Volume 12 Number 21 1984 Nucleic Acids Research

Possible structures of homopurine-homopyrimidine S1-hypersensitive sites

Charles R.Cantor and Argiris Efstratiadis

Department of Human Genetics and Development, Columbia University, 701 West 168th Street, New York, NY 10032, USA

Received 8 August 1984; Revised and Accepted 4 October 1984

ABSTRACT

S1 nuclease hypersensitive sites in the 5' flanking regions of eukaryotic genes, present in small artificial supercoiled DNA circles, reside in homopurine-homopyrimidine stretches. The hierarchical behavior which these sites exhibit is consistent with the notion that they act as sinks of torsional free energy. By employing DMS as a single-strand-specific reagent, we show that these sites (despite their S1 sensitivity) are regions of duplex DNA. A simple thermodynamic treatment indicates that the high torsional stress in the small DNA circles is almost certain to be relieved by the formation of alternate DNA structures. The same treatment places some constraints on the types and sizes of the regions with alternate conformation. While no definitive structural conclusions can be drawn, left-handed helices seem most consistent with the extent and the pattern of sensitivity to S1 nuclease.

INTRODUCTION

S1 nuclease is a sensitive enzymatic probe which recognizes novel structural features in DNA regions proximal to expressing genes in active chromatin (1,2). Such S1 hypersensitive sites (SHS) are also present on naked supercoiled (but not relaxed or linear) DNA (1,3). Since the alternate structures appear only under torsional stress, it is not easy to use traditional biophysical methods to study them. However, the following generalizations can be made from the available biochemical data:

a) The SHS exhibit hierarchical behavior; the presence of one such site in a supercoiled molecule can mask the appearance of another site (lower hierarchical order), which is revealed only if the sequence corresponding to the first site is removed (3). Intermediate situations are also possible, in which a strong and a weak site co-exist in the population of molecules and can be revealed simultaneously from their relative nicking intensities (4). Because of this hierarchical behavior a strategy of in vitro fine mapping of SHS has been devised in which the nicking patterns of hypersensitivity to the nuclease are examined in artificially generated supercoiled circles derived from overlapping , relatively short, restriction fragments of DNA (3).

b) The hierarchical behavior of the SHS is consistent with the notion that they play the role of sequence-dependent sinks of torsional free energy.
c) The sites reside in (pure or predominating) homopurine-homopyrimidine DNA stretches. Though a clear-cut consensus sequence is not evident, the appearance of alternating purines or of repeats of purine combinations (on the homopurine strand) is frequent (ref.4 and other references therein).
d) In some of the sites only one of the DNA strands (usually, but not exclusively, the homopyrimidine stretch) is nicked by the nuclease (3,4).
e) The nicking patterns of the SHS are stable over a wide range of temperature (0^0 to 65^0) (3).
f) The sequences of the regions surrounding the SHS are inconsistent with these sites corresponding to cruciforms or Z-DNA/B-DNA junctions, both of which are known to be S1 nuclease-sensitive (5-7). Also the sites are not the borders of homopurine-homopyrimidine B-DNA and an alternate flanking DNA structure, because they retain their hypersensitivity when transplanted into completely foreign DNA (4).

Here we address two questions raised by these results. First, what types and lengths of alternate DNA structure are likely to be favored by the amounts of torsional free energy stored in a small DNA circle? Second, what kinds of structure might explain the observed patterns of S1 hypersensitivity?

The following possibilities exist for a DNA sequence which can assume an altered (non-B) conformation under torsional stress and can serve as an energy sink:
a) The SHS DNA is an underwound right-handed helical duplex.
b) The DNA aquires extensive sharp bending or kinking.
c) An SHS is a bubble of single-stranded DNA.
d) An SHS is left-handed non-Z DNA.

The first two possibilities seem very unlikely. All known non-B right-handed or kinked or bent duplex structures differ only moderately in their twist from B-DNA. Thus, a long region of altered structure would be required to act as a torsional energy sink. However, the observed SHS are quite short. Therefore, they could not correspond directly to the alternate structure nor could they be the boundaries of that structure and normal B-DNA because they are transplantable.

The last two possibilities seem much more plausible and the thrust of this paper is whether we can discriminate between them. The unilateral nicking of some of the SHS, especially of a $(G/C)_{18}$ structure (3), and other indirect arguments suggested (but did not prove) that the sites are regions of duplex

DNA, despite their sensitivity to an enzyme which is presumably single-strand-specific. Here we present direct experimental evidence which indicates that the SHS are indeed regions of double-stranded DNA. We then use this information and a simple thermodynamic treatment to place some constraints on the types and sizes of the alternate S1-hypersensitive conformations. While it may take time before the unusual structures are understood in molecular detail, our analysis, which narrows down considerably the range of possibilities, could serve as a starting point for more incisive thinking about their behavior and its functional implications.

RESULTS AND DISCUSSION

DMS Cannot Methylate the N-3 of Cytosines Present in SHS

DMS is a useful chemical probe for the mapping of single-stranded regions in duplex DNA at the nucleotide level, because it can be used for single-strand-specific cytosine methylation (8). DMS methylates guanine N-7 in both single- and double-stranded structures, but attacks adenine N-1 and cytosine N-3 only in single-stranded regions, because these positions are involved in hydrogen bond formation in duplexes. Treatment of an end-labeled single-strand with DMS alone or sequentially with DMS and hydrazine, prior to the piperidine reaction of the chemical sequencing procedure (9), results in the appearance of G and G+C patterns, respectively, following the display of the degradation products on a denaturing urea-polyacrylamide gel (8). One limitation of the method is that the intensity of the gel bands resulting from C-methylation is approximately one-third of the intensity of the bands from G-methylation. A more serious limitation is that the window between signal and noise is very narrow because of the significant reactivity of unmethylated cytosines with hydrazine. Keeping these limitations in mind, we reacted supercoiled circles carrying the -55 SHS of the chicken β-globin gene (3,4) with DMS, in parallel with the linear fragment from which the circles were derived, in single- and double-stranded form.

Fig. 1a shows that there is practically no difference between the patterns derived from the methylation of the supercoiled molecules (lanes E and F) and the linear duplex (lanes C and D), while the signal from the C-methylation of the single-strands (lane B) is significantly higher than the background (lane A). This suggests that there is no single-stranded region in the supercoiled circles. Even if the SHS DNA were only breathing, reaction of the cytosines would have been observed. This is because the overall methylation of the guanines in the supercoiled molecules was quite significant as evidenced by

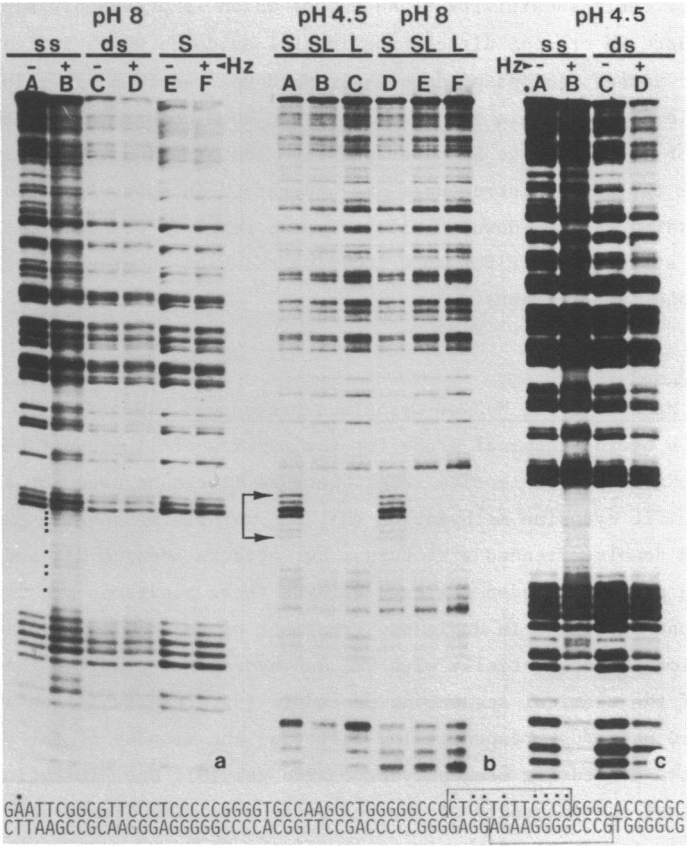

Fig. 1. a) Autoradiogram of a 10% urea-polyacrylamide gel displaying the G and and G+C patterns following methylation by DMS of linear single-stranded (ss), linear double-stranded (ds), and supercoiled (S) material. The samples were treated sequentially with DMS and piperidine (control lanes A, C and E) or with DMS, hydrazine (Hz) and piperidine (lanes B, D and F). For these experiments we used a construction containing the chicken β-globin gene -55 SHS in an inverted orientation (3,4), so that the cytosines in the SHS (dots in the staggered brackets on the DNA sequence at the bottom) could be visualized (dots in lane B) when the 5' end of the Eco RI site at position -127 is labeled (asterisk). The 479 bp fragment we used (Tha I sites at positions -127/+352), which has Eco RI linker ends (3), was first 5' end-labeled , and then either circularized (3) to produce supercoiled circles or cleaved with Pvu II, which has a cutting site at position +192. The largest of the two uniquely labeled Eco RI-Pvu II fragments (319 bp) was isolated and treated with DMS, either directly (ds material) or after alkaline denaturation (ss material). The supercoiled circles were first treated with DMS and then digested with Eco RI and Pvu II to isolate the 319 bp fragment. DMS reactions were in DMS buffer (9) for 1 min at 37^0. All other reactions were performed as described (8). b) Autoradiogram of a gel as in panel a. DMS reactions were performed either in DMS buffer (pH 8) or in S1 buffer (pH 4.5). DMS-treated supercoiled circles (of 5' end-labeled fragments at the Eco RI ends prior to circularization) were

divided into two aliquots. One of them (S material, lanes A and D) was exposed to S1 nuclease under standard conditions (3) for 1 min at 37°. The other aliquot was digested with Pvu II to produce linearized material carrying the radioactivity at the internal Eco RI site (SL material, lanes B and E). Pvu II-linearized material (L) was also produced from untreated supercoiled circles, and then methylated with DMS. The SL and L samples were digested with S1 for 30 min at 37°, and then cleaved with Eco RI. Sample S was digested with Eco RI and Pvu II. The 319 bp fragment from each sample was isolated and run on the gel. Nicking of the methylated supercoiled circles produces a pattern characteristic for the -55 SHS (bracket with arrows). Note that in this case the G-pattern in all lanes was produced by heating of the samples at 65°, even without piperidine treatment. c) Lanes A-D correspond to the same lanes as in panel a, except that the reactions were performed in S1 buffer.

the difficulty of subsequent restriction enzyme digestion of the circles (see Legend to Fig.1 for details). Digestion for an extended period of time with a thousand-fold excess of enzyme resulted in cleavage of approximately 50% of the molecules. However, because this result is intrinsically negative, we supported our conclusion about the duplex character of the SHS by an independent experiment.

We reasoned that if only a small minority of the cytosines are attacked by DMS the bands resulting from their methylation might not be evident on the gel over the background. However, these methylated C's should remain unpaired upon linearization of the circles, and should then be recognized by S1 as single-stranded regions of a linear duplex. Fig.1b shows that S1 nuclease acting on a supercoiled circle, methylated either in the usual DMS reaction buffer (pH 8.0) or in S1 buffer (pH 4.5), produced a nicking pattern characteristic for the site (lanes A and D). In contrast, no nicking by the enzyme was detectable in the linear duplex derived from the methylated supercoiled molecules (lanes B and E), which exhibited a pattern identical to that obtained from a methylated linear duplex control (lanes C and F).

An argument against this (also negative) result would be that S1 nuclease cannot easily recognize single unpaired C's. It is known, however, that S1 recognizes single base mismatches. For example, the enzyme cleaves a heteroduplex between wild-type SV40 and mutant tsA30 DNAs (10), though the cleavage is more limited in extent than that observed with deletion/insertion heteroduplexes. Sequence analysis revealed that the tsA30 mutant has a point mutation at the site of S1 cleavage (Tom Shenk, personal communication). Moreover, a minimum of 5% cleavage was observed using synthetic single base-mismatch heteroduplex substrates (11). We believe that even such a low degree of digestion would have been evident among the labeled products in our experiment if any of the several C's in the SHS we examined were available for methylation. This is consistent with the following result (not shown): The same supercoiled circles

as in Fig.1 were linearized with Pvu II (a digestion that leaves the radioactivity at an internal site). The fragments were then dissociated and a portion of them was treated with DMS. Treated and untreated single-strands were then reannealed, and the resulting duplexes were exposed to S1 nuclease. Electrophoresis of the reaction products resulted in the appearance of a smear from the DMS-treated material, while the untreated sample was S1-insensitive.

We note that DMS methylates unpaired C's with the same efficiency at pH 8.0 and 4.5 (Fig.1c). We also note that a strategy identical to that of the experiment in Fig.1b has been followed previously, using bromoacetaldehyde as a presumably single-strand-specific reagent (12). However, in our hands, the use of bromoacetaldehyde results in non-specific nicking of the small supercoiled DNA circles (data not shown).

Conformational Changes in Large and Small Supercoiled Domains: A Simple Thermodynamic Model

A detailed statistical mechanical theory has been developed (13-16) to describe conformational changes driven by torsional stress (5-7,17-22) in typical supercoiled plasmids. Thermodynamic treatments of such changes have also been described which apply to particular cases (22,23). Here we present a simple, but general, all or none thermodynamic description of a supercoiled molecule with a single region of potential alternate structure. The results are compact and straightforward. An equivalent treatment appeared in print (24) subsequent to the completion of this work.

The torsional free energy stored in a closed circular DNA molecule underwound (or overwound) by i+d additional turns has been shown experimentally (25,26) to be:

$$G_{init} = K(i+d)^2 \qquad [1]$$

where i is an integer (the number of supercoils) and d represents the partial additional turn needed to close a linear duplex containing a non-integral number of total turns. K is the supercoiling force constant, and has units of Kcal/mol.

Considerable torsional energy can be stored in a supercoiled molecule. In certain cases it will be energetically favorable for the molecule to convert a small region into an alternate structure which is less twisted and thereby relieve some of the torsional strain. Producing an alternate structure with n bp will require an increase in free energy $\Delta G_{nuc}+n\Delta G_{gro}$, where ΔG_{nuc} is the energy required for nucleation of the structure and ΔG_{gro} is the energy per bp required for its extension. However, this increase will be offset by a larger decrease in free energy arising from the decrease in torsional stress. The

normal B-DNA helix is 10.5 bp per turn. If the alternate structure has x bp per turn, each bp converted from B-helix to the alternate will result in an unwinding of f = 1/10.5 - 1/x turns. We use the convention that right-handed turns are positive and left-handed turns are negative. Production of an alternate structure with n bp will involve an untwisting of the DNA by nf turns. Therefore, the resulting torsional free energy will be reduced to $K(i+d-nf)^2$.

The total free energy of the final state containing the alternate structure is

$$G_{fin} = \Delta G_{nuc} + n\Delta G_{gro} + K(i+d-nf)^2 \qquad [2]$$

Thus, the total free energy change, ΔG_c, for conversion of the initial supercoiled molecule to one with the alternate structure is

$$\Delta G_c = G_{fin} - G_{init} = \Delta G_{nuc} + n\Delta G_{gro} + K(i+d-nf)^2 - K(i+d)^2$$
$$= n^2Kf^2 + n[\Delta G_{gro} - 2K(i+d)f] + \Delta G_{nuc} \qquad [3]$$

Equation [3] is a simple quadratic in the length, n, of the alternate structure. It is important to note some of the simplifications in this equation. It assumes, as do all previous similar calculations, that the torsional stiffness of alternate DNA structures is the same as that of a B-helix. It also assumes that only one region of altered structure is present per DNA molecule. Furthermore, it ignores the possibility of any extra untwisting at the boundaries between normal B-DNA and the altered structure. Such untwisting may be quite significant for the junctions between B- and Z-DNA (16). The effect of untwisting at the boundaries will alter the values of i+d and ΔG_{nuc} in [3] but it will not change the form of the equation. Finally, an all or none theory, as formulated in equation [3], artificially sharpens the appearance of conformational changes and does not represent accurately the energetics of alterenate structures smaller than the limits allowed by a particular sequence. We believe that these simplifications are not likely to have any significant influence on our analysis.

Since rough estimates for the various thermodynamic and structural parameters in equation [3] are available, one can calculate the roots of the quadratic, by setting $\Delta G_c = 0$. The lengths of potentially stable alternate structures are given by the values of n between the two roots. Thus, one can evaluate the likelihood of DNA assuming various alternate structures, provided that the sequence has the capacity to assume such structures. For a given type of alternate structure one can also find the length of the potentially most stable structure; it will be the value of n halfway between the two roots.

Experiments show that for large DNAs, with N bp, the supercoiling force

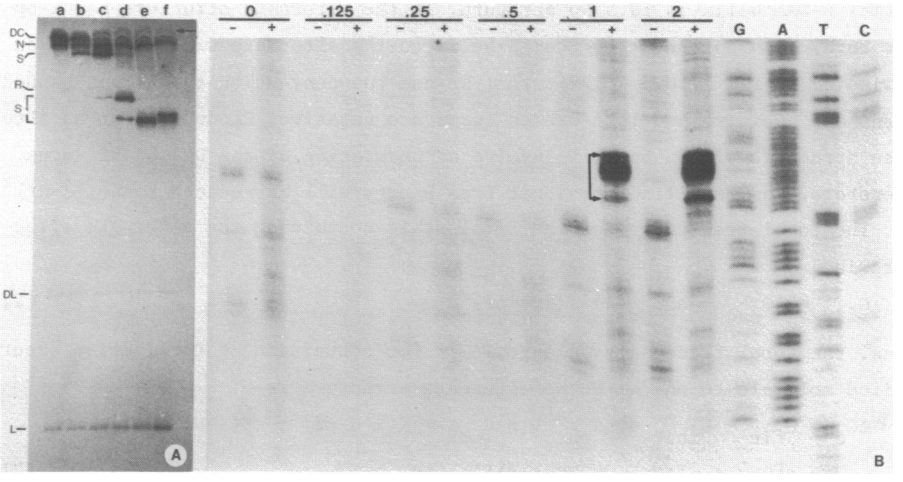

Fig. 2. A) Autoradiogram of a thin non-denaturing polyacrylamide gel (3,28) used to resolve the molecular species generated by ligation of a 5' end-labeled 430 bp Sau 3A fragment (1 µg/ml DNA) carrying the chicken β-globin gene promoter (3). A series of circularization reactions in the presence of 0, 0.125, 0.25, 0.5, 1 and 2 µg/ml ethidium bromide (lanes a-f, respectively) is shown. The species are: DC, double-length circles; N, nicked circles; S, topoisomers of supercoiled circles; R, covalently closed relaxed circles; DL, double-length linears; and L, unit length linears. An arrow indicates the origin. B) S1 nicking patterns of the supercoiled circles isolated from the reactions shown in panel A. The S1 nicking products (+) and their controls (-) were displayed on a denaturing gel, in parallel with their corresponding sequencing ladders. Though in this photographic reproduction the nicking pattern (bracket with arrows) is evident only at the two highest ethidium bromide concentrations, it is also clearly present in the 0.5 µg/ml sample on the original autoradiogram. The non-specific nicking in all lanes is due to an activity contaminating restriction endonuclease Sau 3A.

constant K is proportional to 1/N (18). This reflects the fact that the larger the DNA the more the torsional stress can be distributed and the smaller the angular deformation at any base pair. However, when the length of the DNA is below about 2 kb the value of K increases more rapidly than 1/N (27,28). Thus, for example, at ambient temperature, K for a typical 4.3 kb plasmid DNA circle is only about 0.19 Kcal/mol, while for a 430 bp circle, such as the one previously described (3), K rises to about 5.1 Kcal/mol. This has profound influences on both the supercoiling behavior of the DNA and its propensity to assume alternate structures. As a specific example we will evaluate equation [3] for a large (4.3 kb) and a small (430 bp) DNA circle and for two different conformational changes; melting of a single n bp stretch of B-DNA helix to untwisted single-strands or conversion of that stretch to a Z-DNA helix.

Table 1. Topological and thermodynamic parameters for conformational changes in supercoiled circles

Parameter	Type of circle: Large (4.3 kb)	Small (430 bp)
K (Kcal/mol)*	0.19	5.1
i	20	3
d	0	0.05
	Alternate structure: Z-DNA	Single-strands
f	0.18	0.095
ΔG_{nuc} (Kcal/mol)	10	7
ΔG_{gro} (Kcal/mol)	0.33	0.7

* For circles of N bp, K can be estimated as follows: If N>2,000, K=800/N. This is the average of the values reported in ref. 27 and 28. If N<2,000, K=(2590-0.89N)/N. This is a simple empirical fit to the data in ref. 27. If values from ref. 28 were used instead, the value of K for a 430 bp circle would be 5.8 Kcal/mol.

The values for all of the parameters used in the calculations (Table 1) were selected as follows: Choosing i=20 for the 4.3 kb plasmid provides a superhelix density of -0.05, a value typically observed in natural plasmids. The choice of d is immaterial since d must be less than one, and therefore is insignificant compared with i. The value i=3 for the 430 bp circle is based on our observation that S1 hypersensitivity appears with the third topoisomer (Fig.2). For a 430 bp circle, d will be the difference between 430/10.5 and the nearest integer. Thus, the value of d, 0.05, is negligible in this case also. The value of f for melting simply reflects the unwinding of a single bp of a 10.5 bp/turn B-helix. The larger value of f for the B to Z transition reflects the untwisting of one bp of B-helix and continued untwisting to form one bp of a -12 bp/turn left-handed Z-helix.

Although some of the thermodynamic parameters in equation [3] are not known accurately, the general results are not particularly sensitive to their values. Under typical biological conditions, an AT-rich region will always be the preferred site of melting of a B-helix to single-strands. A typical AT-rich region will have an enthalpy of melting per bp, ΔH_{gro}, of 5 Kcal/mol. This can be used to estimate the free energy of melting for an all-or-none model as $\Delta G_{gro} = \Delta H_{gro}(1 - T/T_m)$, where T is the absolute temperature and T_m is the melting temperature. If the sample is studied 50^0 below T_m, ΔG_{gro} is 0.7 Kcal/mol. This value could easily be halved for a pure AT stretch, if one existed. The value for ΔG_{nuc} is similar to ones used in previous calculations (13).

For conversion from B-helix to Z, we use some of the results obtained

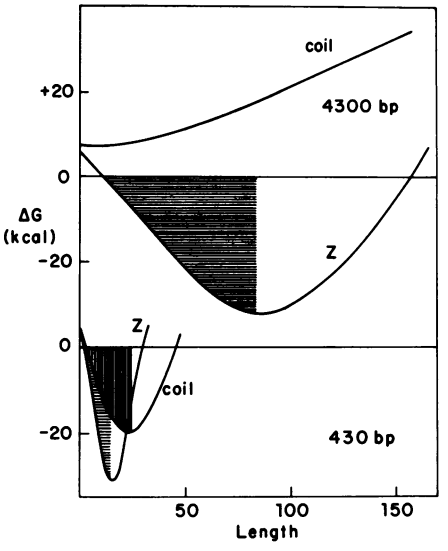

Fig. 3. Calculated free energy changes (equation[3]) for forming altered structures in supercoiled DNA circles. The parameters used for the calculations are shown in Table 1. For each curve the shaded area indicates the lengths of alternate structures that can form if the actual DNA sequence allows for melting of an AT-rich region or conversion of a $(GC)_n$ region into Z-DNA.

before (16) from DNA circles containing alternating dC-dG stretches. The thermodynamic parameters of Z-DNA formation from dC-dA stretches are not the same (17), but the magnitude of the difference is not enough to affect the character of our results.

Using the values in Table 1 we evaluated equation [3] for the minimum and the maximum length (in bp) for which the formation of the alternate structure is thermodynamically favored. We also evaluated the length of the most stable possible alternate structure (that having the lowest possible free energy), by setting the derivative of equation [3] equal to zero and solving for

$$n_e = [2K(i+d)f - \Delta G_{gro}]/2f^2K \qquad [4]$$

This length is independent of the thermodynamics of nucleation of the alternate structure. Insertion of the value of n_e into [3] yields the free energy, ΔG_c, of formation of the most stable possible alternate structure.

The results of the calculations are shown diagrammatically in Fig.3 and summarized in Table 2. For large DNA circles they are consistent with previous calculations (15-17,19) and with experiments. Supercoils can drive the B to Z transition readily, but are much less likely to melt a duplex. For the parti-

Table 2. Calculated characteristics of alternate conformations

	bp in alternate structure			ΔG_c^*
	Minimum	Maximum	Most stable	Kcal/mol
4.3 kb				
Z-DNA	11	157	84	-32
single-strands	no stable alternate structure			+7
430 bp				
Z-DNA	2	30	16	-31
single-strands	3	45	24	-20

* for the most stable structure

cular example considered, a region between 11 and 84 bp in length, which has Z potential, will be converted completely into the Z-helix. However, the length of the Z region cannot exceed 84 bp (even if the potential exists in a larger stretch) because ΔG_c starts to become less negative beyond this limit. If higher superhelical densities are allowed the length of the most stable Z region increases, and Z-DNA formation remains vastly more favorable energetically than melting.

For small DNA circles the calculation shows a new and unexpected conclusion. The energy stored in even a few supercoils is enough, not only to switch small stretches of B-helix to Z-DNA, but also to melt short AT-rich stretches. The energies of both types of altered structure are comparable, but the most stable melted stretch is longer than the most stable Z form. Since all small DNA circles will have sequences capable of one or the other conformational change, the calculations imply that such structural transitions will always occur unless there is some other energy sink in the molecule.

Possible Structures of SHS

We will use the thermodynamic model we presented to argue that the SHS are most likely left-handed helices. We will then attempt to explain the observed patterns of nicking.

In view of our experimental results, melting is very unlikely to be responsible for the observed S1-hypersensitivity. This conclusion is consistent with our analysis (Table 2), because the lengths of the SHS (9-17 bases) are considerably shorter than the lengths expected from the most stable melted regions. The same argument makes it unlikely that the SHS are achiral double-stranded helices (29). The patterns of S1 nicking are either unilateral or have a staggered appearance in the population of molecules (whenever each of the two strands can be independently nicked). Both patterns are very difficult to reconcile if an SHS is simply a melted bubble

The only other known underwound alternate structures are left-handed

helices. As soon as a few supercoils are introduced into a small DNA circle, ΔG_{gro} becomes almost insignificant in equation [3] compared with (i+d)Kf. Thus, the structure with the largest value of f is sure to be favored, once the cost of nucleation can be overcome. This means that left-handed helices, even those with structures too unstable ever to be seen in isolation, could easily be produced in a small supercoiled circle.

Further insight can be gained by eveluating equation [4] for plausible values of f and ΔG_{gro} for the case of a 430 bp circle with 2-4 supercoils. It turns out that the two terms of this equation partially compensate. Thus, it is possible to assign the likely length of the most stable alternate structure as 8-18 bp, if the structure is left-handed Z-like DNA, or as 16-29 bp, if the structure is not twisted at all, or as an even larger length if the structure is some kind of partially unwound right-handed helix. The lengths found experimentally for the homopurine-homopyrimidine SHS are 9-17 bp (when the outer limits of hypersensitivity are considered for both strands together, in the cases of bilateral nicking). This strengthens the argument that the alternate structure is left-handed.

The structure of the putative left-handed helix assumed by these DNA stretches under torsional stress is mostly a matter for speculation. The structure must presumably form smooth junctions with B-DNA, because, according to our observations, the borders of the SHS are not S1-hypersensitive, in contrast to the Z-DNA/B-DNA junctions. A Z-type conformation is the only extensively characterized left-handed structure, and the known Z-DNA helices have alternating purine-pyrimidine sequences. Other types of left-handed helices are consistent with the diffraction pattern of stretched calf thymus DNA (30). Conformational calculations indicate that left-handed analogs of B- and A-helices should also be stereochemically possible (31,32). When large complementary single-stranded DNA circles are hybridized, the amount of left-handed helix produced in the resulting structures (form V DNA), which are under extreme torsional stress, seems too high to be accounted for by alternating purine-pyrimidine Z-DNA alone (33). However, it is not clear whether the sequence requirements of Z-DNA are relaxed in such structures or if other types of left-handed helices are formed.

Since the structure of the putative left-handed helix is unknown, and the mechanism of action of S1 nuclease is also unknown, it is very difficult to suggest a reason for the sensitivity of the structure to nicking.

The staggered pattern of S1 nicking seen in two SHS (3,4) is readily explained if the enzyme binds to a groove of left-handed DNA and cuts at

symmetrically situated sites above and below the plane of a particular base pair. However, the problem of S1 strand-specificity for other SHS still remains. The putative left-handed helix might have a structure in which one strand has a different conformation than the other. Evidence for such a structure for a right-handed helix has been described for the homopurine-homopyrimidine sequence $(dA\text{-}dT)_n$ (34), and has been inferred for other sequences from the patterns of nuclease and chemical cleavage (35). Alternatively, the left-handed structure might be highly bent or coiled. Steric hindrance caused by DNA bending could explain the strand-preference and the nicking profiles, because the bulkiness of DNA adjacent to the bends could shield nearby residues from the nuclease.

One final issue left unresolved by all of the available data is whether the bases in the SHS are protonated. The SHS still appear even when nicking conditions close to neutrality are used (3). Moreover, the results shown in Fig.1 indicate that the cytidines in an SHS are not accessible to methylation either at pH 4.5 or pH 8.0. However, homopurine-homopyrimidine DNAs with regularly alternating sequences are known to form triple-helices containing one strand with protonated cytidines. When the cytidine is methylated these helices are stable even at neutral pH (36). Such structures may be yet another clue to the nature of the underwound helix of the SHS.

ACKNOWLEDGEMENTS

We thank James Wang and Robert Baldwin for helpful discussions and for making available manuscripts prior to publication; James Wang for suggesting the use of DMS; Wilma Olson and Tom Shenk for discussions; Struther Arnott for critical comments on the manuscript; and Eric Schon and Todd Evans for help with the experiments and the figures. Work in our laboratories was supported by grants from the National Institutes of Health.

REFERENCES

1. Larsen, A., and Weintraub, H. (1982) Cell 29, 609-622.
2. Weintraub, H. (1983) Cell 32, 1191-1203.
3. Schon, E., Evans, T., Welsh, J., and Efstratiadis, A. (1983) Cell 35, 837-848.
4. Evans, T., Schon, E., Gora-Maslak, G., Patterson, J., and Efstratiadis, A. (1984) Accompanying paper.
5. Lilley, D.M. (1982) Cold Spr. Harb. Symp. Quant. Biol. 47, 101-112.
6. Panayotatos, N., and Wells, R.D. (1981) Nature 289, 466-470.
7. Singleton, C.K., Klysik, J., Stirdivant, S.M., and Wells, R.D. (1982) Nature 299, 312-316.
8. Kirkegaard, K., Buc, H., Spassky, A., and Wang, J. (1983) Proc. Natl. Acad. Sci. USA 80, 2544-2548.
9. Maxam, A.M., and Gilbert, W. (1980) Meth. Enzymol. 65, 499-560.

10. Shenk, T.E., Rhodes, C., Rigby, P.J.W., and Berg, P. (1975) Proc. Natl. Acad. Sci. USA 72, 989-993.
11. Dodgson, J.B., and Wells, R.D. (1977) Biochem. 11, 2374-2379.
12. Kohwi-Shigematsu, T., Gelinas, R., and Weintraub, H. (1983) Proc. Natl. Acad. Sci. USA 80, 4389-4393.
13. Benham, C.J. (1981) J. Mol. Biol. 150, 43-68.
14. Benham, C.J. (1982) Biopolymers 21, 679-696.
15. Benham, C.J. (1982) Cold. Spr. Harb. Symp. Quant. Biol. 47, 219-227.
16. Peck, L.J., and Wang, J. (1983) Proc. Natl. Acad. Sci. USA 80, 6206-6210.
17. Haniford, D.B., and Pulleyblank, D.E. (1983) Nature 302, 632-634.
18. Mizuuchi, K., Mizuuchi, M., and Gellert, M. (1982) J. Mol. Biol. 156, 229-243.
19. Nordheim, A., Peck, L.J., Lafer, E.M., Stollar, B.D., Wang, J., and Rich, A. (1982) Cold Spr. Harb. Symp. Quant. Biol. 47, 93-100.
20. Nordheim, A., and Rich, A. (1983) Proc. Natl. Acad. Sci. USA 80, 1821-1825.
21. Stirdivant, S.M., Klysik, J., and Wells, R.D. (1982) J. Biol. Chem. 257, 10159-10165.
22. Wang, J., Peck, L.J., and Becherer, K. (1982) Cold. Spr. Harb. Symp. Quant. Biol. 47, 85-91.
23. Courey, A.J., and Wang, J. (1983) Cell 33, 817-829.
24. Frank-Kamenetskii, M.D., and Vologodskii, A.V. (1984) Nature 307, 481-482.
25. Depew, R.E., and Wang, J. (1975) Proc. Natl. Acad. Sci. USA 72, 4275-4279.
26. Pulleyblank, D.E., Shure, M., Tang, D., Vinograd, J., and Vosberg, H.-P. (1975) Proc. Natl. Acad. Sci. USA 72, 4280-4284.
27. Horowitz, D.S., and Wang, J. (1984) J. Mol. Biol. 173, 75-91.
28. Shore, D., and Baldwin, R.L. (1983) J. Mol. Biol. 170, 983-1007.
29. Arnott, S., Bond, P.J., and Chanrasekaran, R. (1980) Nature 287, 561-564.
30. Arnott, S., Chandrasekaran, R., Hall, I.H., Puigjaner, L.C., Walker, J.K., and Wang, M. (1982) Cold Spr. Harb. Symp. Quant. Biol. 47, 53-65.
31. Gupta, G., Bansal, M., and Sasisekharan, V. (1980) Proc. Natl. Acad. Sci. USA 77, 6486-6490.
32. Sasisekharan, V. (1982) Cold Spr. Harb. Symp. Quant. Biol. 47, 45-52.
33. Brahms, S., Vergne, J., Brahms, J.G., DiCapua, E., Bucher, P., and Keller, T. (1982) Cold Spr. Harb. Symp. Quant. Biol. 47, 119-124.
34. Arnott, S., Chandrasekaran, R., Hall, I.H., and Puigjaner, L.C. (1983) Nuc. Acids Res. 11, 4141-4155.
35. Drew, H.R., and Travers, A.A. (1984) Cell 37, 491-502.
36. Lee, J.S., Wordsworth, M.L., Latimer, L.J., and Morgan, A.R. (1984) Nuc. Acids Res. 12, 6603-6614.