1594–1601 *Nucleic Acids Research, 1996, Vol. 24, No. 9*

The structure of 4-way DNA junctions: specific binding of bis-intercalators with rigid linkers

Mark L. Carpenter, Gordon Lowe[1] and Peter R. Cook*

The Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK and [1]Dyson Perrins Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QY, UK

Received February 19, 1996; Revised and Accepted March 21, 1996

ABSTRACT

During replication and recombination, two DNA duplexes li.e. side by side. We have developed reagents that might be used to probe structure during these critical processes; they contain two intercalating groups connected by a rigid linker that forces those groups to point in opposite directions. If their stereochemistry proves appropriate, such structure-specific agents should intercalate specifically into adjacent duplexes in the Y- and X-shaped structures (i.e. 3- and 4-way junctions, now known as 3H and 4H junctions) found at replication and recombination sites. We prepared DNA structures in which four duplexes were arranged in all possible combinations around 2- and 4-way junctions and then probed the accessibility to DNase I of all their phosphodi.e.ster bonds. In the absence of any bis-intercalators, 7–9 nucleotides (nt) in each of the strands in 4-way junctions were protected from attack; protected regions were significantly offset to the 3′ side of the junction in continuous strands, but only slightly offset, if at all, in exchanging strands. All the intercalators decreased accessibility throughout the structure, but none did so at specific points in the two adjacent arms of 4-way junctions. However, one bis-intercalator—but not its sister with a shorter linker—strikingly increased access to a particular CpT bond that lay 9 nt away from the centre of some 4-way junctions without reducing access to neighbouring bonds. Binding was both sequence and structure specific, and depended on complementary stereochemistry between bis-intercalator and junction.

INTRODUCTION

During many critical moments in the life of DNA, two duplexes li.e. side by side at, for example, the replication fork or during recombination and topoisomerase action. We have developed reagents that might be used to probe DNA structure during these critical stages (Fig. 1A; refs 1,2). They contain two binding moi.e.ti.e.s—intercalating groups—connected by a rigid linker that forces the binding moi.e.ti.e.s to point in opposite directions. As a result, they should intercalate into adjacent duplexes that li.e. the appropriate distance apart.

Such bis-intercalators with rigid linkers would be expected to have different binding properti.e.s from their counterparts with flexible linkers, many of which have been studied previously (3–8). Although flexible linkers will, in principle, permit binding to adjacent duplexes, the resulting complex is energetically less stable than that formed when the intercalating groups bind to the same duplex; once one intercalating group has bound, the other is more likely to bind to a neighbouring site on the same duplex rather than to one on another duplex (Fig. 1B, top). The antitumour antibiotic, luzopeptin A, provides the prototypic example of this behaviour; the inter-molecular cross-links seen in crystals revert to intra-molecular cross-links in solution (6,8–10). On the other hand, a bis-intercalator with a rigid linker that points the two binding groups in different directions might bind stably to two different duplexes, unless, of course, the same duplex can fold back on itself. As electrostatic repulsion between two duplexes is so strong, such cross-linking is more likely to occur where the two duplexes are forced to li.e. close together, for example in the X- and Y-shaped structures (i.e. 4- and 3-way junctions, now known as 4H and 3H junctions; 11) found at sites of recombination and replication (Fig. 1B, bottom).

We have previously shown that all the compounds illustrated in Figure 1A unwind DNA in the manner characteristic of intercalators, but only the bifunctional compounds (i.e. **1**, **2**, **3** and **4**) can knot and catenate DNA, which suggests that only they can cross-link different duplexes, a property expected of bis-intercalators (1,2). However, the degree of cross-linking was small, which is to be expected since the entropic factor involved in bringing two DNA molecules together is so large. They should be more effective cross-linking agents where two duplexes are forced to li.e. in close proximity, for example in X- and Y-shaped structures. Although formal proof of bis-intercalation requires the demonstration—probably by X-ray crystallography or NMR spectroscopy (8,12,13)—that both groups intercalate simultaneously, we have now compared the binding of these compounds to different junctions containing the same sequences.

DNase I is used to monitor structure both at the junction and the site of binding; regions close to the junction or obscured by bound intercalator are protected from attack. We soon found that although all the compounds decreased accessibility throughout the structure, none reduced access at specific points on adjacent arms of 3- and 4-way junctions in the manner expected of a bis-intercalator. To our surprise, we found that one, **2** (APEPA), specifically increased accessibility at two sites that were symmetrically arranged about

* To whom correspondence should be addressed

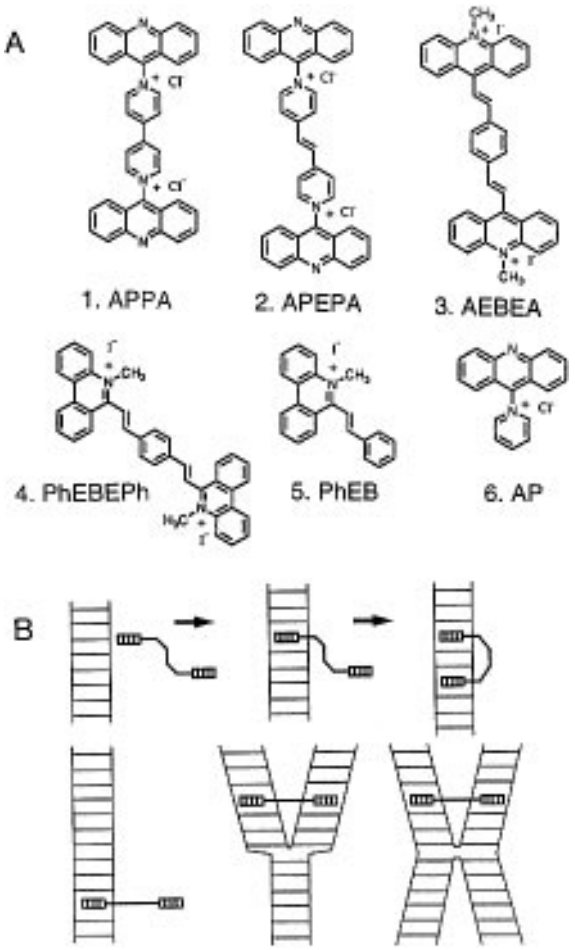

Figure 1. Compounds and their mode of intercalation. (**A**) Structures of the various compounds studied. Compounds are named by their constituent parts, where intercalating groups are A (acridinium) or Ph (phenanthridinium), and linkers contain P (pyridinium), E (ethylene) and B (benzene) groups. Thus, **4** (PhEBEPh) contains two phenanthridine groups connected by an ethylene-benzene-ethylene linker. (**B**, **top**). Stages in the binding of a bis-intercalator with a flexible linker to a duplex. After one intercalating group binds, the linker allows the second to bind to neighbouring sites which are inevitably present; only rarely will the compound bind to two adjacent duplexes. (**B**, **bottom**). A rigid linker prevents the two intercalating groups of a bis-intercalator from binding to the same local region of a duplex, but will allow the two groups to bind to Y- and X-shaped structures. Linker length and shape should dictate exactly where such binding occurs in branched DNA, and then regions around the binding sites on the two adjacent—and now cross-linked—arms should be protected from attack by DNase I.

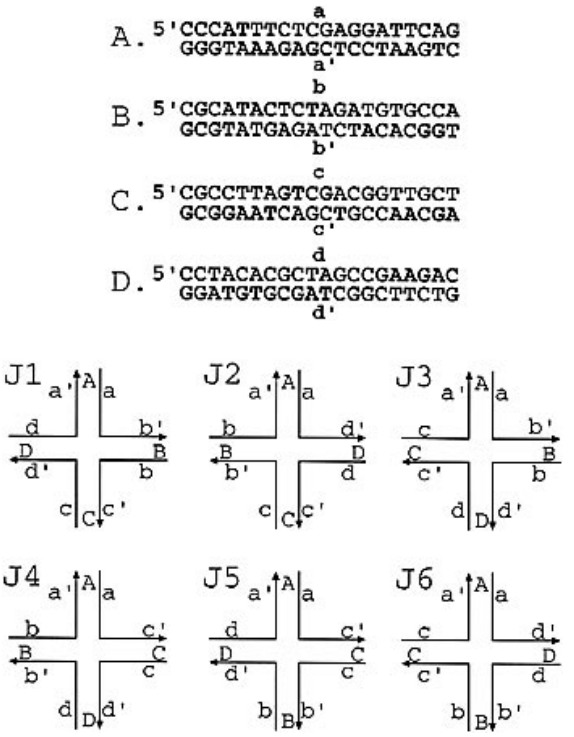

Figure 2. Sequences used to make different 4-way junctions. Four duplexes (A, B, C and D), containing strands a and a′, b and b′, c and c′ and d and d′, respectively, form four arms of the 4-way junction, J1. We adopt the following conventions: (i) each (duplex) arm of a junction is labelled with a capital letter, A–D, (ii) each strand forms two arms and is named by the two lower-case letters (a–d) that make corresponding arms, (iii) the second lower-case letter li.e.s 3′ to the first and is indicated by ′ (eg. ab′) and (iv) an arrowhead designates the 3′ end of each strand. The four arms in J1 can be arranged in five other ways around the junction to give J2–J6. J1–J6 are made by synthesizing 12 strands (i.e. ab′, ac′, ad′, ba′, bc′, bd′, etc.) and annealing sets of four together. All possible arrangements of the same arms in duplexes (with two arms) and Y-shaped structures (with three arms) are made by annealing appropriate sets of two or three of the same 12 strands.

the centre of the first 4-way junction tested. This led us to analyze binding to all possible 4-way junctions that could be formed by rearranging the four arms about the original junction; binding was both sequence and structure specific, and depended on complementary stereochemistry between bis-intercalator and junction.

MATERIALS AND METHODS

Construction of junctions and duplexes

Oligonucleotides were synthesized (applied Biosystems Model 394 DNA synthesizer) using β-cyanoethyl-phosphoramidites, deprotected, purifi.e.d by electrophoresis (10% polyacrylamide gels containing 8 M urea), bands excised, DNA eluted into TE (10 mM Tris, 0.1 mM EDTA, pH 8.0), ethanol precipitated and redissolved in TE. 5′ ends were labelled using [γ-^{32}P]ATP and T4 polynucleotide kinase (New England Biolabs; reactions terminated by heating to 65°C for 10 min) before DNA was precipitated with ethanol.

For each 3- (or 4-way) junction, three (or four) 40mer oligonucleotides were synthesized (see legend to Fig. 2). Each 40mer contained the top strand of one arm joined to the bottom strand of another. Strands complementary to each strand in the junction were also synthesized for duplex construction. Junctions and duplexes containing one labelled strand were formed by annealing (30 min; 85°C) the appropriate strands (1 mol labelled strand: 2 mol each unlabelled strand) and then slowly cooling them to room temperature over several hours. The junctions were purifi.e.d using non-denaturing polyacrylamide gels (8%; 20 h; 180 V at 4°C or 100 V at room temperature), labelled bands excised, eluted into TE buffer, precipitated, redissolved in TE and stored frozen.

DNase I footprinting

DNase I (Sigma; stock solution 7200 U/ml in 0.15 M NaCl, 1 mM $MgCl_2$) was diluted to 0.03 U/ml in 2 mM $MnCl_2$, 2 mM $MgCl_2$,

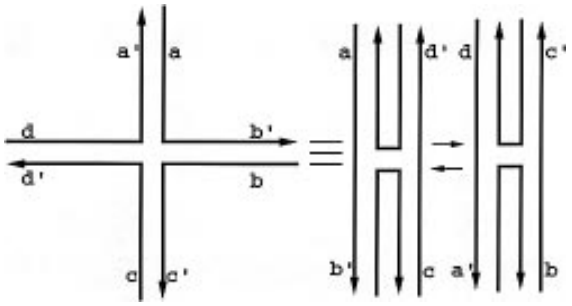

Figure 3. Two of the four isomers of the 4-way junction, J1. Four strands (i.e. ab′, bc′, cd′ and da′) make J1 (left), which can form four isomeric X-shaped structures by the pairwise coaxial stacking of helical arms. In one form (middle), strands ab′ and cd′ form effectively continuous helices, whereas strands da′ and bc′—which are not labelled here—exchange between helices at the junction; in another (right), strands da′ and bc′ are continuous, whereas ab′ and cd′—which, again, are not labelled here—exchange. In both, continuous strands are anti-parallel. Two other forms, with parallel continuous strands, are less stable and are not shown.

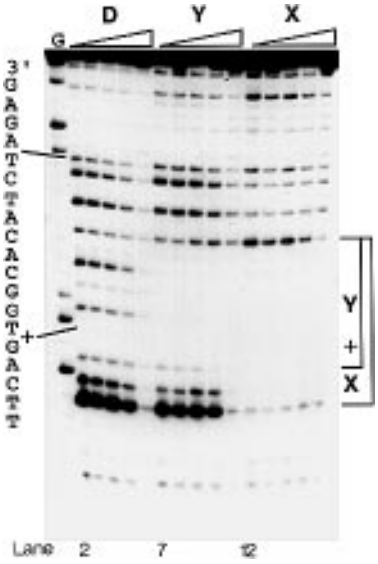

Figure 4. Probing structure with DNase I: the nuclease cuts less effici.e.ntly close to the junction in Y- or X-shaped structures, and is inhibited by high concentrations of **3**. One strand (ab′) was labelled at its 5′ end and used to prepare a duplex (with strand ba′), a Y-shaped structure (with strands ca′ and bc′) and an X-shaped structure, J1 (with strands bc′, cd′ and da′). Different structures were incubated with 0, 1, 5, 25 or 125 µM **3** (lanes 2–6, 7–11 and 12–16, respectively), treated with DNase I, the resulting fragments run on a denaturing gel and an autoradiograph prepared. Fragments produced by treating the duplex with the Maxam–Gilbert G-reaction were loaded on lane 1 (G); they have a slightly greater mobility than those produced by DNase I. The relevant part of the sequence of the labelled strand ab′ is shown on the left (+ marks the point that separates the two arms at the centre of the junction). Brackets on the right indicate regions protected from the nuclease in Y- and X-junctions.

20 mM NaCl immediately before use. Approximately 1 pmol junction in 2 µl TE was incubated (15 min; room temperature) with 2 µl compound in 10 mM Tris, 10 mM NaCl, 5 mM $MgCl_2$ (pH 8.0), 2 µl DNase I added, and the reaction stopped after 4 min by addition of 5 µl 80% formamide containing 10 mM EDTA, 0.1% bromophenol blue. Samples were denatured (100°C; 3 min) before electrophoresis (2 h) through 14% polyacrylamide sequencing gels containing 8 M urea, before gels were fixed (10% acetic acid) and transferred to Whatman 3MM paper, dri.e.d (80°C under vacuum) and autoradiographs prepared at –70°C.

RESULTS

Structures of 4-way junctions

The 4-way junction, J1, contains four double-stranded arms (A–D) arranged around the centre; five other 4-way junctions have the same arms arranged in different ways (Fig. 2; bottom). Duplexes (with two arms) and Y-shaped structures (with three arms) complete a set of structures that have the same arms arranged around a central junction. Such structures are made by synthesizing strands with the appropriate sequences and annealing them together in pairs (to give duplexes), threes (3-way junctions) or fours (4-way junctions). Note that the branch points in the 4-way junctions involved in recombination can normally translate along the double-helical arms; in the structures studied here, sequence asymmetry prevents branch migration and so resolution of the complexes into duplexes.

A 4-way junction like J1 can, in principle, form four isomeric X-shaped structures by the pairwise coaxial stacking of its helical arms (revi.e.wed in refs 14,15; Fig. 3). For example, in one form of J1, strands ab′ and cd′ run continuously along a helical axis, whereas strands da′ and bc′ exchange between helices at the junction (Fig. 3, middle). In a second form, strands da′ and bc′ are continuous, whereas ab′ and cd′ exchange (Fig. 3, right). In both forms, continuous strands are anti-parallel. Two other forms (not shown) having parallel continuous strands are theoretically possible but have not been detected in solution; they will not be considered further.

The sequence at the point of strand exchange governs whether arm A stacks on arm B or D, and whether arm C stacks on arm D or B (Fig. 3, middle and right). DNase I can be used to determine which is the preferred structure (16,17). The pair of stacked helical arms can rotate about each other like the blades of a pair of scissors and, as in a pair of scissors, the two sides are dissimilar and so present different aspects to the nuclease. Regions of all four strands at the point of strand exchange are inaccessible, and on continuous strands, but not on exchanging strands, the centre of the inaccessible region is to the 3′ side of the exchange point. Inaccessible regions also extend further away on exchanging strands.

Mapping regions protected from attack by DNase I

The structures of 2-, 3- and 4-way junctions were probed with DNase I (Fig. 4). One strand (ab′) was labelled with ^{32}P at its 5′ end and incorporated into each of the three structures, which were then incubated (without any intercalator) with DNase I, using suffici.e.nt for ‘single-hit’ cleavage; then the resulting fragments were denatured, run on a gel and an autoradiograph prepared. The cleavage pattern given by the duplex reflects the sequence dependence of DNase I scission, giving a ladder of bands of differing intensiti.e.s (Fig. 4, lane 2). [DNase I interacts with the minor groove of the target duplex over about one full turn of the double helix (18; see refs 19–21 for a discussion of the rules governing such cutting).] Any differences from the duplex pattern given by 3- and 4-way junctions then indicates where access of the enzyme is inhibited or enhanced. A region on the 3′ side of the exchange point in the Y-shaped structure (Fig. 4, lane 7), and a larger region on both sides in the X-shaped structure (Fig. 4,

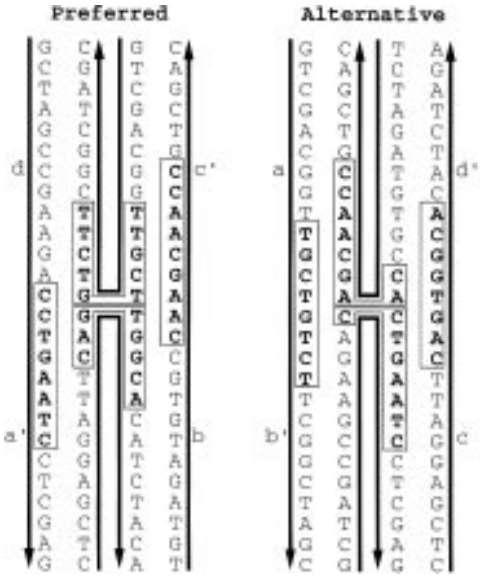

Figure 5. Sequences in J1 protected from DNase I attack. Sequences at junctions of two isomers ('preferred' and 'alternative') with anti-parallel continuous strands are shown. Nucleotides shown in bold were protected from nuclease attack. The boxed region contains both the protected sequence in strand ab′ illustrated in Figure 4, plus those in the three other strands (determined using each end-labelled strand in turn). The most stable isomer ('preferred') was judged to be the one in which protected sequences in both continuous strands were most offset 3′ to the centre of the junction.

Figure 6. Sequences at exchange points in the 'preferred' and 'alternative' isomers of J1–J6. DNase I was used to distinguish 'preferred' from 'alternative' structures, as exemplifi.e.d for J1 in Figures 3–5. Areas protected from attack by DNase I are also shown boxed in Figure 10.

lane 12) were protected from cutting. These results are consistent with those obtained previously (16,17).

The accessibility of the other three strands in J1 was then determined in three separate experiments like those in Figure 4; the results are summarized in Figure 5. The sequences of the two stable isomers with the anti-parallel arrangement of continuous strands are illustrated; regions inaccessible to the nuclease li.e. within the boxes. The rules established previously—from a limited number of examples—could not be applied unambiguously to decide which was the 'preferred' isomer (16,17). However, one rule—that the 'preferred' form always contained protected sequences in both continuous strands which were offset 3′ from the centre of the junction—could be applied here, and to all the other cases studied (see below). But then a second rule—that exchanging strands were more protected—could not apply in this case (or in many others; see below). As this second rule was also not applicable to all the original cases, we concluded that it was superfluous.

The accessibility of all four strands in each of the other 4-way junctions, J2–J6, was determined similarly (see also Fig. 10); the results for all six junctions are summarized in Figure 6, where the 'preferred' arrangement is selected using the first rule.

Intercalators have complex effects on DNase I accessibility

Having probed structure in the absence of any intercalator, we next screened the various compounds to see if any specifically prevented access. Figure 4 illustrates how **3** (AEBEA) affects the cutting of strand ab′ in 2-, 3- and 4-way junctions. Increasing concentrations progressively decrease the intensity of all bands given by the duplex (Fig. 4, lanes 3–6), the Y- (lanes 8–11) and X-shaped structures (lanes 13–16); all regions seem to bind this bis-intercalator equally, preventing cutting. This bis-intercalator does not have the appropriate stereochemistry to bind specifically at any site.

Such an experiment was then repeated with each compound in turn (not shown), allowing two general observations to be made: (i) on a molar basis, bis-intercalators tended to reduce access to all types of junction more than mono-intercalators, and (ii) lower concentrations of bis-intercalators were required to abolish cutting in 4-way junctions (compared with 2-way junctions) despite a doubling of the number of binding sites. None of the compounds had any specific effects on the cleavage pattern of duplex or 3-way junctions, but some clearly affected cleavage at specific points in strand ab′ in J1 (Fig. 7). In the absence of any compound, the protection region around the junction is visible in the centre of the autoradiogram (control lanes, C). The bis-intercalator **1** (APPA) had little effect at any concentration, but its sister with a longer linker (**2**) made one site particularly hypersensitive to cutting, especially at 5 and 25 µM (marked by *); presumably, binding so distorts the structure that one particular phosphodi.e.ster bond that is 9 nucleotides (nt) away from the junction becomes more accessible. This hypersensitivity was associated with protection of some bonds 2–3 nt away on both sides. These effects were not found when the same strand is incorporated into a duplex or a Y-shaped structure (e.g. Figs 8 and 9). As shown previously, **3** non-specifically reduces cutting. Five µM **4** (PhEBEPh), but not its mono-intercalating cousin **5** (PhEB), generates hypersensitive sites within the protected region close to the junction (closed triangles); again, these effects are not found with the corresponding duplex or Y-shaped structure (not shown). These results suggest that **2** and **4** bind specifically where the stereochemistry of the X-shaped structure

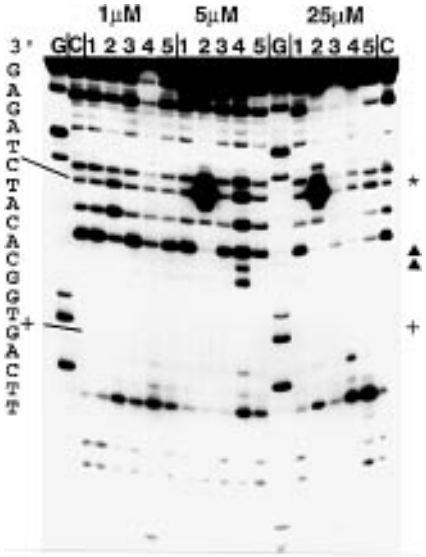

Figure 7. The effects of various intercalating agents on the accessibility of strand ab′ in J1 to DNase I. J1 was prepared using end-labelled strand ab′, incubated without (C) or with 1, 5 or 25 μM **1–5**, treated with DNase I, the resulting fragments run on a denaturing gel and an autoradiograph prepared. Fragments (G) produced by treating a duplex (containing the same end-labelled strand) with the Maxam–Gilbert G-reaction were also run; they have a slightly greater mobility than those produced by DNase I. The relevant sequence of the labelled strand is shown on the left. +, Centre of junction. *, Hypersensitive site induced by 5 and 25 μM **2**. Filled triangles, hypersensitive sites induced by 5 μM **4**; this track is overloaded on this particular gel, but these hypersensitive sites were also seen on other gels (not shown).

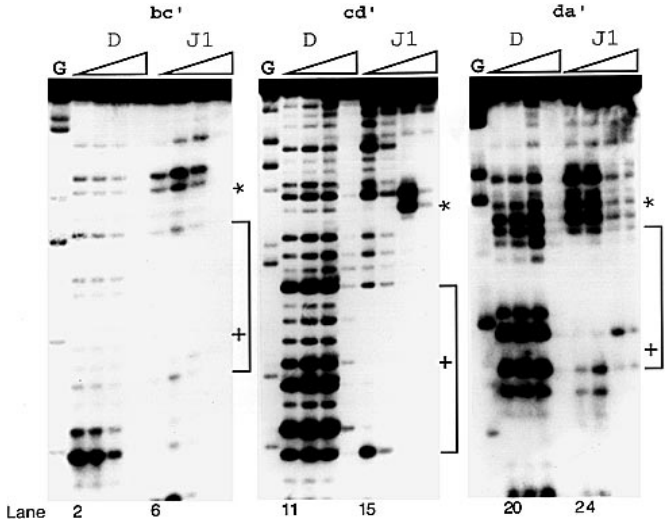

Figure 8. Structure **2** generates a hypersensitive site in strand cd′ (arm D), but not in strands bc′ (arm C) and da′ (arm A) of J1. Four end-labelled strands (ab′, bc′, cd′ and da′) were used to prepare forms of J1 which differed only in which strand was labelled, as well as the four corresponding duplexes (D). Results from three of these pairs are illustrated here; see Figure 7 for the fourth which illustrates the hypersensitive site in strand ab′ (arm B). The different structures were incubated with 0, 1, 5 or 25 μM **2** (increasing concentrations indicated by triangles), treated with DNase I, the resulting fragments run on a denaturing gel and an autoradiograph prepared. Fragments produced by treating the duplex with the Maxam–Gilbert G-reaction were run in lanes 1, 10 and 19 (G). *, Equivalent position of the hypersensitive site seen in Figure 7; hypersensitivity is seen only in lanes 17 and 18 (i.e. on strand cd′ on arm D). Brackets indicate regions protected in the X-junctions, and + the centre of the junction.

is appropriate. [Note that all these compounds produce other subtle changes but only the most striking are considered here.]

Structure-specific binding of 2

As **2** generated a prominent hypersensitive site in strand ab′ of J1, we investigated whether it also did so in the other three strands; in each case, the same strands in duplexes served as controls (Fig. 8). Increasing concentrations progressively blocked access to strands bc′ and da′, whether they were in duplexes or X-shaped structures (lanes 2–9 and 20–27). The prominent hypersensitive site seen earli.e.r in strand ab′ (its position relative to the junction is marked by *) was also found in the equivalent, symmetrically opposite, position in strand cd′ in the X-shaped structure (lane 17), but not in the other two (continuous) strands. Again, hypersensitivity was associated with protection 2–3 nt away on both sides, but, importantly, similar regions on the adjacent arms were not protected from attack.

We went on to determine whether the prominent site (and protected regions) in strand cd′ were also present in the equivalent position (on different strands) on one arm of J2–J6 (Fig. 9). In this seri.e.s, the labelled strands in the duplex and in six different junctions shared a common 3′ sequence (i.e. d′), and the hypersensitive site (marked by *) within it is only seen in some contexts (i.e. in J1, J2, J4, and weakly in J5) but not in others (i.e. J3 and J6). Hypersensitivity was always associated with protection of some bonds ≥2 nt away on each side. [Note that in each case, the appropriate duplex was included for comparison, but only one is shown in Figure 9.]

Finally, we completed the analysis by monitoring hypersensitivity in all arms of all junctions. The results of such a complete analysis, which involves end-labelling 12 different strands and incorporating them into 12 different duplexes and six different junctions, are summarized in Figure 10. The hypersensitive site (marked by *) was found only at a CpT, and only in some contexts but not others. Although hypersensitivity was always associated with protection of flanking regions on the same strand, no protection of equivalent positions on the adjacent arm was ever seen.

Specific binding of other compounds

We have concentrated on one striking hypersensitive site induced by one bis-intercalator (and not by its mono-intercalating parent or its bis-intercalating sisters) in some, but not all, X-shaped structures (and not in duplexes or Y-shaped structures). We also found (using the same approach) that each of the compounds studied induced its own characteristic pattern of resistance/sensitivity in the different structures; some are highlighted in Figure 7 and the most striking, which tend to be found in the protected region around the junction, are summarized in Figure 10.

DISCUSSION

During replication and recombination, two duplexes li.e. side by side. We have developed various reagents that might be used to probe structure during these critical processes (Fig. 1A). They contain two intercalating groups connected by a rigid linker that forces those groups to point in opposite directions. If their stereochemistry proves appropriate, such agents should interca-

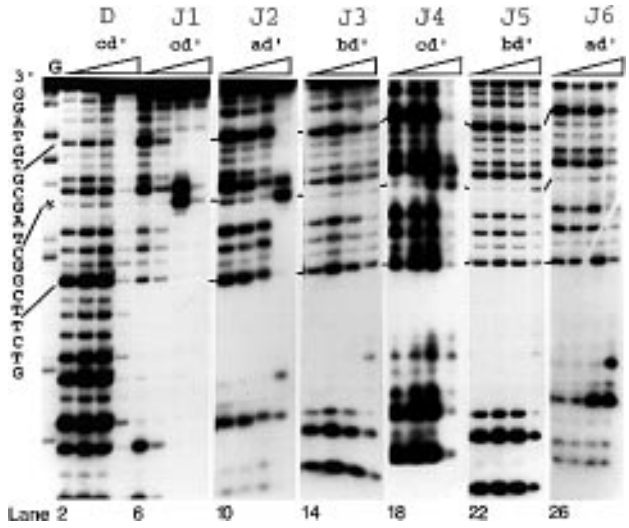

Figure 9. Structure **2** generates hypersensitive sites on an arm when it is in some contexts but not others. Three strands, all with the same 3′ half (i.e. cd′, ad′, bd′) were end-labelled and incorporated into J1–J6 (see Fig. 2); one (i.e. cd′) was incorporated into a duplex (D). The different structures were incubated with 0, 1, 5 or 25 μM **2** (increasing concentrations indicated by triangles), treated with DNase I, the resulting fragments run on a denaturing gel and an autoradiograph prepared. Fragments produced by treating the duplex with the Maxam–Gilbert G-reaction were run in lane 1 (G). The common sequence of the 3′ end of the labelled strand is shown on the left. *, Position of hypersensitive sites induced by **2** in J1, J2, J4 and J5.

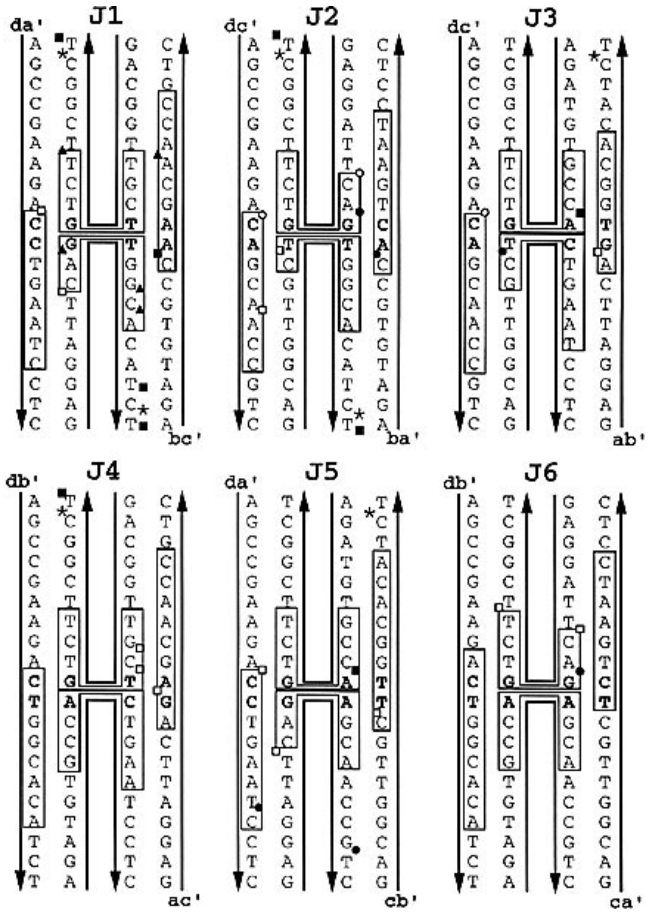

Figure 10. Hypersensitive sites induced by various intercalators in J1–J6; only sequences around the junctions are depicted. DNase I was used to distinguish the 'preferred' from the 'alternative' structure of J1 (Figs 3–5); this involved a 'complete' analysis in which each of the four strands was end-labelled in turn. Such a 'complete' analysis was then repeated for all possible junctions that can be made by rearranging the four arms of J1 to give J2–J6 (Figs 2 and 9). The boxed regions contain protected sequences detected as in Figure 5. The 8 nt at the centre of the junction are shown in bold. *, Strong hypersensitive sites found only with **2**. Filled squares, triangles and circles: hypersensitive sites found only with **2**, only with **4**, or with both **2** and **4**, respectively. Open squares and circles, weak hypersensitive sites found with both bis- and mono-intercalators, or only with mono-intercalators, respectively.

late specifically into adjacent duplexes in the 3- and 4-way junctions found at replication and recombination sites (Fig. 1B). They should be structure-specific, rather than sequence-specific, reagents. We have previously shown that these compounds can unwind, knot and catenate DNA in the manner expected of agents that can intercalate into adjacent duplexes, so cross-linking them (1,2); we have now studied their binding to synthetic replication and recombination sites. We prepared DNA structures in which four duplexes were arranged in all possible combinations around 2- and 4-way junctions and then probed the accessibility to DNase I of all their phosphodi.e.ster bonds. Such a complete analysis has not been carri.e.d out previously on any junction. We first established the accessibility of the nuclease to the junctions in the absence of any intercalators.

Inaccessible regions in 3- and 4-way junctions

DNase I has been widely used to probe the structure of 3- and 4-way junctions; regions on the 3′ side of junctions in Y-shaped structures, and larger regions on both sides in X-shaped structures are protected from nucleolytic attack (16,17; revi.e.wed in refs 14,15). Our results are entirely consistent with these general conclusions (e.g. Fig. 4, lanes 7 and 12).

In principle, a 4-way junction can form four isomeric X-shaped structures by the pairwise coaxial stacking of its helical arms but, in practice, only one is found (Fig. 3; revi.e.wed in ref. 14). Various empirical rules are being established—as yet based on only a limited number of cases—to enable us to predict which form is the most stable. The most important rule concerns strand polarity: the two isomers which have parallel continuous strands must be energetically unstable as they are not seen in solution. DNase I can then be used to determine which of the other two (with the anti-parallel arrangement) is the preferred structure, using another set of rules (16,17): the preferred isomer contains (i) protected sequences in both continuous strands which are offset 3′ from the centre of the junction and (ii) exchanging strands which are more protected than the others. Although all four strands at the centre of our six junctions were inaccessible to the nuclease, two were not invariably more protected (Fig. 10). As inspection of the original results showed that this second 'rule' did not always hold, and as cutting by DNase I is known to depend in a complex way on sequence (18–21), we chose the preferred isomer using only the first rule. Then all our junctions could be arranged in a strikingly similar way (Fig. 10). The 7–9 protected nucleotides in both continuous strands were offset 3′ from the centre of the junction so that there were only 1–2 protected nucleotides on the 5′ side. In addition, the 7–9 protected nucleotides in the

exchanging strands were symmetrically arranged around the junction, or again offset slightly towards the 3′ side.

This asymmetric cutting can be explained by the asymmetry of the enzyme (16,22,23); although it can cut immediately 5′ to a protected region—whether created by drug-binding or the centre of a 4-way junction—2–3 bonds on the 3′ side are immune from attack. (Note that the prominent hypersensitive site induced by **2** in J1 is flanked on both sides by other sensitive sites; therefore it is unlikely that a drug molecule can be bound immediately 3′ or 2–3 nt 5′ to this hypersensitive site.)

The bases at the point of strand exchange seem to determine which isomer is the most stable, as certain base-pairs prefer to stack on others (24–30). Of the 128 possible combinations, only a few have been analyzed to date. Four of the combinations studied had the same eight bases at the junction as others studied previously, and in each case there is agreement as to which is the 'preferred' arrangement (Fig. 6; J3–J6 are equivalent to $J1^c$, J1, L1 and L3 described in ref. 31). These eight at the junction are probably the sole determinants of the 'preferred' structure, as the adjacent bases in J3–J6 were different from these others.

Binding of bis-intercalators

None of our compounds showed any sequence-specific binding to duplex DNA; this is expected as related acridine and ethidium derivatives lack such specificity under our conditions (22). We expected that structure-specific binding of a bis-intercalator to a 3- or 4-way junction would reduce access of DNase I at specific points on adjacent arms (Fig. 1B, bottom). Although all our intercalators non-specifically inhibited access of the nuclease, none of the bis-intercalators acted specifically in the expected manner; rather, our most surprising result is that one agent (**2**) increased access to one particular CpT bond in some 4-way junctions. Although this hypersensitivity was associated with protection of some bonds ≥2 nt away on the same strand, the equivalent region on the adjacent arm was not protected. This induction of hypersensitivity (and protection) in a region that was generally accessible to DNase I was highly specific. It was seen only with one bis-intercalator, but not with a sister compound that had the same intercalating groups and a shorter linker (i.e. **1**), or with a mono-intercalating cousin (i.e. **6**, AP). Hypersensitive sites were only found in 4-way junctions, but not in 2- or 3-way junctions containing the same sequences. All strands studied contained at least one CpT, but only two (i.e. b′ and d′) contained a CpT 9 nt away from the centre of the junction, and it was only these two that were hypersensitive in some 4-way junctions. (Strand a′ contains an unaffected CpT 8 nt from the junction.) Both bonds were hypersensitive in J1 and J2 where they were symmetrically arranged on opposite helical stacks; one or other bond was hypersensitive in J3 and J4 (and weakly in J5), and neither in J6. Clearly, binding of **2** is both sequence and structure dependant. [Note that the term 'hypersensitive' is used relatively here; the CpT could truly become more sensitive, but it might become (absolutely) less sensitive as neighbouring sites became more resistant (perhaps because they were covered by relatively more ligand).]

As **2** does not reduce access of DNase I to two adjacent arms, it seems unlikely that it binds in the way illustrated in Figure 1B (bottom). Moreover, we can only speculate on how it might generate the hypersensitive site. One possibility is that it might mono-intercalate into one arm in the X-shaped structure so that the other (projecting) intercalating moi.e.ty 'kinks' the adjacent arm to make the CpT hypersensitive; but then we might expect bonds around the binding site on the first arm to be protected from DNase I. Alternatively, the lack of significant protection impli.e.s that binding might occur in the already-protected region (i.e. at the junction). There is ample precedent for such binding by intercalators (revi.e.wed in ref. 32), including methidium (33,34), porphyrins (35) and cyanines (36). If binding is right at the junction, then some distortion must be transmitted through almost one complete turn of the double helix; although there is again a precedent for action over such a distance, the end effect is slight (37). If binding is close to the centre of the junction, it must still induce some conformational change that promotes DNase I cutting at the remote CpT. If binding is at the outer edge of the protected region (perhaps even by bis-intercalation), then an additional 2–3 nt 3′ to the drug would be protected (as is seen) and the conformational change would need to be propagated through fewer base pairs. Whichever binding mode turns out to be correct, it must depend on the presence of two intercalating groups connected by a linker of appropriate length and a 4-way junction with complementary stereochemistry, and it must be associated with some startling effects.

Discussion has concentrated on one hypersensitive site induced by **2**; however, each of the compounds studied induced its own pattern of resistance/sensitivity in the different structures (Fig. 10). Clearly, a more detailed understanding of the binding of these reagents to such superficially simple structures awaits further experimentation.

ACKNOWLEDGEMENTS

We thank the Wellcome Trust and the E. P. Abraham Trust for support.

REFERENCES

1 Annan, N.K., Cook, P.R., Mullins, S.T. and Lowe, G. (1992) *Nucleic Acids Res.* **20**, 983–990.
2 Mullins, S.T., Annan, N.K., Cook, P.R. and Lowe, G. (1992) *Biochemistry* **31**, 842–849.
3 LePecq, J-B., Le Bret, M., Barbet, J. and Roques, B. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2915–2919.
4 Cannellakis, E.S., Shaw, Y.H., Hanners, W.E. and Schwartz, R.A. (1976) *Biochem. Biophys. Acta* **418**, 277–298.
5 Gao, X. and Patel, D.J. (1988) *Biochemistry* **27**, 1744–1751.
6 Searle, M.S., Hall, J.G., Denny, W.A. and Wakelin, L.P.G. (1989) *Biochem. J.* **259**, 433–441.
7 Zon, G., and Wilson, W.D. (1990) *Biochemistry* **29**, 10918–10927.
8 Zhang, X. and Patel, D.J. (1991) *Biochemistry* **30**, 4026–4041.
9 Arnold, E. and Clardy, J. (1981) *J. Am. Chem. Soc.* **103**, 1243–1244.
10 Huang, C-H., Mirabelli, C.K., Mong, S. and Crooke, S.T. (1983) *Cancer Res.* **43**, 2718–2724.
11 Lilley, D.M.J., Clegg, R.M Diekmann, S., Seeman, N.C., von Kitzing, E. and Hagerman, P.J. (1995) *Eur. J. Biochem.* **230**, 1–2.
12 Peek, M.E., Lipscombe, L.A., Bertrand, J.A., Gao, Q., Roques, B.P., Garbay-Jaureguiberry, C. and Williams, L.D. (1994) *Biochemistry* **33**, 3794–3800.
13 Jacobsen, J.P., Pedersen, J.B., Hansen, L.F. and Wemmer, D.E. (1995) *Nucleic Acids Res.* **23**, 753–760.
14 Lilley, D.M.J. and Clegg, R.M. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22**, 299–328.
15 Seeman, N.C. and Kallenbach, N.R. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23**, 53–86.
16 Lu, M., Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1989) *J. Biol. Chem.* **264**, 20851–20854.
17 Murchi.e., A.I.H., Carter, W.A., Portugal, J. and Lilley, D.M.J. (1990) *Nucleic Acids Res.* **18**, 2599–2906.

18 Suck, D., Lahm, A. and Oefner, C. (1988) *Nature* **332**, 464–468.
19 Drew, H.R. (1984) *J. Mol. Biol.* **176**, 535–557.
20 Drew, H.R. and Travers, A.A. (1984) *Cell* **37**, 491–502.
21 Herrera, J.E. and Chaires, J.B. (1994) *J. Mol. Biol.* **236**, 405–411.
22 Fox, K.R. and Waring, M.J. (1987) *Nucleic Acids Res.* **15**, 491–507.
23 Ward, B., Rehfuss, R., Goodisman, J. and Dabrowiak, J.C. (1988) *Nucleic Acids Res.* **16**, 1359–1369.
24 Duckett, D.R., Murchi.e., A.I.H., Diekmann, S., von Kitzing, E., Kemper, B. and Lilley, D.M.J. (1988) *Cell* **55**, 79–89.
25 Chen, J.-H., Churchill, M.E.A., Tullius, T.D., Kallenbach, N.R. and Seeman, N.C. (1988) *Biochemistry* **27**, 6032–6038.
26 Kimball, A., Guo, Q., Lu, M., Cunningham, R.P., Kallenbach, N.R., Seeman, N.C. and Tullius, T.D. (1990) *J. Biol. Chem.* **265**, 6544–6547.
27 von Kitzing, E., Lilley, D.M.J. and Diekmann, S. (1990) *Nucleic Acids Res.* **18**, 2671–2683.
28 Lu, M., Guo, Q. and Kallenbach, N.R. (1991) *Biochemistry* **30**, 5815–5820.
29 Zhang, S. and Seeman, N.C. (1994) *J. Mol. Biol.* **238**, 658–668.
30 Fu, T.-J., Tse-Dinh, Y.-C. and Seeman, N.C. (1994) *J. Mol. Biol.* **236**, 91–105.
31 Kallenbach, N.R. and Zhong, M. (1994) *Curr. Opin. Struct. Biol.* **4**, 365–371.
32 Lu, M., Guo, Q. and Kallenbach, N.R. (1992) *Crit. Rev. Biochem. Mol. Biol.* **27**, 157–190.
33 Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1989) *Biochemistry* **28**, 2355–2359.
34 Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1990) *Biochemistry* **29**, 570–578.
35 Lu, M., Guo, Q., Pasternak, R.F., Wink, D.J., Seeman, N.C. and Kallenbach, N.R. (1990) *Biochemistry* **29**, 1614–1624.
36 Lu, M., Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1990) *Biochemistry* **29**, 3407–3412.
37 Bailly, C., Gentle, D., Hamy, F., Purcell, M. and Waring, M.J. (1994) *Biochem J.* **300**, 165–173.