Unusual DNA Structures, Chromatin and Transcription

Kensal van Holde and Jordanka Zlatanova

Accepted 27 September 1993

Summary

Extensive studies of DNA secondary structure during the past decade have shown that DNA is a dynamic molecule, whose structure depends on the underlying nucleotide sequence and is influenced by the environment and the overall DNA topology. Three major non-B-DNA structures have been described (Z-DNA, triplex DNA and cruciform DNA) which are stabilized by unconstrained negative supercoiling and can be formed under physiological conditions. In this essay we summarize the DNA primary structure features that are pertinent to the formation of these conformers and present data concerning the occurrence of these sequences in the eukaryotic genome. The evidence in favor of the existence of these unusual DNA structures *in vivo* is discussed. The effect of alternative non-B-DNA structures on the way DNA is organized in chromatin is considered, and this is followed by evaluation of the data relating these structures to eukaryotic transcription. Some possible mechanisms by which the effect of non-B structures on transcription might be exerted are proposed.

Introduction

It is now known that local structure of DNA is nucleotide-sequence dependent and is very sensitive to changes in molecular environment and overall DNA topology. The structural polymorphism of DNA may range from minor distortions to major deviations from the B-DNA structure, in which the DNA helix may change its trajectory, the number of strands in the helix, its handedness or mode of base pairing. Under physiological conditions these alternative structures are generally thermodynamically unfavorable as compared to B-DNA and require an input of free energy for their formation and stabilization. In many cases this energy can be supplied by the unconstrained supercoiling of the genome. While the prokaryotic genome is known to be under torsional stress, the bulk of the eukaryotic genome is believed not to be stressed, as the inherent supercoiling is largely accommodated by the writhing of the DNA around histone octamers in nucleosomes. The actively transcribing portion of the genome, however, could contain unconstrained supercoiling, part of which can be attributed to the process of transcription *per se*[1]. A number of additional mechanisms may operate to create localized or transient torsional stress in eukaryotic DNA; these may include nucleosomal displacement, binding of transcription factors or other proteins, activity of helix-tracking proteins, looping of DNA by protein binding at two distinct locations, histone acetylation, and last but not least, gyrase activity of topoisomerases in the presence of additional factors. It is beyond the scope of this review to describe all of these possible mechanisms, but even the simple enumeration of such possibilities makes it clear that the local level of supercoiling in eukaryotic DNA is probably an extremely dynamic feature, and that this, in turn, may lead to dynamic alterations in the structure of the double helix, including the generation and removal of non-B structures. Can such structural alterations actually be employed to regulate transcriptional activity in the eukaryotic nucleus? This question is the focus of this paper.

The repertoire of alternative non-B-DNA structures now recognized is impressive (for a recent review, see ref. 2). We will concentrate here on only three major types of structures, all of which are stabilized by unconstrained negative supercoiling and can be formed under physiological conditions: Z-DNA, triplex DNA and cruciform DNA. Other non-B structures, such as quadruplexes[2], seem to be stable structures in very specialized regions of the genome (telomeres) and hence do not directly relate to the questions raised here.

DNA sequences with potential to form alternative non-B structures are abundant in the genome

Left-handed Z-DNA

Left-handed Z-DNA forms at alternating purine (pur)†.pyrimidine (pyr) sequences, most easily in CG dinucleotide repeats, but also in TG repeats. The latter are abundant in eukaryotic DNA (~10^5 copies of $(TG)_n$ with n>25 in the human genome) but are virtually absent from prokaryotic DNAs[3,4]. Sequences capable of forming Z-DNA have also been detected scattered throughout the genome in plants[5]. A recent mapping of such sequences in human genes revealed a non-random distribution, with a strong bias towards locations near the site of transcription initiation[6]. It should be noted, however, that such studies identify only regions with the potential to form Z-DNA. In a number of cases it has been shown that the sequences do adopt the Z-DNA conformation *in vitro*, when constrained in negatively supercoiled plasmids. Whether these actually form Z-DNA *in vivo* is a very different, and in most cases unresolved, question. A study by Rodriguez-Campos *et al.*[7] is instructive in this context: a potential Z-DNA forming sequence $[(TG)_{30}]$ cloned into the transcriptional leader region of the late viral SV40 genes was shown *not* to be predominantly in the Z-DNA form *in vivo*. It should be noted, however, that conditions known to stabilize Z-DNA, such as methylation of C5 of cytosine, the presence of certain cations including polyamines, and the binding of specific proteins, do exist *in vivo*. Evidence for the existence of Z-DNA under physiological conditions will be summarized below.

†Abbreviations: AB, antibody; bp, base pairs; pur, purine; pyr, pyrimidine.

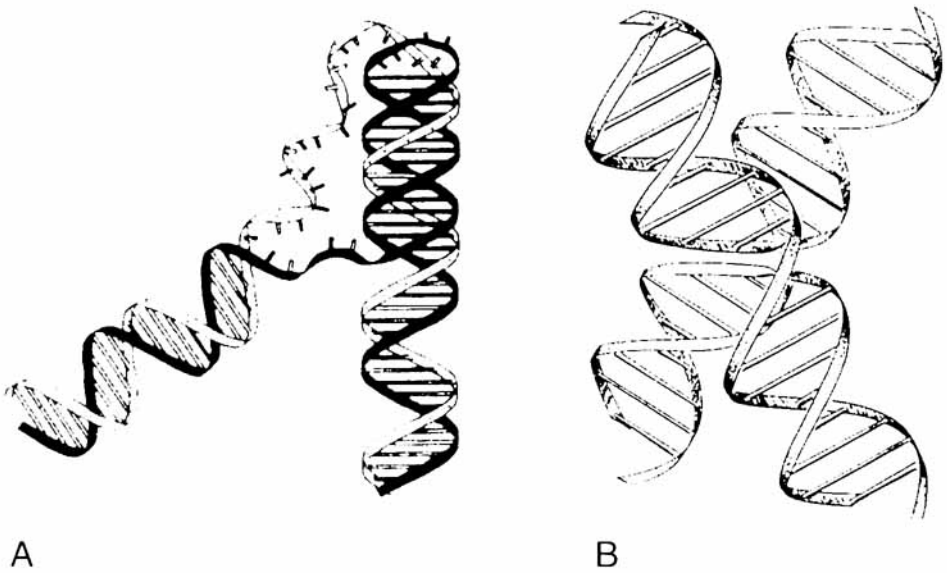

Fig. 1. Schematic representation of: (A) intramolecular triplex DNA (reproduced from Sinden, R. R. and Wells, R. D. (1992). *Current Opinion in Biotechnology* **3**, 612-622, with permission) and (B) cruciform DNA (reproduced from Lilley, D. M. J. (1990). *Nucleic Acids and Molecular Biology*, vol. 4 (ed. J. Eckstein and D. M. J. Lilley), pp. 55-77, with permission).

Triplex-DNA structures

Triplex-DNA structures (triplexes) (Fig. 1) can form at certain homopur.homopyr stretches under conditions of low pH and negative superhelicity. More important, triplexes can also form at physiological pH under higher superhelical density. One and the same sequence can form either pur.pyr.pur or pur.pyr.pyr (H-form) triplex structures, depending on the environmental conditions – mainly the availability of protons and Mg^{2+} ions (e.g. ref. 8). The actual H-DNA conformation itself is determined by the alternative use of the 3′ or 5′ half of the pyr strand as the donated strand and depends on the level of supercoiling[9]. Tracts of 10 or more contiguous pur or pyr residues are found four times more frequently than expected by chance in a 67 kbp region of the human β-globin gene[10]. Similar over-representation was found for 24 eukaryotic viral sequences, whereas no such enrichment was observed in bacteriophage or *E. coli* sequences[11]. Dot blot and plaque hybridization assays reveal that $(TC)_n$ sequences are relatively abundant and highly dispersed in the genomes of rodents and primates[12]. In general, sequences capable of forming triple-stranded helices are found in eukaryotes on the average of one for every 150 kbp[9].

A possible biological significance of the non-B-structures adopted by homopur.homopyr stretches was suggested by experiments mapping a number of S1 nuclease-hypersensitive sites to pur.pyr regions upstream of eukaryotic genes in active chromatin regions[13]. Since S1 is a single-strand specific nuclease, it will readily cleave the unpaired single strand found in any triplex structure. As in the case for Z-DNA, however, we cannot in general conclude that because a sequence exists, it adopts the alternative form *in vivo*. Indeed, cases have been reported of homopur.homopyr sequences *not* adopting triplex helices *in situ*[14].

Cruciforms

Cruciforms (Fig. 1B) can be formed at palindromic sequences. Their participation in genomic recombination has been known for years and a number of enzymes (both prokaryotic and eukaryotic) have been described that specifically cleave the branched stems as an intermediate step in the strand exchange process (reviewed in ref. 15). Sequences capable of adopting cruciforms have been recognized in a number of eukaryotic genes.

Do these alternative DNA structures exist *in vivo*?

Despite their abundance, non-B-DNA forming sequences would be of little physiological significance unless they could actually adopt the alternative structures *in vivo*. The data for eukaryotes have come mainly, if not exclusively, from antibody (AB) binding studies. Many of these studies must be viewed with caution because of two potential artefacts. The first results from the use of acetic acid as a fixative; it extracts histones and may dramatically increase the torsional stress in DNA by converting the constrained DNA supercoils into unconstrained torsion. The free energy released in this way could certainly create non-B structures. Another potential pitfall in using ABs is that simply binding to the respective antigenic sites may itself shift the equilibrium between the B- and the non-B conformations in the direction of the alternative structure. The mere existence of catalytic antibodies, which stabilize the distorted transition states of substrates, argues that this may happen. Thus, it may be taken that most of the AB binding studies only reveal the presence of DNA segments with *potential* for unusual structure formation.

Some of the older and rather controversial data on the presence of Z-DNA in polytene chromosomes of *Drosophila* and *Chironomus* (reviewed in ref. 16) can probably be explained in the above manner. However, several studies have used alternative fixatives that would be less likely to extract proteins (e.g. methanol, ethanol, acetone), and other experiments have used unfixed materials. Thus, Morgenegg et al.[17] have reported positive anti-Z-DNA immunostaining in

rat tissues fixed with several nonacidic agents. Similarly, native transcriptionally active polytene chromosomes of *D. hydei* showed Z-DNA immunological reactivity[18]. Importantly, the cytological correlation between Z-DNA and active transcription in regions of dispersed chromatin was found to be restricted only to specific genes, indicating that only certain genes might be regulated via Z-DNA structures. Data linking Z-DNA immunoreactivity and transcriptional status have been reported using interphase nuclei as well. Lipps et al.[19] have shown that anti-Z-DNA ABs stain the transcriptionally active macronucleus but not the transcriptionally inert micronuclei of *Stylonychia mytilus*. It has also been shown[19] that DNAse I digestion reduces the immunofluorescence, suggesting that supercoiling is involved in stabilizing the conformations that stain for Z-DNA. That the level of Z-DNA in metabolically active, permeabilized mammalian nuclei is also regulated by torsional stress has been demonstrated in the studies of Wittig et al.[20,21]. These authors found a plateau of constant AB binding over a wide range of AB concentrations; this was taken as a measure of 'pre-existing' Z-DNA in nuclei. The plateau level was dependent on the level of torsional stress, for inhibiting topo I by the drug camptothecin led to higher plateau levels; conversely, addition of nicking amounts of DNase I led to a complete loss of AB binding. Additional AB binding could be induced by increasing AB concentration, implying that high concentrations of AB can in fact perturb the B/Z equilibrium. The plateau level of binding was augmented when transcription was experimentally increased, whereas it changed much less during DNA replication. The authors favor the idea that the 'pre-existing' Z-DNA is generated by negative supercoiling in the wake of the passing polymerase, as envisaged by the twin domain model of Liu and Wang[1]. Whether this interpretation is correct has yet to be determined, for we remain uncertain as to how much of such superhelical stress is actually present in cells with active topoisomerases. Indeed, a carefully performed study aimed at determining the level of DNA supercoiling *in vivo* showed that, at least in the bacterial cell, this level is probably determined by topoisomerase activity, not by transcription[22]. Despite these numerous reservations, it is our opinion that current evidence is in favor of the existence of some level of Z-DNA structure in eukaryotic nuclei.

The immunochemical detection of triplex DNA *in vivo* has proved to be even more difficult, as triplex DNA turned out to be only weakly immunogenic and the acidic fixation conditions strongly favor this conformer. A monoclonal AB against $d(Tm^5C)_n$, shown to possess the required specificity for triplex versus duplex DNA, stained both metaphase chromosomes and interphase nuclei from mouse and human cells[23,24]. The staining was unaffected by addition of *E. coli* DNA but was obliterated in the presence of competing triplex DNA. Substitution of cold acetone fixation did not abolish the staining; more importantly, unfixed isolated chromosomes also gave positive immunoreaction. Furthermore, the same antibody reacted positively with polytene chromosomes of *Chironomus* and *Drosophila* that had been formaldehyde-fixed before the acetic acid fixation step; alternative methods of fixation gave comparable results[25]. It is important to note that transcriptionally active nucleolar organizer regions and Balbiani rings were immunonegative. This observation goes well with functional studies (see below) indicating that triplex DNA inhibits transcription. A strong case for the *in vivo* existence of triplex structures has also been presented by Kohwi and Kohwi-Shigematsu[26] (see below).

Labeling of mammalian nuclei with an anti-cruciform monoclonal AB produced a non-uniform pattern of fluorescence in cells arrested at the G_1/S boundary; this pattern changed upon release from the cell cycle block[27]. Although this study seems to correlate cruciform structure mainly with DNA replication, it shows quite convincingly that cruciforms are present in the eukaryotic nuclei. Quantitation establishes that there are more cruciforms than origins of replications, indicating that some cruciforms are involved in other nuclear processes as well. In contrast to anti-Z-DNA and anti-triplex ABs, both of which stain metaphase chromosomes, no immunofluorescence was observed when such chromosomes were treated with the monoclonal AB against cruciform DNA[28].

Specific proteins bind to non-B-DNA structures

Given the existence of non-B structures in eukaryotic nuclei, it is not surprising that a number of proteins have been identified that can be classified as either Z-DNA, triplex, or cruciform binding proteins (Table 1). Some of these studies are to be viewed cautiously because of potential artefacts. Thus, Rohner et al.[29] have noted that the identification of Z-DNA binding proteins critically depends on the proper choice of ligands: while nonhistone proteins HMG1 and 2 bound to brominated synthetic oligonucleotides in the Z-conformation, they failed to preferentially bind to Z-DNA inserts in supercoiled plasmids. (In this context, it should be remembered that the ABs used for detection of Z-DNA *in vivo* were elicited against similar synthetic structures.) Another unexpected artefact was reported by McGhee and colleagues[30,31]. These authors found that some Z-DNA binding proteins were in fact phospholipid-binding proteins which reacted with Z-DNA because of a fortuitous structural similarity of the phospholipid-binding cavity to a site that would interact with Z-DNA. Despite all of these caveats, it must be emphasized that at least some of the Z-DNA binding proteins have been tested using rigorous criteria and seem to be truly specific for Z-DNA (e.g. ref. 32). No major artefacts have been reported in the identification of triplex and cruciform binding proteins.

We should also consider the possible existence of proteins that specifically bind to potential non-B sequences but only in the B-form. Such proteins would be expected to inhibit the B-Z transitions.

Do non-B-DNA structures exist in nucleosomes?

One way in which the presence of local regions of non-B structures might influence transcription is by modifying chromatin organization. But is the nucleosomal structure of chromatin, in fact, sensitive to such conformational changes in the DNA?

Table 1. *Examples of non-B-DNA binding proteins*

Source	Designation	Purification	Properties	Proposed role	Reference*
Z-DNA binding proteins					
SV40 mini-chromosomes		Z-DNA affinity chromatography	High preference for Z-over B-DNA; binding to negatively supercoiled plasmids containing Z-DNA stretches; binding is not due to supercoiling *per se*; possible sequence specificity	B-Z-DNA intraconversions in regulatory regions of SV40	1
S. cerevisiae	Zuotin	Conventional and Z-DNA affinity chromatography	Binds poly(dG-m^{5}dC), oligo-(dG-Br^{5}dC)$_{22}$ and negatively supercoiled pUC19 containing a Z-DNA insert	Chromatin condensation	2
Drosophila		Combination of B- and Z-DNA affinity chromatography	High preference for Z-over B-DNA; binding to negatively supercoiled plasmids containing Z-DNA stretches; possible sequence specificity	B-Z-DNA intraconversions and stabilization of Z-DNA	3
Drosophila	Topoisomerase II	Conventional chromatography	Higher affinity to minicircles containing Z-DNA inserts; binds preferentially to Z-DNA but cleaves elsewhere	Potential feedback mechanism for regulation of chromatin topology	4
Chicken erythrocytes	BGP1	Conventional and poly(dG-dC) chromatography	Binds to any contiguous set of 7 or more Gs; 1 molecule/string; requires Zn^{++}	Binding to the G-string may modulate the conformational switch and/or affect nucleosome positioning	5, 6
Chicken blood cells	Zα	Conventional chromatography	Binds to brominated (dC-dG)$_{35}$; binding is competed by Z- but not by B-DNA and by supercoiled plasmids containing Z-DNA insert; binding constant ~ 10^{10} M^{-1}	Stabilization of Z-DNA	7
Rat neurons		Combination of B- and Z-DNA affinity chromatography	Strong preference to Z- versus B-DNA	Z-DNA binding and induction	8
Rabbit uterus	Estrogen receptor	Sucrose density centrifugation	Relative affinity to poly(dG-dT)·poly-(dA-dC) in the presence of polyamines ~ 20-fold higher than in their absence	Regulation of gene expression via B- to Z-DNA transitions in regulatory elements	9
HeLa and BHK21 cells		Ammonium sulphate precipitation or ion-exchange chromatography (cytoplasmic extracts)	Cause gel retardation of Z-DNA (poly d[(G-C)] in the presence of cobalt ions)	Regulation of events involving specific recognition	10
Wheat germ		Z-DNA affinity chromatography	Affinity for Z-DNA 10^5× greater than for B-DNA; all other properties as of the proteins from SV40 minichromosomes (see above)	B-Z-DNA intraconversions and stabilization of Z-DNA	11
Triplex-DNA binding proteins					
HeLa cells		Triplex DNA affinity chromatography	Affinity 2- to 5-fold higher for TAT triplex than for AT duplex	Association of two distantly located AT duplexes to loop intervening DNA	12
Human	NSEP-1	Recombinant protein; conventional chromatography	Binds to CT-rich elements with strong pur/pyr strand asymmetry; binds also to the C-rich strand	Helps form non-B structures *in vivo*	13
Cruciform-DNA binding proteins†					
Human lympho-blasts and other cell lines		Single-stranded DNA-cellulose	Forms structure-specific complexes with synthetic four- and three-way junctions		14
Rat liver	HMG1	Conventional and affinity chromatography	Forms structure-specific complexes with synthetic four- and three-way junctions	Recognition, stabilization and protection of cruciforms	15

†Cruciform-cutting endonucleases that resolve Holliday junctions in the process of homologous recombination are not included here. Such enzymes have been isolated from a variety of eukaryotic sources (see ref. 16, for references).
*References:

1 Azorin, F. and Rich, A. (1985). *Cell* **41**, 365-374.
2 Zhang, S. et al. (1992). *EMBO J.* **11**, 3787-3796.
3 Nordheim, A. et al. (1982). *Proc. Natl. Acad. Sci. USA* **79**, 7729-7733.
4 Glikin, G. C. et al. (1991). *Nucl. Acids Res.* **19**, 7139-7144.
5 Lewis, C. D. et al. (1988). *Genes Dev.* **2**, 863-873.
6 Clark, S. P. et al. (1990). *Nucl. Acids Res.* **18**, 5119-5126.
7 Herbert, A. G. et al. (1993). *Proc. Natl. Acad. Sci. USA* **90**, 3339-3342.
8 Kuenzle, C. C. et al. (1983). *Cold Spring Harbor Symp. Quant. Biol.* **48**, 493-499.
9 Thomas, T. and Kiang, D. T. (1988). *Nucl. Acids Res.* **16**, 4705-4720.
10 Leith, I. R. et al. (1988). *Nucl. Acids Res.* **16**, 8277-8289.
11 Lafer, E. M. et al. (1985). *Biochemistry* **24**, 5070-5076.
12 Kiyama, R. and Camerini-Otero, R. D. (1991). *Proc. Natl. Acad. Sci. USA* **88**, 10450-10454.
13 Kolluri, R. et al. (1992). *Nucl. Acids Res.* **20**, 111-116.
14 Elborough, K. M. and West, S. C. (1988). *Nucl. Acids Res.* **16**, 3603-3616.
15 Bianchi, M. E. et al. (1989). *Science* **243**, 1056-1059.
16 Kleff, S. et al. (1992). *EMBO J.* **11**, 699-704.

Z-DNA

The question as to whether Z-DNA can be contained within nucleosomes has been controversial. Nickol et al.[33] found that whereas histones would reassociate with B-form poly(dGm^{5}dC)·poly(dGm^{5}dC) to form nucleosomes, the same polymer when switched to the Z-form by addition of $Co(NH_3)_6^{3+}$ produced only poorly defined aggregates. Furthermore, the B-Z transition could not be induced on the DNA when it was incorporated into nucleosomes. This result was challenged by Miller et al.[34], who criticized both the method of nucleosome reconstitution (direct mixing in 200 mM salt) and the use of $Co(NH_3)_6^{3+}$ to induce the Z-form. Employing instead a trout testis nucleosome assembly factor and $MgCl_2$ for Z-stabilization, they reported evidence for the formation of 'core particles' containing Z-DNA and claimed that nucleosomes formed on B-DNA remained stable upon B-Z conversion. The demonstration of a Z-DNA-like CD spectrum for a nucleosome-like particle seemed to support these arguments.

However, more recent studies strongly question the conclusions of Miller et al.[34]. Garner and Felsenfeld[35] reconstituted nucleosomes onto plasmids carrying a potential Z-forming test sequence. When the plasmid was relaxed (and the sequence was in the B-form) core particles could readily form on the test sequence. But supercoiling, with accompanying conversion of the test sequence to the Z-form, greatly inhibited nucleosome formation thereon. Another aspect of the problem was re-examined in a study where poly(dGm^{5}dC)·poly(dGm^{5}dC) was reconstituted into oligonucleosomes by salt-gradient dialysis[36]. Digestion with micrococcal nuclease produced well defined core particles. However, when the Mg^{2+} concentration was increased to induce the B-Z transition, substantial dissociation was observed. In fact, the observed change in circular dichroism could be accounted for *entirely* by the assumption that only DNA that had dissociated from histones had undergone the B-Z transition.

In our opinion, the balance of evidence is against the idea that the Z-conformation is compatible with nucleosomal structure. This is reasonable on theoretical grounds as well, for the positioning and spacing of phosphate groups is different in the two conformations, and Z-DNA is much stiffer than B-DNA. An order-of-magnitude calculation indicates that the difference in stiffness alone would disfavor Z-nucleosomes by about 12 kcal/mol.

This does not imply, of course, that eukaryotic sequences with *potential* for Z-DNA formation cannot lie within nucleosomes. Indeed, Gross et al.[37] show that most such sequences are intranucleosomal in mouse cell nuclei, and Rodriguez-Campos et al.[7] find that a $d(CA/GT)_n$ sequence inserted into SV40 is, *in vivo*, mostly intranucleosomal, and not in the Z form.

In summary, the bulk of the evidence implies that regions of *actual* Z-DNA structure in chromatin will lie in non-nucleosomal regions, either in linker DNA or in DNase I-hypersensitive regions.

Triplex DNA

There have been, to our knowledge, no experimental studies of the relationship of triplex DNA to nucleosome formation or stability. It would seem, *a priori*, highly unlikely that such a very different structure could be accommodated on the specialized surface of the nucleosomal core. Thus, we would predict the triplexes will also exist only in extranucleosomal regions.

Cruciforms

The earliest study[38] actually utilized a synthetic stem-loop heteroduplex structure, which, unlike a palindrome, does not require superhelical torsion for formation or stabilization. Reconstitution of a control homoduplex with histone octamer produced extensive protection against restriction endonucleases, yet the stem-loop structure in the heteroduplex remained wholly unprotected, implying that it was excluded from nucleosomes. In further experiments, Nobile et al.[39] mapped the positions of nucleosomes on reconstitutes lacking or containing the stem-loop. Here the results were unequivocal: when the stem-loop was present, nucleosome positions were rearranged so as to specifically avoid it.

A seemingly contradictory result was reported by Caffarelli et al.[40], who examined bacterial plasmids containing potential cruciforms after reconstitution with varying levels of histones. They found that addition of histone octamers, which should relax superhelical torsion, failed to completely eliminate S1 cleavage of the supercoiled plasmid. Additional studies using psoralen cross-linking were interpreted to indicate a *preference* of the cruciform-forming regions to nucleosomes. The results with stem-loops are clear; those with the cruciform require further study. It seems unlikely to us, however, that a nucleosome could better tolerate a whole cruciform than a single stem-loop.

In summary, we propose that none of the non-B structures described above can reside easily (or at all) within the DNA wrapped about a histone octamer. The core particle and non-B DNA then represent *alternative* options for any sequence capable of undergoing such a transition.

Proteins known to interact with the linker DNA in chromatin recognize and bind specifically to non-B-DNA structures

The majority of studies on this issue concern the nonhistone proteins HMG1 and 2. These proteins are abundant, known to bind to the linker DNA and to be significantly enriched in transcriptionally active chromatin regions (reviewed in ref. 41). They have recently attracted much attention as they contain two homologous DNA-binding domains, each of about 80 amino acid residues in length. Variations of these domains, termed the HMG boxes, are present in many additional DNA binding proteins, like transcription factors, the sex-determining factor and yeast proteins involved in mating-type control (for a review, see ref. 42).

Initially, the interaction of HMG1 with non-B structures was studied in supercoiled plasmids known to contain regions of alternative conformation. It was shown that the preferential binding to supercoiled versus linear DNA, which had been known for years, must involve the presence of S1-sensitive sites, for these were selectively protected against the nuclease

by HMG binding (e.g. as in ref. 43). A hierarchy of binding to different conformations (B-Z junctions>cruciforms>other S1-sensitive structures) was observed[43]. This conclusion was, however, challenged by Sheflin and Spaulding[44], who explain the observed protection against single-strand-specific nucleases by the dramatic reduction in the superhelical density caused by HMG1 binding, and hence the loss of alternative structures. The overcoming by HMG1 of the transcriptional block caused by Z-DNA or cruciforms in supercoiled DNA[45,46] can also be explained by the same mechanism: HMG1 binds to supercoiled DNA, absorbs its tension and causes the non-B structures to flip back to the B-conformation.

More recently, the possibility of HMG1 interacting with some non-B structures has been directly demonstrated in experiments that did not involve superhelicity. Bianchi et al.[47] used stable four-way junctions, formed by annealing appropriate oligodeoxyribonucleotides, to show that HMG1 binds preferentially to cruciforms. Each of the two HMG boxes described above retained this specificity of binding[48]. Recent NMR studies on the three-dimensional folding of one of the HMG boxes suggest that binding of HMG1 to the four-way junction of the cruciform is determined by the specific L-shaped form of the box[49]. It should be noted that the binding of HMG1 to cruciforms might be merely fortuitous, and that it might actually prefer binding to highly bent DNA stretches, which four-way junctions may imitate well. An interesting speculation based on studies of SRY, the sex-determining factor in humans, envisages that due to its high concentration in the nucleus, HMG1 should be able to saturate DNA sites that for a variety of reasons are transiently bent. In so doing they will channel binding of proteins like SRY to their sequence-specific sites; in the absence of HMG1 these will bind nonspecifically to bent regions, which they seem to prefer over straight regions[50]. Alternatively, binding to cruciforms can occur because they resemble the natural substrates of these proteins formed by the wrapping or looping of DNA, as in chromatin[42]. What the functional significance of HMG1 binding to cruciforms may be remains to be established.

The other major linker DNA binding protein in chromatin is histone H1. We have recently shown that it also binds to stable synthetic cruciforms[51], as expected in view of the geometrical similarity of this structure to cross-overs of double stranded DNA to which the histone shows high preference[52]. Histone H1 seems also to be a Z-DNA binding protein[53,54]. In confirmation of this, H1 has been purified by Z-DNA affinity chromatography (unpublished results of K. Lowenhaupt and A. Rich, cited in ref. 55). Interestingly, zuotin, the Z-DNA binding protein recently isolated from *S. cerevisiae* (see Table 1), possesses a region of high homology to histone H1[55]. These initial findings have to be further extended in order to understand whether and how the linker DNA binding proteins can influence transcription through binding to non-B structures.

Alternative non-B-DNA structures and eukaryotic transcription

Data on the effect of alternative non-B-DNA structure on eukaryotic transcription are still scarce and rather controversial. Furthermore, in evaluating data we must consider whether the effects are on transcriptional elongation or initiation.

Synthetic polymers in the Z-DNA conformation can be transcribed, albeit less efficiently than B-DNA, by wheat germ RNA polymerase II[56]. Hipskind and Clarkson[57] have studied the effect of 5′-flanking sequences on the expression of *Xenopus* tRNA genes in a homologous *in vitro* transcription system. A potential Z-DNA forming sequence (9 alternating pur.pyr), situated in the 5′ flank of a poorly expressed variant methionine tRNA gene, was inhibitory to transcription, possibly by forming Z-DNA. Studies using synthetic constructs containing Z-DNA-forming sequences at different positions from a reporter gene have provided conflicting evidence. Thus, cloned $d(TG)_n$ segments showed a strong inhibitory effect on *in vivo* transcription in *Xenopus* oocytes when placed in the intermediate flanking sequences or when located between the two portions of the internal split promoter of *Caenorhabditis elegans* $tRNA^{Pro}$ gene[58]. At the same time, placing the d(TG)n segments far upstream or downstream from the gene did not affect its transcription. The data, as discussed by the authors themselves, do not allow firm conclusions about the role of Z-DNA in transcription. The puzzling observation that the transcriptional efficiency remained the same after linearization of the plasmid suggested that the inhibitory effect was determined by the DNA sequence itself rather than by the secondary structure. A more recent study[59] using the long stretch of alternating pur.pyr sequences from the rat prolactin gene linked to the CAT reporter gene also reported inhibition of *in vivo* transcription.

Synthetic $d(TG)_n$ sequences, or human genomic fragments containing such tracts ligated to expression plasmids, were used to study transcription in transfected CV-1 or HeLa cells[60]. In contrast to the findings in ref. 58, the TG tracts enhanced gene expression and acted from a distance and in an orientation-independent manner. However, the sequence $d(CG)_n$, which is known to form Z-DNA more easily than $(TG)_n$, did not enhance transcription. This rather unexpected result was explained by the possible presence of sequence-specific Z-DNA binding proteins, such that would bind to TG-, but not to CG-repeats (see Table 1 for examples of sequence-specific Z-DNA binding proteins).

There are some additional indirect indications that $(TG)_n$ might, in fact, increase the transcriptional rate of certain genes. *In situ* hybridization to polytene chromosomes of different *Drosophila* species revealed that $(TG)_n$ repeats are nonrandomly distributed[61,62]. The enrichment of the X chromosome in these sequences suggested their involvement via enhanced transcription, in dosage compensation in the male (the X chromosome compensates for the different dosages of its genes in the two sexes by modification of its transcriptional activity). As in the transcription studies above, however, it remains to be established whether the presumed effects on transcription are due to the formation of the alternative non-B structures at these sequences or to some other factors.

The data on the effect of triplex DNA on eukaryotic tran-

scription come almost exclusively via the use of synthetic oligonucleotides that form *intermolecular* triplexes with selected target sites in gene sequences. Cooney et al.[63] reported that a 27 bp-long oligodeoxyribonucleotide bound to duplex DNA within the 5′ flank of the human c-myc gene to form a triple helix which inhibited *in vitro* transcription in HeLa cell extracts. The oligonucleotide was also shown to inhibit transcription in transfected living HeLa cells[64]. Importantly, the inhibition was both site- and oligodeoxyribonucleotide-specific. These, and similar studies (e.g. refs 65 and 66) are of enormous interest in view of possible clinical use of sequence-specific oligonucleotides to repress transcription of oncogenes. The mechanism of repression is thought to involve both blocking of promoter assembly into initiation complexes[66] and interference with binding of transcription factors to their cognate sequences[67-69]. These studies also suggest that *intermolecular* triplexes formed by short RNAs or RNPs can represent a functionally important way of regulation of transcriptional activity *in vivo*.

Is there any evidence whatsoever to implicate *intramolecular* triplex formation in transcriptional regulation? A possible triplex was implicated in the activity of a proto-oncogene promoter, for transient expression assays showed that deletion mutants removing this sequence possessed reduced transcriptional activity, which correlated with reduced protein binding[70]. Whether or not this sequence actually adopts the triplex structure *in vivo* remains to be established. A more direct study of the participation of triplex DNA in transcriptional regulation was that of Kohwi and Kohwi-Shigematsu[26]. The authors demonstrated that a poly(dG).poly(dC) sequence, which can adopt triplex conformation under superhelical torsion *in vitro*, strongly affected the *in vivo* expression of genes when placed 5′ to a promoter. The effect was strikingly dependent on the length of the sequence: dG tracts up to 30 bp augmented transcription, whereas tracts of more than 35 bp had no effect. The augmenting tracts competed *in vivo* for a *trans*-acting factor, while the longer tracts failed to compete. Importantly, the observed activities *in vivo* correlated with the presence of triplex structures *in vitro*. Technical difficulties precluded *in situ* probing of the DNA structure in the recipient cells. The results indicated active participation of the non-B-DNA forming sequences in transcriptional regulation *in vivo*: the presence of G-strings in double helical conformation activates transcription, probably by binding of specific protein factors; the B-DNA to triplex transition leads to inhibition of transcription via interference with factor binding.

How might non-B-DNA structures regulate transcription?

The continuing evidence for the existence of regions of non-B structures in eukaryotic genomes, coupled with the growing list of proteins with specific affinity for such structures, poses the question: what is all this *for*? Does it function to regulate transcription?

One obvious possibility is that the binding of transcription factors to such sequences might be either augmented or inhibited by change in the conformation. However, such explanations cannot suffice in all cases, especially when the sequences act in a distance-independent manner.

A number of years ago, Weintraub[71] pointed out the possibility of local interactions between chromatin structure and non-B DNA conformations. He suggested that steric hindrance in nuclear structure might *localize* transient changes in DNA topology to a region of a few nucleosomes. That is, it is unlikely that torsional strain can be quickly dissipated over a large chromatin region.

Consider, then, the following hypothetical scenario, as depicted in Fig. 2. The promoter region of a gene is blocked by a nucleosome in such a way that sites for transcription factors are inaccessible. There exists in an adjacent linker region a site for potential non-B structure (let us say a potential Z-DNA site). If Z-DNA binding proteins are available, this site will be at least periodically occupied, and switched from the B-state to the Z-state. When this happens, strong *positive* torsional strain must be generated in the immediately adjacent DNA. If this is confined to, say, a 200 bp region, one turn of B-DNA converting to Z-DNA will impose a superhelical density of about +0.05 to +0.1. This might be relaxed by topoisomerases; alternatively, an adjacent nucleosome might be dissociated with concomitant relaxation of the positive tension. If this nucleosome has been covering transcription factor binding sites and *if* (but only if) the appropriate factors are available in the cell, the relaxed state can be stabilized by their binding, and gene activation will be accomplished.

As Fig. 2 shows, the initial step could equally well be the temporary dissociation of the nucleosome, followed by a B-Z transition to relieve the negative stress created; this was the model suggested by Weintraub[71]. In either case, the intermediate states designated by (II) in Fig. 2 are to be regarded as transient, energetically unfavorable states.

Note that the Z-DNA binding protein only *potentiates* transcriptional activation. It is not a sufficient condition, for only if transcription factors bind in the interval before a nucleosome core rebinds will the system be stable in its activated state. The process can be viewed in another way: in a topologically constrained environment, a nucleosome and a nearby non-B structure represent alternative ways of accommodating negative superhelicity. One may imagine a dynamic equilibrium between the two, whose direction can be shifted by the presence (or absence) of proteins specific for non-B structures. We can thus imagine that promoter regions harboring potential non-B sequences will be periodically 'scanned' by the appropriate binding proteins. Only if the correct transcription factors are present will the open state be fixed. The mechanism envisaged does not depend upon either the precise location or the orientation of the non-B sequence. We refer to this proposed mechanism as '*conformational compensation*'.

Another possibility for controlling the transcriptional process could be through the alternative binding of histone H1 and HMG1/2 to non-B structures in the linker DNA. These proteins might be able to discriminate among subtle structural features of otherwise closely related non-B conformers and bind them specifically. In view of what is known about the effect of these two groups of proteins on transcrip-

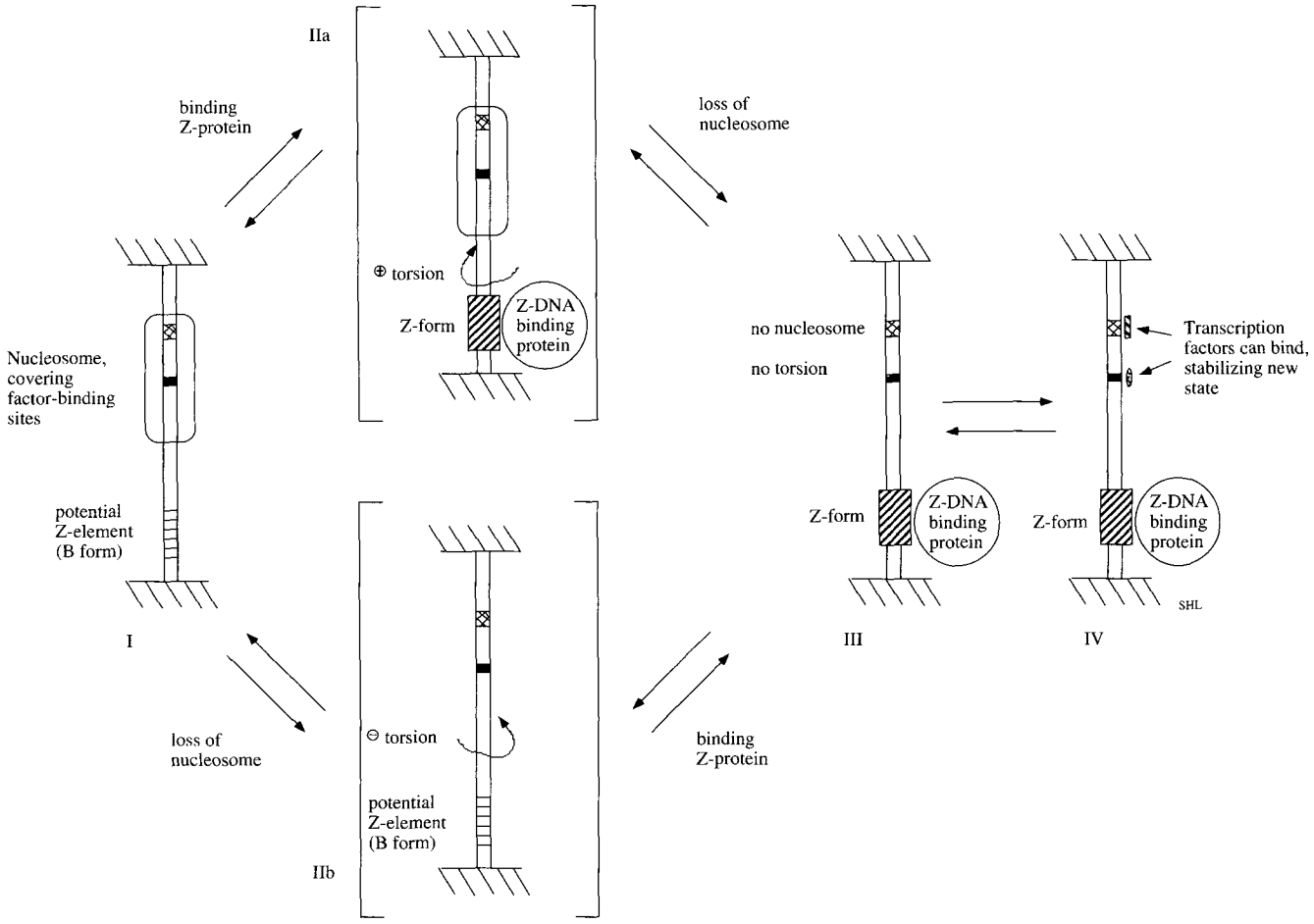

Fig. 2. Conformational compensation as a potential model for gene activation. A region of chromatin is constrained by surrounding structures in the nucleus. Here for specificity we imagine a segment containing one nucleosome in the promoter regions of a gene with a potential Z-element in an adjacent linker (state I). Either binding of a Z-protein or transient dissociation of the nucleosome lead to energetically unfavorable transition states (IIa or IIb) with DNA under torsion. This can be relaxed as shown in III, a state which could then be stabilized by the binding of transcription factors to uncovered sites (state IV).

tion, it is reasonable to suggest that binding of H1 would inhibit, whereas binding of HMG1 would stimulate transcription.

It is also possible that non-B structures play a role in regulating transcriptional elongation. It now seems clear that the transcriptional process alone generates levels of negative supercoiling high enough to drive the formation of alternative structures in the wake of elongating polymerases, at least in bacterial cells[16]. It is also clear that polymerases tend to stop elongating when they encounter Z-DNA, cruciform or H-DNA structures[45,46]. An attractive feedback mechanism has been suggested to us by Dr G. Schroth, wherby polymerase activity could be metered by the formation of non-B-DNA structures within transcribed regions of genes (also discussed recently in ref. 71). If transcriptional activity in a given gene becomes very high, such that local topoisomerase activity is not able to adequately relax the supercoils generated by transcription, then the buildup of negative supercoils behind the polymerase could drive the formation of non-B structures within the gene. These structures would cause elongating polymerase complexes to pause until local topoisomerase activity could lower the superhelical density into a range that would destabilize the non-B structures.

In conclusion, it seems very probable that non-B-DNA structures are involved in the control of eukaryotic transcription. However the mechanisms by which this is affected remain purely speculative and require further experimental studies.

Acknowledgements

We thank Dr G. P. Schroth for critical reading of the manuscript and useful discussions, and Dr S. H. Leuba for help with the figures. K.v.H. is the recipient of an American Cancer Society Research Professorship. This research was supported by NIH Grant GM22916 and NIEHS Grant ES04766 to K. v. H.

References

1 Liu, L. F. and Wang, J. C. (1987). Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. USA* **84**, 7024-7027.

2 Palecek, E. (1991). Local supercoil-stabilized DNA structures. *Crit. Rev. Biochem. Mol. Biol.* **26**, 151-226.

3 Hamada, H. and Kakunaga, T. (1982). Potential Z-DNA forming sequences are highly dispersed in the human genome. *Nature* **298**, 396-398.

4 Gross, D. S. and Garrard, W. T. (1986). The ubiquitous potential Z-forming sequence of eukaryotes, $(dT\text{-}dG)_n{\cdot}(dC\text{-}dA)_n$, is not detectable in the genomes of eubacteria, archaebacteria, or mitochondria. *Mol. Cell. Biol.* **6**, 3010-3013.

5 Delseny, M., Laroche, M. and Penon, P. (1983). Detection of sequences with Z-DNA forming potential in higher plants. *Biochem. biophys. Res. Commun.* **116**, 113-120.

6 Schroth, G. P., Chou, P.-J. and Ho, P. S. (1992). Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. *J. Biol. Chem.* **267**, 11846-11855.

7 Rodriguez-Campos, A., Ellison, M. J., Perez-Grau, L. and Azorin, F. (1986). DNA conformation and chromatin organization of a $d(CA/GT)_{30}$ sequence cloned in SV40 minichromosomes. *EMBO J.* **5**, 1727-1734.

8 Kohwi, Y. (1989). Cationic metal-specific structures adopted by the poly(dG) region and the direct repeats in the chicken adult β^A globin gene promoter. *Nucl. Acids Res.* **17**, 4493-4502.

9 Htun, H. and Dahlberg, J. E. (1989). Topology and formation of triple-stranded H-DNA. *Science* **243**, 1571-1576.

10 Behe, M. J. (1987). The DNA sequence of the human β-globin region is strongly biased in favor of long strings of contiguous purine or pyrimidine residues. *Biochemistry* **26**, 7870-7875.

11 Beasty, A. M. and Behe, M. J. (1988). An oligopurine sequence bias occurs in eukaryotic viruses. *Nucl. Acids Res.* **16**, 1517-1528.

12 Manor, H., Rao, B. S. and Martin, R. G. (1988). Abundance and degree of dispersion of genomic $d(GA)_n{\cdot}d(TC)_n$ sequences. *J. Mol. Evol.* **27**, 96-101.

13 Larsen, A. and Weintraub, H. (1982). An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. *Cell* **29**, 609-622.

14 Glaser, R. L., Thomas, G. H., Siegfried, E., Elgin, S. C. R. and Lis, J. T. (1990). Optimal heat-induced expresson of the *Drosophila* hsp26 gene requires a promoter sequence containing $(CT)_n{\cdot}(GA)_n$ repeats. *J. Mol. Biol.* **211**, 751-761.

15 Duckett, D. R., Murchie, A. I. H., Bhattacharyya, A., Clegg, R. M., Diekmann, S., von Kitzing, E. and Lilley, D. M. J. (1992). The structure of DNA junctions and their interaction with enzymes. *Eur. J. Biochem.* **207**, 285-295.

16 Hill, R. J. (1991). Z-DNA: a prodrome for the 1990s. *J. Cell Sci.* **99**, 675-680.

17 Morgenegg, G., Celio, M. R., Malfoy, B., Leng, M. and Kuenzle, C. C. (1983). Z-DNA immunoreactivity in rat tissues. *Nature* **303**, 540-543.

18 Lancillotti, F., Lopez, M. C., Arias, P. and Alonso, C. (1987). Z-DNA in transcriptionally active chromosomes. *Proc. Natl. Acad. Sci. USA* **84**, 1560-1564.

19 Lipps, H. J., Nordheim, A., Lafer, E. M., Ammermann, D., Stollar, B. D. and Rich, A. (1983). Antibodies against Z-DNA react with the macronucleus but not the micronucleus of the hypotrichous ciliate *Stylonychia mytilus*. *Cell* **32**, 435-441.

20 Wittig, B., Dorbic, T. and Rich, A. (1989). The level of Z-DNA in the metabolically acitive, permeabilized mammalian cell nuclei is regulated by torsional strain. *J. Cell Biol.* **108**, 755-764.

21 Wittig, B., Dorbic, T. and Rich, A. (1991). Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. *Proc. Natl. Acad. Sci. USA* **88**, 2259-2263.

22 Cook, D. N., Ma, D., Pon, N. G. and Hearst, J. E. (1992). Dynamics of supercoiling by transcription in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **89**, 10603-10607.

23 Lee, J. S., Burkholder, G. D., Latimer, L. J. P., Haug, B. L. and Braun, R. P. (1987). A monoclonal antibody to triplex DNA binds to eukaryotic chromosomes. *Nucl. Acids Res.* **15**, 1047-1061.

24 Burkholder, G. D., Latimer, L. J. P. and Lee, J. S. (1988). Immunofluorescent staining of mammalian nuclei and chromosomes with a monoclonal antibody to triplex DNA. *Chromosoma* **97**, 185-192.

25 Burkholder, G. D., Latimer, L. J. P. and Lee, J. S. (1991). Immunofluorescent localization of triplex DNA in polytene chromosomes of *Chironomus* and *Drosophila*. *Chromosoma* **101**, 11-18.

26 Kohwi, Y. and Kohwi-Shigematsu, T. (1991). Altered gene expression correlates with DNA structure. *Genes Dev.* **5**, 2547-2554.

27 Ward, G. K., McKenzie, R., Zannis-Hadjopoulos, M. and Price, G. B. (1990). The dynamic distribution and quantification of DNA cruciforms in eukaryotic nuclei. *Exp. Cell Res.* **188**, 235-246.

28 Ward, G. K., Shihab-el-Deen, A., Zannis-Hadjopoulos, M. and Price, G. B. (1991). DNA cruciforms and the nuclear supporting structure. *Exp. Cell Res.* **195**, 92-98.

29 Rohner, K. J., Hobi, R. and Kuenzle, C. C. (1990). Z-DNA-binding proteins. Identification critically depends on the proper choice of ligands. *J. Biol. Chem.* **265**, 19112-19115.

30 Krishna, P., Kennedy, B. P., van de Sande, J. H. and McGhee, J. D. (1988). Yolk proteins from nematodes, chickens, and frogs bind strongly and preferentially to left-handed Z-DNA. *J. Biol. Chem.* **263**, 19066-19070.

31 Krishna, P., Kennedy, B. P., Waisman, D. M., van de Sande, J. H. and McGhee, J. D. (1990). Are many Z-DNA binding proteins actually phospholipid-binding proteins? *Proc. Natl. Acad. Sci. USA* **87**, 1292-1295.

32 Herbert, A. G., Spitzner, J. R., Lowenhaupt, K. and Rich, A. (1993). Z-DNA binding protein from chicken blood nuclei. *Proc. Natl. Acad. Sci. USA* **90**, 3339-3342.

33 Nickol, J., Behe, M. and Felsenfeld, G. (1982). Effect of the B-Z transition in $poly(dG\text{-}m^5dC){\cdot}poly(dG\text{-}m^5dC)$ on nucleosome formation. *Proc. Natl. Acad. Sci. USA* **79**, 1771-1775.

34 Miller, F. D., Dixon, G. H., Rattner, J. B. and van de Sande, J. H. (1985). Assembly and characterization of nucleosomal cores on B- vs. Z-form DNA. *Biochemistry* **24**, 102-109.

35 Garner, M. M. and Felsenfeld, G. (1987). Effect of Z-DNA on nucleosome placement. *J. Mol. Biol.* **196**, 581-590.

36 Ausio, J., Zhou, G. and van Holde, K. (1987). A reexamination of the reported B-Z DNA transition in nucleosomes reconstituted with $poly(dG\text{-}m^5dC){\cdot}poly(dG\text{-}m^5dC)$. *Biochemistry* **26**, 5595-5599.

37 Gross, D. S., Huang, S.-Y. and Garrard, W. T. (1985). Chromatin structure of the potential Z-forming sequence $(dT\text{-}dG)_n{\cdot}(dC\text{-}dA)_n$. Evidence for an 'alternating-B' conformation. *J. Mol. Biol.* **183**, 251-265.

38 Nickol, J. and Martin, R. G. (1983). DNA stem-loop structures bind poorly to histone octamer cores. *Proc. Natl. Acad. Sci. USA* **80**, 4669-4673.

39 Nobile, C., Nickol, J. and Martin, R. G. (1986). Nucleosome phasing on a DNA fragment from the replication origin of simian virus 40 and rephasing upon cruciform formation of the DNA. *Mol. Cell. Biol.* **6**, 2916-2922.

40 Caffarelli, E., Leoni, L., Sampaolese, B. and Savino, M. (1986). Persistence of cruciform structure and preferential location of nucleosomes on some regions of pBR322 and ColE1 DNAs. *Eur. J. Biochem.* **156**, 335-342.

41 Bustin, M., Lehn, D. A. and Landsman, D. (1990). Structural features of HMG chromosomal proteins and their genes. *Biochim. Biophys. Acta* **1049**, 231-243.

42 Lilley, D. M. J. (1992). HMG has DNA wrapped up. *Nature* **357**, 282-283.

43 Hamada, H. and Bustin, M. (1985). Hierarchy of binding sites for chromosomal proteins HMG1 and 2 in supercoiled deoxyribonucleic acid. *Biochemistry* **24**, 1428-1433.

44 Sheflin, L. G. and Spaulding, S. W. (1989). High mobility group protein 1 preferentially conserves torsion in negatively supercoiled DNA. *Biochemistry* **28**, 5658-5664.

45 Waga, S., Mizuno, S. and Yoshida, M. (1988). Nonhistone protein HMG1 removes the transcriptional block caused by left-handed Z-form segment in a supercoiled DNA. *Biochem. Biophys. Res. Commun.* **153**, 334-339.

46 Waga, S., Mizuno, S. and Yoshida, M. (1990). Chromosomal protein HMG1 removes the transcriptional block caused by the cruciform in supercoiled DNA. *J. Biol. Chem.* **265**, 19424-19428.

47 Bianchi, M. E., Beltrame, M. and Paonessa, G. (1989). Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**, 1056-1059.

48 Bianchi, M. E., Falciola, L., Ferrari, S. and Lilley, D. M. J. (1992). The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins. *EMBO J.* **11**, 1055-1063.

49 Weir, H. M., Kraulis, P. J., Hill, C. S., Raine, A. R. C., Laue, E. D. and Thomas, J. O. (1993). Structure of the HMG box motif in the B-domain of HMG1. *EMBO J.* **12**, 1311-1319.

50 Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. and Bianchi, M. E. (1992). SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J.* **11**, 4497-4506.

51 Varga-Weisz, P., van Holde, K. and Zlatanova, J. (1993). Preferential binding of histone H1 to four-way helical junction DNA. *J. Biol. Chem.*,**268**, 20699-20700.

52 Krylov, D., Leuba, S., van Holde, K. and Zlatanova, J. (1993). Histones H1 and H5 interact preferentially with cross-overs of double helical DNA. *Proc. Natl. Acad. Sci. USA* **90**, 5052-5056.

53 Mura, C. V. and Stollar, B. D. (1984). Interactions of H1 and H5 histones with polynucleotides of B- and Z-DNA conformations. *Biochemistry* **23**, 6147-6152.

54 Russell, W. C., Precious, B., Martin, S. R. and Bayley, P. M. (1984). Differential promotion and suppression of Z-B transitions in poly[d(G-C)] by histone subclasses, polyamino acids and polyamines. *EMBO J.* **2**, 1647-1653.

55 Zhang, S., Lockshin, C., Herbert, A., Winter, E. and Rich, A. (1992). Zuotin, a putative Z-DNA binding protein in *Saccharomyces cerevisiae*. *EMBO J.* **11**, 3787-3796.

56 Durand, R., Job, C., Zarling, D. A., Teissiere, M., Jovin, T. M. and Job, D. (1984). Comparative transcription of right-and left-handed poly[d(G-C)] by wheat germ RNA polymerase II. *EMBO J.* **2**, 1707-1714.

57 Hipskind, R. A. and Clarkson, S. G. (1983). 5′-flanking sequences that inhibit *in vitro* transcription of a *Xenopus laevis* tRNA gene. *Cell* **34**, 881-890.

58 Santoro, C., Costanzo, F. and Ciliberto, G. (1984). Inhibition of eukaryotic tRNA transcription by potential Z-DNA sequences. *EMBO J.* **3**, 1553-1559.

59 Naylor, L. H. and Clark, E. M. (1990). $d(TG)_n{\cdot}d(CA)_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. *Nucl. Acids Res.* **18**, 1595-1601.

60 Hamada, H., Seidman, M., Howard, B. H. and Gorman, C. M. (1984). Enhanced gene expression by the poly(dT-dG)·poly(dC-dA) sequence. *Mol. Cell. Biol.* **4**, 2622-2630.

61 Huijser, P., Henning, W. and Dijkhof, R. (1987). Poly(dC-dA/dG-dT) repeats in the *Drosophila* genome: a key function for dosage compensation and position effects? *Chromosoma* **95**, 209-215.
62 Pardue, M. L., Lowenhaupt, K., Rich, A. and Nordheim, A. (1987). $(dC\text{-}dA)_n\cdot(dG\text{-}dT)_n$ sequences have evolutionarily conserved chromosomal locations in *Drosophila* with implications for roles in chromosome structure and function. *EMBO J.* **6**, 1781-1789.
63 Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. and Hogan, M. E. (1988). Site-specific oligonucleotide binding represses transcription of the human c-myc gene *in vitro*. *Science* **241**, 456-459.
64 Postel, E. H., Flint, S. J., Kessler, D. J. and Hogan, M. E. (1991). Evidence that a triplex-forming oligodeoxyribonucleotide binds to the c-myc promoter in HeLa cells, thereby reducing c-myc mRNA levels. *Proc. Natl. Acad. Sci. USA* **88**, 8227-8231.
65 Grigoriev, M., Praseuth, D., Robin, P., Hemar, A., Saison-Behmoaras, T., Dautry-Varsat, A., Thuong, N. T., Helene, C. and Harel-Bellan, A. (1992). A triple helix-forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor via inhibition of NF κB binding to interleukin-2 receptor α–regulatory sequence. *J. Biol. Chem.* **267**, 3389-3395.
66 Maher, L. J., III, Dervan, P. B. and Wold, B. (1992). Analysis of promoter-specific repression by triple-helical DNA complexes in a eukaryotic cell-free transcription system. *Biochemistry* **31**, 70-81.
67 Maher, L. J., III, Wold, B. and Dervan, P. B. (1989). Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation. *Science* **245**, 725-730.
68 Postel, E. H., Mango, S. E. and Flint, S. J. (1989). A nuclease-hypersensitive element of the human c-myc promoter interacts with a transcription initiation factor. *Mol. Cell. Biol.* **9**, 5123-5133.
69 Gee, J. E., Blume, S., Snyder, R. C., Ray, R. and Miller, D. M. (1992). Triplex formation prevents SP1 binding to the dihydrofolate reductase promoter. *J. Biol. Chem.* **267**, 11163-11167.
70 Johnson, A. C., Jinno, Y. and Merlino, G. T. (1988). Modulation of epidermal growth factor receptor proto-oncogene transcription by a promoter site sensitive to S1 nuclease. *Mol. Cell. Biol.* **8**, 4174-4184.
71 Weintraub, H. (1983). A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. *Cell* **32**, 1191-1203.
72 Wittig, B., Wolfl, S., Dorbic, T., Vahrson, W. and Rich, A. (1992). Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. *EMBO J.* **11**, 4653-4663.

Kensal van Holde* and **Jordanka Zlatanova** are at the Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331-7305, USA. **Jordanka Zlatanova** is on sabbatical leave from the Institute of Genetics, Bulgarian Academy of Sciences, Sofia 1113, Bulgaria.
*Author for correspondence.