Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 5884–5888, August 1986
Biochemistry

Consecutive A·T pairs can adopt a left-handed DNA structure

(supercoiling/S1 nuclease/supercoil relaxation studies/B–Z junctions/energetics)

MICHAEL J. MCLEAN, JOHN A. BLAHO, MICHAEL W. KILPATRICK, AND ROBERT D. WELLS

Department of Biochemistry, The University of Alabama at Birmingham, Schools of Medicine and Dentistry, Birmingham, AL 35294

Communicated by Richard E. Dickerson, April 30, 1986

ABSTRACT **The capacity of six sequences with different numbers and orientations of A·T pairs flanked by alternating C·G pairs to adopt left-handed structures was evaluated in recombinant plasmids. A series of synthetic oligodeoxynucleotides were cloned into the *Bam*HI site of pRW790, a small plasmid (≈2 kilobases) prepared especially for conformational studies of this type. Supercoil relaxation studies by two-dimensional gel electrophoresis on topoisomers of each plasmid revealed the energetics and structures of the left-handed helices. Also, the presence of supercoil-induced altered DNA conformations within the inserts of topoisomer populations of the plasmids was detected by reaction with S1 nuclease followed by restriction mapping of the cleavage sites. We conclude that consecutive T·A base pairs, whether alternating (TATA) or contiguous (TTTT), can adopt a left-handed conformation (presumably Z) when flanked by reasonably short runs of alternating (C-G)$_n$ (n = 3–5). Thus, these results substantially broaden the range of DNA sequences that can adopt left-handed Z conformations.**

Left-handed Z-DNA occurs in suitable sequences in recombinant plasmids, restriction fragments, and DNA polymers and oligomers (reviewed in refs. 1–5). Supercoiling at densities known to exist under physiological conditions induces the B-to-Z transition (6, 7).

The types of sequences proven to adopt left-handed conformations under appropriate conditions are generally of the strictly alternating purine-pyrimidine type and include C-G (reviewed in refs. 1–5), 5-methyl-C-G (8–10), (T-G)·(C-A) (11–13), and mixtures of C-G and T-G (14). Also, we demonstrated (15) that a *Bam*HI site (GGATCC) within a tract of 58 base pairs (bp) of C-G adopted a non-B DNA structure, which may be fully left-handed, as stabilized by negative supercoiling. Further, the intervening sequence 2 (IVS2) sequences (14) contain GTTTG and GACTG sequences, which were proven to adopt left-handed (presumably Z) structures by supercoil-induced topoisomer relaxation studies [on two-dimensional (2D) gels], S1 nuclease mapping of B–Z junctions, and Z-DNA antibody binding studies. Thus, an alternating purine-pyrimidine sequence is not *necessary* for a Z helix.

A number of laboratories have unsuccessfully attempted to convert the strictly alternating A and T duplex DNA polymer from a right-handed to a left-handed Z form with a variety of environmental conditions; similarly, this sequence seems unable to form Z-DNA in recombinant plasmids (16–18). Hence, a strictly alternating purine-pyrimidine sequence is not *sufficient* for Z helix formation. However, several left-handed structures have been proposed (19, 20) for (I-C)$_n$·(I-C)$_n$, and the strictly alternating T and deoxydiaminopurine duplex appears to adopt a Z helix (21).

Thus, if an alternating purine-pyrimidine sequence is neither necessary nor sufficient for Z-helix formation, what are the requirements? We have cloned a series of synthetic oligomers containing consecutive A·T pairs of different lengths and orientations flanked by relatively short tracts (6, 8, or 10 bp) of C-G to further evaluate the sequence requirements, the properties of the alternate conformations, and the energetics and mechanisms of the supercoil-induced structural changes.

MATERIALS AND METHODS

Synthetic Oligonucleotides. Oligomers were prepared on an Applied Biosystems 380A DNA synthesizer and purified (22). GATC (C-G)$_4$TA(C-G)$_4$ was synthesized and generously provided by Gene Brown (Genetics Institute).

Plasmid Preparation. pRW790 is pBRN3 (23) with a 50-bp *Eco*RI–*Hin*dIII fragment containing multiple cloning sites inserted between its *Eco*RI–*Hin*dIII sites. The multiple cloning site fragment was isolated from *Eco*RI–*Hin*dIII-digested M13 mp10 replicative form and purified from low-melting temperature agarose. Duplex form synthetic oligonucleotides were treated with kinase and ligated into the *Bam*HI site of pRW790, which had been treated with calf intestinal phosphatase (Boehringer Mannheim). Clones resulting from the transformation of HB101 were screened for the presence of the inserts by standard procedures (24). pRW1001-1101 were characterized by Maxam–Gilbert sequencing (25) of the *Eco*RI–*Hin*dIII fragments and the observed sequences of the inserts are shown in Fig. 1. In the case of pRW1101, the synthetic insert was originally cloned into the *Bam*HI site of pBR322, sequenced, and then transferred into the *Bam*HI site of pRW790.

Other Methods. Topoisomeric samples of the plasmids (26), 2D gel electrophoresis (4, 13, 14, 27, 28), and S1 nuclease reactions (6, 13–15) were as described.

RESULTS

Sequences of Plasmids. The sequences of the inserts in the *Bam*HI site of pRW790, which is an ≈2000-bp vector containing a 50-bp polylinker, are shown in Fig. 1. The general rationale, except for the pRW1001 and pRW1002 controls, was to maintain an overall insert length of approximately two turns of helix consisting of C-G tracts with different numbers and orientations of A·T pairs in the center. Hence, the base pairs at the vector–insert junctions were the same in all cases.

The pRW1001 control plasmid contains 18 bp of C-G and pRW1002 contains 8 bp. Since the smallest block of alternating C-G that had been shown previously to adopt a left-handed conformation under the influence of supercoiling was (C-G)$_5$ (15) and there was no evidence to suggest that the relatively long alternating T-A sequence in pRW1009 [(TA)$_3$] would form a Z helix, it was of interest to determine what conformation an isolated (C-G)$_4$ sequence would adopt as a function of supercoiling. pRW1101 is interesting since the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: bp, base pair(s); 2D, two-dimensional.

pRW 790 A B C D E F G H

pRW 1001 A B C D $(\frac{CG}{GC})_9$ D E F G H

pRW 1002 A B C D $(\frac{CG}{GC})_4$ D E F G H

pRW 1003 A B C D $(\frac{CG}{GC})_5$ $\frac{TT}{AA}$ $(\frac{CG}{GC})_5$ D E F G H

pRW 1004 A B C D $(\frac{CG}{GC})_5$ $\frac{TTTT}{AAAA}$ $(\frac{CG}{GC})_5$ D E F G H

pRW 1007 A B C D $(\frac{CG}{GC})_5$ $\frac{TA}{AT}$ $(\frac{CG}{GC})_5$ D E F G H

pRW 1008 A B C D $(\frac{CG}{GC})_5$ $\frac{TATA}{ATAT}$ $(\frac{CG}{GC})_5$ D E F G H

pRW 1009 A B C D $(\frac{CG}{GC})_5$ $\frac{TATATA}{ATATAT}$ $(\frac{CG}{GC})_4$ D E F G H

pRW 1101 A B C D $(\frac{CG}{GC})_4$ $\frac{TAC}{ATG}$ $(\frac{CG}{GC})_3$ D E F G H

FIG. 1. List of plasmids showing the multiple cloning site and positions and sequences of inserts. A, *Eco*RI; B, *Sst* I; C, *Sma* I; D, *Bam*HI; E, *Xba* I; F, *Hin*cII; G, *Pst* I; H, *Hin*dIII.

C-G tracts are only 6 and 8 bp in length and the purine-pyrimidine alternating sequence is disrupted with a C-C dimer; pRW1003 and pRW1004 also contain this type of disruption but with TTC and TTTTC sequences, respectively.

In four cases (pRW1001, -1003, -1007, and -1101), the chemically synthesized oligomers were longer than the inserts found in the recombinant plasmids; the observed deletions were $(C\text{-}G)_3$, C-G, C-G, and a single G, respectively. This deletion behavior, in the inserts that can adopt Z structures (at least *in vitro*), has been documented (29–31).

Supercoil-Induced B-to-Z Transitions. The B-to-Z transition in a segment of a plasmid alters its number of supercoil turns since the supercoil number (writhe) is sensitive to conformational changes in the primary helix (twist) because no covalent bonds are broken (no change in linking number) in this process (refs. 6, 10–15, 29, 32, and reviewed in refs. 1, 3–5). Assuming that the vector DNA remains in a right-handed B helix (10.4 bp per turn) and the insert adopts a left-handed Z helix (12.0 bp per turn), a transition of n bp from B to Z would remove $(n/10.4 + n/12.0)$ supercoils from the plasmid.

The nine plasmids shown in Fig. 1 were analyzed by a 2D gel electrophoretic procedure (4). For pRW790, Fig. 2 shows that the topoisomers follow a smooth curve as the linking number progressively decreases until the topoisomers with 10 and 11 negative supercoils (topoisomers −10 and −11) that migrate slower than the preceding topoisomers. The extent of relaxation is in agreement with that calculated for the appearance of the major cruciform of pBR322, from which this plasmid was derived. By contrast, the gel pattern for pRW1009 (Fig. 2*B*) shows a smooth progression until topoisomer −8, and then a sharp break in mobility is observed such that topoisomer −9 migrates in between topoisomers −5 and −6 in the first dimension. Topoisomer −10 runs slightly faster than topoisomer −9 but is still running in between topoisomers −5 and −6. Also, this gel reveals that the relaxation associated with cruciform formation (17, 18, 27, 33–35) now appears at topoisomer −13.

Fig. 2*C* shows a plot of relative mobility in the first dimension for each topoisomer. The total amount of relaxation observed for pRW1009 was 4.3 supercoils, which is in excellent agreement with that calculated (4.28) for the transition of 24 bp from a right-handed B to a left-handed Z helix. Similar analyses were performed for all of the plasmids and, in each case, the relaxation observed was in agreement with that expected for a transition from a right-handed to a left-handed helix.

However, it is possible that the left-handed helices formed are not perfect Z forms with 12.0 bp per turn but instead exist in some alternative left-handed form. Assuming that all of the n bp of each insert adopt this left-handed structure, then the number of base pairs per turn of helix can be calculated from the observed reduction in twist by the relationship:

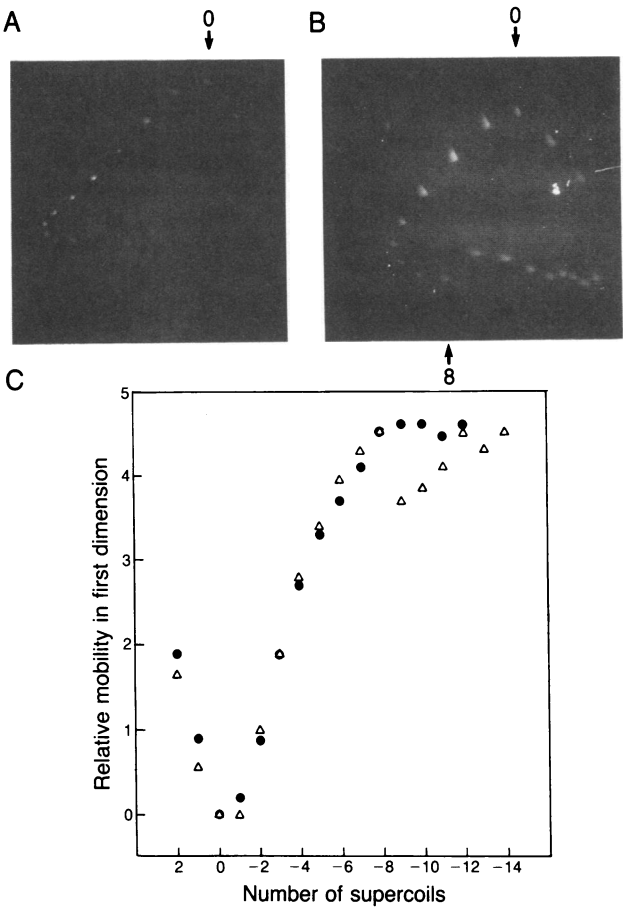

FIG. 2. 2D gel electrophoretic analyses of topoisomers of pRW790 (*A*) and pRW1009 (*B*). 0, Relaxed topoisomer. Topoisomers to the right of the relaxed topoisomer contain positive supercoils and those to the left contain negative supercoils in the first dimension. 8, Topoisomer with 8 negative supercoil turns (*B*). (*C*) Relative mobilities in the first dimension of topoisomers of pRW790 and pRW1009. Mobilities were determined relative to the most relaxed topoisomer. Topoisomer mobilities from different experiments were normalized by using the distance between the most relaxed topoisomer and the topoisomer with 4 negative supercoils (topoisomer − 4) as a normalization factor. ●, pRW790; △, pRW1009.

$$\Delta\tau = n\left[\frac{1}{10.4} + \frac{1}{Nx}\right]$$

or

$$\frac{1}{Nx} = \frac{\Delta\tau}{n} - \frac{1}{10.4},$$

in which Nx is the number of base pairs per turn of left-handed helix and $\Delta\tau$ is the reduction in twist observed on a 2D gel. This analysis was performed for all of the plasmids (Table 1). With the exceptions of pRW1002 and pRW1101, the value of Nx is very close to that expected for a Z helix in each case.

Another possibility is that only the C-G blocks adopt a left-handed conformation. If so, the extra decrease in twist observed in each case would be due to unwinding of the primary helix at the B–Z junctions. The amount of such unwinding can be calculated by:

$$\Delta\tau = n_{BZ}\left[\frac{1}{10.4} + \frac{1}{12.0}\right] = n_J\Delta_J$$

or

$$\Delta\tau = 0.179n_{BZ} + n_J\Delta_J,$$

where n_J is the number of B–Z junctions and Δ_J is the

Table 1. Calculation of the number of base pairs per turn of left-handed helix formed by the full length of each insert, given as Nx

Plasmid	$\Delta\tau$	n	$\Delta\tau/n$	Nx
pRW1001	3.3	18	0.183	11.5
pRW1002	1.6	8	0.200	9.6
pRW1003	4.0	22	0.182	11.6
pRW1004	4.3	24	0.179	12.1
pRW1007	4.0	22	0.182	11.6
pRW1008	4.3	24	0.179	12.1
pRW1009	4.3	24	0.179	12.1
pRW1101	3.2	17	0.188	10.9

$\Delta\tau$ is the number of superhelical turns relaxed for each DNA, and n is the number of base pairs in each insert that undergo the transition. Nx is calculated from $1/Nx = [(\Delta\tau/n) - (1/10.4)]$.

unwinding per junction, which would be expected to be a constant. However, when the data obtained from a study of the plasmids listed in Fig. 1 were subjected to such a treatment, Δ_J ranged from 0.035 to 0.267. Thus, the relaxation observed is due to a transition of the entire insert from B to Z in each case.

Coelectrophoresis of topoisomers of pRW1004 and pRW1008 highlighted the differences in the behaviors of the inserts with TTTT and TATA sequences (Fig. 3). pRW1008 topoisomers showed a sharp break in mobility between −8 and −9 corresponding to the loss of ≈4 supercoils, whereas topoisomer −9 of pRW1004 showed a reduction in mobility equivalent to the loss of about 1.5 supercoils. By topoisomer −11, the B-to-Z transition was complete for both plasmids and they comigrated.

For pRW1002, a relaxation of 1.6 supercoils was observed at topoisomer −9, which is in agreement with that calculated for the conversion of $(C\text{-}G)_4$ from a right-handed to a left-handed helix. However, at topoisomer −10 an additional relaxation was observed, and by −11, the total number of supercoils lost was 4.25. It is conceivable that cruciforms and segments of Z-DNA can coexist in the same molecule (6, 10, 13, 15, 27–32). The total relaxation observed for this plasmid was greater than the calculated total of the relaxations due to formation of a Z-helix and formation of a cruciform. The reason for this difference is uncertain but may be due to the shortness of the $(C\text{-}G)_4$ Z helix; if one more base pair in the *Bam*HI sites on both ends of the insert adopts a Z structure for pRW1002, the total expected relaxation is 4.16, in good agreement with the observed value.

Energetics of B-to-Z Transition. The above data show that the entire length of the inserts in all plasmids adopts a Z helix. However, significant differences in the amount of supercoiling needed to complete the transitions for different plasmids were observed.

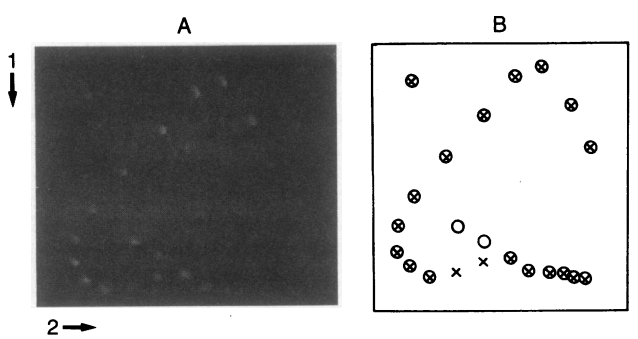

FIG. 3. 2D gel electrophoretic analyses of mixed topoisomer populations of pRW1004 and pRW1008. (*A*) Gel photograph. (*B*) Schematic representation of gel photograph showing the differing mobilities of the two plasmids at the superhelical densities at which the B-to-Z transitions occur. ×, pRW1004; ○, pRW1008.

Several procedures were used to obtain thermodynamical data on B-to-Z transitions (10, 15, 36–39). The data obtained from plots for each plasmid (such as that shown in Fig. 2) were used to calculate the superhelicity at the midpoints (Table 2) for pRW1001–pRW1101. For example, $\tau_{1/2}$ for pRW1001 = −7.2. However, this calculation does not take into account the relaxation of supercoils associated with the conversion of the insert to the Z form, which would change the value of $\tau_{1/2}$ to (−7.2 + 3.21) = −3.99. Therefore, the averages of these two sets of values for the midpoints of transitions were calculated (36) and are presented in Table 2 as $\tau_{1/2}$ (corrected).

Previous determinations of ΔG_{BZ} and ΔG_j by this and other similar methods have focused on different lengths of homogeneous sequences and have assumed that both of these parameters are constant when the only variable is the length of the insert (10, 15, 36–39). This approach is not viable when considering sequences that are heterogeneous, since ΔG_{BZ} now becomes an average of the contributions of each base pair to the overall energy of formation of the Z helix. For each insert shown in Fig. 1, however, the sequences at the junctions are the same and so ΔG_j was assumed to be a constant in each case.

The calculated (36) values for ΔG_{BZ} are presented in Table 2. Considerable differences exist between the value for the insert in pRW1001, which contains a perfect run of C-G, and those inserts that contain interruptions in this run.

The extra energy required to convert the interrupted sequence into a Z helix can be calculated by comparing the overall energy required ($n\Delta G_{BZ}$) with the energy necessary if the sequence were perfect (C-G). This latter value is readily obtained by multiplying ΔG_{BZ} for pRW1001 by the n bp of the sequence in question. Subtracting this from $n\Delta G_{BZ}$ gives the energy penalty resulting from the B-to-Z transition for the interrupting sequence. Table 3 shows that as the length of alternating T-A sequence increases in the middle of a run of C·G base pairs, the energy required to convert the whole region into a left-handed helix also increases. This energy increment is more pronounced when the T residues are contiguous rather than alternating and may be the result of being forced to assume a *syn* conformation in order to maintain a regular Z helix.

Table 3 also shows that ΔG_{BZ} for pRW1002 is negative, whereas the corrected $\tau_{1/2}$ necessary to effect this transition is higher than that found for the other sequences in this series. Thus, this extremely short sequence may be outside the limit

Table 2. Free energies of formation of left-handed Z helices for different sequences

Plasmid	$\tau_{1/2}$	$\tau_{1/2}$ (corrected)	n	ΔG_{BZ}, kcal·mol·bp^{-1}
pRW1001	7.2	5.6	18	0.09
pRW1002	8.7	7.9	8	−0.34
pRW1003	8.9	6.9	22	0.35
pRW1004	9.5	7.4	24	0.43
pRW1007	8.6	6.6	22	0.31
pRW1008	8.5	6.4	24	0.32
pRW1009	8.75	6.6	24	0.35
pRW1101	9.1	7.6	17	0.29

ΔG_{BZ} is the free energy required to convert 1 bp of the insert from a B to a Z structure and is calculated (36) from $\Delta G_{BZ} = [\tau_{1/2}/(N/400RT)] - 2\Delta G_j/n$ is the free energy of formation of one B–Z junction, which is taken to be 5.0 kcal·mol^{-1} (1 cal = 4.18 J). $\tau_{1/2}$ is the number of negative superhelical turns at the midpoint of the B-to-Z transition. n is the length of the inserts (Fig. 1) that undergoes the B-to-Z transition. The negative value for ΔG_{BZ} for pRW1002 is discussed in *Results*.

Table 3. Relative free energies of formation of Z helices in inserts

Plasmid	$n\Delta G_{BZ}$	$n(\Delta G_{BZ}^{1001})$	Energy increment
pRW1001	1.62	1.62	0
pRW1002	(−2.72)	0.72	(−3.44)
pRW1003	7.70	1.98	5.72
pRW1004	10.32	2.16	8.16
pRW1007	6.82	1.98	4.84
pRW1008	7.68	2.16	5.52
pRW1009	8.40	2.16	6.24
pRW1101	4.93	1.53	3.40

$n\Delta G_{BZ}$ is the free energy required to convert the n bp of each insert from a B to a Z structure, and $n(\Delta G_{BZ}^{1001})$ is the free energy that would be required for this conversion for n bp of alternating C-G, assuming that linear extrapolation of the data from pRW1001 is valid. The difference between these values is defined as the energy increment required for the entire length of the insert to adopt a left-handed Z helix with reference to a perfect tract of C-G (pRW1001). The negative values for pRW1002 are discussed in *Results*.

of applicability of this treatment (36). These calculations assume that the B–Z junctions occupy zero space inside the Z helix—that is, all of the base pairs in the insert are in a perfect left-handed Z conformation. In a long Z-forming insert, any departure from this approximation would be small compared with the overall structure of the region and would thus not have a profound effect on any thermodynamic calculations involving this assumption. If, on the other hand, the length of the Z helix is small compared to the lengths of the two junctions (or even smaller than the junction region), it may be expected that significant departures from ideal behavior would occur.

S1 Nuclease Cleavage. S1 nuclease recognizes and cleaves the structural distortions at B–Z junctions (6, 13–15, 40). Populations of topoisomers of all of the plasmids (pRW1001–pRW1101) with different amounts of negative supercoiling were treated with S1 nuclease followed by digestions with *Hae* II and *Bgl* I. For the control pRW790, two S1 nuclease-specific fragments were found at about 420 and 340 bp, which represented S1 nuclease cleavage at the major inverted repeat (cruciform), which lies within the 760-bp *Hae* II–*Bgl* I fragment. For all of the other plasmids, at least one other S1 nuclease-specific fragment was observed at ≈900 bp, which represented cleavage at or very near the synthetic insert, which lies within the 1250-bp *Hae* II–*Bgl* I fragment.

Fig. 4 shows a representative gel electrophoretic analysis of an S1 nuclease sensitivity study with pRW1101 as the substrate at supercoil densities below and above the B-to-Z transitions and with pRW790 as a control.

Thus, all of the plasmids have an S1 nuclease-sensitive site at or near the synthetic inserts and this sensitivity depends upon the negative superhelical density of the plasmids, in agreement with the predictions from the 2D gel analyses.

DISCUSSION

This systematic study shows that consecutive A·T pairs can adopt Z helices irrespective of sequence (TTTT or TATA). The runs of A·T pairs were flanked by C-G runs of 6, 8, or 10 bp in length. The energy required to cause the B-to-Z transitions varied in a sequence-dependent manner, as indicated by the amount of negative supercoiling required; longer tracts of A-Ts required more energy and alternating T-A sequences adopted Z structures more easily than homopurine·homopyrimidine (i.e., TTTT) runs.

$(CG)_4$ in a recombinant plasmid was shown herein to adopt a Z structure. This is the shortest Z-DNA sequence known in a plasmid. Interestingly, the insert in pRW1009 adopted a Z

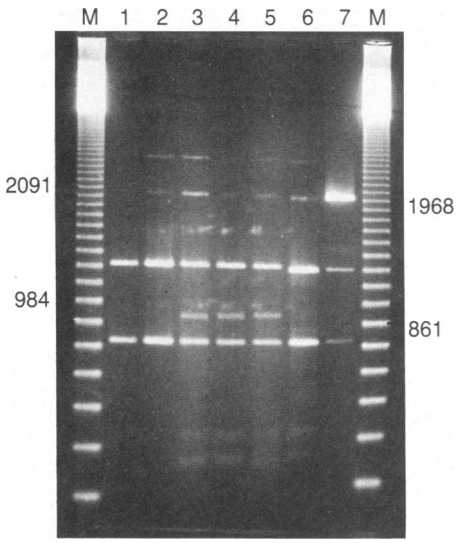

FIG. 4. Reactions of topoisomeric samples of pRW1101 with S1 nuclease followed by cleavage with *Hae* II and *Bgl* I. Lanes M, marker lanes of a 123-bp ladder (Bethesda Research Laboratories); lanes 1–5, topoisomer populations possessing mean negative superhelical densities of 0.008, 0.031, 0.055, 0.072, and 0.087, respectively, before S1 nuclease treatment; lane 6, the same reactions on pRW790 at native superhelical density; lane 7, an *Hae* II/*Bgl* I digest of pRW790. Sizes are indicated in base pairs.

structure, indicating that the left-handed structure potentials of $(C\text{-}G)_5$ and $(C\text{-}G)_4$ were not greatly destabilized by the relatively long intervening TATATA. Since prior work has proven that longer tracts of appropriate sequences more easily adopt Z structures than short tracts (6, 10, 15, 28, 32), the $(C\text{-}G)_4$ region in pRW1101 must adopt the Z helix most readily with the structure being transmitted through the TACC sequence to the $(C\text{-}G)_3$ region; a 6-bp tract has not been shown previously to adopt a Z structure. Also, three supercoil-dependent enzyme inhibitions [*Bam*HI cleavage (15) of the GGATCC sites flanking the insert; M*Hha* I methylation (28, 41) of the GCGCs; *Bss*HII cleavage of the GCGCGC site (42)] provided further evidence that the entire 17-bp insert in pRW1101 adopted a left-handed structure (results not shown).

All of the inserts (Fig. 1) also had the potential to form cruciforms (33–35, 43) but instead adopted Z-helices (from 2D gels). This result is particularly intriguing since others (17, 18) have shown that longer runs of T-A adopt cruciforms under the influence of negative supercoiling.

The nonpaired loops of cruciforms potentially adopted by pRW1008 and pRW1009 and the T-A tracts (17, 18) are similar, in principle. Hence, we conclude that (*i*) the substantial Z-DNA potential of the flanking C-G runs influences the neighboring T-A regions and (*ii*) cruciforms are not intermediates between the B and Z structures since the cruciform branch points would serve as barriers for the interconversion of writhe and twist, thus isolating the C-G duplexes in the hypothetical cruciform stems from the influence of negative supercoiling. Hence, if the inserts shown in Fig. 1 ever adopted cruciforms, Z helices would be prohibited.

Thus, in agreement with prior studies (14, 15, 44, 45), alternating purine-pyrimidine sequences are neither necessary nor sufficient for Z-DNA. Our results greatly broaden the sequence range for this unusual structure and hence make the *in vivo* existence and function of Z-DNA more plausible.

This work was supported by grants from the National Institutes of Health (GM-30822 and P30 CA 13148) and the National Science Foundation (83-08644).

1. Wells, R. D., Brennan, R., Chapman, K. A., Goodman, T. C., Hart, P. A., Hillen, W., Kellogg, D. R., Kilpatrick, M. W., Klein, R. D., Klysik, J., Lambert, P. F., Larson, J. E., Miglietta, J. J., Neuendorf, S. K., O'Connor, T. R., Singleton, C. K., Stirdivant, S. M., Veneziale, C. M., Wartell, R. M. & Zacharias, W. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 77–84.
2. Dickerson, R. E., Drew, H. R., Conner, B. N., Kopka, M. L. & Pjura, P. E. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 13–24.
3. Rich, A. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 1–12.
4. Wang, J. C., Peck, L. J. & Becherer, K. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 85–91.
5. Jovin, T. M., van de Sande, J. H., Zarling, D. A., Arndt-Jovin, D. J., Eckstein, F., Fuldner, H. H., Greider, C., Greider, I., Hamori, E., Kalisch, B., McIntosh, L. P. & Robert-Nicoud, M. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 143–154.
6. Singleton, C. K., Klysik, J., Stirdivant, S. M. & Wells, R. D. (1982) *Nature (London)* **299**, 312–316.
7. Peck, L. J., Nordheim, A., Rich, A. & Wang, J. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4560–4564.
8. Behe, M. & Felsenfeld, G. (1980) *Proc. Natl. Acad. Sci. USA* **78**, 1619–1623.
9. Behe, M., Zimmerman, S. & Felsenfeld, G. (1981) *Nature (London)* **293**, 233–235.
10. Klysik, J., Stirdivant, S. M., Singleton, C. K., Zacharias, W. & Wells, R. D. (1983) *J. Mol. Biol.* **168**, 51–71.
11. Nordheim, A. & Rich, A. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1821–1825.
12. Haniford, D. B. & Pulleyblank, D. E. (1983) *Nature (London)* **302**, 632–634.
13. Singleton, C. K., Kilpatrick, M. W. & Wells, R. D. (1984) *J. Biol. Chem.* **259**, 1963–1967.
14. Kilpatrick, M. W., Klysik, J., Singleton, C. K., Zarling, D., Jovin, T. M., Hanau, L. H., Erlanger, B. F. & Wells, R. D. (1984) *J. Biol. Chem.* **259**, 7268–7274.
15. Singleton, C. K., Klysik, J. & Wells, R. D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2447–2451.
16. Vorlicova, M., Sklenar, V. & Kypr, J. (1983) *J. Mol. Biol.* **166**, 85–92.
17. Greaves, D. R., Patient, R. K. & Lilley, D. M. J. (1985) *J. Mol. Biol.* **185**, 461–478.
18. Haniford, D. B. & Pulleyblank, D. E. (1985) *Nucleic Acids Res.* **13**, 4343–4363.
19. Mitsui, Y., Langridge, R., Shortle, B. E., Cantor, R., Grant, R. C., Kodama, M. & Wells, R. D. (1970) *Nature (London)* **228**, 1166–1169.
20. Drew, H. R. & Dickerson, R. E. (1982) *EMBO J.* **1**, 663–667.
21. Gaffney, B. L., Marky, L. A. & Jones, R. A. (1982) *Nucleic Acids Res.* **10**, 4351–4361.
22. Lo, K. M., Jones, S. S., Hackett, N. R. & Khorana, H. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2285–2289.
23. Blank, R. D. & Wilson, D. B. (1982) *Plasmid* **7**, 287–289.
24. Klein, D., Selsing, E. & Wells, R. D. (1980) *Plasmid* **3**, 88–91.
25. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
26. Singleton, C. K. & Wells, R. D. (1982) *Anal. Biochem.* **122**, 253–257.
27. Kang, D. S. & Wells, R. D. (1985) *J. Biol. Chem.* **260**, 7783–7790.
28. Zacharias, W., Larson, J. E., Kilpatrick, M. W. & Wells, R. D. (1984) *Nucleic Acids Res.* **12**, 7677–7692.
29. Klysik, J., Stirdivant, S. M., Larson, J. E., Hart, P. A. & Wells, R. D. (1981) *Nature (London)* **290**, 672–677.
30. Klysik, J., Stirdivant, S. M. & Wells, R. D. (1982) *J. Biol. Chem.* **257**, 10152–10158.
31. Peck, L. J., Nordheim, A., Rich, A. & Wang, J. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4560–4564.
32. Stirdivant, S. M., Klysik, J. & Wells, R. D. (1982) *J. Biol. Chem.* **257**, 10159–10165.
33. Panayotatos, N. & Wells, R. D. (1981) *Nature (London)* **289**, 466–470.
34. Singleton, C. K. & Wells, R. D. (1982) *J. Biol. Chem.* **257**, 6292–6295.
35. Singleton, C. K. (1983) *J. Biol. Chem.* **258**, 7661–7668.
36. Nordheim, A., Lafer, E. M., Peck, L. J., Wang, J. C., Stollar, B. D. & Rich, A. (1982) *Cell* **31**, 309–318.
37. Frank-Kamenetskii, M. D. & Vologodskii, A. V. (1984) *Nature (London)* **307**, 481–482.
38. Peck, L. J. & Wang, J. C. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6206–6210.
39. Cantor, C. R. & Efstratiadis, A. (1984) *Nucleic Acids Res.* **12**, 8059–8072.
40. Hayes, T. E. & Dixon, J. E. (1985) *J. Biol. Chem.* **260**, 8145–8156.
41. Vardimon, L. & Rich, A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3268–3272.
42. Azorin, F., Hahn, R. & Rich, A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5714–5718.
43. Germann, M. W., Schoenwaelder, K. & van de Sande, J. H. (1985) *Biochemistry* **24**, 5698–5702.
44. Wang, A. H.-J., Hakoshima, T., van der Marel, G., van Boom, J. H. & Rich, A. (1984) *Cell* **37**, 321–331.
45. Ellison, M. J., Kelleher, R. J., III, Wang, A. H.-J., Habener, J. F. & Rich, A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8320–8324.