were obtained in minimal cell-free transcription systems. The addition of cAMP-CRP already shifts the optimal number of negative superturns, albeit toward fewer supercoils.[323] The ratios of open complex formation rates of different lac promoter mutants gets closer to the *in vivo* rates, as superhelicities approach their *in vivo* value of 4%[42] (the rate of abortive initiation being measured). Borowiec and Gralla propose, based on kinetic parameters of open complex formation, that a stressed complex forms as an intermediate, which is partly unwound, and bent, due to superhelical strain. This stressed, partly unwound complex may be related to the intermediate complex derived by kinetic analysis of the RNA polymerase-promoter interaction[60] (see also Straney and Crothers[470]).

* An alternative explanation can be that superhelicity exerts its effect through wrapping of DNA around transcription machinery; this, in its simplest version, would only aggravate the disentanglement problem. Surface wrapping can be accompanied by changes in the degree of primary winding (White and Bauer, 1989). If underwound, as in nucleosomes, the wrapped DNA would consume additional otherwise-available superturns.

The large activation energies needed for transition into the fully open state were interpreted to mean that the intermediate can be topologically unwound but still base stacked conformation.[6]

What is the state of DNA in the open region? Several experimental approaches have been used to obtain information on the state of DNA in the open region (''initial transcription bubble'') of bacterial genes. N3-Cyt methylation studies have shown that the RNA polymerase protects the region between −9 and +3 of the lac UV5 promoter upon formation of the open initiation complex,[233,324,448] slightly more than one turn of the helix. These experiments were carried out with purified polymerase-linear DNA complexes, so that more extensive opening *in vivo* cannot be excluded.

Amouyal and Buc[6] show, however, that a considerable amount of unwinding is present already in the closed state (binary complex). They show that reactivity toward the copper phenanthroline reagent precedes considerably strand separation, indicating the presence of topologically unwound yet unseparated DNA in the promoter region. The possibility that the negative superturns are taken up by wrapping around the polymerase is also considered, and is consistent with present thinking on the role of bending in promoter action. Alternatively, one of the structures shown in Figure 8 could participate. Of the symmetric structures, only the cruciform can be formed in the lac promoter region, but repeated probing with nucleases and chemical reagents does not favor the notion that a cruciform is recognized.[356,451,512] The symmetry of the lac operator may rather be connected with the recognition of dyadic features of the lac repressor. Although bacterial transcription can be carried out in the absence of the multitude of transcription factors needed in higher organisms, some protein factors such as SSB may still play an enhancing role, by forming a complex of distinct structure with unwound DNA segments. Two-dimensional NMR of the transcription complex, when feasible, will hopefully shed light on which of the different possible DNA structures actually participate in transcription-associated unwinding, during both initiation and elongation steps.

**Eukaryotes** — The efficiency of transcription of eukaryotic genes is also dependent on the state of template supercoiling. Transcription of late SV40 genes and polyoma virus genes is considerably more efficient from circular negatively supercoiled templates than from linearized viral DNA.[275] Luchnik et al.[291] have shown that the transcriptionally active fraction of SV40 nucleoprotein complexes is in a relaxable supercoiled state. A high template efficiency of a negative supercoiled relative to a linearized template, up to 500-fold, is observed during the transcription of thymidine kinase from a plasmid construct injected into oocytes.[187] The effect is most profound in plasmids that contain enhancer sequences.[523] Cell-free transcription of *X. laevis* 5S DNA is 5 times faster with supercoiled template than with relaxed template.[458] Considerable acceleration was also observed when the 5S RNA template was injected into oocyte germinal vesicles,[422] or incubated with a germinal vesicle supernatant fraction.[163] The enhancement was explained as the result of an ATP-driven negative supercoiling activity present in the germinal vesicle. This activity causes the conversion of part of the chromatin to a highly supercoiled, unrestrained, ''dynamic'' state, which is active in transcription. A correlation between the extent of transcription and the degree of supercoiling in oocyte extracts supports this picture.[242,435] The results with oocyte extracts were later challenged by Wolffe and co-workers,[535] who found that the linear form of 5S DNA is transcribed most efficiently (see References 241 and 537). The higher template efficiency of linear DNA is nevertheless topologically plausible, because no impediment to template unwinding exists. Transcription of circular templates may still be favored by negative superhelical turns and supported by the activity of topoisomerases, as evidenced in a silk gland-derived transcription system.[200]

### 2. The Size of the Unwound Region

The size of the unwound section was first measured by Saucier and Wang[429] and, more accurately, using gel topoisomer analysis, by Wang et al.[514] An unwinding angle of 240° (6 to 7 bases) per *E. coli* RNA polymerase bound to the promoter regions of bacteriophage fd was found at 37°. No dNTPs were added, so that unwinding

accompanying initiation complex formation was measured. Up to 15 polymerases were bound per phage (100 bases unwound).

Detailed measurements of the unwinding angle of initiation complexes was carried out by Gamper and Hearst.[149,150] *E. coli* RNA polymerase was added to circular SV40 DNA under a variety of conditions. The complex formed was relaxed by addition of topoisomerase, and the change in migration was determined on agarose gel. The observed shift in the topoisomer distribution leads to a figure of 17 base pairs unwound per RNA polymerase molecule bound, i.e., a bit more than 1.7 turns of the helix. A similar-sized region of 17 bp was found to be unwound for binary RNA polymerase-DNA complexes (no dNTP present), for initiation complexes (3 NTP present), or for elongation complexes (all 4 nucleotides present). On the basis of this figure and previous knowledge that the nascent RNA hybridizes with DNA for a constant length of 12 to 15 bases, the transcription model depicted in Figure 19 is proposed. The polymerase complex is suggested to possess, in addition to the RNA polymerizing activity, an unwindase activity acting ahead of the active site, as well as a "rewindase" activity trailing the active site. The transcription machinery (unwound bubble) can proceed by simultaneous unwinding of DNA not yet transcribed and rewinding of the DNA template already transcribed. This requires, however, that the transcription bubble with the attached nascent RNA rotate concomitantly with the progress of the transcription machinery, relative to the DNA ahead of and behind the bubble. It has not yet been possible to establish whether the bubble-RNA complex rotates around the fixed DNA template, or the rest of the DNA template rotates around its helical axis, the nascent RNA remaining fixed. In a closed DNA template, as in a virus or plasmid, DNA rotation cannot occur if it is anchored to any of the cell structural components. A nicking-closing at a suitable location activity can, however, provide a swivel point.

Seventeen bases are unwound for the passage of a single RNA polymerase complex. What happens when many polymerases simultaneously transcribe a gene during high initiation frequencies? One possibility is that the nascent RNA complexes , of increasing lengths, all rotate along the static DNA molecule. This may be a bit hard to envisage from the energy viewpoint. Alternatively, the whole DNA circle or segment may rotate, section by section, as long as no anchoring point interferes or topoisomerase-generated swivel points operate at appropriate points. The successive advancing RNA polymerases may move either as separate transcription bubbles of 17 bases each, or merge to one continuous unwound region. In the first case, DNA sections between two bubbles would rotate individually to rewind the DNA of the leading bubble and to unwind the trailing bubble. In the case of a continuous unwound region, DNA is likely to remain strand separated, at least one strand of the DNA being associated with the RNA polymerase and possibly with other transcription "factors".* A paranemic duplex may also be intermittently formed between passing polymerases. Future studies will have to resolve the exact behavior of the different DNA sections during elongation. The possibility that the length of the unwound region is somehow predetermined during preinitiation steps, and thus serves to control the number of polymerases simultaneously transversing a gene, is an attractive possibility, but has yet to be substantiated.

**Eukaryotes** — The length of the DNA region unwound during transcript elongation on deproteinated SV40 minichromosomes (i.e., bare supercoiled SV40 DNA with only mammalian polymerase II attached) has been measured by Choder and Aloni.[80] A transcription bubble of a size similar to that formed by the prokaryotic polymerase (2 turns) was found for the native

* The symmetric structures, or other structures of Figure 8 may serve to store linking deficit near the promoter or other suitable location. An unwound region can readily be transmitted from one region on a gene or plasmid to another by first supercoiling the whole domain or plasmid, then opening the supercoils while unwinding the other region, transmitting the linking deficit to where it is required. The transmission of a linking deficit along a closed DNA has been nicely demonstrated by Lilley and co-workers (Furlong et al., 1989), by showing that a low-melting, A,T-rich DNA region can accelerate the formation of a cruciform by a nearby dyad symmetric region. This mechanism may operate over longer distances to transfer the unwound region from the end of a transcribed region, where one batch of polymerase terminates, back to the initiation (promoter) region.

mammalian RNA polymerase. The relatively short unwound region is not in line with the formation of an extended unwound region but, again, protein factors or matrix elements necessary to keep the transcribed region unwound may not be present in the studied cell-free system. A longer unwound region was obtained upon removing the polymerase from intact circular DNA, as a result of nascent RNA-DNA hybrid formation (prior to relaxation); in the nicked SV40 circle, the RNA transcript was rapidly displaced leading to DNA duplex reformation. Zhang and Gralla[554] have used potassium permanganate, a reagent for exposed cytosines (Section II.F), to examine the late SV40 initiation complex. About 28 bases upstream from the initiation point were hyperreactive to permanganate when applied to SV40-infected cells, late in the infection. This high sensitivity is unexpected in view of the low percentage of actually transcribing genomes in infected cells. Apparently, most SV40 minichromosomes assume some activated state in the promoter regions.

### 3. Unwinding Caused by Transcription: The Twin Supercoiling Domains Model

An early indication that transcription can determine the extent of supercoiling (rather than being determined by it) comes from the observation of Pruss and Drlica that the degree of negative supercoiling of pBR322 can be very high in topoisomerase I-less bacteria. The degree of supercoiling was dependent on the tetracycline gene being transcribed, at least to some extent. When the promoter was eliminated or the entire tet gene was absent, no excessive supercoiling was observed.[391,392] A clear demonstration that intermediate unwinding does take place during transcription came from a series of experiments with closed bacterial plasmids by Liu and Wang.[286] Changes in supercoiling were studied in a pBR322-related plasmid in which the two main copied genes are transcribed from a single promoter region in opposite directions, in the absence of either negative or positive supercoil relaxing activities.[540] Any lag in the revolution of the polymerase complexes around the template must generate positive supercoils ahead of RNA polymerase, i.e., between the 3′ ends of the two genes (to compensate for the unwinding of the right-handed template) as well as negative superturns behind the polymerase, between the two 5′ ends (because of rewinding of the unwound sections). The "twin" positive and negative supercoiled domains thus formed could be demonstrated by having the plasmid transcribed in transfected cells in the presence of the DNA gyrase inhibitor novobiocin, so that positive supercoils are not removed. Negative turns are still removed by cellular *E. coli* topoisomerase I, which relaxes only negative supercoils. At least 10 *positive* superturns were observed in the deproteinated transcribing plasmids. In topoisomerase I-less mutants, highly negative supercoils appeared.

The positive supercoils were subsequently shown to be also formed in a cell-free transcription system, when supplied with *E. coli* topoisomerase I.[481] The appearance of the positive superturns is entirely transcription dependent, because the addition of rifampicin prevents their appearance. The positive supercoil accumulation (after removal of negative superturns) was not observed when RNAse was present. This effect of the nascent RNA chain removal favors the explanation that the drag on the nascent chain creates a lag between the unwinding rotation and the progress of the transcription machinery, and is thus responsible for the generation of the positive superturns.

It should be pointed out that the number of superturns observed can account for only a fraction of the unwinding expected: the tetracycline-resistance gene of pBR322 is about 1190 bp long, i.e., 119 turns must be unwound and rewound for a single passage of polymerase. During a transcription experiment, hundreds of polymerases track along the template, so that thousands of superturns must be generated. The positive supercoiling observed therefore can represent only a residual delay between the accumulation of superturns during transcription and their subsequent relaxation. This delay can be the result of the lag between the rotating movement of the transcription machinery and its linear progress along the DNA.

The conclusion that the observed supercoiling may be the result of transcription rather than its cause is supported by a study on the induction

of several photosynthetic genes in the facultative aerobic organism *R. capsulata*.[90] A direct assay of the degree of supercoiling, using a psoralene derivative, did not reveal a change in superhelicity during six- to eightfold induction of the photosynthetic genes, indicating the supercoiling per se is not a necessary step of the induction process. The inhibitory effect of novobiocin on the induction of several photosynthetic genes in *R. capsulata* is explained in terms of a gyrase acting at a suitable swivel point to relieve excess supercoiling, extending the twin supercoiling concept to anchored genes.[90] Most recently the accumulation of negative superturns could be substantiated in experiment where a Gal4 binding site was attached to a T7 RNA polymerase. This polymerase was added to a plasmid that has both a T7 promoter and a Gal4 binding site 5′ to it. The simultaneous binding of the T7 RNA polymerase to both binding sites should restrain topologically the DNA region in between and impede polymerase rotation. Positive supercoils were observed indicating that relative rotatory movement does accompany the transcription process.[363]

**Eukaryotes** — Up to 58 negative turns were detected by two-dimensional electrophoresis of a yeast plasmid when actively transcribed in a topoisomerase I negative yeast cell[56] (see Reference 157). Brill and Sternglantz[56] found that the negative supercoils appear concomitantly with RNA elongation, which argues against the preformation of an unwound region. Another indication that changes in supercoiling are related to transcription comes from the observation that inhibition of transcription of the hsp70 gene *in vivo*, by actinomycin D or by DRB, inhibits topoisomerase cleavage at the gene.[255] Excessive positive supercoiling has also been established during T antigen movement of the SV40 origin region in a cell-free system, when the bacterial topoisomerase I was added.[547] The twin supercoiling domains phenomenon is thus not restricted to *E. coli* and related bacterial systems, and may be valid for eukaryotic cells as well.

## 4. The Role of Topoisomerases

An early indication that topoisomerases are involved in transcription was provided by experiments in which drugs that inhibit topoisomerase II, such as coumermycin[256] and novobiocin[3,181,422] (see also References 121 and 285) inhibited transcription of various cellular genes (novobiocin may exert its effect not through topoisomerases).[474,520] Coumermycin, as well as naldixic and oxolinic acids, inhibitors of *E. coli* DNA gyrase subunit B, were found to inhibit transcription in the bacterial cell.[512,256,360] Similar effects were later observed with campthotecin, an inhibitor of eukaryotic topoisomerase I.[159,204,255,466] Topoisomerase cleavage points mapped during transcription were found to correlate with the progress of transcription, indicating a direct requirement of topoisomerase I activity during RNA chain elongation.[465,553] Most recently, oxolinic acid-mediated cleavage sites by DNA gyrase were mapped along the entire *E. coli* genome. The cleavage sites were found to be concentrated in about 100 clusters (''toposites'') and the cleavage was shown to be transcriptionally modulated, alluding to a function of DNA gyrase activity during transcription.[88]

Topoisomerases are also necessary ingredients of cell-free transcription systems, if transcription is to proceed beyond the early abortive stage.[481] The effect of topoisomerase inhibition was initially interpreted as an effect of the ''torsional strain'' on the transcription frequency, operating via a change in nucleoprotein chromatin structure.[422,423] It was later found[242] that the effect of novobiocin, for example, on the transcription of 5S RNA from *X. laevis* chromatin, is independent of the state of supercoiling (which can be changed by incubation with topoisomerase I). Therefore, it was proposed that gyrase-like activity is necessary as part of the transcription process itself (as, for example, in the twin supercoiling domains model). A similar conclusion was arrived at when the effect of novobiocin on *R. capsulata* was studied.[90]

The active role of topoisomerases in transcription is supported by studies with mutated or disrupted topoisomerase genes. This has been shown to be the case for transcription both in *E. coli*[117,198,304,392] and in yeast.[546] A nicking activity is indispensable, if both DNA template and the rotating RNA transcript are anchored to some structural element. Topoisomerases ought also to

be stimulatory in case the rotatory movement of the nascent RNA rather then the linear advance of the transcription machinery are rate limiting. The topoisomerase may become even more essential when many polymerases have to transverse a gene simultaneously, i.e., at high initiation frequencies, so that sufficient negative superhelicity is no longer available in the transcribed domain. This may well be the case in highly active eukaryotic genes, in particular specialized genes like actin, myosin, globin, or immunoglobulin. Studies in defined cell-free systems will have to be made to decide whether topoisomerase is particularly stimulative at high transcription frequencies.

### 5. Enhancers: Initiation Points for Unwinding?

The transcription frequency of many eukaryotic genes is highly dependent on the activity of enhancer and/or upstream activating (UAS) DNA elements. The way enhancer elements function is yet unclear.[189,302,343] Two major mechanisms have been considered: one is the "looping" mechanism, in which a DNA bends to form a loop between the enhancer and promoter regions, so that an enhancer-associated protein factor, or a chain of factors, can directly interact with the promoter region, transmitting the enhancer effect. The other mechanism is the "scanning" or "entry rate" mechanism, in which the enhancer serves as entry point for the RNA polymerase, permitting it to scan the intervening DNA until it finds a promoting region.[338] At present, the bulk of the available evidence tends to favor looping mechanisms.[393,430] We would still like, in this section, to examine the possibility that enhancer elements can serve as unwinding initiation elements. Unwinding initiation can be transmitted to the promoter either by the scanning mechanism or via a super coiling-uncoiling mechanism.

An unwinding role of transcription enhancers could explain several of their well-known properties:

1. Enhancers can operate at a distance from the transcription initiation site. Indeed, once a section of DNA is unwound, it can migrate, for example, as a bubble, to the transcription start (with or without a polymerase).
2. The enhancing activity is direction independent. An unwound region should be able to migrate in either direction.
3. Enhancers need not be located 5′ to the initiation site; they can also be at the end of a gene (e.g., β-globin).[404] Unwound regions have to migrate anyway, after the passage of a single or a batch of polymerases, from the 3′ end back to the 5′ start of the gene. Therefore, it should not matter from the topological standpoint, whether the initial unwound region is formed at the 5′ or 3′ end of the gene.
4. Enhancers are modular, and are composed of distinct subelements. The enhancing effect of the enhancer increases considerably when a subelement is present in multicopies.[379,381] This is readily understandable, because each additional element can enlarge the unwound region, facilitating the entry of more and more successive polymerases.
5. Enhancer subelements can substitute for each other, provided the respective activating protein is present.[430] For instance, the B subelement of the 72-bp SV40 enhancer can be substituted by a second A element, or vice versa.[196] It should be added that the B element has a very high purine-pyrmidine asymmetry, a motif we have seen to be conducive to SSN sensitivity: GGTGTGGAAAG TCCCC AGG CTCCCC. The last 6 bases have indeed been found to be highly S1 sensitive.[130] The A element contains 14/18 R: AGG C AGAAG TATGCAAAG. The last 8 bases are part of the octamer motif, and have been found to be mungbean sensitive.[212] The authors explain their result as due to a 3:5 loop and stem structure. Examination of proper base substituents could distinguish between the two possibilities (in this connection, the AAATAAAAAAAAT motif is also R rich; see Section IV.A.3).
6. Enhancers operate in a cell-type specific fashion,[381] i.e., immunoglobulin enhancers activate the gene in B lymphocytes and not

in other cell lineages. The activation requires that an active form of protein NF-κB interact with the octamer subelement of the Igκ enhancer. A model where a cell-specific factor facilitates (or antiprevents) the opening and/or the migration of the unwound region to the initiation site can be readily envisaged.

7. Experiments in which the movement of an unwound bubble from enhancer to promoter is blocked favor at least certain versions of the scanning model. The block was achieved either by the binding of *lex* repressor to an inserted *lex* operator[53] or by psoralene addition.[92] Muller et al.[343] nevertheless demonstrate that an enhancing effect can be transmitted through a biotin streptavidin bridge. This is not compatible with the idea that an unwound bubble propagates from enhancer to promoter. The linking deficit can, however, be transferred by first supercoiling the entire circle, then de-supercoiling it, unloading the deficit at the promoter. The same result can also be achieved by rotating the DNA intervening between enhancer and promoter. The enhancing effect of a linear construct,[343] nevertheless, cannot be explained by the unwinding effect.

In summary, while there is at present no compelling evidence for a role of enhancers in solving the topological problem, until a fully satisfactory mechanism for enhancer action is available, promotion of unwinding should be considered a mechanism deserving further experimental testing. The various techniques described in Section II, including chemical probing and topoisomer analysis, in the absence or the presence of activating factors, can be helpful.

## C. Recombination and the Paranemic Joint

Genetic recombination is another process where extensive strand unwinding is to be expected. Homologous recombination involves alignment and pairing of two separate DNA molecules (followed by strand exchange) and is hard to envisage without some form of strand unwinding. One of the more thoroughly studied homologous recombination system is at present the *E. coli rec* system (for reviews see Cox and Lehman,[94] Griffith and Harris,[174] and Kowalczykowski[250]). The first step in recA-supported recombination is the coating of a single- or (double)-stranded DNA by recA protein, to form a "presynaptic complex". In the presence of ATP or analog, at optimal conditions, one recA molecule is associated for every 3 bases. The coated DNA, ss or ds, was observed to form a partly unwound helical structure of up to 18 nucleotides per turn and of a pitch of 95 Å.[112,462]

In the next step, the recA-coated molecule interacts with a second, homologous DNA molecule, ds or ss, to form a joint complex. In case either the single strand or the double strand has an open homologous end, the single strand will find a complementary region on the double-stranded DNA to form a "plectonemic complex". Eventually, the single strand may displace its homolog on the double strand, leading to strand exchange. In the intact cell, other proteins of the system, including the RecBCD complex, will help to complete the process and lead to strand exchange between chromosomes.

If, however, both the ssDNA molecule and its complementary strand on the dsDNA are circularly closed, strand intertwining is impossible, and a plectonemic complex cannot be formed. A recA-associated DNA-DNA complex is still detected, in particular if the dsDNA is negatively supercoiled.[102,412] This topologically unwound complex is termed by Radding and co-workers,[33] following Watson and Crick,[519] "paranemic joint". The paranemic joint can be observed with linearized DNA when the slowly hydrolyzable ATP analog ATPγS is present, or when the homologous DNA region is inserted between two nonhomologous free ends. The unwound state of the DNA in these complexes was initially deduced from the appearance of D loops in electron micrographs of the recA heteroduplex complexes.[73,102,443] The unwinding of the ds DNA in the complex was verified by agarose gel analysis after nicking the ds DNA in a nonhomologous region and religating in the presence of recA protein; fast migrating, highly negatively supercoiled bands appeared, termed Form X DNA[294,539] and form P DNA[434] indicating a highly unwound

state at the time of ligation. Unwinding was further verified by the application of topoisomerase, which led to a hemicatenated three-stranded DNA when free homologous ends were absent.[33]

The paranemic joint can be observed when both the ds- and ssDNA components are circularly closed. Thus, joined regions of the length of 300 ± 100 bases are seen by EM following glutaraldehyde fixation of closed superhelical ds plasmid M13mp7 incubated with closed phase M13 in single-stranded form.[81] This necessitates the unwinding of 30 to 40 turns, an amount expected topologically in the 7238-bp-long plasmid at 5% superhelicity. Superhelical turns are no longer observed by EM once the paranemic joint is formed. A similar extent of unwinding was observed recently by topoisomer analysis when a ϕX174 DNA containing the supercoiled duplex DNA interacted with a complementary gapped duplex in the presence of recA protein;[89] 15 to 17 superhelical turns of the circular duplex became unwound upon complex formation. Negative superhelicity thus favors paranemic joint formation, but a joint can be formed even by a relaxed circle, possibly by the introduction of positive superturns.[33]

A clear helical feature of the thick filament of the recA-coated "synaptic complex" is evident on electron micrographs under proper conditions.[342] Little variation in thickness is observed whether ss, ds, or even three strands are present in the joint, which alludes to an intimate structure between the three strands. Resolution of the protein-coated complex is not sufficient, however, to permit determining whether only occasional base contacts occur, whether the complementary strand forms duplexes alternating between the two participating homologous strands, or whether a defined triple-stranded structure is formed. It should be stressed that if a hydrogen-bonded triplex structure is formed it must be structurally distinct from the triplexes discussed in Section III.G, because the incoming single strand must run parallel to its homologous strand, rather than antiparallel as in the structures considered by Lyamichev et al.[294] or Cooney et al.[91]

The triple-stranded regions in the paranemic joint have been reanalzyed recently by EM following photochemical fixation with the psoralene derivative AMT, which cross-links aligned DNA strands but forms no DNA-protein cross-links, so that the recA protein can be removed prior to analysis. Regions of DNA derived alternatingly from the double- and single-stranded components, extending over hundreds of bases, were observed.[489] In addition, triple-stranded regions extending over dozens of bases could also be observed; the absence of the coating recA protein led to the conclusion that triple helixes are present. It is suggested that the triplex is right handed,[489] but it is not clear how the right-handed turns are topologically compensated in the unwound joint.

If indeed a defined unwound structure is present in the paranemic joint, then any of the paranemic structures shown in Figure 8 can be present. The similarity of the joint to form V DNA (Section III.F) has been pointed out by Wu and co-workers;[539] form V DNA arises, like the paranemic joint, from the alignment of complementary strands by topologically restricted DNA molecules. Both the paranemic joint and form V DNA may contain alternating left- and right-handed sections (Figure 8e), so that $\Delta Lk = 0$. Of the other paranemic structures, the three symmetric structures (repeated, dyadic, and mirror) are less likely to contribute, because of the very limited occurrence of symmetric regions in M13 or ϕX174, the phages used to study the paranemic joints. This leaves either loose single-stranded alignments, or a version of the paranemic nonhelical duplex proposed by us[545] as structural alternatives to alternating right and left DNA duplexes. That paranemic duplex has highly exposed base pairs, readily available for interaction with a third or even fourth strand. Substantial evidence for the long-proposed four-stranded joint, also necessarily unwound, has recently been brought.[89,283] Evidence for a role of RNA in stabilizing a paranemic joint has to the best of my knowledge not been reported. The role of recA protein in stabilizing the paranemic joint is a clear example of the ability of proteins to stabilize unwound DNA forms. In summary, while the topologically unwound nature of the paranemic joint is clear, the question of what particular structure DNA and protein assume in the joint awaits further elucidation.

## D. Conclusions

Evidence on the participation of unwound DNA forms in all three major DNA template driven processes is now available. The details of these unwound forms — the arrangement of bases, ranges of existence, participation of auxiliary factors, and other details — have yet to be determined. No less important, the dynamics of their formation, propagation, and interaction with the various components of replication, transcription, or recombination (and repair) mechanisms remain a challenge, and will doubtlessly be taken up intensively in the near future; in particular, the questions in which systems are unwound forms of DNA just transient intermediates and under what conditions can temporally extended paranemic forms be expected. The recently developed techniques reviewed in Section II will certainly be of help in answering these questions and hopefully lead to the ultimate resolution of the topological dilemma raised by Watson and Crick in 1953.[519]



## ACKNOWLEDGMENTS

The thorough comments of Dr. W. R. Bauer, in particular on the topological aspects discussed, are gratefully acknowledged, as is the additional information supplied by Dr. T. Kohwi-Shigematsu and Dr. H. Htun. The author thanks his colleagues, Dr. Chaim Manor, Dr. David Reisman, Dr. Violet Daniel, and Dr. Ezra Yagil, for their helpful comments on the manuscript. The expert editorial assistance of Ms. Malvine Baer is greatly appreciated. The support of the Helen and Milton Kimmelman Center for Biomolecular Structure is gratefully acknowledged.

## ABBREVIATIONS

| | |
|---|---|
| ARS | Autonomous replication sequences |
| bp | Base pairs |
| SSN | Single-strand specific nuclease |
| SSB | Single-strand binding (protein) |
| EM | Electron microscope(y) |
| DMS | Dimethyl sulfate |
| DEPC | Diethyl pyrocarbonate |
| R | Purines |
| Y | Pyrimidines |
| R · Y | Oligopurine · oligopyrimidine DNA tracts |
| R-Y | Alternating oligo(purine-pyrimidine) tracts |
| ss | Single-stranded |
| ds | Double-stranded |

## REFERENCES

1. **Abdul-Monem, M. and Hoffmann-Berling, H.,** Enzymic unwinding of DNA. I. Purification and characterization of the ATPase activity from *E. coli*, *Eur. J. Biochem.*, 65, 431, 1976.
2. **Adachi, Y., Kas, E., and Laemmli, U. K.,** Preferential cooperative binding of DNA topoisomerase II to scaffold associated regions, *EMBO J.*, 8, 3997, 1989.
3. **Akrigg, A. and Cook, P. R.,** DNA gyrase stimulates transcription, *Nucleic Acids Res.*, 8, 845, 1980.
4. **Alberts, B. M. and Frey, L.,** T4 Bacteriophage gene 32: a structural protein in the replication and recombination of DNA, *Nature*, 227, 1313, 1970.
5. **Albiser, G. and Premillat, S.,** A critical analysis of a left handed double helix model for B-DNA fibers, *Nucleic Acids Res.*, 10, 4027, 1982.
6. **Amouyal, M. and Buc, H.,** Topological unwinding of strong and weak promoters by RNA polymerase: a comparison between the lac wild-type and the UV5 sites of *E. coli*, *J. Mol. Biol.*, 195, 795, 1987.
7. **Anshelevich, V. V., Vologodskii, A. V., and Frank-Kamenetskii, M. D.,** A theoretical study of formation of DNA noncanonical structures under negative superhelical stress, *J. Biomol. Struct.Dyn.*, 6, 247, 1988.
8. **Antao, V. P., Gray, D. M., and Ratliff, R. L.,** CD of six different conformational rearrangements of poly{d(A-G)d(C-T)} induced by low pH, *Nucleic Acids Res.*, 16, 719, 1988.
9. **Arndt-Jovin, D. J., Robert-Nicoud, M., Zarling, D., Greider, C., Weimer, E., and Jovin, T. M.,** Left handed Z DNA in bands of acid fixed polytene chromosomes, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 4344, 1983.
10. **Arnott, S., Bond, P. J., and Chandrasekharan, R.,** Visualization of an unwound DNA complex, *Nature*, 287, 561, 1980.
11. **Arnott, S., Bond, P. J., Selsing, E., and Smith, P. J. C.,** Models of triple-stranded polynucleotides with optimised sterochemistry, *Nucleic Acids Res.*, 3, 2459, 1974.

12. **Arnott, S., Chandrasekaran, R., Hall, I. H., and Puigjaner, L. C.,** Heteronomous DNA, *Nucleic Acids Res.*, 11, 4141, 1983.
13. **Arnott, S. and Selsing, E.,** Structures of poly dA · poly dT and poly dA · poly dT · poly dA, *J. Mol. Biol.*, 88, 505, 1974.
14. **Azorin, F., Hahn, R., and Rich, A.,** Restriction endonucleases can be used to study B-Z junctions in supercoiled DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 5714, 1984.
15. **Baer, B. W. and Rhodes, D.,** Eukaryotic RNA polymerase II binds to nucleosomes arrays form transcribed genes, *Nature*, 301, 482, 1983.
16. **Baker, T. A., Bertsch, L. L., Bramhill, D., Sekimizu, K., Wahle, E., Yung, B., and Kornberg, A.,** Enzymatic mechanism of initiation of replication from the origin of the *E. coli* chromosome, *Cancer Cells*, 6, 19, 1988.
17. **Baker, T. A., Funnell, B. E., and Kornberg, A.,** Helicase action of dnaB protein during replication from the *E. coli* chromosomal origin in vitro, *J. Biol. Chem.*, 262, 6877, 1987.
18. **Baker, T. A. and Kornberg, A.,** Transcriptional activation of initiation of replication from *E. coli* chromosomal origin: a RNA-DNA Hybrid near oriC, *Cell*, 55, 113, 1988.
19. **Baran, N., Lapidot, A., and Manor, H.,** Unusual sequence element found at the end of an amplicon, *Mol. Cell. Biol.*, 7, 2636, 1987.
20. **Baran, N., Lapidot, A., and Manor, H.,** Formation of DNA triplexes accounts for arrests of DNA synthesis at $d(TC)_n$ and $d(GA)_n$ tracts, *Proc. Natl. Acad. Sci. U.S.A.*, 88, 507, 1991.
21. **Barsoum, J. and Berg, P.,** Simian virus 40 minichromosomes contain torsionally stressed DNA, *Mol. Cell. Biol.*, 5, 3048, 1985.
22. **Bauer, W. and Vinograd, J.,** The interaction of closed circular DNA with intercalative dyes. I. The superhelix density of SV40 DNA in the presence and absence of dye, *J. Mol. Biol.*, 33, 141, 1968.
23. **Bauer, W. and Vinograd, J.,** Circular DNA, in *Basic Principles in Nucleic Acids Research*, Vol. 2, Ts'o, P.O.P., Ed., Academic Press, New York, 1974, 265.
24. **Bauer, W. R.,** Structure and reactions of closed duplex DNA, *Annu. Rev. Biophys. Bioeng.*, 7, 287, 1978.
25. **Bauer, W. R. and White, J. H.,** Surface Linking and helical repeat of protein wrapped DNA, *Nucleic Acids and Mol. Biol.*, 4, 39, 1990.
26. **Bauer, W. R., Crick, F. H. C., and White, J. H.,** Supercoiled DNA, *Sci. Am.*, 243, 100, 1980.
27. **Baur, C.-P. and Knippers, R.,** Protein-induced bending of the simian virus 40 origin of replication, *J. Mol. Biol.*, 203, 1009, 1988.
28. **Been, M. D., Burgess, R., and Champoux, J. J.,** Nucleotide sequence preference of rat liver or wheat germ type 1 DNA topoisomerase breakage sites in duplex SV40 DNA, *Nucleic Acids Res.*, 12, 3097, 1984.
29. **Belotserkovskii, B. P., Veselkov, A. G., Filippov, S. A., Dobrynin, V. M., Mirkin, S. M., and Frank-Kamenetskii, M. D.,** Formation of intramolecular triplex in homopurine-homopyrimidine nirror repeats with point substitutions, *Nucleic Acids Res.*, 18, 6621, 1990.
30. **Benyaiti, C. and Worcel, A.,** Isolation, characterization and structure of the folded interphase genome of *D. melanogaster*, *Cell*, 9, 393, 1976.
31. **Bernues, J., Beltran, R., Casasnova, J. M., and Azorin, F.,** Structural polymorphism of homopurine-homopyrimidine sequences: the secondary DNA structure adopted by a $d(GA \cdot CT)_{22}$ sequence in the presence of zinc ions, *EMBO J.*, 8, 2087, 1989.
32. **Bernues, J., Beltran, R., Casasnova, J. M., and Azorin, F.,** DNA-Sequence and metal-ion specificity of the formation of H* DNA, *Nucleic Acids Res.*, 18, 4067, 1990.
33. **Bianchi, M., DasGupta, C., and Radding, C. M.,** Synapsis and the formation of paranemic joints by *E. coli* recA protein, *Cell*, 34, 931, 1983.
34. **Biggin, M. D. and Tjian, R.,** Transcription factors that activate the ultrabithorax promoter in developmentally staged extracts, *Cell*, 53, 699, 1988.
35. **Blaho, J. A. and Wells, R. D.,** Left handed Z DNA and genetic recombinaton, *Prog. Nucleic Acids Res. Mol. Biol.*, 37, 107, 1989.
36. **Bliska, J. B. and Cozzarelli, N. R.,** Use of size specific recombination as a probe of DNA structure and metabolism in vivo, *J. Mol. Biol.*, 194, 205, 1987.
37. **Bloomers, M. J. J., Walters, J. A. L. I., Haasnot, C. A. G., Aelen, J. M. I., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W.,** Effects of base sequence on loop folding in DNA hairpins, *Biochemistry*, 28, 7491, 1989.
38. **Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr.,** *Physical Chemistry of Nucleic Acids*, Harper & Row, New York, 1974, 373.
39. **Boles, T. C. and Hogan, M. E.,** DNA structure equilibria in the human c-myc gene, *Biochemistry*, 26, 367, 1987.
40. **Borowiec, J. A., Dean, F. B., Bullock, P. A., and Hurwitz, J.,** Binding and unwinding: how T antigen engages the SV40 origin of DNA replications, *Cell*, 60, 181, 1990.
41. **Borowiec, J. A. and Gralla, J. D.,** Supercoiling response of the lac $p^s$ promoter in vitro, *J. Mol. Biol.*, 184, 587, 1985.
42. **Borowiec, J. A. and Gralla, J. D.,** All three elements of the lac $p^s$ promoter mediate its transciptional response to DNA supercoiling, *J. Mol. Biol.*, 195, 89, 1987.
43. **Borowiec, J. A. and Hurwitz, J.,** Localized melting and structural changes in the SV40 origin of replication induced by T-antigen, *EMBO J.*, 7, 3149, 1988.

44. **Borowiec, J. A., Zhang, Li., Sasse-Dwight, S., and Gralla, J. D.,** DNA supercoiling promotes formation of a bent repression loop in lac DNA, *J. Mol. Biol.*, 196, 101, 1987.
45. **Botchan, P.,** An electron microscopic comparison of transcription on linear and superhelical DNA, *J. Mol. Biol.*, 105, 161, 1976.
46. **Botchan, P., Wang, J. C., and Echols, H.,** Effect of circularity and superhelictiy on transcription from bacteriophage lambda DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 70, 3077, 1973.
47. **Brahmachari, S. K., Shouche, Y. S., Cantor, C. R., and McClelland, M.,** Sequences that adopt non B-DNA conformation in form V DNA as probed by enzymatic methylation, *J. Mol. Biol.*, 193, 201, 1987.
48. **Brahms, J. G., Dargouge, O., Brahms, S., Ohara, Y., and Vagner, V.,** Activation and inhibition of transcription by supercoiling, *J. Mol. Biol.*, 181, 455, 1985.
49. **Brahms, S., Vergne, J., Brahms, J. G., Di Capua, E., Bucher, P., and Koller, T.,** DNA formed by reassociation of complementary single-stranded circles from natural DNA is shown to contain left- and right-handed double helices, Cold Spring Harbor Symp., 47, 112, 1982.
50. **Brahms, S., Vergne, J., Brahms, J. G., Di-Capua, E., Bucher, Ph., and Koller, Th.,** Natural DNA sequences can form left handed helices in low salt solution under conditions of topological constraint, *J. Mol. Biol.*, 162, 473, 1983.
51. **Bramhill, D. and Kornberg, A.,** Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome, *Cell*, 52, 743, 1988.
52. **Bramhill, D. and Kornberg, A.,** A model for initiation at origins of DNA initiation, *Cell*, 54, 915, 1988.
53. **Brent, R. and Ptashne, M.,** A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene, *Nature*, 312, 612, 1984.
54. **Breslauer, F. J., Frank, R., Blocker, H., and Marky, L. A.,** Predicting DNA duplex stability from the base sequence, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 3746, 1986.
55. **Brill, S. J., Dinardo, S., Voelkel-Meiman, K., and Sternglantz, R.,** Need for topisomerase activity as a swivel for DNA replication and transcription of ribosomal RNA, *Nature*, 326, 414, 1987.
56. **Brill, S. J. and Sternglanz, R.,** Transcription-dependent DNA supercoiling in yeast DNA topoisomerase mutants, *Cell*, 54, 403, 1988.
57. **Broitman, S. L., Im, D. D., and Fresco, J. R.,** Formation of the triple stranded polynucleotide helix poly(A · A · U), *Proc. Natl. Acad. Sci. U.S.A.*, 84, 5120, 1987.
58. **Brown, D. M.,** Chemical reactions of polynucleotides and nucleic acids, in *Basic Principles of Nucleic Acid Chemistry*, Vol. 2, Ts'o, P.O.P., Ed., Academic Press, New York, 1974, 1.
59. **Brown, D. M., Gray, D. M., and Patrick, M. H.,** Photochemical demonstration of stacked C · $C^+$ base pairs in novel DNA secondary structure, *Biochemistry*, 24, 1676, 1985.
60. **Buc, H. and McClure, W. R.,** Kinetics of open complex formation between *E. coli* RNA polymerase and the lac UV5 promoter: evidence for a sequential mechanism involving three steps, *Biochemistry*, 24, 2712, 1985.
61. **Bucher, P. and Yagil, G.,** The occurence of polypurine · polypyrimidine tracts in eukaryotic and prokaryotic genes, *DNA Sequence*, 1, 157, 1991.
62. **Budowsky, E. I.,** The mechanism of the mutagenic action of hydroxylamines, *Prog. Nucleic Acids Res.*, 16, 125, 1976.
63. **Bullock, P. A., Seo, Y. S., and Hurwitz, J.,** Initiation of simian virus 40 DNA replication in vitro: pulse-chase experiments identify the first labeled species as topologically unwound, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 3944, 1989.
64. **Burkhoff, A. M. and Tullius, T. D.,** The unusual conformation adopted by the adenine tracts in kinetoplast DNA, *Cell*, 48, 935, 1987.
65. **Burton, K. and Riley, W. T.,** Selective degradation of thymidine and thymine deoxynucleotides, *Biochem. J.*, 98, 70, 1965.
66. **Caddle, M. S., Lussier, R. H., and Heintz, N. H.,** Intramolecular DNA triplexes, bent DNA and DNA unwinding elements in the initiation region of an amplified dihydrofolate reductase replicon, *J. Mol. Biol.*, 211, 19, 1990.
67. **Cairns, J.,** The bacterial chromosome, *Sci. Am.*, 214, 12, 1966.
68. **Campos, J. L. and Subirana, J. A.,** The influence of $Mg^{++}$ and $Zn^{++}$ on polypurine and polypyrimidines multistranded helices, *J. Biomol. Struct. Dyn.*, 8, 793, 1991.
69. **Cantor, C. R. and Efstratiadis, A.,** Possible structures of homopurine-homopyrimidine S1-hypersensitive sites, *Nucleic Acids Res.*, 12, 8059, 1984.
70. **Cao, S.-X. and Schechter, A. N.,** Nuclease hypersensitivity in the beta-globin gene region of 562 cells, *Eur. J. Biochem.*, 173, 517, 1988.
71. **Capranico, G., Kohn, K. W., and Pommier, I.,** Local sequence requirements for DNA cleavage by mammalian topoisomerase II in the presence of doxorubicin, *Nucleic Acids Res.*, 18, 6611, 1990.
72. **Carty, M. and Menzel, R.,** Inhibition of DNA gyrase activity in an in vitro transcription-translation system stimulates gyrA expression in a DNA concentration-dependent manner: evident for the involvement of factors which may be titrated, *J. Mol. Biol.*, 214, 397, 1990.
73. **Cassuto, E., West, S. C., Podell, J., and Howard-Flanders, P.,** Genetic recombination: recA promotes homologous pairing between duplex DNA molecules

without strand unwinding, *Nucleic Acids Res.*, 9, 4201, 1981.
74. **Cha, T. A. and Alberts, B. M.,** Studies on DNA Helicase-RNA primase unit from bacteriophage T4, *J. Biol. Chem.*, 261, 7001, 1986; *Cancer Cells*, 6, 1, 1986.
75. **Chalberg, M. D. and Kelly, T. J.,** Animal virus DNA replication, *Annu. Rev. Biochem.*, 58, 671, 1989.
76. **Chamberlin, M.,** in *RNA Polymerase*, Losick, R. and Chamberlin, M., Eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1976, 17, 159.
77. **Champoux, J. J. and Dulbecco, R.,** An activity from mammalian cells that untwists superhelical DNA: a possible swivel for DNA replication, *Proc. Natl. Acad. Sci. U.S.A.*, 69, 143, 1972.
78. **Champoux, J. J. and McConaughy, B. L.,** Priming of superhelical SV40 DNA by *E. coli* RNA polymerase for in vitro DNA synthesis, *Biochemistry*, 14, 307, 1975.
79. **Chase, J. W. and Williams, K. R.,** Single-stranded DNA binding proteins required for DNA replication, *Annu. Rev. Biochem.*, 55, 103, 1986.
80. **Choder, M. and Aloni, Y.,** RNA polymerase II allows unwinding and rewinding of the DNA and thus maintains a constant length of the transcription bubble, *J. Biol. Chem.*, 263, 12994, 1987.
81. **Christiansen, G. and Griffith, J.,** Visualization of the paranemic joining of homologous DNA molecules catalyzed by the recA protein of *E. coli*, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 2066, 1986.
82. **Christophe, D., Cabrera, B., Bacolla, A., Targovnik, H., Pohl, V., and Vassart, G.,** An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene, *Nucleic Acids Res.*, 13, 5127, 1985.
83. **Clark, S. P., Lewis, C. D., and Felsenfeld, G.,** Properties of BGP1, a poly(dG)-binding protein from chicken erythrocytes, *Nucleic Acids Res.*, 18, 5119, 1990.
84. **Cockerill, P. N. and Garrard, W. T.,** Chromosomal anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites, *Cell*, 44, 273, 1986.
85. Cold Spring Harbor Symposium on Eukaryotic DNA replication, Kelly, T. and Stillman, B., Eds., *Cancer Cells*, 6, 1, 1988.
86. **Collier, D. A., Griffin, J. A., and Wells, R. D.,** Non-B right-handed DNA conformations of homopurine · homopyrimidine sequences in the murine immunoglobulin Cα switch region, *J. Biol. Chem.*, 263, 7397, 1988.
87. **Collier, D. A. and Wells, R. D.,** Effect of length, supercoiling and pH on intramolecular triplex formation, *J. Biol. Chem.*, 265, 10652, 1990.
88. **Condemine, G. and Smith, C. L.,** Transcription regulates oxolinic acid induced DNA-gyrase cleavage at specific sited on the *E. coli* chromosome, *Nucleic Acids Res.*, 18, 7389, 1990.
89. **Conley, E. C. and West, S. C.,** Underwinding of DNA associated with duplex-duplex pairing by recA protein, *J. Biol. Chem.*, 265, 10156, 1990.
90. **Cook, D. N., Armstrong, G. A., and Hearst, J. E.,** Induction of anaerobic gene expression in *Rhodobacter capsulatus* is not accompanied by a local change in chromosomal supercoiling as measured by a novel assay, *J. Bacteriol.*, 171, 4836, 1989.
91. **Cooney, M., Czeruncziewicz, G., Postel, E. H., Flint, S. J., and Hogan, M. E.,** Site-specific oligonucleotide binding represses transcription of three human c-*myc* gene in vitro, *Science*, 241, 456, 1988.
92. **Courey, A. J., Plon, S. E., and Wang, J. C.,** The use of psoralene-modified DNA to probe the mechanism of enhancer action, *Cell*, 45, 567, 1986.
93. **Courey, A. J. and Wang, J. C.,** Influence of DNA sequence and supercoiling on the process of supercoil formation, *J. Mol. Biol.*, 202, 35, 1988.
94. **Cox, M. M. and Lehman, I. R.,** Enzymes of general recombination, *Annu. Rev. Biochem.*, 56, 229, 1987.
95. **Crick, F.,** General model for the chromosomes of higher organisms, *Nature*, 234, 25, 1971.
96. **Crick, F. H. C., Wang, J. C., and Bauer, W. R.,** Is DNA really a double helix?, *J. Mol. Biol.*, 129, 449, 1979.
97. **Crothes, D. M.,** The kinetics of DNA denaturation, *J. Mol. Biol.*, 9, 712, 1964.
98. **Crute, F., Morawski, G., and Lehman, I. R.,** A DNA helicase induced by herpes simplex type 1, *Nucleic Acids Res.*, 16, 6585, 1988.
99. **Culotta, V. and Sollner-Webb, B.,** Sites of topoisomerase I action on *X. laevis* ribosomal chromatin: transcriptionally active rDNA has a 200 bp repeating unit, *Cell*, 52, 585, 1988.
100. **Cunningham, R. P., Shibata, T., DasGupta, C., and Radding, C. M.,** Single strands induce recA protein to unwind duplex DNA for homologous pairing, *Nature*, 281, 191, 1979.
101. **Cyriax, B. and Gaeth, R.,** The conformation of double-stranded DNA, *Naturwissenschaften*, 65, 106, 1978.
102. **DasGupta, C., Shibata, T., Cunningham, R. P., and Radding, C. M.,** The topology of homologous pairing promoted by recA protein, *Cell*, 22, 437, 1980.
103. **DasGupta, S., Masukata, H., and Tomizawa, J.,** Multiple mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H, *Cell*, 51, 1113, 1987.
104. **Davis, T. L., Firulli, A. B., and Kinniburgh, A. J.,** Ribonucleoprotein and protein factors bind to a H-DNA forming c-*myc* element forming DNA: possible regulators of the c-*myc* gene, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 9682, 1989.
105. **Davison, P. F.,** The rate of strand separation in alkali treated DNA, *J. Mol. Biol.*, 22, 97, 1966.

106. **de los Santos, C., Rosen, M., and Patel, D.,** NMR studies of DNA (R+)n · (Y−)n · (Y+)n triple helices in solution: imino and amino proton markers of T · A · T and G · C · G+ base-triple formation, *Biochemistry*, 28, 7282, 1989.
107. **Dean, F. B., Borowiec, J. A., Ishimi, Y., Deb, S., Tegtmeyer, P., and Hurwitz, J.,** Simian virus 40 large tumor antigen required three core replication origin domains for DNA unwinding and replication in vitro, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 8267, 1987.
108. **Dean, F. B., Bullock, P., Murakami, Y., Wobbe, C. R., Weissbach, L., and Hurwitz, J.,** Simian virus 40 (SV40) DNA replication: SV40 large T antigen unwinds DNA containing the SV40 origin of replication, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 16, 1987.
109. **Dean, F. B., Dodson, M., Borowiec, J. A., Ishimi, Y., Goetz, G. S., Bullock, P., Matson, S. W., Echols, H., and Hurwitz, J.,** SV40 origin-dependent DNA unwinding and nucleoprotein complex formation by SV40 origin-dependent DNA unwinding and nucleoprotein complex formation by SV40 T antigen, *Cancer Cells*, 6, 113, 1988.
110. **Delbruck, M.,** On the replication of DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 40, 783, 1954.
111. **Depew, R. E. and Wang, J. C.,** Conformational fluctuations of DNA helix, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 4275, 1975.
112. **Di Capua, E., Schnarr, M., Ruigrok, R. W. H., Linder, P., and Timmins, P. A.,** Complexes of recA protein in solution: a study by small angle neutron scattering, *J. Mol. Biol.*, 214, 557, 1990.
113. **Di Capua, E., Stasiak, A., Koller, Th., Brahms, S., Thomae, F., and Pohl, F. M.,** Torsional stress induces left handed helical stretches in DNA of matched base sequence: circular dichroism and antibody binding, *EMBO J.*, 2, 1531, 1983.
114. **Dickerson, R. E.,** The DNA helix and how it is read, *Sci. Am.*, 249 (6), 94, 1983.
115. **DiGate, R. J. and Marians, K. J.,** Identification of a potent decatenating enzyme from *E. coli*, *J. Biol. Chem.*, 263, 13366, 1988.
116. **DiGate, R. J. and Marians, K. J.,** Molecular cloning and DNA sequence analysis of *E. coli* topB, the gene encoding topoisomerase III, *J. Biol. Chem.*, 264, 17924, 1989.
117. **DiNardo, S., Voelkel, K., Sternglanz, R., Reynolds, A. E., and Wright, A.,** *E. coli* DNA topoisomerase I mutants have compensatory mutations at or near DNA gyrase genes, *Cold Spring Harbor Symp. Quant. Biol.*, 42, 779, 1983.
118. **Dodson, M., Dean, F. B., Bullock, P., Echols, H., and Hurwitz, J.,** Unwinding of duplex DNA from the SV40 origin of replication by the T antigen, *Science*, 238, 964, 1987.
119. **Dodson, M., Echols, H., Wickner, S., Alfano, C., Mensa-Wilmot, K., Gomes, B., LeBowitz, J., Roberts, J. D., and McMacken, R.,** Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: localized unwinding of duplex DNA by a six-protein reaction, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 7638, 1986.
120. **Dodson, M., McMacken, R., and Echols, H.,** Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda, *J. Biol. Chem.*, 264, 10719, 1989.
121. **Drlica, K. and Franco, R. J.,** Inhibitors of DNA topoisomerases, *Biochemistry*, 27, 2253, 1988.
122. **Drlica, K. and Rouviere-Yaniv, J.,** Histonelike proteins of bacteria, *Microbiol. Rev.*, 51, 301, 1987.
123. **Dulbecco, R. and Vogt, M.,** Evidence for a ring structure of polyoma virus DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 50, 236, 1963.
124. **Dybvig, K., Clark, Ch. D., Aliperti, G., and Schlesinger, M. J.,** A Chicken repetitive DNA sequence that is highly sensitive to single-strand specific endonucleases, *Nucleic Acids Res.*, 11, 8495, 1983.
125. **Earnshaw, W. C., Hallison, B., Cooke, C. A., Heck, M. M. S., and Liu, L. F.,** Structural components of mitotic chromosome scaffolds, *J. Cell Biol.*, 100, 1706, 1985.
126. **Earnshaw, W. C. and Heck, M.,** The cell biology of topoisomerase II, *Cancer Cells*, 6, 279, 1988.
127. **Eguchi, Y. and Tomizawa, J.,** Complex formed by complementary RNA stem-loops and its stabilization by a protein: function of ColE1 rom protein, *Cell*, 60, 199, 1990.
128. **Ellison, M. J., Fenton, M. J., Ho, P. S., and Rich, A.,** Long-range interactions of multiple DNA structural transitions within a common topological domain, *EMBO J.*, 6, 1513, 1987.
129. **Evans, T. and Efstratiadis, A.,** Sequence-dependent S1 nuclease hypersensitivity of a heteronomous DNA duplex, *J. Biol. Chem.*, 261, 14771, 1986.
130. **Evans, T., Schon, E., Gora-Maslak, G., Patterson, J., and Efstratiadis, A.,** S1-hypersensitive sites in eukaryotic promoter regions, *Nucleic Acids Res.*, 12, 8043, 1984.
131. **Feldman, M. Ya.,** Reaction of nucleic acids and nucleoproteins with formaldehyde, *Prog. Nucleic Acids Mol. Biol.*, 13, 1, 1974.
132. **Felsenfeld, G. and Miles, H. T.,** The physical and chemical properties of nucleic acids, *Annu. Rev. Biochem.*, 36, 407, 1967.
133. **Felsenfeld, G. and Rich, A.,** Studies on the formation of two and three stranded polyribonucleotides, *Biochim. Biophys. Acta*, 26, 457, 1957.
134. **Fiers, W. and Sinsheimer, R. L.,** The structure of the DNA of the bacteriophage φx174, *J. Mol. Biol.*, 5, 408, 1962.
135. **Financsek, I., Tora, L., Kelemen, G., and Hidvegi, E.,** Supercoil induced D1 hypersensitive sites in the rat and human ribosomal RNA genes, *Nucleic Acids Res.*, 14, 3263, 1986.
136. **Finer, M. H., Aho, S., Gerstenfeld, L. C., Boedtger, H., and Doty, P.,** Unusual DNA se-

quences located within the promoter region and the first intron of the chicken pro-alfa(I) collagen gene, *J. Biol. Chem.*, 262, 13323, 1987.

137. **Finer, M. H., Fodor, Boedtger, H., and Doty, P.,** Endonuclease S1 sensitive site in chicken pro-$\alpha_2$(I) collagen 5′ flanking gene region, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 1659, 1984.
138. **Fleischmann, G., Pflugfelder, G., Steiner, E. K., Javaherian, K., Howard, G. C., Wang, J. C., and Elgin, S. C. R.,** Drosophila DNA topol in association with transcriptionally active regions in the genome, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 6958, 1984.
139. **Fowler, R. F. and Skinner, D. M.,** Eukaryotic DNA diverges at a long and complex pyrimidine · purine tract that can adopt altered conformations, *J. Biol. Chem.*, 261, 8994, 1986.

139a. **Fox, K. R.,** Long (dA)n · (dT)n tracts can form intramolecular triplexes under superhelical stress *Nucleic Acids Res.*, 18, 5387, 1990.

140. **Fox, J. J. and Shugar, D.,** Spectrophotometric studies of nucleic acid derivatives and related compounds. I. Natural and synthetic pyrimidine nucleosides, *Biochim. Biophys. Acta*, 9, 369, 1952.
141. **Francois, J.-C., Saison-Behmoaras, T., and Helene, C.,** Sequence-specific recognition of the major groove of DNA by oligodeoxynucleotides via triple helix formation. Footprinting studies, *Nucleic Acids Res.*, 16, 11431, 1988.
142. **Frank-Kamenetskii, M. and Lazurkin, Y.,** Conformational changes in DNA molecules, *Annu. Rev. Biophys. Bioeng.*, 3, 127, 1974.
143. **Frank-Kamenetskii, M.,** The turn of the quadruplex?, *Nature*, 342, 737, 1989.
144. **Fraser, M. J.,** Purification and properties of *N. crassa* endo-exonuclease — an enzyme which can be converted into a single strand specific endonuclease, *Meth. Enzymol.*, 65, 255, 1980.
145. **Furlong, J. C. and Lilley, D. M. J.,** Highly selective chemical modification of cruciform loops by diethyl pyrocarbonate, *Nucleic Acids Res.*, 14, 3995, 1986.
146. **Furuichi, Y. and Miura, K.,** A blocked structure at the 5′ terminus of RNA from cytoplasmic polyhedrosis virus, *Nature*, 253, 374, 1975.
147. **Galazka, G., Palecek, E., Wells, R. D., and Klysik, J.,** Site-specific $OsO_4$ modification of the B-Z junctions formed at the $(dA\text{-}dC)_{32}$ region in supercoiled DNA, *J. Biol. Chem.*, 261, 7093, 1986.
148. **Gamow, G.,** Topological properties of coiled helical systems, *Proc. Natl. Acad. Sci. U.S.A.*, 41, 7, 1955.
149. **Gamper, H. B. and Hearst, J. E.,** A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation and ternary complexes, *Cell*, 29, 81, 1982.
150. **Gamper, H. B. and Hearst, J. E.,** Size of the unwound region of DNA in *E. coli* RNA polymerase and Calf thymus RNA polymerase II ternary complexes, *Cold Spring Harbor Symp. Quant. Biol.*, 42, 455, 1983.
151. **Gasser, S. M. and Laemmli, U. K.,** A glimpse at chromosomal order, *Trends Genet.*, 3, 16, 1987.
152. **Gellert, M.,** DNA topoisomerases, *Annu. Rev. Biochem.*, 50, 879, 1981.
153. **Gellert, M., Mizuuchi, K., O'Dea, M. H., and Nash, H. A.,** DNA gyrase: an enzyme that introduces superhelical turns into DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 73, 3872, 1976.
154. **Gellert, M., O'Dea, M. H., and Mizuuchi, K.,** Slow cruciform transitions in palindromic DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 5545, 1983.
155. **Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M., and Chambon, P.,** Folding of the DNA double helix in chromatin like structures from SV40, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 1843, 1975.
156. **Giaever, G. N., Snyder, L., and Wang, J. C.,** DNA supercoiling in vivo, *Biophys. Chem.*, 29, 7, 1988.
157. **Giaever, G. N. and Wang, J. C.,** Supercoiling of intracellular DNA can occur in eukaryotic cells, *Cell*, 55, 849, 1988.
158. **Gierer, A.,** Model for DNA and protein interactions and the function of the operator, *Nature*, 212, 1480, 1966.
159. **Gilmour, D. and Elgin, S.,** Localization of specific topoisomerase I interactions within the transcribed regions of heat shock genes using campthotecin as inhibitor, *Mol. Cell. Biol.*, 7, 141, 1987.
160. **Gilmour, D. S., Pflugfelder, G., Wang, J. C., and Lis, J. T.,** Topoisomerase I interacts with transcribed regions in Drosophila cells, *Cell*, 44, 401, 1986.
161. **Glaser, R. L., Thomas, G. H., Siegfried, E., Elgin, S. C. R., and Lis, J. T.,** Optimal heat-induced expression of the Drosophila hsp26 gene requires a promoter sequence containing $(CT_n \cdot (GA)_n$ repeats, *J. Mol. Biol.*, 211, 751, 1990.
162. **Glikin, G. C., Gargiulo, G., Rena-Descalzi, L., and Worcel, A.,** *Escherichia coli* single-strand binding protein stabilizes specific denatured sites in superhelical DNA, *Nature*, 303, 770, 1983.
163. **Glikin, G. C., Ruberti, I., and Worcel, A.,** Chromatin assembly in *Xenopus* oocytes: in vitro studies, *Cell*, 37, 33, 1984.
164. **Glover, J. N. M. and Pulleyblank, D. E.,** Protonated polypurine/polypyrimidine DNA tracts that appear to lack the single-stranded pyrimidine loop predicted by the "H" model, *J. Mol. Biol.*, 215, 653, 1990.
165. **Godding, C. R. and Russel, W. C.,** S1 sensitive sites in adenovirus DNA, *Nucleic Acids Res.*, 11, 21, 1983.
166. **Goetz, G. S., Dean, F. B., Hurwitz, J., and Matson, S. W.,** The unwinding of duplex region in DNA by the SV40 large tumor antigen-associated DNA helicase activity, *J. Biol. Chem.*, 263, 383, 1988.
167. **Gogos, J. A., Karayiorgou, M., Aburatani, H., and Kafatos, F. C.,** Detection of single base mismatches of T and C residues by $KMnO_4$ and hy-

droxylamine in presence of tetraalkyl ammonium salts, *Nucleic Acids Res.*, 18, 6807, 1990.
168. **Gorski, F.,** Critical review of the mechanisms separating the strands of DNA double helices during cell division, *Folia Biol.*, 23, 81, 1975.
169. **Goto, T. and Wang, J. C.,** Cloning of yeast TOP1, the gene encoding topoisomerase I and and construction of mutants defective in both DNA topoisomerase I and topoisomerase II, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 7178, 1985.
170. **Gotoh, O.,** Prediction of melting profiles and local helix stability for sequenced DNA, *Adv. Biophys.*, 16, 1, 1983.
171. **Gray, D. M. and Cui, T.,** Circular dichroism measurements show that C · $C^+$ base pairs can coexist with A · T base pairs between antiparallel strands of an oligodeoxynucleotide double-helix, *Nucleic Acids Res.*, 12, 7565, 1984.
172. **Gray, D. M., Morgan, A. R., and Ratliff, R. L.,** A comparison of CD spectra of synthetic DNA sequences of the homopurine · homopyrimidine types, *Nucleic Acids Res.*, 5, 3679, 1978.
173. **Greaves, D. R. and Borst, P.,** Trypanosoma brucei variant-specific glycoprotein gene chromatin is sensitive to single-strand-specific endonuclease digestion, *J. Mol. Biol.*, 197, 471, 1987.
174. **Griffith, J. D. and Harris, L. D.,** DNA strand exchanges, *Crit. Rev. Biochem.*, 23, S61, 1988.
175. **Gromova, M. I., Buchman, V. L., Abagyan, R. A., Ulyanov, A. V., and Bronstein, I. B.,** Sequence-dependent modulating effet of camptothecin on the DNA-cleaving activity of the calf thymus type I topoisomerase, *Nucleic Acids Res.*, 18, 637, 1990.
176. **Gross, D. S. and Garrard, W. T.,** Poising chromatin for transcription, *Trends Biochem. Sci.*, 12, 293, 1987.
177. **Guttierrez, C., Guo, Z.-S., Roberts, J., and DePamphilis, M. L.,** Simian virus 40 origin auxiliary sequences weakly facilitate T-antigen binding but strongly facilitate DNA unwinding, *Mol. Cell. Biol.*, 10, 1719, 1990.
178. **Hamada, H. and Bustin, M.,** Hierarchy of binding sites for chromosomal proteins HMG 1 and 2 in supercoiled deoxyribonucleic acid, *Biochemistry*, 24, 1428, 1985.
179. **Hames, B. D. and Glover, D. M.,** *Transcription and Splicing*, IRL Press, London, 1989.
180. **Han, S., Udvardy, A., and Schedl, P.,** Transcriptionally active chromatin is sensitive to *N. crassa* and S1 nucleases, *J. Mol. Biol.*, 179, 469, 1984.
181. **Han, S., Udvardy, A., and Schedl, P.,** Novobiocin blocks the Drosophila heat shock response, *J. Mol. Biol.*, 183, 13, 1985.
182. **Haner, R. and Dervan, P. B.,** Single-strand DNA triplex-helix formation, *Biochemistry*, 29, 9761, 1990.
183. **Haniford, D. B. and Pulleyblank, D. E.,** Facile transition of poly{d(TG)d(CA)} into a left-handed helix in physiological conditions, *Nature*, 302, 632, 1983.
184. **Hanvey, J. C., Klysik, J., and Wells, R. D.,** Influence of DNA sequence on the formation of non-B right-handed helices in oligopurine · oligopyrimidine inserts in plasmids, *J. Biol. Chem.*, 263, 7386, 1988.
185. **Hanvey, J. C., Shimizu, M., and Wells, R. D.,** Intramolecular DNA triplexes in supercoiled plasmids, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 6292, 1988.
186. **Hanvey, J. C., Shimizu, M., and Wells, R. D.,** Intramolecular DNA triplexes in supercoiled plasmids. II. Effect of base composition and noncentral interruptions on formation and stability, *J. Biol. Chem.*, 264, 5950, 1989.
187. **Harland, R. M., Weintraub, H., and McKnight, S. L.,** Transcription of DNA injected into *Xenopus* oocytes is influenced by template topology, *Nature*, 302, 38, 1983.
188. **Hartman, K. A. and Rich, A.,** The tautomeric form of helical polyribocytidylic acid, *J. Am. Chem. Soc.*, 87, 2033, 1965.
189. **Hatzopoulos, A. K., Schlokat, U., and Gruss, P.,** Enhancers and other cis-acting regulatory sequences, in *Transcription and Splicing*, Hames, B. D. and Glover, D. M., Eds., IRL Press, London, 1989, 143.
190. **Hayes, T. E. and Dixon, J. E.,** Z-DNA in the rat somatostatin gene, *J. Biol. Chem.*, 260, 8145, 1985.
191. **Helinski, D. R. and Clewell, D. B.,** Circular DNA, *Annu. Rev. Biochem.*, 40, 899, 1971.
192. **Henderson, E., Hardin, Ch. C., Walk, S. K., Tinoco, I., Jr., and Blackburn, E. H.,** Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine · gunaine base pairs, *Cell*, 51, 899, 1987.
193. **Hendrickson, T. A., Fritze, C. E., Folk, W. R., and DePamphilis, M. L.,** The origin of bidirectional DNA replication in polyoma virus, *EMBO J.*, 6, 2011, 1987.
194. **Hentschel, Ch. C.,** Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to $S_1$ nuclease, *Nature*, 295, 714, 1982.
195. **Herr, W.,** Diethyl pyrocarbonate: a chemical probe for secondary structure in negatively supercoiled DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 8009, 1985.
196. **Herr, W. and Clarke, J.,** The SV40 enhancer is composed of multiple functional elements that can compensate for one another, *Cell*, 45, 461, 1986.
197. **Herrick, G., Deluis, H., and Alberts, B.,** Single-stranded DNA structure and DNA polymerase activity within presence of nucleic acid helix unwinding proteins from calf thymus, *J. Biol. Chem.*, 251, 2142, 1976; (Cf., however, Chase and Williams, 1986, pp. 128–132.)
198. **Higgins, C. F., Dorman, C. J., Stirling, D. A., Waddell, L., Booth, I. R., May, G., and Bremer, E.,** A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*, *Cell*, 52, 569, 1988.
199. **Hinton, D. M. and Nossal, N. G.,** Bacteriophage T4 primase-helicase. Characterization of oligomer

synthesis by T4 61 protein alone and in conjunction with the T4 41 protein, *J. Biol. Chem.*, 262, 10873, 1987.

200. **Hirose, S. and Suzuki, Y.,** In vitro transcription of eukaryotic genes is affected differently by the degree of DNA supercoiling, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 718, 1988.
201. **Hoffman, E. K., Trusko, S. P., Murphy, M., and George, D. L.,** An S1 nuclease-sensitive homopurine/homopyrimidine domain in the c-Ki-ras promoter interacts with a nuclear factor, *Proc. Natl. Acad. Sci. U.S.A.*, 87, 2705, 1990.
202. **Hoheisel, J. D. and Pohl, F. M.,** Searching for potential Z DNA in genomic *E. coli* DNA, *J. Mol. Biol.*, 193, 447, 1987.
203. **Honigberg, S. M. and Radding, C. M.,** The mechanics of winding and unwinding helices in recombination: torsional stress associated with strand transfer promoted by RecA protein, *Cell*, 54, 525, 1988.
204. **Hsiang, Y.-H., Hertzberg, R., Hecht, S., and Liu, L.,** Campthotecin induces protein linked DNA breaks via mammalian topoisomerase I, *J. Biol. Chem.*, 260, 14783, 1985.
205. **Htun, H. and Dahlberg, J. E.,** Single strands, triple strands, and kinks in H-DNA, *Science*, 241, 1791, 1988.
206. **Htun, H. and Dahlberg, J. E.,** Topology and formation of triple-stranded H-DNA, *Science*, 243, 1571, 1989.
207. **Htun, H. and Dahlberg, J. E.,** The structure and formation of triple stranded H-DNA, in *Structure and Methods*, Vol. 3, *DNA & RNA* Sarma, M. H. and Sarma, R. H., Eds., 1990, 185.
208. **Htun, H., Lund, E., and Dahlberg, J. E.,** Human U1 RNA genes contain an unusually sensitive nuclease S1 cleavage site within the conserved 3′ flanking region, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 7288, 1984.
209. **Htun, H., Lund, E., Westin, G., Pettersson, U., and Dahlberg, J. E.,** Nuclease S1-sensitive sites in multigene families: human U2 small nuclear RNA genes, *EMBO J.*, 4, 1839, 1985.
210. **Huber, H. E., Bernstein, J., Nakai, H., Tabor, S., and Richardson, C. C.,** Interactions of DNA replication proteins of bacteriophage T7, *Cancer Cells*, 6, 11, 1988.
211. **Husain, I., van Houten, B., Thomas, D. C., Abdul-Monem, M., and Sancar, A.,** Effect of DNA polymerase I and DNA helicase II on the turnover rate of UvrABC excision nuclease, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 6774, 1985.
212. **Iaconno-Connors, L. and Kowalski, D.,** Altered conformations in gene regulatory regions of torsionally stressed SV40 DNA, *Nucleic Acids Res.*, 14, 8949, 1986.
213. **Igo-Kemenes, T., Horz, W., and Zachau, H.,** Chromatin, *Annu. Rev. Biochem.*, 51, 102, 1982.
214. **Inman, R. B.,** Transition of DNA homopolymes, *J. Mol. Biol.*, 9, 624, 1964.
215. **Inman, R. B. and Schnos, M.,** Partial denaturation of thymic and 5 Br uracil containing lamda DNA in alkali, *J. Mol. Biol.*, 49, 93, 1970; 28, 103, 1967.
216. **Javaherian, K. and Liu, L. F.,** Association of eukaryotic topoisomerase I with nucleosomes and chromosomal proteins, *Nucleic Acids Res.*, 11, 461, 1983.
217. **Jaxel, C., Kohn, K. W., and Pommier, Y.,** Topoisomerase I interaction with SV40 DNA in the presence and absence of camptothecin, *Nucleic Acids Res.*, 23, 11166, 1988.
218. **Johnson, A. C., Jinno, Y., and Merlino, G. T.,** Modulation of EGF proto-oncogene transcription by an a promoter site sensitive to S1 nuclease, *Mol. Cell. Biol.*, 8, 4174, 1988.
219. **Johnson, P. F. and McKnight, S. L.,** Eukaryotic transcriptional regulatory proteins, *Annu. Rev. Biochem.*, 58, 799, 1989.
220. **Johnston, B. H.,** The S1-sensitive form of $d(C-T)_n \cdot d(A-G)_n$: chemical evidence for a three-stranded structure in plasmids, *Science*, 241, 1800, 1988.
221. **Johnston, B. H. and Rich, A.,** Chemical probes of DNA conformation: detection of Z-DNA at nucleotide resolution, *Cell*, 42, 713, 1985.
222. **Jong, A. Y.-S. and Campbell, L.,** Isolation of the gene encoding yeast single stranded nucleic acid binding protein 1, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 877, 1986.
223. **Kalambet, Yu. A., Borovik, A. S., Lyamichev, V. I., and Lyubchenko, L.,** Electron microscopy of the melting of sequenced DNA, *Biopolymers*, 24, 359, 1985.
224. **Kane, C. M.,** Renaturase and Ribonuclease H: a novel mechanism that influences transcript displacement by RNA polymerase II in vitro, *Biochemistry*, 27, 3187, 1988.
225. **Kang, D. S. and Wells, R. D.,** B-Z DNA junctions contain few, if any, nonpaired bases at physiological superhelical densities, *J. Biol. Chem.*, 260, 7783, 1985.
226. **Kavenoff, R., Klotz, L. C., and Zimm, D. H.,** On the nature of chromosome sized DNA molecules, *Cold Spring Harbor Symp. Quant. Biol.*, 38, 1, 1973.
227. **Keller, W.,** Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 4876, 1975.
228. **Kenny, M. K., Lee, S.-H., and Hurwitz, J.,** Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 9757, 1989.
229. **Kilpatrick, M. W., Klysik, J., Singleton, Ch. K., Zarling, D. A., Jovin, T. M., Hanau, L. H., Erlanger, B. F., and Wells, R. D.,** Intervening sequences in human fetal globin genes adopt left-handed Z helices, *J. Biol. Chem.*, 259, 7268, 1984.
230. **Kilpatrick, M. W., Torri, A., Kang, D. S., Engler, J. A., and Wells, R. D.,** Unusual DNA structures

in the adenovirus genome, *J. Biol. Chem.*, 261, 11350, 1986.

231. **Kilpatrick, M. W., Wei, C.-F., Gray, B., Jr., and Wells, R. D.,** BAL 31 nuclease as a probe in concentrated salt for the B-Z DNA junction, *Nucleic Acids Res.*, 11, 3811, 1983.
232. **Kinniburgh, A. J.,** A cys acting transcription element of the c-myc gene can assume an H-DNA conformation, *Nucleic Acids Res.*, 17, 7771, 1989.
233. **Kirkegaard, K., Buc, H., Spassky, A., and Wang, J. C.,** Mapping of single-stranded regions in duplex DNA at the sequence level: single-strand-specific cytosine methylation in RNA polymerase-promoter complexes, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 2544, 1983.
234. **Kirkegaard, K. and Wang, J. C.,** Mapping the topography of DNA wrapped around gyrase by nucleolytic and chemical probing of complexes of unique DNA sequences, *Cell*, 23, 721, 1981.
235. **Kirshenbaum, M. R., Tribolet, R., and Barton, J. K.,** $Rh(DIP)_3^{(3+)}$: a shape-selective metal complex which targets cruciforms, *Nucleic Acids Res.*, 16, 7943, 1988.
236. **Kjeldsen, E., Mollerup, S., Thomsen, B., Bonven, B. J., Bolund, L., and Westergaard, O.,** Sequence-dependent effect of campthotecin on human topoisomerase I DNA cleavage, *J. Mol. Biol.*, 202, 333, 1988.
237. **Klausing, K., Scheidtmann, K. H., Baumann, E. A., and Knippers, R.,** Effects of *in vitro* dephosphorylation on DNA binding and DNA helicase activities of SV40 large tumor antigen, *J. Virol.*, 62, 1258, 1988.
238. **Klug, A., Jack, A., Visamitra, M. A., Kennard, O., Shakked, Z., and Steitz, T. A.,** A hypothesis on a sequence specific-dependent conformation of DNA and its relation to the binding of the lac repressor protein, *J. Mol. Biol.*, 131, 669, 1979.
239. **Klug, A. and Lutter, L.,** The helical periodicity of DNA on the nucleosome, *Nucleic Acids Res.*, 9, 4267, 1981.
240. **Klug, A. and Travers, A. A.,** The helical repeat of nucleosome wrapped DNA, *Cell*, 56, 11, 1989.
241. **Kmiec, E. B.,** Transcription factor IIIA and 5S DNA supercoiling in vitro, *Cell*, 54, 919, 1988.
242. **Kmiec, E. B., Ryoji, M., and Worcel, A.,** Gyration is required for 5S RNA transcription from a chromatin template, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 1305, 1986.
243. **Kmiec, E. G. and Worcel, A.,** The positive transcription factor of the 5S RNA gene induces a 5S DNA specific gyration in *Xenopus* oocyte extracts, *Cell*, 41, 945, 1985.
244. **Kohwi, Y. and Kohwi-Shigematsu, T.,** Magnesium ion-dependent triple-helix structure formed by homopurine-homopyrimidine sequences in supercoiled plasmid DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 3781, 1988.
245. **Kohwi-Shigematus, T., Gelinas, R., and Weintraub, H.,** Detection of an altered DNA conformation at specific sites in chromatin and supercoiled DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 4389, 1983.
246. **Kohwi-Shigematsu, T. and Kohwi, Y.,** Poly(dG)-Poly(dC) sequences, under torsional stress, induce an altered DNA conformation upon neighboring DNA sequences, *Cell*, 43, 199, 1985.
247. **Konopka, A. K., Reiter, J., Jung, M., Zarling, D. A., and Jovin, T. M.,** Concordance of experimentally mapped or predicted Z- sites with positions of selected alternating purine-pyrimidine tracts, *Nucleic Acids Res.*, 13, 1683, 1985.
248. **Koo, H.-S. and Crothes, D. M.,** Calibration of DNA curvature and a unified description of sequence-directed bending, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 1763, 1988.
249. **Kornberg, R. D. and Klug, A.,** The nucleosome, *Sci. Am.*, 244 (2), 52, 1981.
250. **Kowalczykowski, S. C.,** Mechanistic aspects of the DNA strand exchange activity of *E. coli* recA protein, *Trends Biochem. Sci.*, 12, 141, 1987.
251. **Kowalski, D.,** Changes in site specificity of single-strand-specific endonucleases on supercoiled PM2 DNA with temperature and ionic environment, *Nucleic Acids Res.*, 12, 7071, 1984.
252. **Kowalski, D. and Eddy, M. J.,** The DNA unwinding element: a novel, cis acting component that facilitates opening of the *E. coli* replication origin, *EMBO J.*, 8, 4335, 1989.
253. **Kowalski, D., Kroecker, W. D., and Laskowski, M., Sr.,** Mungbean nuclease. I. Physical, chemical and catalytic properties, *Biochemistry*, 15, 4457, 1976.
254. **Kowalski, D., Natale, D. A., and Eddy, M. J.,** Stable DNA unwinding, not breathing, accounts for single-strand-specific nuclease hypersensitivity of specific A + T rich sequences, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 9464, 1988.
255. **Kroeger, P.E. and Rowe, T. C.,** Interaction of topoisomerase with the transcribed region of the Drosophila HSP 70 heat shock gene, *Nucleic Acids Res.*, 17, 8495, 1989.
256. **Kubo, M., Kano, Y., Nakamura, H., Nagata, A., and Imamoto, F.,** In vivo enhancement of general and specific transcription in *E. coli* by DNA gyrase activity, *Gene*, 7, 153, 1979.
257. **Kusmierek, J. T. and Singer, B.,** Chloroacetaldehyde treated ribo- and deoxyribopolynucleotides. I. Reaction products, *Biochemistry*, 21, 5717, 1982.
258. **Lahue, E. E. and Matson, S. W.,** *E. coli* helicase I catalyzes a unidirectional and highly processive unwiding reaction, *J. Biol. Chem.*, 263, 3208, 1988.
259. **Lapidot, A., Baran, N., and Manor, H.,** $(dT\text{-}dC)_n$ and $(dG\text{-}dA)_n$ tracts inserted in M13, arrest single stranded DNA replication in vitro, *Nucleic Acids Res.*, 17, 883, 1989.
260. **Larsen, A. and Weintraub, H.,** An altered DNA

conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin, *Cell*, 29, 609, 1982.
261. **Lasken, R. S. and Kornberg, A.,** The primosomal protein n′ of *E. coli* is a DNA helicase, *J. Biol. Chem.*, 263, 5512, 1988.
262. **Lebowitz, J., Chaudhuri, A. K., Gonenne, A., and Kitos, G.,** Carbodiimide modification of superhelical PM2 DNA: considerations regarding reaction at unpaired bases and the unwinding of superhelical DNA with chemical probes, *Nucleic Acids Res.*, 3, 1695, 1977.
263. **LeBowitz, J. H. and McMacken, R.,** The *E. coli* dnaB replication protein is a DNA helicase, *J. Biol. Chem.*, 261, 4738, 1986.
264. **Lee, C.-H., Mizusawa, H., and Kakefuda, T.,** Unwinding of double-stranded DNA helix by dehydration, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 2838, 1981.
265. **Lee, F. S. and Bauer, W. R.,** Temperature dependence of the gel electrophoretic mobility of superhelical DNA, *Nucleic Acids Res.*, 13, 1665, 1985.
266. **Lee, J. S.,** The stability of polypurine tetraplexes in the presence of mono- and divalent cations, *Nucleic Acids Res.*, 18, 6057, 1990.
267. **Lee, J. S., Johnson, D. A., and Morgan, A. R.,** Complexes formed by $(pyrimidine)_n \cdot (purine)_n$ DNAs on lowering the pH are three-stranded, *Nucleic Acids Res.*, 6, 3073, 1979.
268. **Lee, J. S., Woodsworth, M. L., Latimer, J. P., and Morgan, A. R.,** Polypurine · polypyrimidine synthetic DNAs containing 5-methyl cytosine forms stable triplexes at neutral pH, *Nucleic Acids Res.*, 12, 6603, 1984.
269. **Lee, M. S. and Marians, K. J.,** *E. coli* replication factor Y, a component of the primosome, can act as a helicase, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 8345, 1987.
270. **Lee, M. S. and Marians, K. J.,** Differential ATP requirements distinguish the DNA translocation and DNA unwinding activities of the *E. coli* PRI A protein, *J. Biol. Chem.*, 265, 17078, 1990.
271. **Legerski, R. J., Hodnett, H. L., and Gray, H. B.,** Extracellular nucleases of Pseudomonas Bal31, *Nucleic Acids Res.*, 5, 1445, 1978.
272. **Leonard, N. J., McDonald, J. J., Henderson, R. E. L., and Eichmann, M. E.,** Reaction of diethyl pyrocarbonate with nucleic acid components. Adenosine, *Biochemistry*, 10, 3335, 1971.
273. **Lerman, L. S., Fischer, S. G., Hurley, I., Silverstein, K., and Lumelsky, N.,** Sequence-determined DNA separations, *Annu. Rev. Biophys. Bioeng.*, 13, 399, 1984.
274. **Lerman, L. S., Fischer, S. G., Hurley, I., and Silverstein, K.,** Computational simulation of DNA melting and its application to density gradient gel electrophoresis, *Meth. Enzymol.*, 155, 482, 1987.
275. **Lescure, B., Chestier, A., and Yaniv, M.,** Transcription of polyoma virus in vitro. Transcription of superhelical and linear polyoma DNA by RNA polymerase II, *J. Mol. Biol.*, 124, 73, 1978.
276. **Levinthal, C. and Crane, H. R.,** On the unwinding of DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 42, 436, 1956.
277. **Levitt, M.,** Computer simulation of double helix dynamics, *Cold Spring Harbor Symp. Quant. Biol.*, 47, 251, 1983.
278. **Lilley, D. M. J.,** The inverted repeat as a recognizable structural feature in supercoiled DNA molecules, *Proc. Natl. Acad. Sci. U.S.A.*, 77, 6468, 1980.
279. **Lilley, D. M. J.,** Dynamic, sequence-dependent DNA structure as exemplified by cruciform extrusion from inverted repeats in negatively supercoiled DNA, *Cold Spring Harbor Symp Quant. Biol.*, 47, 101, 1983.
280. **Lilley, D. M. J.,** The genetic control of DNA supercoiling, and vice versa, in *Regulation of Gene Expression — 25 years On*, Booth, I. R. and Higgins, C. F., Eds., 29th Symp. Soc. Gen. Microbiol., Cambridge University Press, 1986, 105.
281. **Lilley, D. M. J. and Palecek, E.,** The supercoil stabilized cruciform of ColE1 is hyper-reactive to osmium tetroxide, *EMBO J.*, 3, 1187, 1984.
282. **Lilley, D. M. J., Sullivan, K. M., Murchie, A. I. H., and Furlong, J. C.,** in *Unusual DNA Structures*, Wells, and Harvey, Eds., 1988, 55.
283. **Lindsley, J. E. and Cox, M. M.,** On RecA protein-mediated homologous alignment of two DNA molecules: three strands vs. four strands, *J. Biol. Chem.*, 165, 10164, 1990.
284. **Liu, L.,** DNA topoisomerases — enzymes that catalyze the breaking and rejoining of DNA, *Crit. Rev. Biochem.*, 15, 1, 1983.
285. **Liu, L. F.,** DNA topoisomerase poisons as antitumor drugs, *Annu. Rev. Biochem.*, 58, 351, 1989.
286. **Liu, L. F. and Wang, J. C.,** Supercoiling of the DNA template during transcription, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 7024, 1987.
287. **Lockshon, D. and Morris, D. R.,** Sites of reaction of *E. coli* DNA gyrase on pBR322 in vivo revealed by oxolinic acid-induced plasmid linearization, *J. Mol. Biol.*, 181, 63, 1985.
288. **Lohman, T. M., Bujalowski, W., and Overman, L. B.,** *E. coli* singlestrand binding protein: a new look at helix destabilizing proteins, *Trends Biochem. Sci.*, 13, 250, 1988.
289. **Longuet-Higgins, H. C., and Zimm, B. R.,** Calculation of the rate of unwinding of the DNA molecule, *J. Mol. Biol.*, 2, 1, 1960.
290. **Losick, R. and Chamberlin, M., Eds.,** *RNA Polymerase*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1976.
291. **Luchnik, A. N., Bakayev, V. V., Zbarsky, I. B., and Georgiev, G. P.,** Elastic torsional strain in DNA within a fraction of SV40 minichromosomes: relation to transcriptionally active chromatin, *EMBO J.*, 11, 1353, 1982.
292. **Lyamichev, V. I., Mirkin, S. M., Danilevskaya, O. N., Voloshin, O. N., Balatskaya, S. V.,**

**Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D.,** An unusual DNA structure detected in a telomeric sequence under superhelical stress and at low pH, *Nature*, 339, 634, 1989.

293. **Lyamichev, V. I., Mirkin, S. M., and Frank-Kamenetskii, M. D.,** A pH-dependent structural transition in the homopurine-homopyrimidine tract in superhelical DNA, *J. Biomol. Struct. Dyn.*, 3, 327, 1985.
294. **Lyamichev, V. I., Mirkin, S. M., and Frank-Kamenetskii, M. D.,** Structures of homopurine-homopyrimidine tract in superhelical DNA, *J. Biomol. Struct. Dyn.*, 3, 667, 1986.
295. **Lyamichev, V. I., Mirkin, S. M., and Frank-Kamenetskii, M. D.,** Structure of $(dG)_n \cdot (dC)_n$ under superhelical stress and acid pH, *J. Biomol. Struct. Dyn.*, 5, 275, 1987.
296. **Lyamichev, V. I., Mirkin, S. M., Frank-Kamenetskii, M. D., and Cantor, Ch. R.,** A stable complex between homopyrimidine oligomers and the homologous regions of duplex DNAs, *Nucleic Acids Res.*, 16, 2165, 1988.
297. **Lyamichev, V. I., Mirkin, S. M., Kumarev, V. P., Baranova, L. V., Vologodskii, A. V., and Frank-Kamenetskii, M. D.,** Energetics of the B-H transition in supercoiled DNA carrying $d(CT)_n \cdot d(AG)_n$ and $d(C)_n \cdot d(G)_n$ inserts, *Nucleic Acids Res.*, 17, 9417, 1989.
298. **Lyamichev, V. I., Panyutin, I. G., and Frank-Kamenetskii, M. D.,** Evidence of cruciform structures in superhelical DNA provided by two-dimensional gel electrophoresis, *FEBS Lett.*, 153, 299, 1983.
299. **Lyubchenko, Yu. L. and Shlyakhtenko, L. S.,** Early melting of supercoiled DNA, *Nucleic Acids Res.*, 16, 3269, 1988.
300. **Mace, H. A. F., Pelham, H. R. B., and Travers, A. A.,** Association of an $S_1$ nuclease-sensitive structure with short direct repeats 5′ of Drosophila heat shock genes, *Nature*, 304, 555, 1983.
301. **Malan, T. P., Kolb, A., Buc, H., and McClure, W. R.,** Mechanism of CRP-cAMP activation of lac operon transcription initiation: activation of the P1 promoter, *J. Mol. Biol.*, 180, 881, 1984.
302. **Maniatis, T., Goodbourn, S., and Fischer, J. A.,** Regulation of inducible and tissue-specific gene expression, *Science*, 236, 1237, 1987.
303. **Marcaud, H., Gabarro-Arpa, J., Ehrlich, R., and Reiss, C.,** An algorithm for studying cooperative transitions in DNA, *Nucleic Acids Res.*, 13, 551, 1985.
304. **Margolin, P., Zumstein, L., Sternglanz, R., and Wang, J. C.,** The *E. coli* supX locus is topA, the structural gene for DNA topoisomerase I, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 5437, 1985.
305. **Margot, J. B. and Hardison, R. C.,** DNase I and nuclease $S_1$ sensitivity of the rabbit β1 globin gene in nuclei and in supercoiled plasmids, *J. Mol. Biol.*, 184, 195, 1985
306. **Marians, K. J.,** Enzymology of DNA in replication in prokaryotes, *Crit. Rev. Biochem.*, 17, 153, 1984.
307. **Marsden, M. P. F. and Laemmli, U. K.,** Metaphase chromosome structure: evidence for a radial loop model, *Cell*, 17, 849, 1979.
308. **Masukata, H., DasGupta, S., and Tomizawa, J.,** Transcriptional activation of Col E1 DNA synthesis by displacement of nontranscribed strand, *Cell*, 51, 1123, 1987.
309. **Masukata, H. and Tomizawa, J.,** A mechanism of formation of a persistent hybrid between elongating RNA and template DNA, *Cell*, 62, 331, 1990.
310. **Matson, S. W.,** *E. coli* helicase II (uvrD gene product) catalyzes the unwinding of DNA · RNA hybrids in vitro, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 4430, 1989.
311. **Matson, S. W. and George, J. W.,** DNA helicase II of *E. coli*, *J. Biol. Chem.*, 262, 2066, 1987.
312. **Maxam, A. M. and Gilbert, W.,** Sequencing end labeled DNA with base specific chemical cleavages, *Meth. Enzymol.*, 65, 499, 1980.
313. **McCarthy, J. G. and Heywood, S. M.,** A long polypyrimidine/polypurine tract induces an altered DNA conformation on the 3′ coding region of the adjacent myosin heavy chain gene, *Nucleic Acids Res.*, 15, 8069, 1987.
314. **McClellan, J. A. and Lilley, D. M. J.,** A two-state conformational equilibrium for alternating $(A\text{-}T)_n$ sequences in negatively supercoiled DNA, *J. Mol. Biol.*, 197, 707, 1987.
315. **McClellan, J. A., Palecek, E., and Lilley, D. M. J.,** $(A\text{-}T)_n$ tracts embedded in random sequence DNA — formation of a structure which is chemically reactive and torsionally deformable, *Nucleic Acids Res.*, 14, 9291, 1986.
316. **McGhee, J. D. and Felsenfeld, G.,** Nucleosome structure, *Annu. Rev. Biochem.*, 49, 1115, 1980.
317. **McGhee, J. D. and von Hippel, P. H.,** Formaldehyde as probe for DNA structure. III. Equilibrium denaturation of DNA and of synthetic polynucleotides, *Biochemistry*, 15, 3276, 1977.
318. **McKeon, C., Schmidt, A., and de Crombrugghe, B.,** A sequence conserved in both the chicken and mouse α2(I) collagen promoter contains sites sensitive to $S_1$ nuclease, *J. Biol. Chem.*, 259, 6636, 1984.
319. **McLean, M. J., Larson, J. E., Wohlrab, F., and Wells, R. D.,** Reaction conditions affect the specificity of bromoacetaldehyde as a probe for DNA cruciforms and B-Z junctions, *Nucleic Acids Res.*, 15, 6917, 1987.
320. **McLean, M. J. and Wells, R. D.,** The role of DNA sequence in the formation of Z-DNA versus cruciforms in plasmids, *J. Biol. Chem.*, 263, 7370, 1988.
321. **McMaster, G. K. and Carmichael, G. G.,** Analysis of single and double stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange, *Proc. Natl. Acad. Sci. U.S.A.*, 74, 4835, 1977.
322. **Mei, Y.-Y. and Barton, J. K.,** Tris(tetramethylphenantroline)ruthenium(II): a chiral probe that

cleaves -DNA conformations, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 1339, 1988.

323. **Meiklejohn, A. L. and Gralla, J. D.,** Activation of the lac promoter and its variants. Synergistic effects of catabolite activator protein and supercoiling in vitro, *J. Mol. Biol.*, 207, 661, 1989.
324. **Melnikova, A. B., Beabealashili, R., and Mirzabekov, A. D.,** A study of unwinding of DNA and shielding of the DNA grooves by RNA polymerase by using methylation by dimethylsulfate, *Eur. J. Biochem.*, 84, 301, 1978.
325. **Mensa-Wilmot, K., Carroll, K., and McMacken, R.,** Transcriptional activation of bacteriophage lambda DNA replication in vitro: regulatory role of histone-like portein HU of *E. coli*, *EMBO J.*, 8, 2393, 1989.
326. **Menzel, R. and Gellert, M.,** Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling, *Cell*, 34, 105, 1983.
327. *Methods in Enzymology*, Vol. 65, *Single-Strand Specific Nucleases*, 1980, 248.
328. **Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T., and Bode, J.,** Hierarchical binding of DNA fragments derived from scaffold-attached regions: correlation of properties in vitro and function in vivo, *Biochemistry*, 29, 7475, 1990.
329. **Millane, H. B. and Rodley, G. A.,** Stereochemical details of the side by side model for DNA, *Nucleic Acids Res.*, 9, 1765, 1981.
330. **Minden, J. S., Mok, M., and Marians, K. J.,** Studies on the replication of pBR322 DNA in vitro with purified proteins, in *DNA Replication and Recombination*, Alan R. Liss, New York, 1987, 247.
331. **Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D.,** DNA H form requires a homopurine-homopyrimidine mirror repeat, *Nature*, 330, 495, 1987.
332. **Mirkovich, J., Mirault, M.-E., and Laemmli, U. K.,** Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold, *Cell*, 39, 223, 1984.
333. **Mitchell, P. J. and Tjian, R.,** Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins, *Science*, 245, 371, 1989.
334. **Mok, M. and Marians, K. J.,** The *E. coli* preprimosome and dnaB helicase can form replication forks that move at the same rate, *J. Biol. Chem.*, 262, 16644, 1987.
335. **Mooren, M. M. W., Pulleyblank, D. E., Wijmenga, S. S., Blommers, M. J. J., and Hilbers, C. W.,** Polypurine/polypyrimidine hairpins form a triple helix structure at low pH, *Nucleic Acids Res.*, 18, 6523, 1990.
336. **Moreau, J., Marcaud, L., Maschat, F., Kejzlarova-Lepesant, J., Lepesant, J.-A., and Scherrer, K.,** A+T rich linkers define functional domains in eukaryotic DNA, *Nature*, 295, 260, 1982.
337. **Moreau, J., Matjash-Smirniaguina, L., and Scherrer, K.,** Systematic punctuation of eukaryotic DNA by A+T rich sequences, *Proc. Natl. Açad. Sci. U.S.A.*, 78, 1341, 1981.
338. **Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M. P., and Chambon, P.,** The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants, *Nucleic Acids Res.*, 9, 6047, 1981.
339. **Morgan, A. R. and Wells, R. D.,** Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences, *J. Mol. Biol.*, 37, 63, 1968.
340. **Morita, M. and Oka, A.,** The structure of a transcriptional unit on colicin E1 plasmid, *Eur. J. Biochem.*, 97, 435, 1979.
341. **Moser, H. E. and Dervan, P. B.,** Sequence-specific cleavage of double helical DNA by triple helix formation, *Science*, 238, 645, 1987.
342. **Muller, B., Koller, T., and Stasiak, A.,** Characterization of the DNA binding activity of stable recA-DNA complexes: interaction between the two DNA binding sites within RecA helical filaments, *J. Mol. Biol.*, 212, 97, 1990.
343. **Muller, H. P., Sogo, J. M., and Schaffner, W.,** An enhancer stimulates in transcription trans when attached to the promoter via a protein bridge, *Cell*, 58, 767, 1989.
344. **Muller, U. R. and Fitch, W. M.,** Evolutionary selection for perfect hairpin structures in viral DNAs, *Nature*, 298, 582, 1982.
345. **Myers, R. M., Maniatis, T. M., and Lerman, R. S.,** Detection and localization of single base changes by denaturing gradient gel electrophoresis, *Meth. Enzymol.*, 155, 501, 1987.
346. **Nakai, H. and Richardson, C. C.,** The effect of T7 and *E. coli* DNA binding proteins on the replication fork of bacteriophage T7, *J. Biol. Chem.*, 263, 9831, 1988.
347. **Nass, M. M. K.,** Mitochondrial DNA. II. Structure and physicochemical properties of isolated DNA, *J. Mol. Biol.*, 42, 521, 1969.
348. **Naylor, L. H. and Clark, E. M.,** $d(TG)_n \cdot d(CA)_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription, *Nucleic Acids Res.*, 18, 1595, 1990.
349. **Nejedly, K., Matyasek, R., and Palecek, E.,** Site-specific chemical modification of B-Z junctions in supercoiled DNA as detected by nuclease S1 digestion, inhibition of restriction cleavage and nucleotide sequencing, *J. Biomol. Struct. Dyn.*, 6, 261, 1988.
350. **Nelson, W. G., Plientta, K. J., Barrack, E. R., and Coffey, D. S.,** The role of the nuclear matrix in the organization and function of DNA, *Annu. Rev. Biochem. Biophys.*, 15, 457, 1986.
351. **Nickol, J. M. and Felsenfeld, G.,** DNA conformation at the 5′ end of the chicken adult β-globin gene, *Cell*, 35, 467, 1983.
352. **Noller, H. F. and Chaires, J. B.,** Functional mod-

ification of 16S RNA by kethoxal, *Proc. Natl. Acad. Sci. U.S.A.*, 69, 3115, 1972.
353. **Nordheim, A.,** Z-DNA: exploring the biological significance, *Nucleic Acids Mol. Biol.*, 1, 112, 1987.
354. **Nordheim, A. and Meese, K.,** Topoisomer gel retardation: detection of anti-Z-DNA antibodies bound to Z-DNA antibodies bound to Z-DNA within supercoiled DNA minicircles, *Nucleic Acids Res.*, 16, 21, 1988.
355. **Norris, V.,** DNA replication in *E. coli* is initiated by membrane detachment of oriC: a model, *J. Mol. Biol.*, 215, 67, 1990.
356. **Ogata, R. T. and Gilbert, W.,** DNA binding site of lac repressor probed by dimethyl sulfate methylation, *J. Mol. Biol.*, 132, 709, 1979.
357. **Oh, E. Y. and Grossman, L.,** Helicase properties of the *E. coli* uvrAB protein complex, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 3638, 1987.
358. **Ohmori, H., Murakami, Y., and Nagata, T.,** Nucleotide sequences required for c Col E1-type plasmid to replicate in *E. coli* cells with or without RNase H, *J. Mol. Biol.*, 198, 223, 1987.
359. **Olson, W. K.,** Spatial configuration of ordered polynucleotide chains: a novel double helix, *Proc. Natl. Acad. Sci. U.S.A.*, 74, 1775, 1977.
360. **Oostra, B. A., van Vliet, A. J., Ab, G., and Gruber, M.,** Enhancement of ribosomal ribonucleic acid synthesis by deoxyribonucleic acid gyrase activity in *E. coli*, *J. Bacteriol.*, 148, 782, 1981.
361. **Orbons, L. P. M.,** Hairpin and duplex formation of the DNA octamer d(mC-G-mC-G-T-G-mC-G) in solution, *Nucleic Acids Res.*, 14, 4187, 1986.
362. **Orbons, L. P. M., van Bezeukom, T., and Altona, C.,** Conformational and modelbuilding studies of the hairpin form of the mismatched octamer d(mC-G-mC-G-T-G-mC-G), *J. Biomol. Struct. Dyn.*, 4, 965, 1987.
363. **Ostrander, E. A., Benedetti, P., and Wang, J. C.,** Template supercoiling by a chimera of yeast Gal4 protein and phage T7 RNA polymerase, *Science*, 126, 1261, 1990.
364. **Palecek, E., Boublikova, P., Nejedly, K., Galazka, G., and Klysik, J.,** B-Z junctions in supercoiled pRW751 DNA, unpaired bases or non-Watson-Crick base paired?, *J. Biomol. Struct. Dyn.*, 5, 297, 1987.
365. **Panayotatos, N. and Wells, R. D.,** Cruciform structures in supercoiled DNA, *Nature*, 289, 466, 1982.
366. **Panyutin, I. G., Klishko, V., and Lyamichev, V. I.,** Kinetics of cruciform formation and stability of cruciform structure in superhelical DNA, *J. Biomol. Struct. Dyn.*, 1, 1311, 1984.
367. **Panyutin, I. G., Kovalski, O. I., Budowsky, E. I., Dickerson, R. E., Rikhirev, M. E., and Lipanov, A. A.,** G-DNA: a twice folded DNA structure adopted by single-stranded Oligo(dG) and its implication for telomeres, *Proc. Natl. Acad. Sci. U.S.A.*, 87, 867, 1990.
368. **Panyutin, I. G., Kovalsky, O. I., and Budowsky, E. I.,** Magnesium-dependent supercoiling-induced transition in $(dG)_n \cdot (dC)_n$ stretches and formation of a new G-structure by $(dG)_n$ strand, *Nucleic Acids Res.*, 17, 8257, 1989.
369. **Panyutin, I. G., Lyamichev, V. I., and Lyubchenko, Yu. L.,** A sharp structural transition in pA03 plasmid DNA caused by increased superhelix density, *FEBS Lett.*, 148, 297, 1982.
370. **Parada, C. A. and Marians, K. J.,** Transcriptional activation of pBR322 DNA can lead to duplex DNA unwinding catalyzed by the *E. coli* preprimosome, *J. Biol. Chem.*, 264, 15120, 1989.
371. **Parker, C. S.,** Transcription factors, *Curr. Opin. Cell Biol.*, 1, 512, 1990.
373. **Parniewski, P., Galazka, G., Wilk, A., and Klysik, J.,** Complex structural behavior of oligopurine-oligopyrimidine sequence cloned within the supercoiled plasmid, *Nucleic Acids Res.*, 17, 617, 1989.
374. **Parsons, R., Anderson, M. E., and Tegtmeyer, P.,** Three domains in the simian virus 40 core origin orchestrate the binding, melting, and DNA helicase activities of T antigen, *J. Virol.*, 64, 509, 1990.
375. **Patel, D. J., Shapiro, L., and Hare, D.,** NMR-distance geometry studies of helical errors and sequence dependent conformation of DNA in solution, in *Unusual DNA Structures*, Wells, R. C. and Harvey, S. C., Eds., Springer-Verlag, New York, 1988, 115.
376. **Paulson, J. R. and Laemmli, U. K.,** The structure of histone depleted metaphase chromosomes, *Cell*, 12, 817, 1977.
377. **Peck, L. J., Nordheim, A., Rich, A., and Wang, J. C.,** Flipping of cloned $d(pCpG)_n d(pCpG)_n$ DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling, *Proc. Natl. Acad. Sci. U.S.A.*, 79, 4560, 1982.
378. **Peck, L. J. and Wang, J. C.,** Energetics of B-to-Z transition in DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 6206, 1983.
379. **Pettersson, M. and Schaffner, W.,** Synergistic activation of transcription by multiple binding sites for NF-κB even in absence of co-operative factor binding to DNA, *J. Mol. Biol.*, 214, 373, 1990.
380. **Pettijohn, D. E. and Hecht, R.,** RNA molecules bound to the folded bacterial genome stabilize DNA folds and segregate domains of supercoiling, *Cold Spring Harbor Symp. Quant. Biol.*, 38, 31, 1973.
381. **Pierce, J. W., Lenard, M., and Baltimore, D.,** Oligonucleotide that binds nuclear factor NF-κB acts as a lymphoid-specific and inducible enhancer element, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 1482, 1988.
382. **Pilch, D. S., Levenson, C., and Shafer, R. H.,** Structural analysis of the $(dA)_{10} \cdot 2(dT)_{10}$ triple helix, *Proc. Natl. Acad. Sci. U.S.A.*, 87, 1942, 1990.
383. **Platt, J. R.,** Possible separation of the intertwined nucleic acid chains by transfer-twist, *Proc. Natl. Acad. Sci. U.S.A.*, 41, 181, 1955.
384. **Pohl, F. M.,** Ein Modell der DNS Struktur, *Naturwissenschaften*, 54, 616, 1967.

385. **Pohl, F. M. and Jovin, T. M.,** Salt induced cooperative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly(dG-dC), *J. Mol. Biol.*, 67, 375, 1972.
386. **Pommier, Y., Cockerill, P. N., Kohn, K. W., and Garrard, W. T.,** Identification within the simian virus 40 genome of a chromosomal loop attachment site that contains topoisomerase II cleavage sites, *J. Virol.*, 64, 419, 1990.
387. **Pope, L. E. and Sigman, D. S.,** Secondary structure specificity of the nuclease activity of the 1,10-phenantroline-copper complex, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 3, 1984.
388. **Porter, S. E. and Champoux, J. J.,** The basis for camptothecin enhancement of DNA breakage by eukaryotic topoisomerase I, *Nucleic Acids Res.*, 1, 8521, 1989.
389. **Price, C. C., Gaucher, M., Koneru, P., Shibakawa, R., Sowa, K., and Yamaguchi, M.,** Relative reactivities for monofunctional nitrogen mustard alkylation of nucleic acid components, *Biochim. Biophys. Acta*, 166, 327, 1957.
390. **Prives, C.,** The replication functions of SV40 T antigen are regulated by phosphorylation, *Cell*, 61, 735, 1990.
391. **Pruss, G. J. and Drlica, K.,** Topoisomerase I mutants: the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 8952, 1986.
392. **Pruss, G. J. and Drlica, K.,** DNA supercoiling and prokaryotic transcription, *Cell*, 56, 521, 1989.
393. **Ptashne, M. and Gann, A. A. F.,** Activators and targets, *Nature*, 346, 329, 1990.
394. **Pugh, B. F. and Cox, M. M.,** Stable binding of recA protein to duplex DNA, *J. Biol. Chem.*, 262, 1326, 1987.
395. **Pulleyblank, D. E.,** Noncommensurate polypurine/polypyrimidine DNA structures, *Nucleic Acids Res.*, 2, 31, 1988.
396. **Pulleyblank, D. E. and Haniford, D. B.,** Anomalous protonation of polypurine-polypyrimidine DNA sequences: a new form of DNA helix, *J. Biomol. Struct. Dyn.*, 4, 201, 1986.
397. **Pulleyblank, D. E., Haniford, D. B., and Morgan, A. R.,** A structural basis for S1 nuclease sensitivity of double-stranded DNA, *Cell*, 42, 271, 1985.
398. **Pulleyblank, D. E., Shure, M., Tang, D., Vinograd, J., and Vosberg, H.-P.,** Action of nicking-closing enzyme on supercoiled and nonsupercoiled closed circular DNA: formation of a Boltzmann distribution of topological isomers, *Proc. Natl. Acad. Sci. U.S.A.*, 72, 4280, 1975.
399. **Quigley, G. J., Ughetto, G., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., and Rich, A.,** Non-Watson-Crick G · C and A · T base pairs in a DNA-antibiotic complex, *Science*, 232, 1255, 1986.
400. **Radding, C. M., Beattie, K. L., Holloman, W. K., and Wiegand, R. C.,** Uptake of homologous single stranded fragments by superhelical DNA. IV. Branch migration, *J. Mol. Biol.*, 116, 825, 1977.
401. **Radloff, R., Bauer, W., and Vinograd, J.,** A dye-buoyant density method for the detection and isolation of closed circular duplex DNA: the closed DNA of Hela cells, *Proc. Natl. Acad. Sci. U.S.A.*, 57, 1514, 1967.
402. **Rajagopal, P. and Feigon, J.,** Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides $d(G\text{-}A)_4$ and $d(T\text{-}C)_4$, *Nature*, 339, 637, 1989.
403. **Rajagopal, P. and Feigon, J.,** NMR studies of triple strand formation from the homopurinehomopyrimidine deoxyribonucleotides d(GA)4 and d(TC)4, *Biochemistry*, 28, 7859, 1989.
404. **Reitman, M. and Felsenfeld, G.,** Developmental regulation of topoisomerase II sites and Dnase I-hypersensitive sites in the chicken beta-globin locus, *Mol. Cell. Biol.*, 10, 2774, 1990.
405. **Rekosh, D. M. K., Russel, W. C., Bellet, A. J. D., and Robinson, A. J.,** Identification of a protein linked to the ends of adenovirus DNA, *Cell*, 11, 283, 1977.
406. **Reynolds, W. F. and Gottesfeld, J. M.,** Torsional stress induces an S1 nuclease-hypersensitive site within the promoter of the *Xenopus* laevis oocyte-type 5S RNA gene, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 4018, 1985.
407. **Reznikoff, W. S., Siegele, D. A., Cowing, D. B., and Gross, C. A.,** The regulation of transcription initiation in bacteria, *Annu. Rev. Genet.*, 19, 355, 1985.
408. **Rhodes, D. and Klug, A.,** Sequence-dependent helical periodicity of DNA, *Nature*, 292, 378, 1981.
409. **Rich, A., Nordheim, A., and Wang, A. H.-J.,** Z DNA, *Annu. Rev. Biochem.*, 53, 791, 1984.
410. **Richardson, J. P.,** Initiation of transcription of *E. coli* RNA polymerase from supercoiled and non-supercoiled bacteriophage PM2 DNA, *J. Mol. Biol.*, 91, 477, 1975.
411. **Richardson, J. P.,** Attachment of nascent RNA molecules to superhelical DNA, *J. Mol. Biol.*, 98, 565, 1975.
412. **Riddles, P. W. and Lehman, I. R.,** The formation of paranemic and plectonemic joints between DNA molecules by the rec A and single-stranded DNA-binding proteins of *E. coli*, *J. Biol. Chem.*, 260, 165, 1985.
413. **Riley, M., Maling, B., and Chamberlin, M. J.,** Physical and chemical characterization of two and three stranded A-T and A-U homopolymer complexes, *J. Mol. Biol.*, 20, 359, 1966.
414. **Roberts, J. M.,** Simian virus 40 (SV40) large tumor antigen causes stepwise changes in SV40 origin structure during initiation of DNA replication, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 3939, 1989.
415. **Rodley, G. A., Scobie, R. S., Bates, R. H. T., and Lewitt, R. M.,** A possible conformation for double-stranded polynucleotides, *Proc. Natl. Acad. Sci. U.S.A.*, 73, 2959, 1976.

416. **Rokeach, L. A. and Zyskind, J. W.,** RNA terminating within the *E. coli* origin of replication: stringent regulation and control by DnaA protein, *Cell,* 46, 763, 1986.
417. **Rose, D., Thomas, W., and Holm, C.,** Segregation of recombined chromosomes in meiosis I requires DNA topoisomerase II, *Cell,* 60, 1009, 1990.
418. **Rouet, P. and Essigman, J. M.,** Possible role of thymidine glycol in the selective inhibition of DNA synthesis on oxidized DNA templates, *Cancer Res.,* 45, 6113, 1985.
419. **Rubin, C. M. and Schmid, C. W.,** Pyrimidine specific chemical reactions useful for DNA sequencing, *Nucleic Acids Res.,* 8, 4613, 1980.
420. **Ruiz-Carrillo, A.,** The histone H5 gene is flanked by S1 hypersensitive structures, *Nucleic Acids Res.,* 12, 6473, 1984.
421. **Ryder, K., Silver, S., DeLucia, A. L., Fanning, E., and Tegtmeyer, P.,** An altered DNA conformation in origin region I is a determinant for the binding of SV40 large T antigen, *Cell,* 44, 719, 1986.
422. **Ryoji, M. and Worcel, A.,** Chromatin assembly in *Xenopus* oocytes: in vivo studies, *Cell,* 37, 21, 1984.
423. **Ryoji, M. and Worcel, A.,** Structure of the two distinct types of minichromosomes that are assembled on DNA injected in *Xenopus* oocytes, *Cell,* 40, 923, 1985.
424. **Saenger, W.,** *Principles of Nucleic Acid Structure,* Springer-Verlag, New York, 1983.
425. **Saragosti, S., Moyne, G., and Yaniv, M.,** Absence of nucleosomes in a fraction of SV40 chromatin, between the origin of replication and the region coding for the late leader RNA, *Cell,* 20, 65, 1980.
426. **Sarma, R. H.,** Is DNA really a double helix? Its diverse spatial configurations and the evidence for a vertically stabilized double helix, *Biomol. Stereodyn.,* 5, 83, 1987.
427. **Sasisekharan, V., Pattabiraman, N., and Gupta, G.,** Some implications of an alternative structure for DNA, *Proc. Natl. Acad. Sci. U.S.A.,* 75, 4092, 1978.
428. **Sasse-Dwight, S. and Gralla, J. D.,** $KMnO_4$ as a probe for lac promoter DNA melting and mechanism in vivo, *J. Biol. Chem.,* 264, 8074, 1989.
429. **Saucier, J.-M. and Wang, J. C.,** Angular alteration of the DNA helix by *E. coli* RNA polymerase, *Nature,* 239, 167, 1972.
430. **Schatt, M. D., Rusconi, S., and Schaffner, W.,** A single DNA-binding transcription factor is sufficient for activation from a distant enhancer and/or from a promoter position, *EMBO J.,* 9, 481, 1990.
431. **Scheffner, M., Wessel, R., and Stahl, H.,** Sequence independent duplex DNA opening reaction catalysed by SV40 large tumor antigen, *Nucleic Acids Res.,* 17, 93, 1989.
432. **Schnos, M., Zahn, K., Inman, R. B., and Blattner, F. R.,** Initiation protein induced helix destabilization at the lambda origin: a prepriming step in DNA replication, *Cell,* 52, 385, 1988.
433. **Schon, E., Evans, T., Welsh, J., and Efstratiadis, A.,** Conformation of promoter DNA: fine mapping of S1-hypersensitive sites, *Cell,* 35, 837, 1983.
434. **Schutte, B. C. and Cox, M. M.,** Homology-dependent underwinding of duplex DNA in recA protein generated paranemic complexes, *Biochemistry,* 27, 7886, 1988.
435. **Sekiguchi, J. A. M. and Kmiec, E. B.,** Reaction parameters of TFIIIA-induced supercoiling catalyzed by a *Xenopus* laevis cell-free extract, *Nucleic Acids Res.,* 18, 1021, 1990.
436. **Sekimizu, K., Bramhill, D., and Kornberg, A.,** ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the *E. coli* chromosome, *Cell,* 50, 259, 1987.
437. **Sekimizu, K., Yung, B. Y., and Kornberg, A.,** The dnaA protein of *E. coli:* abundance, improved purification, and membrane binding, *J. Biol. Chem.,* 263, 7136, 1988.
438. **Selleck, S. B., Elgin, S. C. R., and Cartwright, I. L.,** Supercoil-dependent features of DNA structure at Drosophila locus 67B1, *J. Mol. Biol.,* 178, 17, 1984.
439. **Shastri, K., Issackson, P. J., Fishback, J., Land, D. M., and Reeck, G. R.,** Influence of nonhistone chromatin protein HMG-1 on the enzymatic digestion of purified DNA, *Nucleic Acids Res.,* 10, 5059, 1982.
440. **Sheflin, L. G. and Kowalski, D.,** Mung bean nuclease cleavage of dA + dE-rich sequence or an inverted repeat sequence in supercoiled PM2 DNA depends on ionic environment, *Nucleic Acids Res.,* 12, 7081, 1984.
441. **Sheflin, L. G. and Kowalski, D.,** Altered DNA conformations detected by mung bean nuclease occur in promoter and terminator regions of supercoiled pBR322 DNA, *Nucleic Acids Res.,* 13, 6137, 1985.
442. **Shen, C.-K. J.,** Superhelicity induces hypersensitivity of a human polypyrimidine · polypurine DNA sequence in the human $\alpha 2$-$\alpha 1$ globin intergenic region to S1 nuclease digestion — high resolution mapping of the clustered cleavage sites, *Nucleic Acids Res.,* 11, 7899, 1983.
443. **Shibata, T., DasGupta, C., Cunningham, R. P., and Radding, C. M.,** Purified *E. coli* rec A protein catalyzes homologous pairing of superhelical DNA and single-stranded fragments, *Proc. Natl. Acad. Sci. U.S.A.,* 76, 1638, 1979.
444. **Shimizu, M., Hanvey, J. C., and Wells, R. D.,** Intramolecular DNA triplexes in supercoiled plasmids. I. effect of loop size on formation and stability, *J. Biol. Chem.,* 264, 5944, 1989.
445. **Shishido, K.,** Relationship between S1 endonuclease-sensitivity and number of superhelical turns in a negatively-twisted DNA, *FEBS Lett.,* 111, 333, 1980.
446. **Skarstad, K., Baker, T., and Kornberg, A.,** Strand separation required for initiation of replication at the chromosomal origin of *E. coli* is facilitated by a distant RNA-DNA hybrid, *EMBO J.,* 9, 2341, 1990.

447. **Shouche, Y. S., Ramesh, N., and Brahmachari, S. K.,** Probing of unusual DNA structures in topologically constrained form V DNA: use of restriction enzymes as structural probes, *Nucleic Acids Res.*, 18, 267, 1990.
448. **Siebenlist, U., Simpson, R. B., and Gilbert, W.,** *E. coli*, RNA polymerase interacts homologously with two different promoters, *Cell*, 20, 269, 1980.
449. **Siegfried, E., Thomas, G. H., Bond, U. M., and Elgin, S. C. R.,** Characterization of a supercoil-dependent S1 sensitive site 5′ to the *Drosophila melanogaster* hsp 26 gene, *Nucleic Acids Res.*, 14, 9425, 1986.
450. **Sigal, N., Delius, H., Kornberg, T., Gefter, M. L., and Alberts, B. M.,** A DNA unwinding protein isolated from *E. coli:* its interaction with DNA and with DNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.*, 69, 3537, 1972.
451. **Sinden, R. R., Broyles, S., and Pettijohn, E.,** Perfect palindromic lac operator DNA sequence as stable cruciform structure in vitro but not in vivo, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 1797, 1983.
452. **Sinden, R. R. and Kochel, T. J.,** Reduced 4,5′, 8-TMP crosslinking of left handed Z DNA, stabilized by DNA supercoiling, *Biochemistry*, 26, 1343, 1987.
453. **Sinden, R. R. and Pettijohn, D. E.,** Chromosomes in living *E. coli* cells are segregated into domains of supercoiling, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 224, 1981.
454. **Singer, B. and Grunberger, D.,** *Molecular Biology of Mutagens and Carcinogens*, Plenum Press, New York, 1983.
455. **Singleton, C. K.,** Effect of salts, temperature and stem length on supercoil induced formation of cruciforms, *J. Biol. Chem.*, 258, 7661, 1983.
456. **Singleton, C. K., Kilpatrick, M. W., and Wells, R. D.,** S1 nuclease recognizes DNA conformational junctions between left-handed helical $(dT\text{-}dG)_n \cdot (dC\text{-}dA)_n$ and contiguous righthanded sequences, *J. Biol. Chem.*, 259, 1963, 1984.
457. **Singleton, C. K., Klysik, J., Stirdivant, S. M., and Wells, R. D.,** Left-handed Z-DNA is induced by supercoiling in physiological ionic conditions, *Nature*, 299, 312, 1982.
458. **Sklar, V.E. F., Yamamoto, M., and Roeder, R. G.,** Molecular structures of eukaryotic class III polymerases, in *RNA Polymerase*, Losick, R. and Chamberlin, M., Eds., Academic Press, New York, 1976, 813.
459. **Smith, C. L., Kubo, M., and Imamoto, F.,** Promoter-specific inhibition of transcription by antibiotics which act on DNA gyrase, *Nature*, 275, 420, 1978.
460. **Spatz, H.-C. and Crothers, D. M.,** The rate of DNA unwinding, *J. Mol. Biol.*, 42, 191, 1969.
461. **Sridhara Rao, B., Manor, H., and Martin, R. G.,** Pausing in simian virus 40 DNA replication by a sequence containing $(dG\text{-}dA)_{27} \cdot (dT\text{-}dC)_{27}$, *Nucleic Acids Res.*, 16, 8077, 1988.
462. **Stasiak, A., Di Capua, E., and Koller, T.,** Elongation of duplex DNA by Rec A protein, *J. Mol. Biol.*, 151, 557, 1981.
463. **Sternglantz, R.,** DNA topoisomerases, *Curr. Opin. Cell Biol.*, 1, 533, 1989.
464. **Stettler, U. H., Weber, H., Koller, Th., and Weissman, H.,** Preparation and characterization of Form V DNA, the duplex DNA resulting from association of complementary circular single-stranded DNA, *J. Mol. Biol.*, 131, 21, 1979.
465. **Stewart, A. F., Herrera, R. E., and Nordheim, A.,** Rapid induction of c-fos transcription reveals quantitative linkage of RNA polymerase II and DNA topoisomerase I enzyme activities, *Cell*, 50, 141, 1990.
466. **Stewart, A. F. and Schutz, G.,** Camptothecin-induced in vivo topoisomerase I cleavages in the transcriptionally active tyrosine aminotransferase gene, *Cell*, 50, 1109, 1987.
467. **Stillman, B.,** Initiation of eukaryotic DNA replication in vitro, *Annu. Rev. Cell Biol.*, 5, 197, 1989.
468. **Stokrova, J., Vojtiskova, M., and Palecek, E.,** Electron microscopy of supercoiled pEJ4 DNA containing homopurine · homopyrimidine sequences, *J. Biomol. Struct. Dyn.*, 6, 891, 1989.
469. **Stonington, O. and Pettijohn, D.,** The folded genome of *E. coli* isolated in a protein-RNA-DNA complex, *Proc. Natl. Acad. Sci. U.S.A.*, 68, 6, 1971.
470. **Straney, D. C. and Crothers, D. M.,** Comparison of open complexes formed by RNA polymerase at the *E. coli* UV5 promoter, *J. Mol. Biol.*, 594, 279, 1987.
471. **Suggs, W. and Warner, R. W.,** Nuclease recognition of an alternating structure in a $d(AT)_{14}$ plasmid insert, *Nucleic Acids Res.*, 14, 3703, 1986.
472. **Sun, J.-S., Francois, J.-C., Montenay-Garestier, T., Saison-Behmoaras, T., Roig, V., Thuong, N. T., and Helene, C.,** Sequence-specific intercalating agents: intercalation at specific sequences on duplex DNA via major groove recognition by oligonucleotide-intercalator conjugates, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 9198, 1989.
473. **Sundquist, W. I. and Klug, A.,** Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops, *Nature*, 342, 825, 1989.
474. **Svaren, J., Inagumi, S., Lovegren, E., and Chalkley, R.,** DNA denatures upon drying after ethanol precipitation, *Nucleic Acids Res.*, 15, 8739, 1987.
475. **Svejstrup, J. Q., Christiansen, K., Andersen, A. H., Lund, K., and Westergaard, O.,** Minimal DNA duplex requirements for topoisomerase I-mediated cleavage in vitro, *J. Biol. Chem.*, 265, 12529, 1990.
476. **Tegtmeyer, P., Deb, S., DeLucia, A. L., Deb, S. P., Koff, A., Tsui, S., Parsons, R., Partin, K., Baur, C. P., Dean, F. B., and Hurwitz, J.,** SV40 origin of replication, *Cancer Cells*, 6, 123, 1988.
477. **Thresher, R. J., Christiansen, G., and Griffith,**

**J. D.,** Assembly of presynaptic filaments: factors affecting the assembly of RecA protein onto single-stranded DNA, *J. Mol. Biol.*, 201, 101, 1988.
478. **Tomizawa, J.,** Control of col E1 plasmid replication: interaction of rom protein with an unstable complex formed by RNA I and RNA II, *J. Mol. Biol.*, 212, 695, 1990.
479. **Trifonov, E. N.,** Curved DNA, *Crit. Rev. Biochem.*, 19, 89, 1985.
480. **Trifonov, E. N. and Ulanovsky, L. E.,** Inherently curved DNA and its structural elements, in *Unusual DNA Structures*, Wells, R. D. and Harvey, S. C., Eds., Springer-Verlag, New York, 1988, 173.
481. **Tsao, Y.-P., Wu, H.-Y., and Liu, L. F.,** Transcription-driven supercoiling of DNA: direct biochemical evidence from in vitro studies, *Cell*, 56, 111, 1989.
482. **Tsurimoto, T., Melendy, T., and Stillman, B.,** Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin, *Nature*, 346, 346, 1990.
483. **Uemura, T., Morino, K., Uzawa, S., and Yanagida, N.,** Cloning and sequencing of S. pombe topoisomerase I gene and effect of gene distruption, *Nucleic Acids Res.*, 15, 9727, 1987.
484. **Uesugi, S., Lee, B. L., Ikehara, M., Kobayashi, Y., and Kyogoku, Y.,** Hybrid oligomer of cyclonucleotides and deoxynucleotides. A high anti left-handed double helical DNA structure, *J. Biomol. Struct. Dyn.*, 3, 339, 1985.
485. **Umek, R. M., Eddy, M. J., and Kowalski, D.,** DNA sequences required for unwinding prokaryotic and eukaryotic replication origins, *Cancer Cells*, 6, 473, 1988.
486. **Umek, R. M. and Kowalski, D.,** Yeast regulatory sequences preferentially adopt a non-B conformation in supercoiled DNA, *Nucleic Acids Res.*, 15, 4467, 1987.
487. **Umek, R. M. and Kowalski, D.,** The ease of DNA unwinding as a determinant of initiation at yeast replication origins, *Cell*, 52, 559, 1988.
488. **Umek, R. M. and Kowalski, D.,** The DNA unwinding element in a yeast replication origin functions independently of easily unwound sequences present elsewhere on the plasmid, *Nucleic Acids Res.*, 18, 6601, 1990.
489. **Umlauf, S. W., Cox, M. M., and Inman, R. B.,** Triple-helical DNA pairing intermediates formed by recA protein, *J. Biol. Chem.*, 265, 16898, 1990.
490. **Usdin, K. and Furano, A. V.,** Rat L (long interspersed repeated DNA) elements contain guanine-rich homopurine sequences that induce unpairing of contiguous duplex DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 4416, 1988.
491. **Van Holde, K. E.,** *Chromatin*, Springer-Verlag, Heidelberg, 1988.
492. **Venkatesam, M., Silver, L. L., and Nossal, N. G.,** Bacteriophage T4 gene 41 protein, required for the synthesis of DNA is also a DNA helicase, *J. Biol. Chem.*, 257, 12426, 1982.
493. **Vinograd, J., Lebowitz, J., Radloff, R., Watson, R., and Laipis, P.,** The twisted circular form of polyoma viral DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 53, 1104, 1965.
494. **Vinograd, J., Lebowitz, J., and Watson, R.,** Early and late helix-coil transitions in closed circular DNA. The number of superhelical turns in polyoma DNA, *J. Mol. Biol.*, 33, 173, 1968.
495. **Vinzce, A., Henderson, R. E. L., McDonald, F. F., and Leonard, N. J.,** Reaction of diethyl pyrocarbonate with nucleic acid components. Bases and nucleosides derived from guanine, cytosine and uracil, *J. Am. Chem. Soc.*, 95, 2677, 1973.
496. **Vizard, D. L. and Ansevin, A. T.,** High resolution thermal denaturation of DNA: thermalites of bacteriophage DNA, *Biochemistry*, 15, 741, 1976.
497. **Vogt, V. M.,** Purification and further properties of single-strand specific nuclease from *A. oryzae*, *Eur. J. Biochem.*, 33, 192, 1973; see also Reference 327.
498. **Vojtiskova, M. and Palecek, E.,** Unusual protonated structure in the homopurine · homopyrimidine tract of supercoiled and linearized plasmids recognized by chemical probes, *J. Biomol. Struct. Dyn.*, 5, 283, 1987.
499. **Vologodskii, A. V., Amirikyan, B. R., Lyubchenko, Y. L., and Frank-Kamenetskii, M. D.,** Allowance for heterogeneous stacking in the DNA helix-coil transition theory, *J. Biomol. Struct. Dyn.*, 2, 131, 1984.
500. **Vologodskii, A. V. and Frank-Kamenetskii, M. D.,** Theoretical study of cruciform states in superhelical DNAs, *FEBS Lett.*, 143, 257, 1982.
501. **Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P., and Frank-Kamenetskii, M. D.,** Chemical probing of homopurine-homopyrimidine mirror repeats in supercoiled DNA, *Nature*, 333, 475, 1988.
502. **Wada, A. and Suyama, A.,** Local stability of DNA and RNA secondary structure and its relation to biological function, *Prog. Biophys. Mol. Biol.*, 47, 113, 1986.
503. **Wada, A., Yabuki, S., and Husimi, Y.,** Fine structure in the thermal denaturation of DNA: high temperature-resolution spectrophotometric studies, *Crit. Rev. Biochem.*, 9, 87, 1980.
504. **Wallis, J. W., Chrebet, G., Brodsky, G., Rolfe, R., and Rothstein, R.,** A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase, *Cell*, 58, 409, 1989.
505. **Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G., and Rich, A.,** Molecular structure of a left handed double helical DNA fragment at atomic resolution, *Nature*, 282, 680, 1979.
506. **Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., and Rich, A.,** The molecular structure of a

DNA-triostin A complex, *Science*, 225, 1115, 1989.
507. **Wang, J. C.,** Variation of the average rotation angle of the DNA helix and the superhelical turns of covalently closed cyclic & lambda DNA, *J. Mol. Biol.*, 43, 25, 1969.
508. **Wang, J. C.,** Degree of superhelicity of covalently closed cyclic DNAs from *Escherichia coli*, *J. Mol. Biol.*, 43, 263, 1969.
509. **Wang, J. C.,** Interaction between DNA and an *E. coli* protein omega, *J. Mol. Biol.*, 55, 523, 1971.
510. **Wang, J. C.,** DNA topoisomerases, *Annu. Rev. Biochem.*, 54, 665, 1985.
511. **Wang, J. C.,** Recent studies of DNA topoisomerases, *Biochim, Biophys. Acta*, 909, 1, 1987.
512. **Wang, J. C., Barkley, M. D., and Bourgeois, S.,** Measurements of unwinding of lac operator by repressor, *Nature*, 251, 247, 1974.
513. **Wang, J. C. and Giaever, G. N.,** Action at a distance along DNA, *Science*, 240, 300, 1988.
514. **Wang, J. C., Jacobsen, J. H., and Saucier, J.-M.,** Physicochemical studies on interactions between DNA and RNA polymerase. Unwinding of the helix by *E. coli* RNA polymerase, *Nucleic Acids Res.*, 4, 1225, 1977.
515. **Wang, J. C., Peck, L. J., and Becherer, K.,** DNA supercoiling and its effects on DNA structure and function, *Cold Spring Harbor Symp.*, 47, 85, 1983.
516. **Wang, J. N. C. and Hogan, M.,** An equilibrium between distorted and undistorted DNA in the adult chicken β^-globin gene, *J. Biol. Chem.*, 260, 8194, 1985.
517. **Wasylyk, B.,** Transcription elements and factors of RNA polymerase B promoters of higher eukaryotes, *Crit. Rev. Biochem.*, 23, 77, 1988.
518. Deleted in proof.
519. **Watson, J. D. and Crick, F. H. C.,** The structure of DNA viruses, *Cold Spring Harbor Symp. Quant. Biol.*, 18, 123, 1953.
520. **Webb, M. L., Maguire, K. A., and Jacob, S. T.,** Novobiocin inhibits initiation of RNA polymerase II-directed transcription of the mouse metallothionein-I gene independent of its effect on DNA topoisomerase II, *Nucleic Acids Res.*, 15, 8547, 1987.
521. **Weil, R. and Vinograd, J.,** The cyclic helix and cyclic coil forms of polyoma DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 50, 730, 1963.
522. **Weintraub, H.,** A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin, *Cell*, 32, 1191, 1983.
523. **Weintraub, H., Chen, P. F., and Conrad, K.,** Expression of transfected DNA depends on DNA topology, *Cell*, 46, 115, 1986.
524. **Weintraub, H. and Groudine, M.,** Chromosomal subunits in active genes have an altered conformation, *Science*, 193, 848, 1976.
525. **Wells, R. D.,** Unusual DNA structures, *J. Biol. Chem.*, 263, 1095, 1988.
526. **Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., and Wohlrab, F.,** The chemistry and biology of unusual DNA structures adopted by oligopurine · oligopyrimidine sequences, *FASEB J.*, 2, 2939, 1988.
527. **Wells, R. D. and Harvey, S. C., Eds.,** *Unusual DNA Structures*, Springer-Verlag, New York, 1988.
528. **Wemmer, D. E., Chou, S. H., Hare, D. R., and Reid, B.,** Duplex-hairpin transitions in DNA: NMR studies on CGCGTATACGCG, *Nucleic Acids Res.*, 13, 3755, 1985.
529. **Wetmur, J. G. and Davidson, N.,** Kinetics of renaturation of DNA, *J. Mol. Biol.*, 31, 349, 1968.
530. **White, J. H., Cozzarelli, R., and Bauer, W. R.,** Helical repeat and linking number of surface wrapped DNA, *Science*, 241, 323, 1988.
531. **Wiegand, R. C., Beattie, K. L., Holloman, W. K., and Radding, C. M.,** Uptake of homologous single stranded fragments by superhelical DNA. III. The product and its enzymic conversion to a recombinant molecule, *J. Mol. Biol.*, 116, 805, 1977.
532. **Wiekowski, M., Schwartz, M. W., and Stahl, H.,** SV40 large T-antigen helicase. Characterization of the ATPase- dependent DNA unwinding activity and its substrate requirements, *J. Biol. Chem.*, 263, 436, 1988.
533. **Wohlrab, F., McLean, M. J., and Wells, R. D.,** The segment inversion site of herpes simplex virus type 1 adopts a novel DNA structure, *J. Biol. Chem.*, 262, 6407, 1987.
534. **Wold, M. S., Li, J. J., Weinberg, D. H., Virshup, D. M., Sherley, J. L., Verheyen, E., and Kelly, T.,** Cellular proteins required for SV40 DNA replication in vitro, *Cancer Cells*, 6, 133, 1988.
535. **Wolffe, A. P., Andrews, M. T., Crawford, E., Losa, R., and Brown, D. D.,** Negative supercoiling is not required for 5S RNA tanscription in vitro, *Cell*, 49, 301, 1987.
536. **Wood, E. R. and Matson, S. W.,** Purification and characterization of a new DNA-dependent ATPase with helicase activity from *E. coli*, *J. Biol. Chem.*, 262, 15269, 1987.
537. **Worcel, A.,** Transcription factor IIIA and 5S DNA supercoiing in vitro, *Cell*, 54, 919, 1988.
538. **Worcel, A. and Burgi, E.,** On the structure of the folded chromosome of *E. coli*, *J. Mol. Biol.*, 71, 127, 1972.
539. **Wu, A. M., Bianchi, M., DasGupta, C., and Radding, C. M.,** Unwinding associated with synapsis of DNA molecules by recA protein, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 1256, 1983.
540. **Wu, H.-Y., Shyy, S., Wang, J. C., and Liu, L. F.,** Transcription generates positively and negatively supercoiled domains in the template, *Cell*, 53, 433, 1988.
541. **Wuthrich, K.,** *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York, 1986.
542. **Xodo, L. E., Manzini, G., and Quadrifoglio, F.,** Spectroscopic and calorimetric investigation on the DNA triplex formed by d(CTCTTCTTTCTTTTCTTTCTTCTC) and

d(GAGAAGAAAGA) at acidic pH, *Nucleic Acids Res.*, 18, 3557, 1990.

543. **Yager, T. D. and von Hippel, P. H.,** Transcript elongation and termination in *E. coli* in *E. coli and S. typhimurium: Cellular and Molecular Biology*, Neidhardt, F. C., Ed., American Society of Microbiology, Washington D.C. 1987, 1241.
544. **Yagil, G.,** Quantitative aspects of protein induction, *Curr. Top. Cell. Regul.*, 9, 187, 1975.
545. **Yagil, G. and Sussman, J. L.,** Structural models for non-helical DNA, *EMBO J.*, 5, 1719, 1986.
546. **Yanagida, M. and Wang, J. C.,** Yeast DNA topoisomerases and their structural genes, *Nucleic Acids Res.*, 2, 196, 1987.
547. **Yang, L., Jessee, C. B., Lau, K., Zhang, H., and Liu, L. F.,** Template supercoiling during ATP-dependent DNA helix tracking: studies with simian virus 40 large tumor antigen, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 6121, 1989.
548. **Yang, L., Rowe, T. C., Nelson, E. M., and Liu, L. F.,** In vivo mapping of DNA topoisomerase II-specific cleavage sites on SV40 chromatin, *Cell*, 41, 127, 1985.
549. **Yerushalmi, A. and Yagil, G.,** The interaction of chromatin with alkylating agents. The monofunctional action of bis(2-chloroethyl)methylamine, *Eur. J. Biochem.*, 103, 237, 1980.
550. **Yu, Y.-T. and Manley, J. L.,** Structure and function of the S1 nuclease-sensitive site in the adenovirus later promoter, *Cell*, 45, 743, 1986.
551. **Yung, B. Y. and Kornberg, A.,** The dnaA initiator protein binds separate domains in the replication origin of *E. coli*, *J. Biol. Chem.*, 264, 6146, 1989.
552. **Zacharias, W., Jaworski, A., Larson, J. E., and Wells, R. D.,** The B to Z equilibrium in vivo is perturbed by biological processes, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 7069, 1988.
553. **Zhang, H., Wang, J. C., and Liu, L. F.,** Involvement of DNA topoisomerase I in transcription of human ribosomal RNA gene, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 1060, 1988.
554. **Zhang, L. and Gralla, J. D.,** In situ nucleoprotein structure at the SV40 major later promoter: melted and wrapped DNA flank the start site, *Genes Dev.*, 3, 1814, 1989.
555. **Zyskind, J. W. and Smith, D. W.,** The bacterial origin of replication, oriC, *Cell*, 46, 489, 1986.
556. **Shimron, and Yagil, G.,** Unwound states of DNA in gene control, in *Recent Advances in Chemistry and Molecular Biology of Cancer*, Dai, Q-H., Ed., Springer-Verlag, in press, 1991.
557. **Palaček, E.,** Local supercoil stabilized structures, *Crit. Rev. Biochem. Mol. Biol.*, 20, 51, 1991.