Highly selective chemical modification of cruciform loops by diethyl pyrocarbonate

Judy C Furlong and David M.J. Lilley*

Department of Biochemistry, The University, Dundee DD1 4HN, UK

Received 20 March 1986; Accepted 11 April 1986

ABSTRACT
Diethyl pyrocarbonate reacts with the single-stranded loops of cruciform structures with great selectivity. Adenine bases are carbethoxylated, as a result of which the backbone may be cleaved with piperidine, and the level of chemical modification at each base may be determined. We have studied the ColE1 and (A-T)$_{34}$ cruciforms of pColIR315 and pXG540. In each case we observe maximal modification at the most central adenosine of the loop, and an overall pattern of modification corresponding to a total loop size of about six bases. The results may be interpreted in terms of a model in which the loop has a defined tertiary structure. No modification was detected at either cruciform four-way junction, suggesting that this region is fully base-paired.

INTRODUCTION
The structure of DNA is very dependent upon local base-sequence. Whilst this is true for torsionally relaxed DNA, the effects become even greater if it is negatively supercoiled. At the level of the single nucleotide, DNA is structurally polymorphic in the crystal (1) and in solution (2,3). Negative supercoiling (4) may stabilise markedly deviant secondary structures if the local DNA sequence is appropriate, including left-handed Z-DNA (5-7), intra-strand paired cruciform structures (8-10), poly-pyrimidine stretches (11-13), slippage structures (14,15) and doubtless many other conformations.

Non-B conformations of DNA may expose the functional groups of nucleoside bases in a manner not normally found in B-DNA. For this reason these polymorphs may act as reactive centres towards chemical reagents to which DNA is normally refractory. For example, cruciform structures possess unpaired loops which include bases having single-stranded character. These have been demonstrated...
to be reactive towards a series of chemical compounds, including bromoacetaldehyde (16), osmium tetroxide (17), glyoxal (18) bisulphite (19) and ozone (20). Similarly the bases of DNA in the Z-conformation exhibit enhanced reactivity towards a number of compounds, including osmium tetroxide (21,22), hydroxylamine and dialkyl sulphates (21), diethyl pyrocarbonate (23), and a series of intercalating chiral transition metal complexes (24,25).

Chemical hyper-reactivity at local structural perturbations, or perhaps at interfacial regions, is significant for two reasons. First, it provides a convenient method of probing the novel conformation. Detailed analysis of the patterns of chemical modification can yield considerable insight into local stereochemical arrangements. Second, compounds which chemically modify DNA are potential mutagens and carcinogens. The haloacetaldehydes, for example, are the ultimate metabolites of the vinyl halides, which are well documented environmental carcinogens (26). Thus it is conceivable that gene control regions may be especially susceptible to chemical reaction by virtue of perturbed DNA conformation.

We have described a number of chemical probes of cruciform structures in supercoiled DNA (16-19). However, it has often been difficult to deduce the initial positions of chemical modifications for two reasons. First, a fundamental problem with reagents such as bromoacetaldehyde and osmium tetroxide (16,17) is propagation of chemical reaction away from the point of initial attack. Thus adduct formation at one position, such as the centre of an inverted repeat, appears to predispose neighbouring bases towards further reaction, resulting in the expansion of a 'bubble' of modified DNA driven by the ever increasing topological relaxation as the bubble grows. Second, S1 nuclease has frequently been employed to detect adduct formation by virtue of impaired base-pairing (16,17). Sodium bisulphite (19) suffers from neither of these problems, but two drawbacks exist nevertheless. The procedure involves transformation into *E. coli* and sequence analysis of DNA purified from individual colonies, which is time consuming, and only C-G base pairs are examined by this approach.
Recently Herr described the use of diethyl pyrocarbonate (ethoxyformic anhydride or DEPC) (23) as a selective chemical probe of Z-DNA. DEPC reacts with adenosine, and to a lesser degree with guanosine, by attack principally at N7, resulting in ring opening at the imidazole ring of the purine nucleus (27,28). DEPC-modified DNA may be reacted with piperidine at 90°C resulting in backbone scission at carbethoxylated adenosine residues, thus providing a very simple procedure for locating positions of elevated chemical reactivity. The conformation of Z-DNA is such as to expose N7 of the syn purine bases, thereby increasing their chemical reactivity. N7 should also be exposed when the base is part of the formally single-stranded loop of a cruciform, and we therefore surmised that DEPC might serve as a chemical probe of cruciform structure. We tested this proposition on two plasmids, indicated in figure 1, containing the ColE1 cruciform (9,29) and an (A-T)$_{34}$ cruciform (30). The results obtained confirm that DEPC is an excellent selective probe of cruciform structure, and the patterns of modification indicates that the cruciform loops may have a distinct structure.

**FIGURE 1** The plasmids employed in these studies. Maps of pColIR315 and pXG540, showing the location of the inverted repeats (filled boxes) relative to the restriction enzyme sites used. The sequences of the two inverted repeats are written below.
EXPERIMENTAL
DNA Plasmids were prepared from *E. coli* K12 HB101 by lysozyme, EDTA, SDS lysis, and two rounds of isopycnic centrifugation in CsCl and ethidium bromide. Supercoiled DNA was recovered by side puncture, extraction with butan-1-ol and extensive dialysis against 10mM Tris pH7.5, 0.1mM EDTA (TE) at 5°C. pColIR315 was heated to 37°C for 5 min in TE to ensure full extrusion of the CoIE1 cruciform (31). Preextrusion is not necessary for the (A-T) cruciform of pXG540 (30).

DEPC reactions DNA (1-5ug) was diluted into 200ul of 50mM Na cacodylate, pH7.1, 1mM EDTA. 1.5ul of DEPC (Sigma) were added and vigorously agitated on a rotor mixer to ensure solution of this relatively imiscible compound. Modification was performed at 20°C (unless otherwise indicated) for 15m, with vigorous mixing at 5m intervals. Reaction was terminated by ethanol precipitation. After a second ethanol precipitation the DNA pellet was dissolved in 100ul of 1M piperidine (BDH) and heated at 90°C for 30 min. The cleaved DNA was then lyophilised three times and finally dissolved in deionised formamide.

Sequencing Reactions and electrophoresis Sequencing reactions were carried out by standard Maxam and Gilbert techniques (32). Electrophoresis was performed on 8% (pColIR315) or 5% (pXG540) gels containing 7M urea. Sequencing gels were dried onto 3MM paper (Whatman) and autoradiographed on Fuji RX X-ray film using Ilford fast-tungstate intensifier screens. Results were quantified by densitometry of autoradiographs using an LKB 2202 Ultrascan laser densitometer.

Enzyme Reactions Restriction enzymes were obtained from Bethesda Research Laboratories, and used as directed. 5' [³²P] labelling of DNA was carried out using exchange labelling procedures (33) and T4 polynucleotide kinase (Amersham).

RESULTS
Gel mobility of supercoiled pColIR315 following reaction with DEPC
Supercoiled pColIR315 was reacted with 50mM DEPC at 20°C for 15m, and the DNA examined by electrophoresis in 1% agarose. The results, shown in figure 2, show that the DEPC reacted DNA was of
FIGURE 2 The mobility of supercoiled pColIR315 on agarose gel electrophoresis is unchanged by DEPC modification. S, supercoiled pColIR315 unreacted; D, supercoiled pColIR315 after reaction with 50 mM DEPC for 15 min at 20°C. The arrows denote the position of supercoiled (S) and open circular (D) pColIR315.

indistinguishable mobility from the unreacted DNA. This behaviour contrasts with either bromoacetaldehyde or osmium tetroxide modification of the same plasmid (16,17) where reaction leads to both retardation and broadening. This indicates that the extent of reaction is probably quite limited, and that there is at most a very small degree of local unwinding.

Nucleotides modified in the ColEl cruciform of pColIR315

Supercoiled pColIR315 was heated to 37°C for 5 min to ensure full cruciform extrusion (31). The DNA was then reacted with 50 mM DEPC at 20°C for 15 min. After ethanol precipitation the modified pColIR315 was cleaved with EcoRI, 5' termini were labelled with [32P] and the DNA re-cleaved with BamHI. The 440bp EcoRI to BamHI fragment was isolated by polyacrylamide gel electrophoresis and electroelution. The phosphodiester backbone of the DNA was broken at the points of base carbethoxylation by treatment with 1M piperidine at 90°C, and the fragments resolved
FIGURE 3  DEPC modification of supercoiled pColIR315, top strand. Supercoiled pColIR315 was reacted with 50mM DEPC, and the EcoRI to BamHI fragment, [\(32^P\)] labelled at the EcoRI site, was cleaved with piperidine. This was electrophoresed on an 8% sequencing gel (D), flanked by Maxam-Gilbert (32) tracks derived from the equivalent EcoRI to BamHI DNA fragment. The extent of the ColE1 inverted repeat is shown by the arrows drawn on the right hand side.
on a sequencing gel. The results are shown in figure 3, where the fragments due to DEPC treatment were electrophoresed against Maxam-Gilbert sequencing ladders derived from an equivalent *EcoRI* to *BamHI* fragment. It can be seen that the DEPC modification was limited tightly to nucleotides in the loop of the ColEl cruciform of pColIR315. The major sites of modification were the three adenosine bases of the AAATG loop (adenosine is intrinsically more reactive towards DEPC than is guanosine), but there was a gradation of reactivity of the adenosines, becoming less reactive in the 5' direction.

We have also examined DEPC modification on the opposite strand, using the *TaqI* site which lies 90 bp to the right of the ColEl inverted repeat. The results are shown in figure 4. The only significant DEPC modification was at the single adenosine in the CATTT loop of the cruciform, just as we would have predicted.

Outside the cruciform loops there was almost no modification by the reagent. It is particularly significant that no reaction was detectable at the bases at the four-way junction of the cruciform. At the 5' end flanking DNA there is a run of adenosines which abut the inverted repeat, yet these remained totally unmodified.

The pattern of DEPC modification in and around the ColEl cruciform of pColIR315 was largely insensitive to the reaction conditions. Variations of DEPC concentration resulted in no discernible alteration in modification sites. Elevation of reaction temperature to 37°C increased the level of background non-specific carbethoxylation, but once again the essential nature of the distribution of modification was unchanged.

*Nucleotides modified in the (A-T)₃⁴ cruciform of pXG540*

We have demonstrated previously that an (A-T)₃⁴ sequence isolated from an intron of the major αT1 globin gene of *Xenopus laevis* (34) forms a well defined cruciform structure when supercoiled to a moderate degree (30). pXG540, which contains this sequence, was reacted with DEPC, and an *EcoRI* [³²P] - *BamHI* fragment isolated and cleaved with piperidine as before.

The sequencing gel of the modified pXG540 alongside corresponding sequence ladders is shown in figure 5. Just as...
FIGURE 4 DEPC modification of supercoiled pColIR315, bottom strand. Supercoiled pColIR315 was reacted with 50mM DEPC, and the TaqI to EcoRI fragment, [32P] labelled at the TaqI site, was cleaved with piperidine. This was electrophoresed on an 8% sequencing gel (D), flanked by Maxam-Gilbert (32) tracks derived from the equivalent TaqI to EcoRI DNA fragment. The extent of the ColE1 inverted repeat is shown by the arrows drawn on the right hand side.
FIGURE 5 DEPC modification of supercoiled pXG540. Supercoiled pXG540 was reacted with 50mM DEPC, and the EcoRI to BamHI fragment, $[^{32}P]$ labelled at the EcoRI site, was cleaved with piperidine. This was electrophoresed on a 5% sequencing gel (D), flanked by Maxam-Gilbert (32) tracks derived from the equivalent EcoRI to BamHI DNA fragment. The extent of the (A-T)$_{34}$ sequence is shown by the arrows drawn on the right hand side.

with the ColE1 cruciform, it may be seen that the extent of DEPC modification is tightly localised to a few bases, and these are at the symmetrical centre of the (A-T)$_{34}$ run. No modification is
seen elsewhere, including the cruciform junctions. When the supercoiling was abolished, by linearisation with a restriction enzyme, then no modification was detectable in the (A-T) region.

DISCUSSION

DEPC as a probe of perturbed DNA secondary structure

These results demonstrate that DEPC is a most useful probe of cruciform structure. Carbethoxylation of nucleotide bases is highly selective, with no indication of 'propagation' of reactivity to contiguous bases, as has been observed for other reagents (16,17). Furthermore, the ability to cause backbone scission at the point of modification with piperidine allows single-base resolution. This has only been possible previously for cruciform structures by using bisulphite (19), requiring a protocol of much greater complexity. In addition, DEPC provides a probe which has a chemical target different from all existing reagents available to date.

The utility of DEPC for studies of Z-DNA has already been demonstrated by Herr (23), and we now show its value for cruciform studies. We believe that DEPC may be of general use in probing altered DNA secondary structures, where the environment of purine N-7 may be altered from that in B-DNA.

Analysis of cruciform structures

The modification patterns due to DEPC reactions with unpaired adenosine in the loops of the ColEl and (A-T) cruciforms.
cruciforms are summarized in figure 6. We have previously estimated optimal loop sizes of cruciforms based upon reactivity towards sodium bisulphite (19), concluding that 4-6 bases are unpaired. This is also fully consistent with thermodynamic studies of a series of DNA hairpins (35). The DEPC modification patterns are broadly in agreement with these estimates. In the (A-T) cruciform, for example, the most central adenosine (A18) exhibits the highest reactivity, and the second most reactive is that to its 5' side (A17). Two points emerge from these data. First, the adenosine to the 3' side of the most reactive base (A19) has 50% reactivity, yet to be unpaired requires the formation of a six base loop at 20°C. Even the adenosine which is two doublets 5' (A16) shows significantly non-zero reactivity. It seems likely that the A-T loop may be looser than those of less A + T rich cruciforms. Second, the adenosine to the 5' side of the most reactive one should be fully unpaired, corresponding to a loop-size of just four bases, yet its reactivity is reduced to 64% maximal. It seems unlikely that this may base-pair even transiently, because of stereochemical considerations, and thus its lower reactivity requires explanation. We suggest that the loop adopts a defined structure in three dimensions, involving base stacking, reducing the availability of the adenosine N7 positions. This point is seen more strikingly in the ColEl cruciform. The AAATG must be fully unpaired on symmetry grounds, yet the three adenosines are not equally modified by DEPC. The central adenosine (ie AAATG) is the most reactive, and the reactivity reduces towards the 5' direction, ie with proximity to the stem. We suggest that these adenosine bases are stacked on to the top of the stem double helix, giving them quasi-double-stranded character and hence reducing their chemical reactivity. This model for the cruciform loop is suggested by the structure of a DNA hairpin solved in solution by Hilbers and colleagues by 2D NMR methods (36), where a loop of thymines exhibits quasi-helical stacking on the 5' side.

The four-way junctions of both the ColEl and (A-T) cruciforms are modified by DEPC to a negligible extent. A very similar observation was made using bisulphite (19). This
strengthens our conclusion that the junction is fully base-paired, at least up to 37°C.

The results obtained by DEPC modification of two cruciforms are in excellent accord with data using other chemical probes and physical studies. A consistent picture is emerging in which the cruciform is characterised by a structured loop of 4 to 6 bases, and a fully paired four-way junction.

POSTSCRIPT

We understand that equivalent DEPC modification of cruciform loops has been demonstrated by Dr Alfred Nordheim (Heidelberg).

ACKNOWLEDGEMENTS

We are very grateful to Dr Winship Herr for communicating his results on DEPC modification of Z-DNA prior to publication. We thank the Cancer Research Campaign, the Medical Research Council and the Royal Society for financial support.

*To whom correspondence should be addressed

REFERENCES


4006