The Effect of Supercoil and Temperature on the Recognition of Palindromic and Non-palindromic Regions in φX174 Replicative Form DNA by S1 and Bal31*

Uwe R. Muller and Carole L. Wilson

From the Department of Microbiology and Immunology, East Carolina University School of Medicine, Greenville, North Carolina 27894-4554

The effect of supercoil and temperature on the topology of φX174 replicative form (RF) DNA was studied using single-strand specific endonucleases S1 and Bal31 as probes for cruciform extrusion and other structural perturbations of the B-helix. Both enzymes were found to recognize specifically and reproducibly over 30 sites, most of which were by cleavage of both enzymes independent of the superhelicity of the genome. A negative superhelical density exceeding 0.06 stabilized a transition in the DNA conformation that generated several new cleavage sites for Bal31. The underlying structures appeared to be only transiently stable and were lost from in vitro supercoiled DNA during brief incubation at 65°C. They were generally absent from in vivo supercoiled RF DNA of equal superhelicity as a consequence of the extraction and storage procedure. Mapping of the cleavage sites suggested that they were preferentially located near the begin-

This allows for the possibility that at least transiently very small cruciforms may be extruded in supercoiled genomes but are generally not recognized because they escape detection by the techniques employed.

We have sought to test this hypothesis by carefully analyzing the topology of supercoiled φX174 RF DNA, and that of a mutant strain with a synthetic palindromic insert, using two different endonuclease probes. We show herein that small

These probes are generally not recognized because they escape detection by the techniques employed.

* This work was supported in part by National Institutes of Health Grant GM31983 (to U. R. M.). We gratefully acknowledge access to BIONET through BIONET Grant 1 U41 RR-01685-03. 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: bp, base pairs; RF, replicative form; RFL, supercoiled replicative form.
cruciforms with 7 and possibly less bp in the stem can apparently be extruded since they are recognized specifically by Bal31, but not S1. However, these structures are very unstable and are lost due to a topological shift in the RF DNA that can be induced by heat or by the special conditions of the S1 reaction. In addition to these supercoil-dependent sites, a variety of other specific cleavage sites exist that are present even in relaxed DNA and are shared by both enzymes.

**Materials and Methods**

**Bacterial and Phage Strains—**Bacteriophage φX174 strain am3 is referred to herein as wild type. Strain de3 (φX174 J-F de3) has a deletion of 27 bp in the J–F intercistronic region (position 979 on the standard φX map (36)), and ins6 (φX174 J-F ins6) has an insertion of 117 bp at the same position. Both strains have been described previously (37, 38). Strain ins4930 was constructed by inserting the palindromic BamHI linker sequence CCGGATCCGG into the F-specific site of strain ins6 (39). Escherichia coli HF4714 was the permissive host (40), and E. coli C was the nonpermissive host.

**Enzymes—**Calf thymus topoisomerase I and nuclease S1 and Bal31 were purchased from Bethesda Research Laboratories (BRL), T4 polynucleotide kinase was from Pharmacia P-L Biochemicals or U. S. Biochemical Corporation, and calf intestinal alkaline phosphatase was from Boehringer-Mannheim. Restriction enzymes were obtained from either International Biotechnologies Inc., New England Biolabs, or the manufacturer or made according to the manufacturer’s specifications.

**Isolation and Purification of RF DNA—**E. coli C cells were infected with phage at a multiplicity of infection of 10, and chloramphenicol (40 μg/ml) was added 7 min after infection to prevent the shift to single-stranded DNA synthesis and thereby increase the yield of RF DNA. Infected cells were incubated for 2 h before harvesting. RF DNA was purified by two CsCl gradients in the presence of restriction enzyme from calf thymus instead of wheat germ. DNA was purified by two CsCl gradients in the presence of topoisomerase from calf thymus instead of wheat germ.

**Enzyme Assays—**Topoisomerase activity was assayed as described by Bernardi (41). The total reactions contained 50 μl of buffer (50 mM NaC1, 30 mM Tris-HCl, pH 8.1), 10 μl of 25 mM MgCl2, 1 unit of polynucleotide kinase, 1 unit of Bal31 or S1, and 5 μg of linear RF DNA. The reactions were incubated for 2 h at 37°C and terminated by adjusting the volume to 2 ml with water. Nuclease assays were performed as described by Hahn and Schowalter (42), with 5 μg of linear RF DNA, 1 μl of 25 mM MgCl2, and 1 unit of S1 in a total volume of 20 μl. The reactions were incubated for 15 min at 37°C, and the DNA was then digested with 1 unit of Bal31.

**Topoisomerase and Nuclease Assays—**Topoisomerase activity was assayed as described by Bernardi (41). The total reactions contained 50 μl of buffer (50 mM NaC1, 30 mM Tris-HCl, pH 8.1), 10 μl of 25 mM MgCl2, 1 unit of polynucleotide kinase, 1 unit of Bal31 or S1, and 5 μg of linear RF DNA. The reactions were incubated for 2 h at 37°C and terminated by adjusting the volume to 2 ml with water. Nuclease assays were performed as described by Hahn and Schowalter (42), with 5 μg of linear RF DNA, 1 μl of 25 mM MgCl2, and 1 unit of S1 in a total volume of 20 μl. The reactions were incubated for 15 min at 37°C, and the DNA was then digested with 1 unit of Bal31.

**RESULTS**

**Probing the Topology of Supercoiled φX174 RF DNA with Single-Strand-Specific Nucleases Bal31 and S1—**Previous analyses of the distribution of cruciforms in the supercoiled φX174 RF DNA using S1 nucleases have led to the conclusion that there were either no (7), or only a single major S1 cleavage site identifiable, which coincided with the 28 base-pair long palindromic in the F-G intercistronic region of this genome (14, 30). Since single-strand-specific endonuclease Bal31 was reported to cleave supercoiled ColEl DNA under more physiological conditions but with similar specificity as S1 (18), we have used this enzyme in addition to S1 to probe for the presence of cruciforms and other non-right-handed secondary structures in supercoiled wild type φX174 RF DNA. The strategy employed was to treat supercoiled DNA with these nucleases under conditions where most of the DNA was converted from the supercoiled to the linear form. The nucleases were then treated with phenol extraction to avoid “nibbling” at the ends of the linear molecules, and the number of S1- or Bal31-sensitive sites was determined by introducing into each molecule a second but unique cut with restriction endonuclease AvaII. Hence, each S1/Bal31 cleavage site should be represented by two unique fragments, one larger than half-genome size, the other smaller, unless the cleavage occurred precisely opposite from the AvaII site. Since AvaII leaves 5′-protruding single-stranded ends, these fragments could be efficiently end-labeled with [32P]ATP. Fig. 1 shows an autoradiograph of Bal31- and S1-generated fragments separated on high resolution agarose gels. The typical products of a Bal31 reaction carried out under our standard conditions were run in lane A. Approximately 30 bands with different intensities can be identified, most of which have been assigned a number that has been used consistently throughout this paper. In the original autoradiograph the bands are very sharp and their intensity and position is reproducible between different experiments and DNA preparations. Thus, Bal31 recognizes specifically, but with different frequency, at least 30 sites in in vivo supercoiled φX174 RF DNA.

An S1 analysis of this DNA is shown in lane D. In general the same bands were generated as with Bal31 but with two different significances. Band 22, a minor band only in the Bal31 digest (and generally absent), appeared as the most predominant band. Mapping of the corresponding cleavage site shows it to correspond to the “major cruciform,” as identified by Lilley (14, 30), which is located almost exactly half a genome length away from the AvaII site. S1 cleavage at that position generates, therefore, two fragments of almost equal size, which comigrate in the same position on this gel. The identity of this cleavage site was confirmed as discussed below. The second major cleavage site (6a) was also only a minor site in the Bal31 digest and will be discussed later.

While the Bal31 pattern was obtained reproducibly with different DNA preparations, the S1 reaction varied, but only with regard to the major band. This is shown in lane C, where the same S1 reaction was carried out with a different RF DNA preparation. With the exception of band 22, the typical S1 pattern was obtained.

Since it is known that the formation of at least fairly large
the DNA, except for the F-G cruciform, suggesting that it is thermodynamically the least stable structure recognized.

As will be shown below, the difference between the DNA preparations in lanes C and D was apparently a result of different superhelical densities to which extrusion of the F-G cruciform is exquisitely sensitive.

Extrusion of the F-G Cruciform Is Promoted by Low pH and S1—The striking difference between the S1 and Bal31 reaction with regard to recognition of the F-G cruciform (as well as band 6a), but not the other cleavage sites, suggested a sensitivity of this structure to the special reaction conditions, i.e. the ionic strength or the low pH of the S1 buffer. Thus, the reactions were repeated by exchanging the reaction buffers and adjusting for the reduced efficiencies of the enzymes with increased enzyme concentrations. Fig. 2A shows that under physiological supercoil S1 recognizes the F-G palindrome only at low pH but not under the Bal31 reaction conditions. Bal31 on the other hand does not cleave this palindrome under either condition. This suggests that low pH as well as the action of S1 itself are required for recognition of the F-G cruciform. Varying the conditions of the Bal31 reaction to include a temperature range of 7-37 °C did not affect the number and type of fragments obtained (data not shown), but increasing the NaCl concentration from 50 to 600 mM changed the results somewhat (Fig. 2B). In general the same number and types of fragments were obtained, but band 12 became the major cleavage site, while it was only a minor band at the lower ionic strength. Such a change in cleavage specificity in response to increased ionic strength is not unexpected as it has been observed by Kowalski (23) for other single-strand-specific enzymes.

Effect of Superhelical Density on the Bal31-Specific Cleavage of φX174 RF DNA—Extrusion of cruciforms is known to be supercoil-dependent (16, 34, 44). Hence, we have investigated the effect of changing the superhelical density on the specificity of Bal31 cleavage. The method of Singleton and Wells (41) was used to either decrease or increase the average superhelical density of purified covalently closed RF DNA preparations. This technique consists of treatment of RF DNA with topoisomerase I in the presence of ethidium bromide and subsequent removal of the enzyme and the dye with phenol. The resulting superhelical density is characteristic for the mean number of supercoils. With this method the reactions were repeated by exchanging the reaction buffers and adjusting for the reduced efficiencies of the enzymes with increased enzyme concentrations. Fig. 2A shows that under physiological supercoil S1 recognizes the F-G palindrome only at low pH but not under the Bal31 reaction conditions. Bal31 on the other hand does not cleave this palindrome under either condition. This suggests that low pH as well as the action of S1 itself are required for recognition of the F-G cruciform. Varying the conditions of the Bal31 reaction to include a temperature range of 7-37 °C did not affect the number and type of fragments obtained (data not shown), but increasing the NaCl concentration from 50 to 600 mM changed the results somewhat (Fig. 2B). In general the same number and types of fragments were obtained, but band 12 became the major cleavage site, while it was only a minor band at the lower ionic strength. Such a change in cleavage specificity in response to increased ionic strength is not unexpected as it has been observed by Kowalski (23) for other single-strand-specific enzymes.

Extrusion of the F-G Cruciform Is Promoted by Low pH and S1—The striking difference between the S1 and Bal31 reaction with regard to recognition of the F-G cruciform (as well as band 6a), but not the other cleavage sites, suggested a sensitivity of this structure to the special reaction conditions, i.e. the ionic strength or the low pH of the S1 buffer. Thus, the reactions were repeated by exchanging the reaction buffers and adjusting for the reduced efficiencies of the enzymes with increased enzyme concentrations. Fig. 2A shows that under physiological supercoil S1 recognizes the F-G palindrome only at low pH but not under the Bal31 reaction conditions. Bal31 on the other hand does not cleave this palindrome under either condition. This suggests that low pH as well as the action of S1 itself are required for recognition of the F-G cruciform. Varying the conditions of the Bal31 reaction to include a temperature range of 7-37 °C did not affect the number and type of fragments obtained (data not shown), but increasing the NaCl concentration from 50 to 600 mM changed the results somewhat (Fig. 2B). In general the same number and types of fragments were obtained, but band 12 became the major cleavage site, while it was only a minor band at the lower ionic strength. Such a change in cleavage specificity in response to increased ionic strength is not unexpected as it has been observed by Kowalski (23) for other single-strand-specific enzymes.

Effect of Superhelical Density on the Bal31-Specific Cleavage of φX174 RF DNA—Extrusion of cruciforms is known to be supercoil-dependent (16, 34, 44). Hence, we have investigated the effect of changing the superhelical density on the specificity of Bal31 cleavage. The method of Singleton and Wells (41) was used to either decrease or increase the average superhelical density of purified covalently closed RF DNA preparations. This technique consists of treatment of RF DNA with topoisomerase I in the presence of ethidium bromide and subsequent removal of the enzyme and the dye with phenol. The resulting superhelical density is characteristic for the mean number of supercoils. With this method the reactions were repeated by exchanging the reaction buffers and adjusting for the reduced efficiencies of the enzymes with increased enzyme concentrations. Fig. 2A shows that under physiological supercoil S1 recognizes the F-G palindrome only at low pH but not under the Bal31 reaction conditions. Bal31 on the other hand does not cleave this palindrome under either condition. This suggests that low pH as well as the action of S1 itself are required for recognition of the F-G cruciform. Varying the conditions of the Bal31 reaction to include a temperature range of 7-37 °C did not affect the number and type of fragments obtained (data not shown), but increasing the NaCl concentration from 50 to 600 mM changed the results somewhat (Fig. 2B). In general the same number and types of fragments were obtained, but band 12 became the major cleavage site, while it was only a minor band at the lower ionic strength. Such a change in cleavage specificity in response to increased ionic strength is not unexpected as it has been observed by Kowalski (23) for other single-strand-specific enzymes.

Extrusion of the F-G Cruciform Is Promoted by Low pH and S1—The striking difference between the S1 and Bal31 reaction with regard to recognition of the F-G cruciform (as well as band 6a), but not the other cleavage sites, suggested a sensitivity of this structure to the special reaction conditions, i.e. the ionic strength or the low pH of the S1 buffer. Thus, the reactions were repeated by exchanging the reaction buffers and adjusting for the reduced efficiencies of the enzymes with increased enzyme concentrations. Fig. 2A shows that under physiological supercoil S1 recognizes the F-G palindrome only at low pH but not under the Bal31 reaction conditions. Bal31 on the other hand does not cleave this palindrome under either condition. This suggests that low pH as well as the action of S1 itself are required for recognition of the F-G cruciform. Varying the conditions of the Bal31 reaction to include a temperature range of 7-37 °C did not affect the number and type of fragments obtained (data not shown), but increasing the NaCl concentration from 50 to 600 mM changed the results somewhat (Fig. 2B). In general the same number and types of fragments were obtained, but band 12 became the major cleavage site, while it was only a minor band at the lower ionic strength. Such a change in cleavage specificity in response to increased ionic strength is not unexpected as it has been observed by Kowalski (23) for other single-strand-specific enzymes.

Effect of Superhelical Density on the Bal31-Specific Cleavage of φX174 RF DNA—Extrusion of cruciforms is known to be supercoil-dependent (16, 34, 44). Hence, we have investigated the effect of changing the superhelical density on the specificity of Bal31 cleavage. The method of Singleton and Wells (41) was used to either decrease or increase the average superhelical density of purified covalently closed RF DNA preparations. This technique consists of treatment of RF DNA with topoisomerase I in the presence of ethidium bromide and subsequent removal of the enzyme and the dye with phenol. The resulting superhelical density is characteristic for the mean number of supercoils. With this method the reactions were repeated by exchanging the reaction buffers and adjusting for the reduced efficiencies of the enzymes with increased enzyme concentrations. Fig. 2A shows that under physiological supercoil S1 recognizes the F-G palindrome only at low pH but not under the Bal31 reaction conditions. Bal31 on the other hand does not cleave this palindrome under either condition. This suggests that low pH as well as the action of S1 itself are required for recognition of the F-G cruciform. Varying the conditions of the Bal31 reaction to include a temperature range of 7-37 °C did not affect the number and type of fragments obtained (data not shown), but increasing the NaCl concentration from 50 to 600 mM changed the results somewhat (Fig. 2B). In general the same number and types of fragments were obtained, but band 12 became the major cleavage site, while it was only a minor band at the lower ionic strength. Such a change in cleavage specificity in response to increased ionic strength is not unexpected as it has been observed by Kowalski (23) for other single-strand-specific enzymes.
linearized with AeuII prior to the Bal31 treatment the same cleavage pattern was observed as shown in lane 0 (data not shown). However, increase of $\sigma$ beyond $-0.06$ (lanes 15-30) resulted in the appearance of at least five major new recognition sites (bands 6, 9, 11, and 22). Several additional bands (15, 15a, 19, and 20) became more intense, but it is not clear whether that was due to generation of new bands of similar size or increase in intensity of existing bands. This suggests that Bal31 recognizes two types of structural perturbations in the closed circular RF DNA. One type is supercoil-independent and is cleaved with relatively low frequency, the other type is supercoil-dependent and, once formed, is recognized with high frequency, leading to very prominent bands in the gel. Formation of this type of structure requires a negative superhelical density of approximately 0.06-0.07. We have estimated this based on two other experiments, where the characteristic bands appeared after treatment of RF DNA with topoisomerase in the presence of 10 $\mu$M ethidium bromide (see Fig. 4), which generates a mixture of topoisomers with 33 $\pm$ 5 supercoils (not shown).

Bal31 but Not S1 Cleaves Small Cruciforms in Supercoiled RF DNA—The supercoil-dependence of several of the Bal31 cleavage sites, and particularly the appearance of band 22 in Fig. 3, suggests that they may be cruciforms. For mapping of the cleavage sites, a larger amount of RF DNA was treated with topoisomerase in the presence of 10 $\mu$M ethidium bromide before Bal31/AeuII cleavage and subsequent end-labeling. After electrophoresis and autoradiography (Fig. 4) the most prominent bands (6, 11, 19, and 22) were cut from the agarose gel using the autoradiograph as a guide. Both the large and corresponding small fragments were electroeluted and further restricted with either HhaI, HaeIII, or HinfI and electrophoresed on 4% polyacrylamide gels (not shown). All fragment sizes were determined graphically as described under “Materials and Methods.” The cleavage sites were thereby mapped with an error of approximately $\pm$30 bp to position 3969 (site 6), 3360 (site 11), 2588 (site 19), and 2335 (site 22) on the standard $\phi$X174 map (36). In each case the cleavage positions overlapped exactly with the location of a perfect (no G-T bp in the stem) or near perfect (G-T bp in the stem) potential hairpin (Fig. 7). However, at least for site 19 the underlying cruciform would only have a stem size of 6 bp, which is thought to be thermodynamically unstable (18). Thus, we have tested more directly whether Bal31 can recognize and cleave small palindromes. The recently constructed mutant strain ins6430 (39) contains a synthetic palindrom with the sequence GCAGCGGATCCGGCTGC, potentially capable of forming a hairpin/cruciform with 7 bp in the stem and 4 bases in the loop. A Bal31 analysis of ins6430 RF DNA and that of the parental strain ins6 is shown in Fig. 4, lanes E and D. A new band is clearly visible in lane E but not lane D, which was mapped by the above described procedure and corresponds to the inserted palindrome. This presents strong evidence that at least certain small palindromes can be extruded from the helical DNA backbone under sufficient torsional stress and can then be recognized and cleaved by Bal31.

When heating these DNAs prior to the Bal31 treatment (lanes B and F) the major bands disappeared, including the one for the inserted palindrome, resulting in the “typical” cleavage pattern shown in Fig. 1. This suggests that the secondary and tertiary structure that $\phi$X174 RF DNA assumes during supercoiling is not the thermodynamically most stable configuration. Apparently, the cruciforms corresponding to bands 6, 11, 19, and 22, as well as the inserted palindrome, are the most predominant structural features in unheated in vitro supercoiled ($\sigma = -0.06$ to $-0.07$) RF DNA, but become only minor sites, or disappear completely, after partial denaturation and renaturation. The sites that remain appear to be those that are recognized by Bal31 even in the absence of supercoil (Fig. 3).

This topological shift can also be induced in the absence of heat by the S1 reaction conditions. In Fig. 1 we have already
shown that S1 recognizes basically the same sites as Bal31 but cleaves in vivo supercoiled RF DNA predominantly at sites 22 and 6a (if sufficiently supercoiled). The same cleavage pattern was obtained with these unheated in vitro supercoiled DNAs (Fig. 4, lanes C and G). Thus, the low pH of the S1 reaction most likely causes partial denaturation of the DNA and loss of the extruded cruciforms, with the exception of the F-G cruciform (site 22), which may be stabilized by the action of S1 itself (Fig. 2).

Loss of Cruciforms during DNA Purification and Storage—The finding that in vitro supercoiled RF DNA is topologically different from in vivo supercoiled DNA of similar superhelical density raises the question whether the topoisomerase reaction introduces topological artifacts or whether such artifacts are introduced into the in vivo supercoiled DNA during the extraction. We have fortuitous evidence suggesting the latter to be true. Fig. 5 shows a Bal31 analysis of several RF DNA preparations that were made at the onset of this study. The digest of wild type RF DNA (lane A) gave a pattern very similar to that of in vitro supercoiled RF DNA (Fig. 3, lane 15), except that band 22 was very weak and band 17 fairly strong. A similar pattern was seen in the mutant digests, but no trace of the band characteristic for the inserted palindromic cruciform was observed in the ins6430 digest. Heating these DNAs (lanes B and E), or storage at 4 °C for several weeks (not shown), caused them to shift into a topological configuration whereby band 17 displaced band 19 as the major Bal31 cleavage site, band 21 emerged, bands 6, 9, 11, and 19 became minor bands, and bands 22 and 15a disappeared completely. This pattern is virtually identical to that in Fig. 1, lanes A and B, and is what we have observed in all subsequent RF DNA preparations as the typical Bal31 cleavage pattern, even though they were not exposed to temperatures above ambient throughout extraction.

We conclude that supercoiling φX174 RF DNA in vitro or in vivo causes the DNA to assume a very specific but unstable configuration, which converts into a more stable but different topological form due to thermal fluctuations. The kinetic
barriers that hinder this conversion are rapidly overcome at elevated temperatures. These different topological isomers can be differentiated with Bal31, but not with S1, since the reaction conditions (and possibly the enzyme itself) apparently lower the above-mentioned kinetic barriers and favor a topological fine structure in which site 6a and site 22 are predominantly recognized. On the other hand S1 recognizes a difference in the structure between heated and unheated DNA that can apparently not be distinguished by Bal31 (Fig. 1).

RF DNA Structure at Maximum Supercoil—The above data show that at physiological superhelical density the conformation of RF DNA with extruded cruciforms is relatively unstable and that an alternate conformation without these structures is apparently a thermodynamically more stable sink for the available supercoil energy. We have wondered whether this alternate conformation was able to absorb additional supercoil energy, or whether a superhelical density could be reached at which cruciforms remained stable structures. Hence, the superhelicity of wild type and mutant RF DNAs was increased to a maximum value of $\sigma = -0.16$ (41) before probing the structure with Bal31 and S1 (Fig. 6). As expected from the data in Fig. 3, bands 8, 15, 15a, 16, and 20 in addition to 6, 19, and 22 were the most prominent bands in the Bal31 reaction (lane A), suggesting that some of them may also result from cruciform extrusion. Apparently several of these survived the melting and renaturation process (lane B), or the destabilization effect of the S1 reaction (lane C), suggesting that sufficient supercoil energy was left for their stabilization. It is also interesting to note that the inserted palindrome as well as some other bands that were dominant near physiological supercoil (bands 9, 11, 22; Fig. 3, lane 15) became very weak or were absent at high superhelicity, apparently being replaced by other competing structures (bands 8 and 20). Thus, stabilization of a given cruciform appears to be limited not only by a minimum superhelical density, which provides the energy for extrusion, but also by a maximum superhelicity, which when exceeded may cause a conformational transition that favors other structures.

Mapping of the Major S1 and Bal31 Cleavage Sites on the $\phi X174$ Genome—The reproducibility of the size and intensity of almost all bands obtained by cleavage of in vivo supercoiled RF DNA with either Bal31 or S1 allowed a preliminary mapping of the corresponding cleavage sites on the $\phi X174$ genome. Two mutant $\phi X174$ strains, ins6 with an insertion of 117 bp and del3 with a deletion of 27 bp (37, 48) were compared to the wild type strain to aid in the mapping. The Bal31/AuAl1 or S1/AuAl1 cleavages were carried out as described for Fig. 1, and the typical cleavage pattern was ob-
tained for all strains (not shown). Any differences between these strains in the migration of bands were due to the insertion or deletion of nucleotides at position 979 on the Sanger map (36). The length of each clearly identifiable band of size greater than half-gene length (2693 bp) was determined graphically ("Materials and Methods"), and plotted clockwise and counter-clockwise from the AaII restriction site (position 5042) around the circular φX174 map. One of these sites was eliminated based on comparison of the wild type to the mutant maps. For an example, comparison of lanes A (wild type) and D (ins6) in Fig. 4 shows migration of one of the two fragments corresponding to cleavage site 19, and separation of dimer band 22, due to the insertion of 117 bp. This clearly indicates which of the fragments spans position 979.

This approach was useless, however, for fragments exceeding 4060 bp in length, and a second analysis, using the unique AaII site (position 162) instead of AaII I was necessary to assign the cleavage position of larger fragments and to confirm the location of other sites (not shown). Nevertheless, because of the multitude of bands, the relative inaccuracy of an agarose gel in the high molecular weight range (±30 bp), and configuration of some bands, not all of the cleavage sites could be mapped with certainty.

Fig. 7 shows the location of the major and several minor Bal31/S1 cleavage sites on the standard φX174 map. At least four more cleavage sites between bp 4500 and 300 have been determined with the AaII digest but could not be confirmed by the AaII digest. The locations of statistically significant palindromes (see legend) have also been included in this figure to determine whether they might be the basis for S1 or Bal31 cleavage. Clearly, several arrows point directly at or very near positions of palindromes, but the inherent inaccuracy of the mapping technique precludes an assignment with any level of confidence based on this technique alone. Yet, the dependence on supercoil for the formation of bands 6, 9, 11, 15a, 19, 20, and 22 (Fig. 3) support the theory that they are cruciforms cleaved by Bal31 in the non-base-paired loop region. At least for the F-G cruciform (site 22) this has been well established (14, 30), and theoretical calculations predict also the extrusion of the two largest perfect palindromes under sufficient torsional stress (49, but see also Ref. 34). Their positions coincide with sites 6 and 15a respectively.

Several other sites (2, 3, 7, 10, and 17b) that map near palindromes appear to be only marginally or not at all favored by supercoiling, suggesting that this overlap may be coincidental or that cruciform extrusion is not required for their recognition by the nuclease.

On the other hand it can be said that several cleavage sites (1, 4, 5, 8, 12, 14a.b, 16, 17a.c, 18, and 21) are sufficiently far removed from the indicated palindromic regions to preclude them as the basis for recognition, even though cleavage by Bal31 of some of these was promoted by supercoiling. However, several of these sites map at or very near palindromes with loop sizes between 11 and 20 bases (not shown).

Comparison of cleavage positions with the base-pair opening probability of the underlying DNA sequence shows that there is little correlation, and apparently increased breathing of DNA at certain positions in the genome is not the reason for cleavage by S1 or Bal31.

It is interesting to note that with the exception of site 11 all supercoil-dependent sites are located near the start sites or ends of genes.

**DISCUSSION**

Computer studies of the φX174 genome have revealed that it contains many more palindromic sequences than mere chance would predict (31-33), particularly in intercistronic regions, and several of these have been speculated to have a biological function (reviewed in Ref. 50). If this putative function requires formation of a cruciform structure in the RFI DNA, rather than a hairpin in mRNA or the single-stranded genomic DNA, one would assume that this structure would be present in at least one or two of the roughly 20 RF molecules in the infected cell. However, most of these palindromes are relatively small, and the idea of small cruciforms being stabilized at physiological supercoil has been generally dismissed, based on some experimental evidence (21), but mostly on theoretical grounds (51). Based on the experimental findings in this report and the following consideration, we argue that cruciforms with stem sizes of 6 to 7 bp may exist at least transiently in a fraction of the DNA molecules at physiological supercoil. Purified φX174 RFI DNA isolated from infected E. coli cells generally has a superhelical density near -0.065, which provides a substantial amount of free energy for the formation of non-right-handed secondary structures.

The free energy cost for cruciform formation has been estimated by various laboratories and was found to range between 17 and 25 kcal/mol (19, 51-54). Based on this one can estimate that the minimum stem length of a cruciform that would be stable in the φX genome must be around 9 bp (44, 55), which was also estimated to be the minimum stable cruciform size in the plasmid ColE1 (18). These calculations assume a mean supercoil of 33.5 negative superhelical turns for φX174, but it is known that a distribution of topoisomers around this mean exists. We have found that a measurable fraction of molecules have 40' and more supercoils, which would decrease the minimum stem size for stable cruciforms to 7 or 8 bp in those topoisomers (loop size assumed to be 4 bases).

Using a mutant φX174 strain from our collection with a
palindromic insert, theoretically capable of forming a cruciform with 7-bp stem size and 4 bases in the loop, we have shown that a measurable fraction of mutant RF DNA molecules were cleaved by Bal31 specifically at the site of insertion, when such molecules were supercoiled in vitro to a negative superhelical density of $\phi$X174 RF DNA genome with the technique employed here has 6 bp in the stem (sites 9 and 19, Fig. 7). This is based on our mapping studies, which place each of the supercoil-dependent cleavage sites at the location of a potential cruciform with at least 6 bp in the stem and not more than 10 bases in the loop. Cleavage of a cruciform with 6-bp stem length by a nuclease was also reported by Sheflin and Kowalski (9).

It is unlikely that all such palindromes in a genome can be extruded into cruciforms, since in addition to stem length there are clearly other factors that determine this transition. This is borne out by the finding that no consistent correlation can be made between band intensity and the stem size of the putative underlying cruciform structure. There is also a level one palindrome (position 1700) for which we have not detected any evidence of cleavage either by Bal31 or S1. We assume that the intensity of a band is directly proportional to the frequency with which a given size is cleaved by the nuclease, i.e. the frequency with which that structure occurs among the population of DNA molecules. This is known to be a function of stem and loop size, as well as sequence (27, 56).

If small cruciforms can exist at physiological supercoils, why are they generally not detected by the variety of techniques that have been employed in the analysis of DNA conformation? Our data strongly suggest that most, if not all, cruciforms in $\phi$X174 RF DNA are only transient structures. They apparently serve as temporary sinks of supercoil energy, but this energy can be stored more permanently in an alternate DNA conformation. We have shown that a substantial difference exists between the conformation of in vitro supercoiled RF DNA and that isolated from infected cells, even though both had similar superhelical densities. Heating the in vitro supercoiled DNA changed its conformation to where it gave the same Bal31 digestion pattern as unheated in vivo supercoiled DNA, while heating the latter was without effect. Thus, partial melting of the DNA allowed the transition into the stable conformation, which is devoid of cruciforms. The loss of small cruciform structures due to heating has been observed by other investigators (22). Apparently the activation energy required for this transition is very low, since in the case where the supercoil-dependent Bal31 cleavage sites were preserved during the DNA extraction, they were subsequently lost upon storage of the DNA for an extended period of time at 4 °C.

Just as important as the effect of temperature and isolation protocol on the structural transitions in DNA is the probe employed for its analysis. We have found a substantial difference between nuclease Bal31 and S1 with regard to cleavage of the supercoil-dependent sites. The cloned palindromic sequence and most other supercoil-dependent Bal31 cleavage sites were not recognized by S1 to any detectable degree, with the exception of the major palindrome (site 22). This structure was the predominant S1 cleavage site in sufficiently supercoiled RF DNA but was recognized only at low pH. The second most frequently cleaved S1 site (6a) was only a minor site in the Bal31 reaction and is apparently not a cruciform structure. We believe that the low pH at which the S1 reaction is generally performed, and possibly the enzyme itself, induces a structural transition in the DNA which destabilizes some structures ( cruciforms) and stabilizes others. Recognition sites other than cruciforms have already been described for S1 (4, 13), and differences between this enzyme and other nucleases in their specificity for alternate DNA secondary structures have been observed (9). In general our data suggest that Bal31 is the better of these two probes when testing for the presence of small cruciforms.

While differences exist between S1 and Bal31 regarding the recognition of supercoil-dependent sites, they both recognize consistently and specifically a series of sites on the $\phi$X174 genome independent on superhelical density, which appeared to be present even in linear DNA. Similar observations have been reported by Kowalski for PM2 DNA (23). Regions with a high base-pair opening probability can be excluded as candidates for recognition as our mapping studies have shown, but there are many supercoil-independent deviations from the B-helix known that are due to sequence and can be recognized by a variety of nucleases (28, 57, 58).

We have discussed our results of the supercoil-dependent Bal31 cleavage sites with a strong bias toward cruciforms as the underlying non-B secondary structure, based on our mapping data. Yet, we can not completely exclude that the overlap of cleavage sites with the location of palindromes is coincidental in at least some cases, and that the structure recognized by the nuclease is actually a Z-DNA to B-DNA junction or a slippage site. But independent on the actual conformation that leads to recognition by Bal31, it is important to note that the supercoil-dependent sites are preferentially located within intercistronic regions or near start or stop sites of genes, while the supercoil-independent sites appear to be randomly distributed in the genome. This supports the notion that these alternate secondary structures are involved in the regulation of gene expression. The labile nature of particularly small palindromes, which are highly sensitive to the superhelical density and any other factors that might redistribute the supercoil energy of a genome, seems to be an important prerequisite for a regulatory region, if it is to respond to small changes in the intracellular environment.

REFERENCES
1. Hau, M. T. (1985) Virology 143, 617–621
2. Lyamin, V. I., Panyutin, I. G., and Frank-Kamenetski, M. D. (1983) FEBS Lett. 153, 298–302
3. Lilley, D. M. J. (1983) Nucleic Acids Res. 11, 3097–3112
4. Cantor, C. R., and Lalouette, A. (1984) Nucleic Acids Res. 12, 8659–8972
5. Lilley, D. M., and Palcev. K. (1984) EMBO J. 3, 1187–1192
6. Johnston, B. H., and Rich, A. (1985) Cell 42, 713–724
7. Panayotatos, N., and Wells, R. D. (1981) Nature 289, 466–470
8. Mizuochi, K., Kenper, B., Hays, J., and Weissberg, R. A. (1982) Cell 29, 357–365
9. Sheflin, L. G., and Kowalski, D. (1985) Nucleic Acids Res. 13, 6137–6154
10. West, S. C., and Kornler, A. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 6454–6459
11. Klug, A. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 1215–1223
12. Rich, A. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 1–12
13. Pulleyblank, D. E., Hanford, D. B., and Morgan, A. R. (1985) Cell 24, 271–280
14. Lilley, D. M. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 6468–6472
15. Deleted in proof
16. Mizuochi, K., Mizuochi, M., and Gelert, M. (1982) J. Mol. Biol. 156, 229–243
17. Wang, J. C. (1984) J. Cell. Sci. (Suppl.) 1, 21–29
18. Lilley, D. M., and Hallam, L. R. (1984) J. Mol. Biol. 180, 179–200
19. Courrey, A. J., and Wang, J. C. (1983) Cell 33, 817–829
20. Sinden, R. R., Browne, S. S., and Pettijohn, D. E. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1797–1801
21. Singleton, C. K. (1983) *J. Biol. Chem.* **258**, 7661–7668
22. Sinden, R. R., and Pettijohn, D. E. (1984) *J. Biol. Chem.* **259**, 6593–6600
23. Kowalski, D. (1984) *Nucleic Acids Res.* **12**, 7071–7086
24. Scheffin, L. G., and Kowalski, D. (1984) *Nucleic Acids Res.* **12**, 7087–7104
25. Caffarelli, E., Franzini, C., Leoni, L., and Savino, M. (1984) *Cell.* **35**, 225–231
26. Lilley, D. M., and Markham, A. F. (1983) *EMBO J.* **2**, 527–533
27. Lilley, D. M. (1985) *Nucleic Acids Res.* **13**, 1443–1465
28. Nickol, J. M., and Felsenfeld, G. (1983) *Cell* **35**, 467–477
29. Muller, U. R., and Turnage, M. A. (1986) *J. Mol. Biol.* **189**, 285–297
30. Muller, U. R., Goodman, T. C., Hillen, W., Horn, G. T., Klein, R. D., Larsson, J. E., Muller, U. R., Neuendorf, S. K., Panayotatos, N., and Stirdivant, S. H. (1980) *Progr. Nucleic Acid Res. Mol. Biol.* **24**, 167–267
31. Hsieh, T., and Wang, J. C. (1975) *Biochemistry* **14**, 527–535
32. Gellert, M., O’Dea, M. H., and Mizuuchi, K. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5545–5549
33. Singleton, C. K., and Wells, R. D. (1982) *J. Biol. Chem.* **257**, 6593–6600
34. Vogt, V. M. (1973) *Eur. J. Biochem.* **33**, 192–200
35. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) in *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
36. Keller, W. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 4876–4880
37. Shure, M., Pulleyblank, D. E., and Vinograd, J. (1977) *Nucleic Acids Res.* **4**, 1183–1205
38. Depew, R. E., and Wang, J. C. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 4275–4279
39. Muller, U. R., and Wells, R. D. (1984) *J. Mol. Biol.* **141**, 1–24
40. Benbow, R. M., Hutchison, C. A., III, Fabricant, J. D., and Sinsheimer, R. L. (1971) *J. Virol.* **7**, 549–558
41. Singleton, C. K., and Wells, R. D. (1982) *Anal. Biochem.* **122**, 253–257
42. Wang, J. C., Peck, L. J., and Becherer, K. (1982) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 85–91