Sequence-dependent S1 Nuclease Hypersensitivity of a Heteronomous DNA Duplex*

Todd Evans and Argiris Efstratiadis

From the Department of Genetics and Development, Columbia University, New York, New York 10032

(Received for publication, June 3, 1986)

Using cloned (dG-dA),-(dC-dT), DNA duplexes ((GA)n) as models of homopurine-homopyrimidine S1-hypersensitive sites, we show that cleavage of the alternate (non-B, non-Z) DNA structure by S1 nuclease is length-dependent, in both supercoiled and linear forms, which are similar because of the identity of their nicking profiles. However, the length of flanking sequences, the presence of borders, and the DNA topology affect the equilibrium between the alternate structure and B-DNA. The B form of (GA)n has a 10.4-base pair helical repeat, but the two phosphodiester backbones have different conformations (heteronomous DNA with a diastride repeat unit). Extension experiments reveal that the alternate structure is also heteronomous, in agreement with the nicking patterns generated by S1 and mung bean nucleases and by venom phosphodiesterase. Sensitivity to the latter enzyme at pH 9.0 indicates that the alternate DNA does not appear only in the low pH of the S1 nuclease reaction. Moreover, Hoogsteen G-CH* base-pairing does not seem to be a prerequisite for the appearance of sensitivity because S1 still recognizes the structure even when all Gs are methylated at N-7. This is consistent with the results of chemical probing of the structure using dimethyl sulfate and diethyl pyrocarbonate at various pH values, which show absence of protection at guanine N-7. However, diethyl pyrocarbonate treatment at low pH results in hyper-reactivity of A residues.

DNA sequences which are hypersensitive to S1 nuclease often map to the 5' flanking regions of transcribing genes in active chromatin domains (see, for example, Refs. 1 and 2, and Ref. 3 for a review). They are also present on naked supercoiled (but usually not relaxed or E. coli DNA (1, 4, 5). Such S1-hypersensitive sites (SHS), which may play a role in transcription or chromatin organization, reside in homopurine-homopyrimidine stretches (4, 5), and, despite their sensitivity to single strand-specific nucleases, they seem to be regions of duplex DNA (6) in an alternate (non-B, non-Z) conformation. Previous work has shown that a primary structure consensus sequence cannot be derived from the examined SHS (5).

* This work was supported by a grant from the National Institutes of Health (to A. E.) and a gift to the laboratory by the Bristol-Myers Co. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The abbreviations used are: SHS, S1-hypersensitive sites; DEP, diethyl pyrocarbonate; bp, base pairs; kb, kilobase pairs; nt, nucleotide; Hepes, 4-(2-hydroxyethyl)-1-piperazineethane sulfonic acid; NTP, nucleotide triphosphate.

However, (dG-dA),-(dC-dT), stretches (henceforth referred to as (GA)n) appear frequently in SHS with heterogeneous sequence (4, 5), while (GA)n, homopolymeric SHS are associated with several eukaryotic genes, encoding sea urchin histones (7), Drosophila histones (8), Drosophila heat shock proteins (9), mouse rens (10), and human U1 RNA (11). A (GA)n, element is present in the 3' flanking region of the mouse C4 heavy chain gene, while a (CT)n, element is present in the second intron of C5 (12). Three tracts of (TC)n, (GA)n, and (TC)n are all contained in the smallest known fragment to exhibit enhancer activity that maps at the 5' flanking region of the mouse MHC Eβ gene (13). Interestingly, a protein purified from Drosophila cultured cells which stimulates transcription binding to (GA)n, stretch present in the spacer region between the Drosophila histone genes H3 and H4. Thus, considering that (GA)n, sequences could be used as models for further investigation of the alternate structure of an SHS, we cloned into pUC9 (14) such duplexes with n = 5, 10, 20, or 38 (pGA5, pGA10, pGA20, and pGA38) and examined them in supercoiled and linear forms using enzymatic and chemical probes, and by primer extension experiments.

EXPERIMENTAL PROCEDURES

Materials

Restriction enzymes, T4 DNA polymerase, Escherichia coli DNA polymerase or its Klenow fragment, and M13 universal sequencing primers were from New England Biolabs; S1 nuclease, T4 polynucleotide kinase, and NACS.52 resin were from Bethesda Research Laboratories; mung bean nuclease, EcoRI linkers, and phosphorylated (dA-dG), and (dC-dT), (henceforth referred to as AG and CT deomers) were from P-L Biochemicals; bacterial alkaline phosphatase, venom phosphodiesterase, and DNase I were from Worthington; reverse transcriptase was from Life Sciences; Dimethyl sulfate was from Aldrich; DEP was from Sigma; 7-methyl-dGTP, synthesized as described (15), was a generous gift from P. Srinivasan; [α-32P]dNTPs (700 Ci/mmol) and [γ-32P]ATP (3,000 Ci/mmol) were from New England Nuclear. A 20-base-long oligonucleotide (5'TGC-

Plasmid and M13 Constructions

pGA38—AG and CT deomers (1 μg each) were hybridized at 15 °C in a 20-μl reaction containing 50 mM Tris, pH 7.5, 10 mM MgCl2, 15 mM NaCl, 1 mM dithiothreitol, 25 μg/ml bovine serum albumin, 1 mM concentration each of the four unlabeled dNTPs, and 10 μM [α-32P]dATP as tracer. After 1 h of hybridization, 10 units of E. coli DNA polymerase were added, and the reaction was incubated for 2 h at the same temperature. Following phenol extraction and ethanol precipitation, the mixture was electrophoresed on a 10% nondenaturing polyacrylamide gel, and molecules ranging in size between 75 and 125 bp were eluted. The ends of the purified molecules were made blunt by using T4 DNA polymerase, and then they were cloned (16) into the EcoRI site of vector pUC9 (14) following attachment of C. Parker, personal communication.
EcoRI linkers. Plasmid pgA38, which was randomly picked, was characterized by DNA sequencing, using the chemical method (17). pgA5, pgA10, and pgA20—AG and CT decamers (1 μg each) were first dephosphorylated and then 5'-end labeled using [γ-32P]ATP and T4 polynucleotide kinase. They were then ligated at 22 °C for 16 h. Concreted DNA was digested with a combination of SacI (GAGCTC) and XbaI (TCTAGA) to eliminate molecules containing AGCT and CTAG junctions formed by ligation. The digestion mixtures were electrophoresed on non-denaturing polyacrylamide gels and fragments with sizes of 10, 20, and 40 bp were cloned into pUC9 as described above. The inserts of clones pgA5, pgA10, and pgA20, which were randomly picked, were characterized by DNA sequencing. mpCT38 and mpGA19—The insert of pgA38 was excised by EcoRI digestion, and cloned into the EcoRI site of the replicative form of M13mp9. A randomly picked plaque was shown by dideoxy sequencing (18) to contain the CT strand of pgA38 in its entirety (mpCT38). However, we were unable to obtain a recombinant phage containing a complete copy of the homopurine strand of pgA38. One clone, presumably arising from a spontaneous deletion, which gave a weakly positive C-test to mpCT38, was shown by sequencing to contain 19 GA doublets (mpGA19).

DNA Labeling

End-labeled duplex DNA was prepared from the pgA plasmids by first linearizing with HindIII and then labeling either at the 5' end by kinasing or at the 3' end by filling-in with reverse transcriptase in the presence of [α-32P]ATP and unlabeled dATP, dGTP, and dCTP. Following secondary restriction enzyme digestion, uniquely labeled fragments were separated on nondenaturing polyacrylamide gels and eluted at room temperature as described (17).

End-labeled single-stranded DNA used for the DNase I experiments described in Fig. 5 was prepared using a template either mpCT38 or mpGA19 (1.5 μg) and as a primer the universal 15-nt 5' end-labeled by kinasing. The primer was extended using the Klenow enzyme and unlabeled dNTPs (0.5 mM each) for 1 h at 22 °C. Following HindIII digestion, the newly synthesized strand was denatured from its template and purified by gel electrophoresis on a urea-polyacrylamide gel.

S1 Nucleases

Circularly labeled circles were prepared as described (4) by circularizing linearized pgA38, which was 5' end-labeled at the HindIII site. For the experiment of Fig. 3, these molecules were linearized with ScaI.

To generate a uniquely labeled EcoRI insert of pgA38, the purified fragment was first labeled at both 5' ends by kinasing. Following denaturation and strand separation on a non-denaturing polyacrylamide gel (17), the eluted GA or CT strand was ligated into M13mp9. A randomly picked plaque was shown by dideoxy sequencing (18) to contain the CT strand of pgA38 in its entirety (mpCT38). However, we were unable to obtain a recombinant phage containing a complete copy of the homopurine strand of pgA38. One clone, presumably arising from a spontaneous deletion, which gave a weakly positive C-test to mpCT38, was shown by sequencing to contain 19 GA doublets (mpGA19).

DNA Electrophoresis

Denaturing urea-polyacrylamide gels were as described (17). All non-denaturing polyacrylamide gels were 29:1 acrylamide: bis-acrylamide, except for those used for strand separation, which were 58:1. Agarose gel electrophoresis was as described (19). The agarose gel of Fig. 3 was fixed in 10% acetic acid and dried prior to autoradiography.

Nuclease Digestions

S1 and mung bean nuclease digestions were performed in 40-μl reactions containing 50 mM sodium acetate, 1 mM ZnSO4, 1 μg of DNA, and 5 or 25 units of enzyme, at NaCl and pH conditions described in the figure legends. The incubation was maintained at 37 °C for 20 min. To stop S1 or mung bean nuclease nicking at nucleotide resolution, 1 μg of supercoiled or linear DNA was nicked for 1 min at 37 °C with 5 units of enzyme in 50 mM sodium acetate, pH 4.5 (or 5.0 for mung bean nuclease), 50 or 300 mM NaCl, and 1 mM ZnSO4. Following digestion with S1, the resulting nicked DNA was treated with HindIII, the molecules were 5'-end labeled, and after digestion with a secondary restriction enzyme, the uniquely labeled fragments were isolated from non-denaturing polyacrylamide gels and electrophoresed on denaturing urea-polyacrylamide gels in parallel with appropriate chemical sequencing ladders. Mock reactions contained dilution buffer instead of nuclease. All of the reactions in Fig. 2c contained also an equal amount of carrier phage λ-DNA (1 μg).

Venom phosphodiesterase digestions were performed in 40-μl reactions containing 20 mM Tris, pH 9.0, 1 mM MgCl2, 1 μg of DNA, and 0.05 unit of enzyme for 5-10 min at 37 °C, or for 1 min for fine mapping.

DNase I digestions of DNA in solution were performed in 50-μl reactions containing labeled DNA, 1 μg of carrier DNA, 10 mM MgCl2, and buffer (100 mM sodium acetate, pH 4.5, 100 mM Tris, pH 7.5, or 100 mM Tris, pH 9.0). Enzyme dilutions and reaction times at 22 or 37 °C were established in preliminary calibration experiments.

DNase I digestions of end-labeled DNA fragments immobilized on hydroxyapatite were performed essentially as described (20, 21). Following digestion for 1 min at 22 °C, the reactions were terminated by adding EDTA to 125 mM, which also dissolved the calcium phosphate crystals. The DNA was purified by using a NACS-52 minicolumn (binding in 0.2 M NaCl, washing with the same salt concentration, and eluting with 0.8 M NaCl).

The position of nicks was indicated according to rules described previously (4).

RESULTS AND DISCUSSION

Chemical Probing

Labeled DNA in the presence of 1 μg of carrier DNA was treated with dimethyl sulfate or DEP in 200 μl of buffer (50 mM sodium acetate, pH 4.5, 50 mM Tris, pH 7.0 or 9.0, or sodium cacodylate, pH 8.0) at 37 °C for 1 min (dimethyl sulfate) or 5 min (DEP). The molecules were then processed as described (17).

Primer Extensions

One pmol of single-stranded AG or CT decamer, 5' end-labeled by an exchange reaction (22), was incubated with 400 ng of template and 15 units of reverse transcriptase for 1 h at 37 °C under standard conditions (16). Alternatively, the extension reactions were performed using Klenow enzyme in 100 mM Hepes, pH 6.9, 10 mM MgCl2, 50 mM dithiothreitol, 90 mM KCl, and 0.5 mM concentration each of cold dNTPs, at 15 °C for 2 h.

RESULTS AND DISCUSSION

Experimental Rationale—The SHS are sensitive not only to S1 nuclease but to other single-strand-specific nucleases as well (2, 23–28). Because of the significant preference of these activities for single strands, and although the mechanism of this action is unknown, it was initially proposed that the SHS are single-strand-specific nucleases generated by "DNA slippage" (7–9). In contrast, based first on indirect evidence (4) and later on direct data (6), we proposed that the SHS are duplex DNA. This interpretation is consistent with the absence of SHS reactivity with E. coli single-strand binding protein (Ref. 8 and data not shown). Thus, the suggestion that the homopurine-homopyrimidine SHS are alternate (non-B, non-Z) DNA is based on the argument that this presumably duplex structure is anomalously sensitive to enzymes that were previously known to recognize mainly single strands and B-DNA/Z-DNA junctions. It is noteworthy that even the B-Z junctions themselves may be fully paired (27).

Nevertheless, since most of the nuclease sensitivity experiments are performed by necessity in vitro, different secondary structures might be appearing under different experimental conditions. Thus, the salt and pH conditions of the digestion reactions become important parameters that should be considered in interpreting the data. However, it is almost impossible to discriminate between effects of the ionic strength or of the pH on the alternate structure itself and effects on the enzymatic activity. For this reason we examine the (GA)n SHS with more than one nuclease, appraise the similarity of structures under different conditions from the results of fine mapping, and complement our enzymatic results with chemical probing and primer extension experiments.

Length-dependent S1 Hypersensitivity of Supercoiled and Linear DNA—It is known that S1 nuclease recognizes an
SHS and first introduces a nick into the DNA molecule. In a second step, it recognizes the nick and cuts across from it (cleavage; Ref. 4). When supercoiled vector pUC9 (control) or pGA5 are exposed to S1 nuclease, linearization does occur, but it is not due to cleavage at a specific site (not shown). However, S1 linearization of pGA10, pGA20, or pGA38, followed by ScaI digestion (which cuts once in pUC9), results in the appearance of two fragments of 1.76 and 0.95 kb, mapping to the GA inserts of the plasmids (Fig. 1, a and b).

These results are in agreement with observations by Htun et al. (11), who studied a (GA)$_n$ element, exhibiting length polymorphism, that maps at a distance of 1.8 kb 3' to the cleavage to the GA inserts of the plasmids (Fig. 1, a and d).

![Fig. 1. Gross mapping of SHS in pGA plasmid DNA. Panel a, supercoiled plasmids pGA38 (lanes 4–8), pGA20 (lanes 11 and 12), or pGA10 (lanes 13 and 14) were first digested for 30 min by S1 nuclease (5 units/μg of DNA) at pH 4.5 in the presence of various concentrations of NaCl as shown. Following purification, the DNA was digested with ScaI and electrophoresed on a 1% agarose gel (ethidium bromide staining). Lanes 9 and 10 are as lane 7, except that 25 units of S1/μg of DNA were used and the pH was 6.0 (lane 9) or 7.0 (lane 10). The DNA markers in lanes 1–3 and 15 are: supercoiled (S) pGA38 (lane 1), ScaI-linearized (L) pGA38 (lane 2), and HindIII fragments of λ-DNA (from top to bottom, 23, 9.4, 6.6, 4.3, 2.3, 2.0, and 0.56 kb, lanes 3 and 15). Lanes 4–14 contain (from top to bottom) DNA which was not cleaved by S1 but was linearized by ScaI (approximately 2.7 kb) and the ScaI products of S1 cleaved material (bands of approximately 1.76 and 0.95 kb, see panel d). The site of S1 cleavage was verified by using other restriction enzymes (not shown). Panel b, as panel a except that the plasmids were first linearized by ScaI and then digested by S1 nuclease. Panel c, lanes 1–3 (markers) are the same as lanes 1–3 in panel a. Lane 4 is as lane 5 in panel a. Lanes 5 and 6 are as lane 4, except that mung bean nuclease was used instead of S1 nuclease, at pH 4.5 and 7.0, respectively. Lanes 7–9 are as lane 4, except that venom phosphodiesterase was used instead of S1 nuclease, for 5, 10, and 15 min, respectively. Panel d, restriction map of pGA plasmids. Black triangle, GA insert; black square, pUC9 polylinker; R, EcoRI site; H, HindIII site. The unique ScaI site and one of the two BglII sites are also indicated. Other sites used for secondary restriction mapping in other experiments (see legend to Fig. 2) are omitted for clarity.
by gel electrophoresis or S1 cleavage at the B/Z junctions. This salt effect on Z-DNA present in supercoiled molecules contrasts the known stabilization in high ionic strength of (GC), not associated with other sequences, and, therefore, lacking borders (30).

The simplest way to interpret our observations is that, in comparison to B-DNA, the SHS are underwound, yet intrinsically unstable structures, which, when short, are destabilized by the presence of intrinsically unstable SHS/B-DNA boundaries. Superhelical tension may cancel this destabilizing effect, by assisting in unwinding the helix and thereby reducing the energy barrier imposed by the boundary. This interpretation is consistent with the observation that a synthetic linear CGC(GA),GCCG duplex, which lacks borders, is S1-sensitive (data not shown).

**Fine Mapping of Nuclease Hypersensitivity**—Fine mapping of the S1 nicks using supercoiled or linearized pGA38 as a substrate (Fig. 2) reveals that nicking occurs throughout the length of the GA insert and on both strands. Under fixed conditions, reproducible nicking profiles are obtained, showing “peak-valley” discrimination between base steps and regional differences in intensity. In the 5′ to 3′ direction, the ApG and TpC steps are preferred over their neighbors. In addition, reproducible, but different nicking profiles, with regard to the distribution of the intensity of hits, are observed in high (0.3 m) and low (0.5 m) salt (Fig. 2, panels a and b, respectively), possibly reflecting subtle shifts within the population of conformational family members, adjusting to the changes of the ionic environment.

The most important observation from this set of experiments is that the profiles of supercoiled and linear molecules are identical (Fig. 2b, compare lanes 1 and 4), indicating that the nuclease recognizes the same alternate structure in both cases. We note that, despite the identity of these profiles, the rate of nicking is faster for the supercoiled molecules (data not shown). This suggests that the DNA topology might influence the equilibrium between B and alternate DNA.

A different fine-mapping profile was reported by Pulleyblank *et al.* (26) for a (GA), stretch present in a supercoiled plasmid. On the CT strand, S1 nuclease nicked almost uniformly the CpT, but not the TpC, step. Data were not shown for the AG strand, but it was mentioned that it is usually S1-insensitive, although it can become sensitive under certain (unspecified) conditions. In contrast, Margot and Hardison (31) reported that a (GA), stretch, present around position -400 from the capping site of the rabbit β-globin gene, is S1-sensitive in supercoiled molecules primarily on the GA strand, yielding a peak-valley profile.

It is unclear whether these differences in nicking profiles are due to different lengths or variations in the environmental conditions, which can obviously influence the outcome of this type of mapping. However, this variability does not affect our main conclusion that the nuclease is recognizing similar conformations in supercoiled and linear molecules under fixed experimental conditions.

Using fine mapping, we also examined (Fig. 2c) the nuclease sensitivity of linear (GA), with shorter flanking sequences than the molecules in Fig. 2b. Surprisingly, a HindIII-BglII fragment of pGA38, which contains the (GA), stretch flanked by 57 bp on the HindIII side and 151 bp on the BglII side, is relatively insensitive to mung bean nuclease (Fig. 2c, lane 2) and almost completely insensitive to S1 nuclease (data not shown). The nicking intensity was increased by reducing the downstream flank of the HindIII-BglII fragment from 151 to 13 bp by HaeIII digestion (Fig. 2c, lane 3) or by increasing the length of the upstream flank by ligating phage λ HindIII fragments (Fig. 2c, lane 4). Since the sequence context around the junctions was not changed in these experiments, we attribute these results to the relationships between the lengths of the two flanks.

When the flanking DNA is essentially removed by releasing the insert of pGA38 with EcoRI digestion, mung bean nuclease nicked the homopolymeric sequence (flanked on either side by 3 bp and a 4-nt single-stranded overhang) efficiently on both strands (Fig. 2d, lanes 2 and 3). We note that the nicking of the homopurine strand by mung bean nuclease does not exhibit just a bond preference for the dinucleotide steps (as S1 does), but it hits exclusively between A and G in the 5′ to 3′ direction. However, the nicking profile of the homopyrimidine strand is the same for both enzymes (summarized in Table 1).

To rule out definitively any intermolecular interactions as the basis of S1 hypersensitivity, we examined the rate of nuclease digestion as a function of DNA concentration. Fig. 3 shows the results of an experiment in which a constant amount of internally labeled ScaI-linearized pGA38 (see “Experimental Procedures” and legend to Fig. 3) was digested with S1 nuclease in the presence of increasing amounts (1X, 10X, and 100X) of the same fragment in unlabeled form. The appearance of a 1.76-kb fragment (see also Fig. 1) is diagnostic of S1 cleavage at the (GA),, insert. As Fig. 3 shows, the DNA concentration has no effect on the rate of cleavage (the...
S1 Nuclease Hypersensitivity of Duplex DNA

Table I

| Nuclease                        | Pattern                                      |
|---------------------------------|----------------------------------------------|
| DNase I                         | 5'...GAGAGA...3'                             |
|                                 | 3'...CTCTCT...5'                             |
| Mung bean nuclease              | 5'...GAGAGA...3'                             |
|                                 | 3'...CTCTCT...5'                             |
| S1 nuclease and venom phosphodiesterase | 5'...GAGAGA...3'                             |
|                                 | 3'...CTCTCT...5'                             |

Fig. 3. Rates of S1 cleavage of pGA38 at different DNA concentrations. A vast excess of unlabeled pGA38 at three different concentrations (1X, 10X, and 100X) was added to ScaI-linearized pGA38 internally labeled at the HindIII site, and each mixture was cleaved with S1 nuclease for the time period shown, under the conditions of Fig. 1a, lane 5. Following purification, the molecules were electrophoresed on a 1% agarose gel. The (-) lane corresponds to a mock reaction incubated for 30 min. The DNA size markers in lane M are HindIII fragments of phage λ.

Table II

| Nuclease Pattern | Hypersensitivity |
|------------------|------------------|
| DNase I          |                  |
| Mung bean nuclease |                |
| S1 nuclease      |                  |

Fig. 4. DNase I digestion patterns of the linear pGA38 SHS immobilized on hydroxyapatite and in solution. Panel a, autoradiogram of a 6% urea-polyacrylamide gel. The 250-bp HindIII-BglI fragment of pGA38 (Fig. 2d), labeled either at the 5' end of the CT strand (lanes 1–4) or at the 5' end of the GA strand (lanes 5–8) was partially digested by DNase I, either in solution (sol; lanes 4 and 7) or after binding to hydroxyapatite (HAP; lanes 2, 3, 5, and 6). Lanes 3 and 6 are the same as lanes 2 and 5, respectively, except that the amount of enzyme was used. Lanes 1 and 8 are control reactions C and G, respectively, of the CT or the GA strand. Panels b and c, densitometric scans (left to right) of lanes 3 and 6 (top to bottom) of panel a, respectively. For each strand, the helical repeat is calculated by dividing the number of bands between peaks by the number of repeat units encompassing these peaks (e.g. 52/5 = 10.4 in panel b). Since DNase I hits every other base step in lanes 5 and 6, the number of peaks in panel c must be doubled for the calculation.

DNA structure using single-strand-specific nucleases was unsuccessful. For this purpose, we used calcium oxalate (37), instead of hydroxyapatite, because the latter is soluble in the low optimum pH of the S1 or mung bean nuclease reactions. Moreover, it is known that phosphate ions inhibit these
nucleases. We found that mung bean nuclease can completely solubilize single-stranded DNA bound to calcium oxalate, whereas the activity of S1 nuclease is inhibited under these conditions. However, the EcoRI fragment of pGA38, which is hypersensitive to mung bean nuclease in solution (see Fig. 2d), is totally resistant to this enzyme when it is bound to calcium oxalate. We think that during binding to a solid surface the B-DNA/alternate DNA equilibrium is drastically shifted to the left.

When HindIII-BglII fragments in solution (not immobilized) are examined with DNase I, they exhibit the same bond sensitivity that is observed with the immobilized molecules (Fig. 4a, lanes 4 and 8). We note that the alternating peak-valley nicking profile of the CT strand is analogous to that detected by the same method using a poly(dA-dT)-poly(dT-dA) substrate (38). In contrast to the B-type form, the homopolyrimidine strand of the alternate structure exhibits a different bond sensitivity to single-strand-specific nucleases, which hit in the 5′ to 3′ direction the T-C step more frequently than the C-T step (summarized in Table I).

Further probing using DNase I showed that this enzyme generates identical nicking profiles from linear and supercoiled molecules (Fig. 5a). However, the (GA)$_n$ inserts in the latter molecules are much less susceptible to digestion than their linear counterparts, presumably because the B-alternate equilibrium has been significantly shifted to the right. Interestingly, in the supercoiled molecules, DNase I also detects alterations in the immediate flanking regions of the GA insert. The flank toward the HindIII site is nuclease-insensitive, whereas the nicking pattern toward the BglII site has changed (Fig. 5a, compare lanes L and S). We note that in the linear molecules the DNase I nicking profile is practically the same at pH 4.5, 7.5, or 9.0 (Fig. 5b), although the nicking intensities in the GA insert region, but not in the flanking DNA, differ, suggesting pH-dependent shifts in the equilibrium between the B-DNA and alternate DNA forms.

To determine whether DNase I is detecting length-dependent features, like S1 does, we examined the linear form of pGA10, which is S1-insensitive when it is flanked by long stretches of B-DNA. Fig. 5c shows that the GA strand is still hit by DNase I exclusively at the ApG step (lane 1), but the CT strand does not have a peak-valley profile (lane 2). In contrast, the S1-hypersensitive synthetic CGC(GA)$_p$GCG duplex, which lacks borders, exhibits the characteristic peak-valley profile of the CT strand when examined with DNase I (data not shown). When the GA stretch is as short as 10 bp and embedded in long flanking DNA (linearized pGA5), the GA strand begins to lose the uniformity of its pattern (Fig. 5d). However, the GpA step is still insensitive to nicking.

Surprisingly, when we used an excess of DNase I to digest single strands (which are much less sensitive than duplexes to this enzyme), we observed that the (GA)$_n$ strand (Fig. 5e, lane 1) retained its digestion profile (hits every other base), while the peak-valley profile of the CT-strand (Fig. 5e, lane 2) was lost. A GpA step outside the GA insert (Fig. 5e, lane 1, arrow) is sensitive to DNase I, an observation consistent with our conclusion that the enzyme is recognizing a homopolymeric GA single-stranded stretch in a unique way. However, a 25-fold (and in some cases 100-fold) higher DNase I concentration is necessary to digest the single-stranded GA or CT sequences than the (dG-dA)$_n$·(dC-dT)$_n$ DNA.

According to these results, the structure of the B form of (GA)$_n$ can be classified as heteronomous DNA (39) by analogy to the proposed structure for (dA)$_n$·(dT)$_n$. We use this term in the broadest possible sense to simply indicate that the two backbones have different conformations. The DNase I diges-

FIG. 5. Probing of pGA SHS by DNase I. Panel a, DNase I nicking profiles at pH 7.5 of the GA strand of linear (lane L) and internally labeled supercoiled (lane S) pGA38. After nicking, the molecules were processed as described in the legend to Fig. 2. Panel b, as lane L in panel a, except that nicking was performed at pH 4.5, 7.5, or 9.0 (lanes 1, 2, and 3, respectively). Panel c, DNase I nicking profiles at pH 7.5 of the GA strand (lane 1) and CT strand (lane 2) of linear pGA10 (linearization with Scal). In this case, in contrast to other pGA clones, the GA strand is 3′ end-labeled and the CT strand is 5′ end-labeled at the HindIII site. Panel d, as panel b lane 2, except that linearized pGA5 was used. The region of the five GA doublets is indicated by the bracket. Panel e, DNase I nicking profiles at pH 7.5 of single-stranded DNA containing either a (GA)$_n$ stretch (lane 1) or a (CT)$_n$ stretch (lane 2). The label is at the 5′ end of the 15-nt universal M13 primer used for synthesis of the strands on M13 templates mpCT38 and mpGA19 (see “Experimental Procedures”). The concentration of DNase I used for these experiments was 25-fold higher than for experiments in panels a–d. An arrowhead indicates a GpA step outside the GA stretch, which is DNase I insensitive. G, G + A, and C + T in various panels indicate chemical sequencing ladders. In all cases, electrophoresis was on 6% urea-polyacrylamide gels.
tion experiments also provide information about the repeat unit, which is obviously a dinucleotide, in agreement with x-ray diffraction data from DNA fibers (40).

We interpret the results of the DNase I digestion results as due primarily to unusual stacking of the purines on the GA strand, which governs the heteronomous B conformation, while the CT strand follows passively the structure that the homopurine strand assumes, depending on length. It is likely that this unusual stacking arrangement, which might be related to the absence of "propeller twist"-dependent steric hindrance (41) in the homopurine-homopyrimidine sequences, leads to concerted changes in the backbone torsion angles and in the orientation about the glycosidic bonds, generating a heteronomous structure which is in equilibrium with alternate DNA, anomalously sensitive to single-strand-specific nucleases. We note that this notion is consistent with limited information from NMR studies on A- and G-containing ribo-oligonucleotides, which indicates the presence of unusual stacking arrangements depending on length and residue positions in the oligonucleotide chain (42). We also note that a model for poly(dG)-poly(dC), based on x-ray crystallographic data from single crystals indicating an A-DNA form (43), suggests that the homopurine strand is stacked, while the homopyrimidine strand is unstacked. Moreover, Raman spectroscopy indicates that, at 30 °C and in 1 M NaCl, poly(dG)-poly(dC) assumes a heteronomous A-form (44). Although the (G/C), structure is S1-hypersensitive in supercoiled form (4), the exact relationship between the (G/C), and (GA), conformations is unknown. It is not unlikely, however, that they both belong to the same conformational family.

Is the Alternate DNA Structure pH-dependent?—The pH optimum for S1 nuclease is 4.5, and the activity of the enzyme is diminished at higher pH values (28). Pulleyblank et al. (28) interpreted the S1 hypersensitivity of a (GA)32 stretch as due to recognition by the nuclease of an alternate structure (induced by the low pH of the enzymatic reaction) in which N-3 protonated cytosines form Hoogsteen or reverse Hoogsteen base pairs with guanines. Lyamichev et al. (45, 46) also interpreted S1 hypersensitivity as a low pH-dependent structural transition, but proposed different models to explain it. Since most of the S1-hypersensitivity experiments are performed at low pH optimum, the contents of these groups are plausible. Although we showed that S1 is still recognizing the structure at pH 7.0, their hypothesis is still viable for the following reason. Unmodified Cs are not protonated above pH 6.0 (47), but the pKa of C in a polynucleotide chain might shift drastically, depending on the stabilities of the putative secondary structures formed with and without the proton. This question cannot be addressed with other commonly used single-strand-specific nucleases, because the pH dependence of mung bean nuclease, for example, is almost identical with that of S1, whereas the Neurospora crassa endonuclease, which has optimal activity at neutral pH, is not a dependable enzymatic probe because its specificity varies depending on the preparation (2, 4; but see Refs. 23 and 24). For these reasons, we decided to examine the sensitivity of (GA), SHS to venom phosphodiesterase, which exhibits optimal activity at pH 9.0, a condition under which C-protonation seems very unlikely.

Venom phosphodiesterase from Crotalus adamanteus is primarily an exonuclease degrading both RNA and DNA (native or denatured) in the 3' to 5' direction to yield 5' mononucleotides. In addition, the enzyme possesses an intrinsic endonuclease activity which exhibits all of the characteristic features of a single-strand-specific nuclease; it degrades preferentially single-stranded DNA and also nicks supercoiled substrates (48).

Fig. 1c (lanes 7–9) shows that the nicking/cleavage of supercoiled pGA38 by venom phosphodiesterase at pH 9.0 maps at the GA insert of the plasmid as in the reactions using S1 (lane 4) or mung bean nuclease (lanes 5 and 6), except that the exonucleolytic activity of the former enzyme shortens the products of endonucleolysis in a time-dependent fashion, as expected. In addition, the fine mapping profile generated by this enzyme (Fig. 2b, lane 3) is similar to that derived by using S1; all dinucleotide steps of both strands are accessible to venom phosphodiesterase, but there is a preference for the ApG and TpC steps in the 5’ to 3’ direction (summarized in Table I). These results seem to exclude protonation as the basis of formation of the alternate DNA structure, under the reasonable assumption that all single-strand-specific nucleases, acting as conformational probes, recognize similar features.

Chemical Probing of the Alternate Structure—Despite the sensitivity of pGA38 to venom phosphodiesterase, our conclusion that the alternate GA structure is pH-independent (at least in one of its possible forms, in case of further polymorphism) was still incompatible with other results reported by Pulleyblank et al. (26). These authors presented evidence according to which approximately half of the Cs in the GA strand were protected against alkylation by dimethyl sulfate at pH 4.5, but not at pH 7.0. On the basis of this result, they proposed a model according to which the N-3 position of cytosine is protonated in low pH, leading to the formation of G-CH+ Hoogsteen base pairs (with the G residue in syn conformation), alternating with Watson-Crick A-T base pairs. Accordingly, we decided to examine the dimethyl sulfate sensitivity of supercoiled or linearized pGA38.

When HindIII-BglII fragments carrying (GA), mostly in its B form are treated with dimethyl sulfate, the chemical methylates the N-7 position of G to the same extent and throughout the length of the GA stretch at pH 4.5, 8.0, or 9.0 (Fig. 6a, lanes 1–3). These observations remain unchanged when the S1-hypersensitive EcoRI excised GA stretch is examined at pH 4.5 or 8.0 in the presence or absence of Mg2+ (Fig. 6c). The profile of supercoiled molecules examined at pH 4.5 is not altered at pH 9.0 (Fig. 6b), and it is similar to that of linear molecules, except that under superhelical stress certain G residues (Fig. 6b, arrowheads) exhibit hyper-reactivity at high pH, while some A residues (Fig. 6b, arrows) also react with dimethyl sulfate.

Because of the negative character of these results that did not confirm the observation reported by Pulleyblank et al. (26), we decided to use diethyl pyrocarbonate (DEP) as an additional probe. DEP carboxylates purines at N-7, and was recently shown to be a sensitive chemical probe for Z-DNA present in supercoiled plasmids (49, 50). Fig. 7a shows that in the (S1-insensitive) HindIII-BglII molecules there is no regional protection to DEP at pH 4.5, 8.0, or 9.0. However, although DEP generates an A > G pattern in mixed sequence DNA at pH 8.0 (51), the pattern is inverted in the (GA), stretch at pH 8.0 or 9.0 (Fig. 7a, lanes 2 and 3, respectively). At pH 4.5, the expected A > G pattern is restored in approximately half of the insert, which, in addition, becomes hyper-reactive (Fig. 7a, lanes 1, compare to lanes 2 and 3 inside and outside the GA stretch). When supercoiled or Scal-linearized pGA38 molecules, in both of which the GA stretch is S1-hypersensitive, were reacted with DEP they yielded identical products. At pH 8.0, the expected A > G sensitivity was observed without regional protection, whereas A hyper-reactivity was detected at pH 4.5 (Fig. 7b, lanes 1–4). Hyper-
reactivity of As is also observed at pH 4.5 and to a lesser extent at pH 7.0 when the EcoRI fragment of pGA38 is reacted with DEP (Fig. 7c, lanes 2 and 3). We conclude that DEP is detecting a feature of the alternate structure related to the conformation of A residues in the helix that is enhanced as the pH values are lowered below 7.0. An analogous hyper-reactivity of As in Z-DNA has been interpreted as reflecting the syn-conformation of these purine residues (49, 50). It is not unlikely that the A residues in S1-hypersensitive (GA), stretches are also in syn. Comparison of the results produced with the control HindIII-BglI linear molecules and the S1-hypersensitive linear or supercoiled molecules indicates that the low pH can alter the conformation or shift the equilibrium toward the alternate structure.

Are Hoogsteen Base Pairs Present in the Alternate Structure?—Although our results using chemical probes did not reproduce the observation of regional protection reported by Pulleyblank et al. (26), and the sensitivity to venom phosphodiesterase at pH 9.0 excluded C-protonation as the basis of the appearance of the alternate structure, we decided to perform an additional experiment that could show directly that Hoogsteen G-CH\(^+\) base pairs are not necessary for the appearance of the alternate structure in the (GA)\(_{38}\) stretch. We reasoned that if all the Gs in the GA strand are methylated at the N-7 position and the structure is still S1-hypersensitive, Hoogsteen base-pairing cannot be invoked as the basis of nuclease sensitivity, because, even if the N-3 of cytosine carries a proton because of the low pH of the S1 reaction, the N-7 of guanine is not available to form a hydrogen bond.

Although generation of test molecules containing 7-methyl-G exclusively in the GA stretch is experimentally very difficult, we designed an experiment that could assay for single-strand nuclease sensitivity present in the HindIII-HaeIII fragment of pGA38 (see Fig. 2c). For this purpose, we used a single-stranded M13 template carrying the CT strand (mpCT38), which had been first linearized by HindIII, following hybridization with a 20-nt-long synthetic oligonucleotide (complementary to positions 6275-6294 of M13mp18; see Fig. 8a). This new primer was purified by electrophoresis on a denaturing urea-polyacrylamide gel, and hybridized back to mpCT38, which had been first linearized by HindIII, following hybridization with a 20-nt-long synthetic oligonucleotide (complementary to positions 6275-6294 of M13mp18; see Fig. 8b). The internally labeled primer was then extended further with unlabeled dNTPs, substituting dGTP with 7-methyl-dGTP, to generate a complementary “run-off” strand. The product was then cleaved with HaeIII to release duplex DNA and bring the label closer to the 5’ end, and the generated fragment was purified by electrophoresis on a nondenaturing gel and used for the experiment shown in Fig. 8c, which shows that the G-methylated duplex still retains sensitivity to S1 nuclease (lane 4). The fact that the duplex was fully methylated was documented by its complete degradation by heating in the presence of piperidine (Fig. 8c, lane 1). We conclude that the alternate structure contains Watson-Crick GC base pairs and that its sensitivity to single-strand-specific nucleases is an intrinsic sequence-dependent property, which does not depend on C-protonation.

Probing the Alternate Structure by Primer Extension—In the past, using dimethyl sulfate as a single-strand-specific
reagent, we concluded that an 18-bp SHS present in a supercoiled circle is duplex DNA (6). However, we thought that we could not address the same question independently by performing a primer extension experiment, because we expected that supercoiled molecules might assimilate single-stranded fragments and form D-loops (52), making the results uninterpretable. Now, because of the presence of SHS in segments of linear molecules, we attempted such primer extension experiments.

We first used as template ScaI-linearized pGA38, in which the GA stretch is S1-hypersensitive. If single-stranded regions are present, it should be possible to anneal a complementary primer to either strand and extend it using a polymerase in the presence of dNTPs. Fig. 9 shows that an end-labeled 10-nt long AG or CT primer could be extended as expected, when denatured linearized pGA38 is used as a template (lanes 2 and 9). However, when the same template is used in a nondenatured duplex form, the CT primer can be extended (lane 10), but the AG primer cannot (lane 3).

The same result (extension of only the CT primer) was detected with another S1-hypersensitive nondenatured substrate: the linear EcoRI fragment of pGA38 (Fig. 9, lanes 6 and 7 and 13 and 14). In contrast, the nondenatured HindIII-
models ("slipped" DNA, single strands, and protonated structures). In addition, the data from enzymatic and chemical probing suggest that the equilibrium between alternate DNA and heteronomous B-DNA can be shifted depending on the length of (GA), or its flanks, the DNA topology, or the environmental conditions (pH and salt). Finally, our data also provide some clues for the eventual resolution of the paradoxical sensitivity of homopurine-homopyrimidine duplexes to single-strand-specific nucleases. Drew (34) has also concluded (although in a different context) that S1 nuclease and DNase I do not distinguish base-pairing arrangements, but the disposition of the sugar-phosphate backbone. According to our results, the heteronomy of the backbone conformation is possibly related to differential stacking of the two strands.

Acknowledgments—We thank C. R. Cantor for helpful discussions and critical reading of the manuscript. H. Drew, M. McCall, and C. Calladine for discussions, W. Herr for making information available before publication, and P. R. Srinivasan for a generous gift of 7-methyl-dGTP.

REFERENCES

1. Larsen, A., and Weintraub, H. (1982) Cell 29, 609–622
2. Selleck, S. B., Elgin, S. C. R., and Cartwright, I. L. (1984) J. Mol. Biol. 178, 17–33
3. Elgin, S. C. R. (1984) Nature 309, 213–214
4. Schon, E. A., Evans, T., Welch, J., and Efstratiadis, A. (1983) Cell 35, 837–848
5. Evans, T., Schon, E. A., Gora-Maslak, G., Patterson, J., and Efstratiadis, A. (1984) Nucleic Acids Res. 12, 8043–8058
6. Cantor, C. R., and Efstratiadis, A. (1984) Nucleic Acids Res. 12, 8059–8072
7. Hentschel, C. C. (1982) Nature 295, 714–716
8. Glikin, G. C., Gargiulo, G., Rena-Descalzi, L., and Worcel, A. (1983) Nature 303, 770–774
9. Mace, H. A. F., Pelham, H. R. B., and Travers, A. A. (1983) Nature 304, 555–557
10. Field, L. J., Philbrick, W. M., Howles, P. N., Dickinson, D. P., McGwen, R. A., and Gross, K. W. (1984) Mol. Biol. 4, 2321–2331
11. Huu, H. N., Lund, E., and Dahlberg, J. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 7288–7292
12. Richards, J. E., Gilliam, A. C., Shen, A., Tucker, P. W., and Blattner, F. (1983) Nature 306, 483–487
13. Gillies, S. D., Folsom, V., and Tonegawa, S. (1984) Nature 310, 594–597
14. Vieira, J., and Messing, J. (1982) Gene (Amst.) 19, 259–268
15. Hendler, S., Furer, E., and Srinivasan, P. R. (1970) Biochemistry 9, 4141–4153
16. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499–560
18. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
19. Sugden, B., DeTroy, B., Roberts, R. J., and Sambrook, J. (1975) Anal. Biochem. 68, 36–46
20. Rhodes, D., and Klug, A. (1980) Nature 286, 573–578
21. Rhodes, D., and Klug, A. (1981) Nature 292, 378–380
22. Beese, K. L., and Folk, W. R. (1977) J. Biol. Chem. 252, 3176–3184
23. Dybvig, K., Clark, C. D., Aliperti, G., and Schlesinger, M. J. (1983) Nucleic Acids Res. 11, 8495–8506
24. Han, S., Udvardy, A., and Schiedl, P. (1984) J. Mol. Biol. 179, 450–469
25. Christophe, D., Cabrer, B., Baccala, A., Targovnik, H., Pohl, V., and Vassart, G. (1985) Nucleic Acids Res. 13, 5127–5144
26. Pulleyblank, D. E., Haniford, D. B., and Morgan, A. R. (1985) Cell 42, 271–280
27. Kang, D. S., and Wells, R. D. (1985) J. Biol. Chem. 260, 7782–7790
28. Vogt, V. M. (1973) Eur. J. Biochem. 33, 192–200
29. Singleton, C. K., Klisic, J., Stridivant, S. M., and Wells, R. D. (1983) Nature 299, 312–316
30. Rich, A., Nordheim, A., and Wang, A. H.-J. (1984) Annu. Rev. Biochem. 53, 791–846
31. Margot, J. B., and Hardison, R. C. (1985) J. Mol. Biol. 184, 195–210
32. Lomono, G. P., Butler, P. J. G., and Klug, A. (1981) J. Mol. Biol. 149, 745–760
33. Galas, D., and Schmitz, A. (1978) Nucleic Acids Res. 5, 3157–3170
34. Drew, H. R. (1984) J. Mol. Biol. 176, 553–557
35. Berek, H. R., and Travers, A. A. (1984) Cell 37, 491–502
36. Fox, K. R., and Waring, M. J. (1984) Nucleic Acids Res. 12, 9271–9285
37. Behe, M., Zimmerman, S., and Felsenfeld, G. (1981) Nature 289, 232–235
38. Klein, J., Jack, A., Viswamitra, M. A., Kennard, O., Shakzd, Z., and Steitz, T. A. (1979) J. Mol. Biol. 131, 669–680
39. Arnott, S., Chandrasekaran, R., Hall, I. H., and Puigjaner, L. C. (1983) Nucleic Acids Res. 11, 4141–4155
40. Leslie, A. G. W., Arnott, S., Chandrasekaran, R., and Ratliff, R. L. (1980) J. Mol. Biol. 143, 49–72
41. Calladine, C. R. (1982) J. Mol. Biol. 161, 343–352
42. Shum, B. W.-K., and Structure, D. M. (1983) Biopolymers 22, 919–933
43. McCall, M., Brown, T., and Kennard, O. (1985) J. Mol. Biol. 183, 385–396
44. Nishimura, Y., Torogoe, C., and Teuboi, M. (1986) Nucleic Acids Res. 14, 2737–2748
45. Lyamachev, V. I., Mirkin, S. M., and Frank-Kamenetski, M. D. (1985) J. Biol. Mol. Biol. 26, 327–338
46. Lyamachev, V. I., Mirkin, S. M., and Frank-Kamenetski, M. D. (1986) J. Biol. Mol. Biol. 3, 667–669
47. Lee, J. S., Woodworth, M. L., Latimer, L. J. P., and Morgan, A. R. (1984) Nucleic Acids Res. 12, 6003–6014
48. Pritchard, A. E., Kowalski, D., and Laskowski, M., Sr. (1977) J. Biol. Chem. 252, 8652–8659
49. Johnston, B. H., and Rich, A. (1985) Cell 42, 713–724
50. Herr, W. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 8009–8013
51. Herr, W., Corbin, V., and Gilbert, W. (1985) Nucleic Acids Res. 10, 6931–6944
52. Beattie, K. L., Wiegand, R. C., and Radding, C. M. (1977) J. Mol. Biol. 116, 783–803
53. Calladine, C. R., and Drew, H. R. (1984) J. Mol. Biol. 178, 773–782