The EMBO Journal vol.14 no.9 pp.2121–2131, 1995

The influence of the 2-amino group of guanine on DNA conformation. Uranyl and DNase I probing of inosine/diaminopurine substituted DNA

Christian Bailly, Niels Erik Møllegaard[1], Peter E.Nielsen[1,2] and Michael J.Waring[2]

Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, UK and [1]Department of Medical Biochemistry and Genetics, The Panum Institute, Blegdamsvej 3c, DK-2200 Copenhagen N, Denmark

[2]Corresponding authors

Communicated by A.Travers

This paper is dedicated to the memory of Professor Ole Buchardt (deceased 5 September 1994) whose insight, energy and enthusiasm provided much of the inspiration which initiated our collaboration.

The conformation of the DNA helix is supposed to be a critical element in site-specific recognition by ligands both large and small. Groove width is one important measure of the conformation which varies with the local nucleotide composition, perhaps because of the presence of a purine 2-amino group on G·C base pairs. We have probed DNA with G→inosine (I) and/or A→diaminopurine (DAP) substitutions to see whether the location of the purine 2-amino group can indeed affect the minor groove width. At acid pH, the reactivity towards uranyl nitrate is modulated in substituted DNA quite differently from natural DNA, consistent with a marked narrowing of the minor groove at sites of G→I substitution and widening at sites of A→DAP replacement. The latter exerts the dominant effect. The expected changes in conformation are equally evident in the patterns of susceptibility to DNase I cleavage, but not to hydroxyl radical attack. Nuclease cleavage is maximal in normal and substituted DNA at regions of inferred moderate groove width which are generally little affected by the nucleotide substitutions. Consistent with models of sequence-dependent cutting by DNase I we find that the presence of a purine 2-amino group on the base pair three places upstream of the cutting site has a profound influence on the rate of reaction.

Key words: DNA structure/uranyl photocleavage/DNase I cleavage/minor groove width/inosine/diaminopurine

Introduction

The sequence-specific binding of proteins and drugs to double-stranded DNA can formally be treated as either a 'digital' or an 'analogue readout' of the DNA sequence (Travers, 1989). A purely 'digital readout' would correspond to a direct interaction of the ligand with each base pair independently of neighbouring base pairs: according to this motif simple consensus recognition sequences, which exist for most DNA-binding proteins, can be determined. A purely 'analogue readout' would correspond to a ligand–DNA shape complementarity whereby the DNA binding domain of the ligand recognizes the global structure/conformation of the DNA, itself determined by the primary base sequence in a fashion not yet fully understood. The digital recognition component seems to be predominant for some ligands (e.g. λ-repressor, *met*-repressor, echinomycin), whereas the analogue component is more important for other ligands such as CRP and distamycin (Steitz, 1993; Travers, 1993). Thus DNA structure/conformation as expressed in terms of such features as groove width, curvature and bendability, and the correlation of such features with DNA base sequence, furnish important parameters for understanding sequence specific DNA recognition.

We have previously found that uranyl-mediated photocleavage of double-stranded DNA at acidic pH (~6.0–6.5) exhibits a very strong modulation which is correlated with minor groove width/electronegative potential. Specifically, we have shown that increased cleavage across the minor groove is observed at AT-rich distamycin binding sites on a pUC19 restriction fragment (Nielsen *et al.*, 1990). Furthermore, the minor groove width along the internal control region (ICR) of the *Xenopus* 5S RNA gene (the TFIIIA binding site) as predicted from uranyl photocleavage analysis agrees with that deduced from DNase I probing (Nielsen *et al.*, 1990). Finally, we have found that the uranyl photocleavage pattern interpreted in terms of minor groove width of the 'Drew–Dickerson dodecamer' sequence (CGCGAATTGCGC) either as a short oligonucleotide or cloned into a larger (250 bp) DNA fragment is fully consistent with the minor groove width determined by X-ray crystallography (Sönnichsen, 1991; S.H.Sönnichsen and P.E.Nielsen, in preparation).

The width of the minor groove of the DNA helix is to a first approximation correlated with its AT/GC content since A–T tracts appear to be associated with a narrowed groove, whereas G–C tracts often have a widened minor groove (Burkhoff and Tullius, 1987, 1988; Nelson *et al.*, 1987; Heinemann and Alings, 1989; Edwards *et al.*, 1992). The major chemical difference in the minor groove between A·T and G·C base pairs is the C2 exocyclic amino group of guanine which protrudes out from the stack of base pairs (Figure 1). Thus by replacing inosine (I) for guanosine and 2,6-diaminopurine (DAP) for adenine, it is possible from the standpoint of a view into the minor groove to 'convert' G·C base pairs into A·T base pairs and *vice versa*, and thereby to assess the contribution of this amino group to overall DNA structure/conformation. We now report the results of such experiments using PCR-generated DNA fragments and examining them with the structure sensitive probes DNase I, UO_2^{2+}/UVA and Fe^{2+}/EDTA.

Our strategy was to choose a 160 bp fragment containing the *Escherichia coli tyrT* promoter. This fragment *tyrT*(A93) differs from the well characterized wild-type

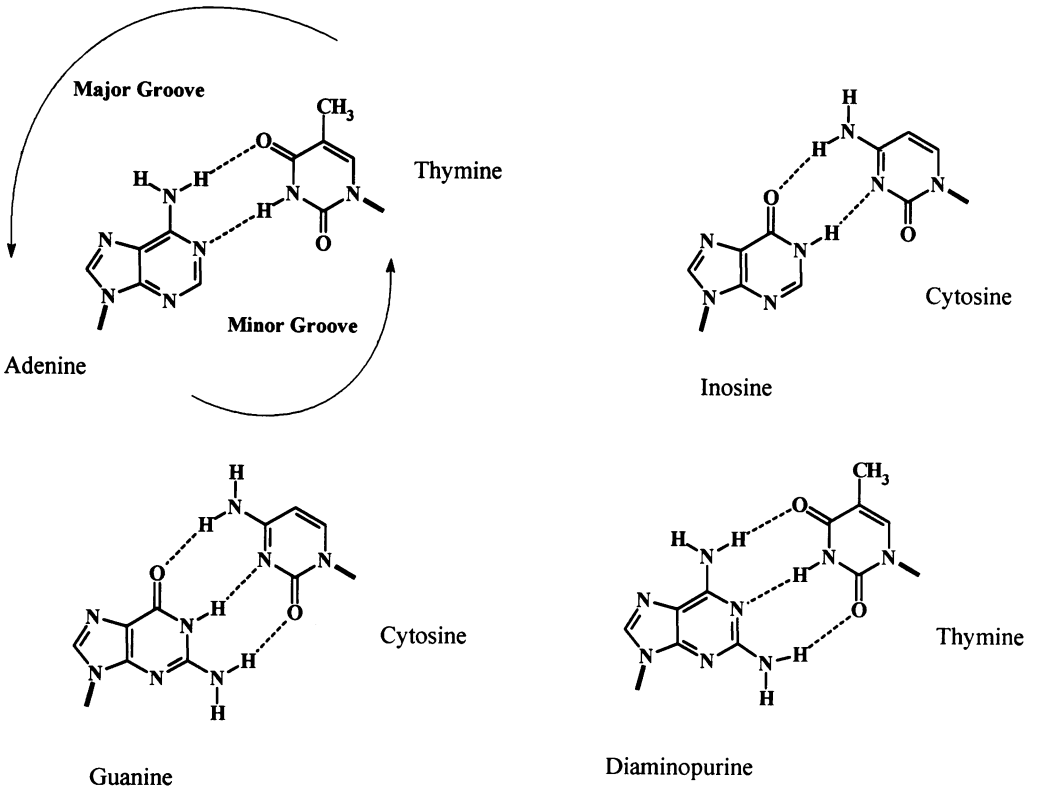


Fig. 1. Structures of purine–pyrimidine base pairs involving unnatural nucleotides, with Watson–Crick base pairs for comparison.

tyrT duplex by a single A·T to T·A transversion at the position numbered 93 according to the Drew and Travers (1984) numbering scheme (used throughout this report), i.e. at position −8 relative to the *in vivo* transcription start site. This transversion corresponds to a strong down-activity mutation (Berman and Landy, 1979; Drew *et al.*, 1985).

Four different 160 bp DNA fragments containing either natural bases or inosine residues in place of guanines (G→I substitution) or 2,6-diaminopurine residues (henceforth abbreviated DAP or D within a sequence for clarity) in place of adenines (A→DAP substitution) or both I and DAP residues were synthesized by PCR amplification. Primers in which the 5′ terminal nucleotide residue bears a 5′-OH or a 5′-NH_2 terminal group were used so as to enable selective labelling of one or the other strand in the PCR product, i.e. the Watson (antisense) strand or the Crick (sense) strand chosen at will. The cleavage patterns of these modified DNAs were analysed and compared with normal DNA using enzymatic and chemical/photochemical probes which are known to be sensitive to the DNA conformation. These were (i) DNase I, (ii) the photoactivated probe uranyl(VI) nitrate and (iii) the hydroxyl radical-generating transition metal complex Fe(II)·EDTA.

Results

The purine 2-amino group strongly influences the uranyl-mediated photocleavage of DNA

The 5′-^{32}P-labelled normal and modified DNA fragments were incubated with uranyl nitrate in a 50 mM sodium acetate buffer adjusted to pH 6.5 and then irradiated with long wavelength UV light (λ~420 nm) to induce UO_2^{2+}-mediated oxidation of the deoxyribose rings which generates breakages in the DNA backbone (Nielsen *et al.*, 1992). The resulting DNA cleavage products were resolved on a sequencing gel. Typical autoradiographs of gels obtained with the normal and modified PCR products are shown in Figure 2. At low pH the extent of photocleavage varies substantially according to the sequence of the DNA whereas at neutral pH (e.g. in 50 mM Tris buffer, pH 7.0) the cleavage patterns show very little variation (data not shown) in accordance with previous results (Nielsen *et al.*, 1990, 1992; S.H.Sönnichsen and P.E.Nielsen, in preparation). Substantial differences between the normal DNA and that containing both inosine and DAP residues can be seen, and are even more clearly evident in the quantitated cleavage plots shown in Figure 3. These plots were obtained from densitometric analyses of phosphorimages of several gels including those shown in Figure 2. The cleavage of normal DNA is most efficient at AT-rich sequences such as occur around positions 32, 50, 66 and 84 on the Watson strand and positions 87, 112 and 128 on the Crick strand. Furthermore, the cleavage maximum is located at the 3′-end of the A/T tract. By contrast, the cleavage at GC-rich sequences (e.g. around positions 75 and 97 on both strands) is considerably weaker than average. The cleavage patterns observed with this *tyrT*(A93) fragment are totally consistent with those previously reported for other natural DNA fragments under similar experimental conditions (Nielsen *et al.*, 1990). The cleavage patterns obtained with the I+DAP DNA appear, to a certain extent, like a mirror image of those seen with normal DNA: upon irradiation the cleavage of the I+DAP DNA by the uranyl ions is clearly stronger

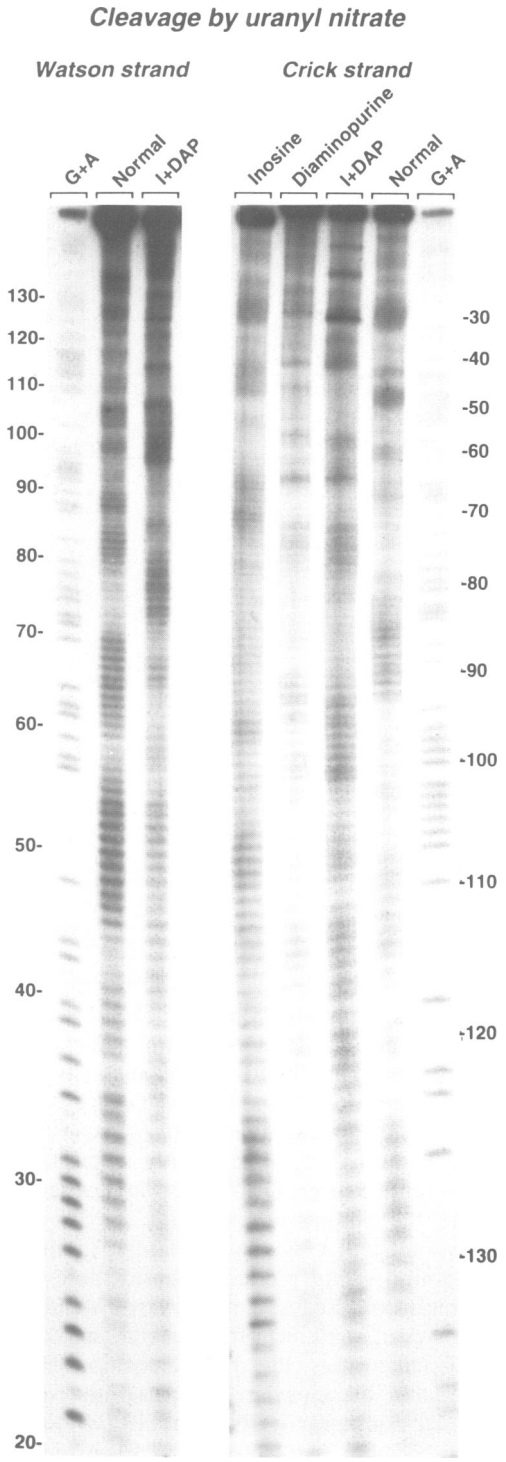


Fig. 2. Autoradiograph showing the uranyl-mediated photocleavage patterns of *tyrT*(A93) DNA fragments containing the four natural nucleotides (Normal DNA), inosine residues in place of guanosine (Inosine DNA), DAP residues in place of adenine (DAP DNA) or inosine and DAP residues in place of guanosine and adenine, respectively (I+DAP DNA). Chemical identities of the digestion products were assigned by reference to the normal sequencing markers (lanes G + A). The scales on the sides of the gels correspond to the standard numbering of the *tyrT* sequence (Drew and Travers, 1984). I and/or DAP residues were introduced by PCR amplification using primers specifically designed to ensure selective 5′-end labelling of the Watson (antisense) or Crick (sense) strand. The cleavage products of the DNase I digestion were resolved on an 8% polyacrylamide gel containing 8 M urea.

at GC-rich sequences and quite poor at AT-rich sequences. At some places the phase is almost perfectly inverted, thus attesting that shifting the 2-amino group from guanine to adenine residues has a considerable influence on the uranyl-mediated photocleavage of DNA. This shows that the structure of the DNA has changed locally and the uranyl ions are very sensitive to these local structural variations.

At first sight, it would seem that the UO_2^{2+} ions bind best to sequences lacking the 2-amino group, i.e. to $(A{\cdot}T)_n$ sites in normal DNA and to $(I{\cdot}C)_n$ sites in the I+DAP DNA. However, this straightforward interpretation is too simple, as is revealed by the cleavage patterns obtained with the inosine-substituted DNA and the DAP DNA. Figure 4 shows a critical comparison of the cleavage plots obtained for both strands of the normal and all three modified DNAs. With the inosine DNA, the modulation of cleavage intensity is fairly attenuated compared with what is observed with normal DNA. Thus the removal of the 2-amino group of guanine significantly reduces—but does not completely abolish—the sequence-dependent variations in the cleavage of the DNA fragment by uranyl. This suggests that there is much less structural variation within the inosine DNA than within the other homologous DNAs. Conversely, the DAP DNA in which all purine residues now bear a 2-amino group exposed in the minor groove exhibits a significant modulation of the uranyl photocleavage, but the cleavage profile observed with the DAP DNA is quite similar to that observed with the I+DAP DNA. The cleavage at GC-rich sites is generally more pronounced than the cleavage at DAP·T sites. Thus the effect of adding a 2-amino group on to the adenines (A→DAP substitution) is to reduce drastically the relative cleavage at AT regions and produce a concomitant relative increase in cleavage at GC-rich sequences despite the continued presence of the 2-amino group at G. In other words, the preferred UO_2^{2+} cleavage has been shifted from AT to GC sites which must be ascribed to DAP-induced structural changes in DNA which are very sensitively perceived by the uranyl probe. Because DAP·T-rich sequences show weaker reactivity towards uranyl nitrate compared with G·C-containing sequences we must further conclude that adding the 2-amino group to adenine residues converts AT-rich tracts into regions which have structural characteristics (e.g. minor groove width) significantly different from those of GC-rich regions. Furthermore, we can establish from this first set of experiments that of the two types of DNA modification we have been investigating, i.e. removing or adding the purine 2-amino group, it is clearly the latter which is dominant and exerts the strongest influence on the uranyl-mediated photocleavage reaction. This observation raises interesting implications for ligand binding.

Shifting, adding or removing the purine 2-amino group has little influence on the hydroxyl radical-induced cleavage of DNA

Hydroxyl radicals can be produced by metal transition complexes such as Fe·EDTA via the Fenton reaction (Tullius, 1988). They break the DNA backbone according to a mechanism which is believed to proceed by attack on the sugar ring at the 4′ position (Stubbe and Kozarich, 1987).

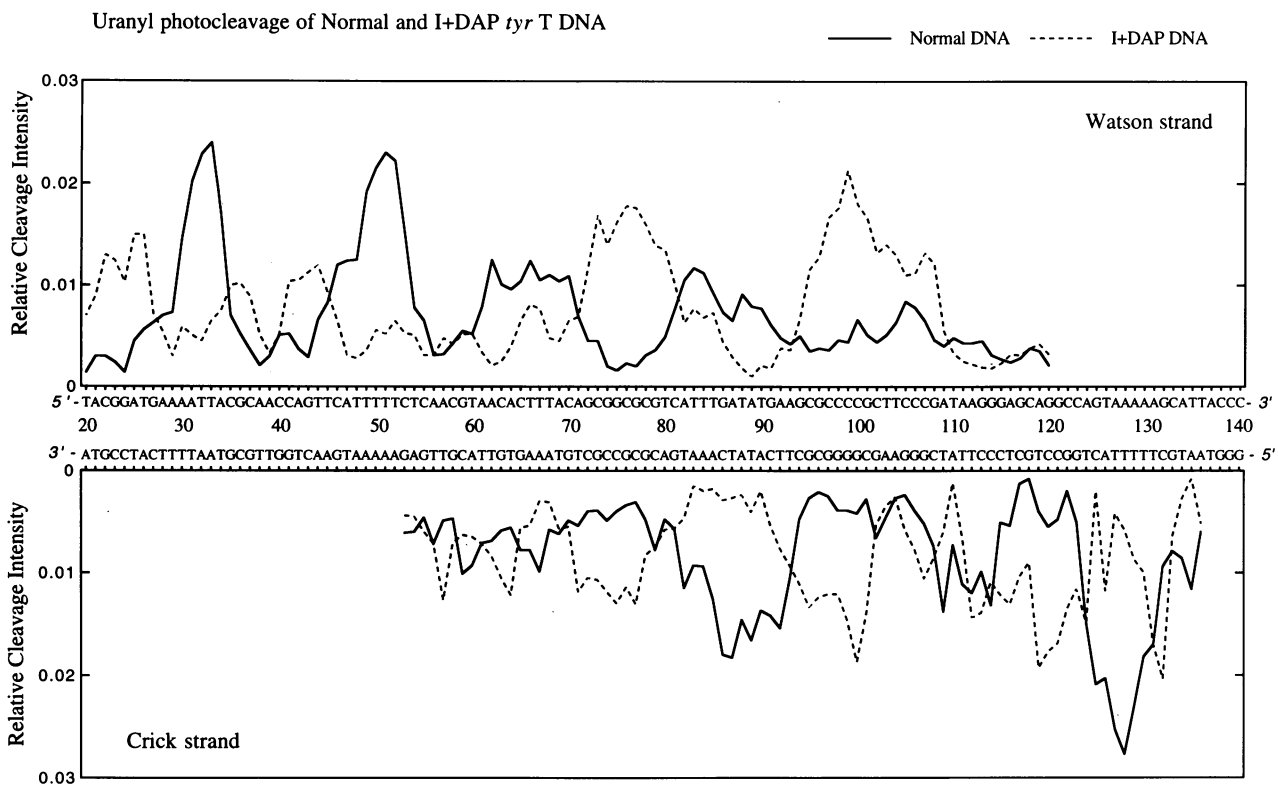


Fig. 3. Cleavage plots comparing the susceptibility of normal and I+DAP DNAs to the uranyl-mediated photocleavage reaction. The relative cleavage intensity values (in arbitrary units) refer to the extent of cleavage of a given bond relative to the sum of the cleavage of all the phosphodiester bonds examined within the sequence. The double-stranded sequence shown above and below the axis corresponds to that of the normal DNA. In the I+DAP DNA, adenine and guanosine residues are replaced by diaminopurine and inosine residues, respectively. The numbers between the two strands indicate the nucleotide position.

Sequencing gels and cleavage plots for normal and modified DNAs subjected to Fe·EDTA-mediated hydroxyl radical attack are shown in Figure 5. The cutting pattern of normal DNA is consistent with the literature data: the cleavage at AT-tracts is slightly weaker than that at GC or mixed sequences. With the doubly substituted DNA containing I and DAP residues the $(DAP{\cdot}T)_n$ tracts are still cleaved less efficiently than the other sequences in the DNA. Apart from some small differences, the same alternation of regions of decreased cutting at $(A{\cdot}T)_n$ or $(DAP{\cdot}T)_n$ sites with regions of normal cutting is observed with any of the modified DNAs. Therefore, whatever the exact nature of the structural changes in DNA introduced by removing, adding or shifting the purine 2-amino group, they are not perceived by hydroxyl radicals. The accessibility of hydroxyl radicals to the C4′ position of the deoxyribose ring cannot be significantly altered.

Endonuclease cleavage of DNA is highly sensitive to the position of the purine 2-amino group

DNase I digestion patterns of normal and modified *tyrT*(A93) DNAs are shown in Figure 6. It is well established that in natural DNA the intensity of cleavage varies considerably according to the target sequence, and the cleavage is highly reduced at DNA regions where the minor groove is significantly narrowed or bendability is impaired (Drew, 1984; Drew and Travers, 1984). Similarly with all three modified DNAs we observed that certain phosphodiester bonds are cleaved very efficiently whereas others are cut poorly, and the cutting pattern of each DNA is quite distinct. The cleavage plots in Figure 7, obtained from densitometric analyses of the gels corresponding to the Watson strand-labelled DNAs, show this more clearly. The differences in the relative band intensity at defined positions between normal and I+DAP DNA are sometimes very pronounced: for example, bonds at positions 20 and 79 are cleaved much more efficiently when the 2-amino groups are shifted from G to A residues whereas at the same time the cutting at positions 23, 36 and 71 is drastically reduced. Obviously the repositioning of the purine 2-amino group produces considerable effects on the recognition and cleavage of DNA by the endonuclease.

It is, however, significant that despite large differences in cleavage intensities, the overall positions of maxima and minima in the cleavage pattern (Figure 7) remain quite similar for normal DNA and I+DAP substituted DNA. Furthermore, we note that the positions of the majority of the maxima in the DNase I cleavage profiles correspond to regions of the DNA where the uranyl photocleavage is comparable for normal DNA and I+DAP substituted DNA, i.e. at the intersections of the curves presented in Figure 3 (e.g. positions 27, 36, 41, 45, 55, 59, 71, 80, 93 and 109). These observations are fully compatible with the structural implications generally inferred from DNase I and the uranyl cleavage analyses. Strong uranyl cleavage is interpreted to betoken a narrowed minor groove, which is found predominantly at AT-rich regions in normal DNA and thus at IC-rich regions in I+DAP DNA. Correspondingly, regions of average (canonical B-DNA) minor groove width would be expected

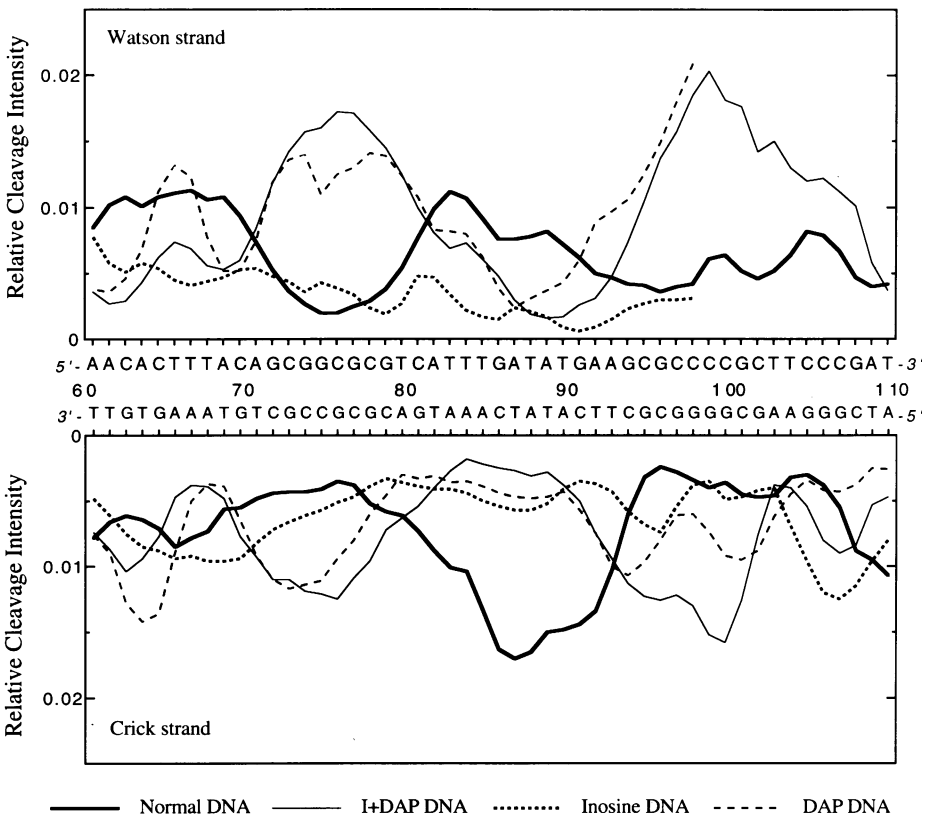


Fig. 4. Cleavage plots comparing the susceptibility of the normal and modified DNAs to the uranyl-mediated photocleavage reaction. The double-stranded sequence shown above and below the axis corresponds to that of the normal DNA. In the modified DNAs, adenine and/or guanosine residues are replaced by diaminopurine and/or inosine residues. The curves have been smoothed by taking a three-bond running average (Drew and Travers, 1984).

to be cleaved equally well by uranyl in normal and I+DAP DNA, and thus show up as intersections in a plot like that shown in Figure 3. By the same token, we might expect these regions in both types of DNA to be most susceptible to attack by DNase I and thus appear as maxima in cleavage plots.

By comparing the cleavage patterns obtained with the doubly substituted I+DAP DNA with those observed with the mono-substituted DNA containing either inosine or DAP residues it is possible to determine in detail whether the modification of the cleavage rate for a given bond is due to G→I or A→DAP substitution or both. For example, the increased cleavage intensity at the 5′-GpT dinucleotide step at position 79 observed with the I+DAP DNA relative to normal DNA seems to be attributable to the G→I rather than to the A→DAP substitution since the enhanced rate of cleavage at this position is also observed with the inosine DNA but not with the DAP DNA. Conversely, the susceptibility of the 5′-GpT dinucleotide step at position 59 is reduced in the I+DAP DNA compared with normal DNA and that seems to be due to the A→DAP rather than to the G→I substitution. To complete the picture, both A→DAP and G→I substitutions seem to contribute more or less equally to the reduced cleavage at the 5′-GpT step on the Crick strand at position 62 (not shown). These confusing observations prompted us to study on a more rigorous basis the influence of the nucleotides adjacent to the cleaved phosphodiester bond (i.e. the bases at positions +1 and −1). The relative cleavage intensity for each of 158 bonds was calculated for all four of the normal and modified DNA fragments. Figure 8 reports the logarithm of the relative cleavage intensity for the ten unique dinucleotide steps present in each DNA. Whichever dinucleotide step is considered, whether it contains natural or modified bases, the cleavage intensity always varies substantially, up to four natural log units. Recently, Herrera and Chaires (1994) used the same analysis to show that the cleavage of the wild-type *tyrT* DNA by DNase I cannot be interpreted on the simple consideration of the nature of the dinucleotide step at the cutting site. Evidently the same holds true for any of the modified DNAs containing inosine and/or DAP residues. Similarly, no clear correlation could be established between the sequence selectivity of the enzyme and the nature of the bases from positions −2 to +2 from the cutting site. Consideration of 6 bp sequences (extending from −3 to +3) does yield information of an interpretable nature.

The hexanucleotide sequences where the rate of cleavage of the central phosphodiester bond differs most between normal and doubly substituted I+DAP DNA are collated in Tables I and II. The entries are ranked in order of the magnitude of the effect, beginning with the most marked example at the top and concluding when the effect is considered to be verging on marginal. Those sequences

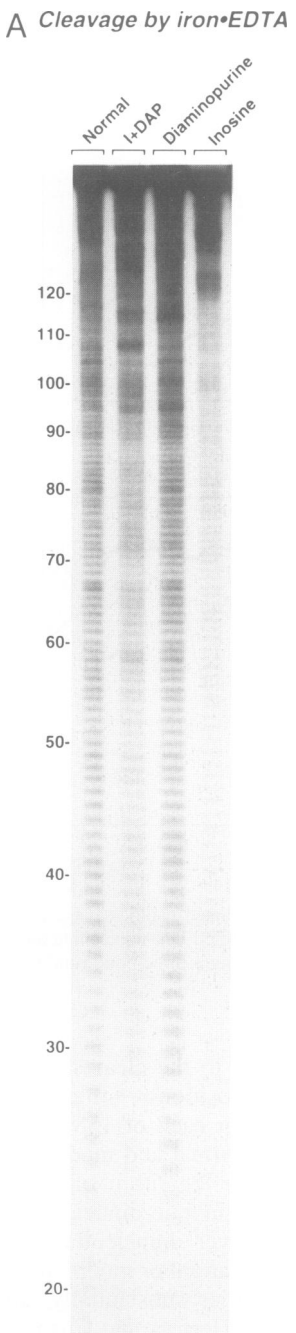
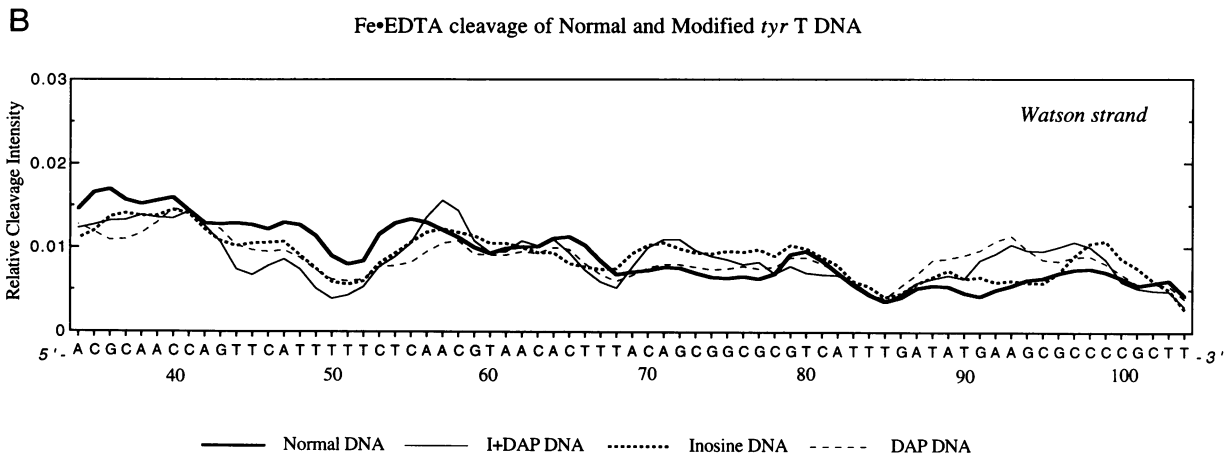


Fig. 5. (**A**) Autoradiograph and (**B**) cleavage plots comparing the susceptibility of the normal and modified DNAs to the Fe·EDTA-mediated hydroxyl radical cleavage reaction. The sequence shown on the axis corresponds to that of the Watson strand of the duplex containing natural bases. In the modified DNAs, adenine and/or guanosine residues are replaced by diaminopurine and/or inosine residues. Other details as for Figure 3.



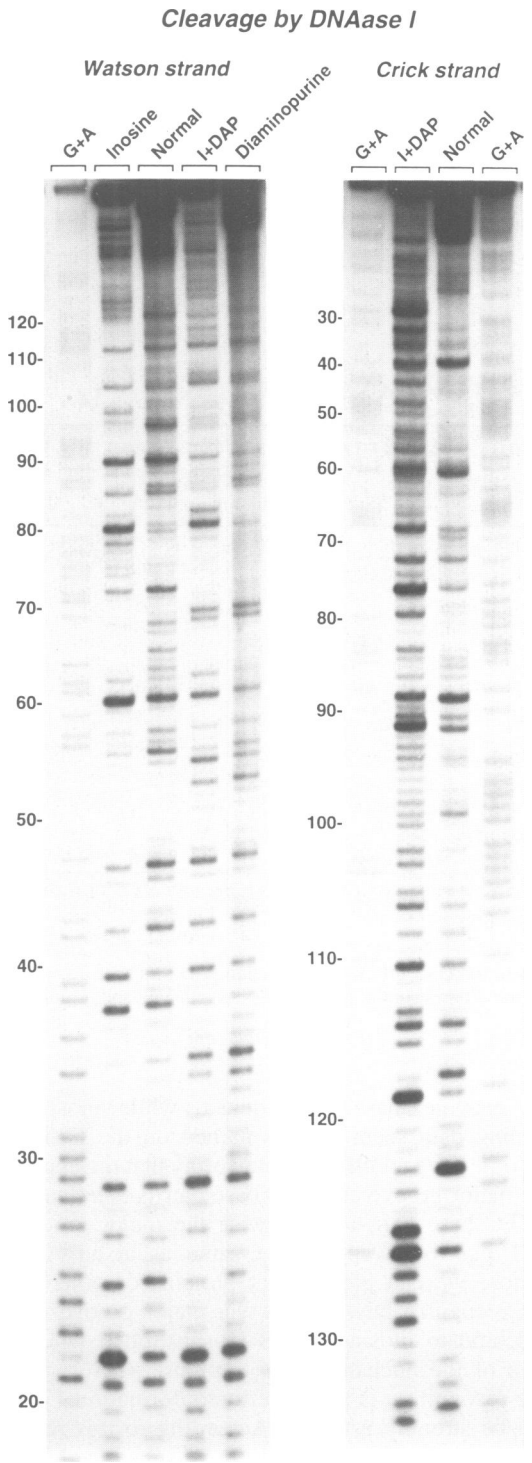


Fig. 6. Autoradiograph showing the DNase I cleavage pattern of *tyrT*(A93) DNA fragments containing natural nucleotides (Normal DNA), inosine residues in place of guanosine (Inosine DNA), DAP residues in place of adenine (DAP DNA) or inosine and DAP residues in place of guanosine and adenine, respectively (I+DAP DNA). Bands in the gels were assigned by reference to the purine markers (lanes G+A), taking into account the differences in mobility of the fragments due to the presence or absence of a 3′-phosphate group. Note that formic acid reacts to a similar extent with A and G whereas it reacts well with DAP and very weakly with inosine in the I+DAP DNA (compare the two lanes labelled G + A on the autoradiograph corresponding to the Crick strand-labelled DNA). Other details as for Figure 2.

where endonuclease cutting is strongly enhanced relative to the normal DNA can be grouped into two categories (Table I). In the first group all sequences without exception have a G·C base pair at position −3. In the second group, all sequences begin with a couple of A·T base pairs and contain at least four such pairs which are often part of a longer $(A{\cdot}T)_n$ tract. Such $(A{\cdot}T)_n$ clusters are known to constitute poor substrates for DNase I (Drew and Travers, 1984), which is generally attributed to their reduced flexibility and their narrow minor groove which disfavour the binding of the enzyme. It thus seems entirely plausible that the A→DAP substitution, which results in a noticeable variation of the minor groove width as evidenced by the uranyl photocleavage experiments detailed above, modifies the recognition and cleavage of these sequences by DNase I. Moreover, A→DAP replacement should prevent the formation of a B' structure at A-tracts (containing ⩾4 adenines in a row, not to be confused with AT-rich regions) which satisfactorily explains the enhanced cleavage by DNAase I at positions 33, 51 and 128. Consistent with this interpretation we observed that the cleavage within the GC-rich sequence at position 119 is considerably reduced when inosine residues are substituted for guanosines (Table II). In this case the G→I substitution is, to a limited extent, comparable with the conversion of a GC-rich sequence to an AT-rich sequence, i.e. to a poor DNase I substrate. It is interesting also that the cleavage at sequences containing one G·C base pair embedded in an AT tract is further reduced when that particular G·C pair is replaced by an I·C pair (see for examples positions 51, 68 in Table I and 62, 64 in Table II). Here the effect of introducing an I·C base pair (which mimics more or less an A·T base pair) is comparable with an extension of the AT tract.

The chief outcome of the data in Table I is that the substitution of an I·C pair for a G·C pair at position −3 strongly potentiates the catalytic activity of the endonuclease. By extension, if the −3 base pair is really critical for DNase I cleavage we should observe a reduced cleavage intensity when a −3 A·T pair is replaced with a DAP·T pair. The data in Table II reveal that this prediction is essentially correct since 12 out of 14 of the sequences reported in this table possess an A·T base pair at position −3. Therefore, we may conclude that indeed the nature of the base pair at position −3 plays a decisive role in the sequence selectivity of the DNase I cleavage reaction. More particularly, we can propose that it is the presence or the absence of a purine 2-amino group in this −3 base pair which, respectively, disfavours or stimulates the cleavage by the enzyme 3 bp ahead. These observations provide definitive confirmation that the base pair at position −3 is critical for the DNase I cleavage of DNA. However, upon careful inspection of the data in Figure 7 and Tables I and II, it seems also clear that while the −3 base pair constitutes one anchor site for the enzyme (perhaps the predominant one) the other proximal base pairs, notably those at positions −1 and +2, are also very important for the enzyme to recognize its target sequence. Our inferences are in all respects compatible with the DNAase I−DNA crystal structures (Suck and Oefner, 1986; Suck *et al.*, 1988; Weston *et al.*, 1992; Suck, 1994) and provide experimental evidence that it is the 2-amino group on the

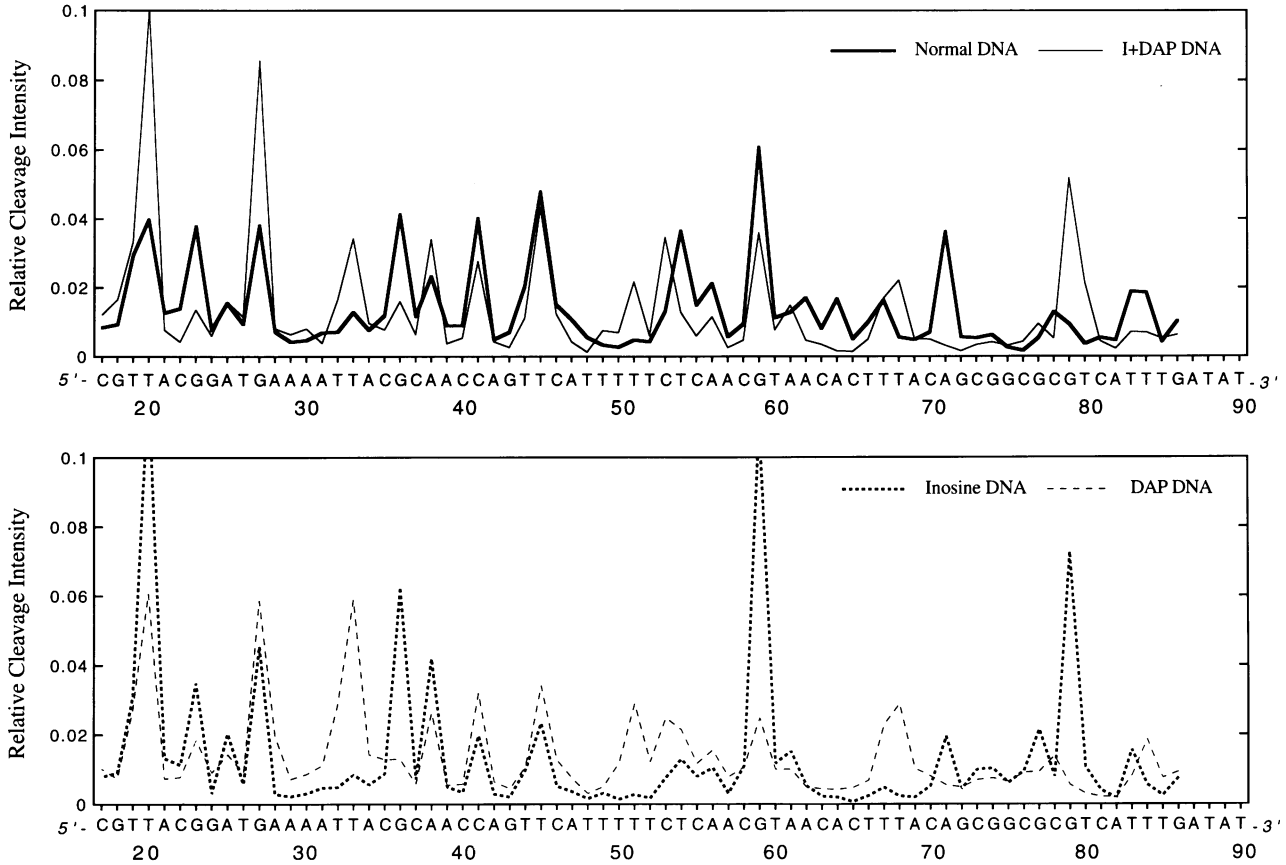


Fig. 7. Cleavage plots comparing the susceptibility of the normal and modified DNAs to the DNase I cleavage reaction. The sequence shown on the axis corresponds to that of the Watson strand of the *tyrT*(A93) fragment containing natural bases. In the modified DNAs, adenine and/or guanosine residues are replaced by diaminopurine and/or inosine residues.

purine nucleotide at position −3 which has a decisive influence on the cutting reaction.

Discussion

The present results show that the 2-amino group of guanine plays a dual role in DNA structure/conformation and recognition. It exerts a structural effect as a hydrogen bond donor in the minor groove, but it can also result in a widened minor groove and thus in a less electronegative potential as compared with sequences of A·T base pairs, where it is not present. Thus it seems to play a significant role in the sequence conformational microheterogeneity of the DNA helix, which agrees with previous studies using base analogues (Diekmann *et al.*, 1987, 1992; Koo and Crothers, 1987). Our observations on the effects of G→I and A→DAP substitution are fully consistent with altered patterns of minor groove width, although they do not rule out the possibility that changes in ligand binding may be directly due to the presence or lack of a hydrogen bond donor group, which is the simplest interpretation of results reported for the binding of several antibiotics (C.Bailly and M.J.Waring, in preparation). The findings with DNase I are at once decisive and complex: a dominant role for the base pair positioned three steps upstream of the cleavage site is shown, most probably involving a direct enzyme−base pair contact, while most of the remaining data seem best explained on the basis of an effect of groove width. It is noteworthy that the appearance of the major groove at an I·C base pair is identical to that at a G·C base pair, and likewise the major groove at a DAP·T pair must appear the same as at an A·T pair. Therefore we can predict that proteins or other ligands which recognize DNA solely via major groove contacts would bind to inosine and DAP-substituted DNA irrespective of the nucleotide substitutions, whereas proteins whose recognition is sensitive to conformational changes would be strongly affected. According to this reasoning it should be possible to use I+DAP-substituted DNA to discriminate between predominantly 'digital' as opposed to 'analogue' readout by major groove binding proteins.

Materials and methods

Chemicals and biochemicals

Ammonium persulphate, Tris base, acrylamide, *bis*-acrylamide, ultrapure urea, boric acid, tetramethylethylenediamine and dimethyl sulphate were from BDH. Formic acid, piperidine and formamide were from Aldrich. Photographic requisites were from Kodak. Bromophenol blue and xylene cyanol were from Serva. The labelled nucleoside triphosphate, [γ-^{32}P]ATP, was obtained from NEN Dupont. Restriction endonucleases *EcoR*I and *Ava*I (Boehringer Mannheim), *Taq* polymerase (Promega), DNase I (Sigma) and T4 polynucleotide kinase (Pharmacia) were used

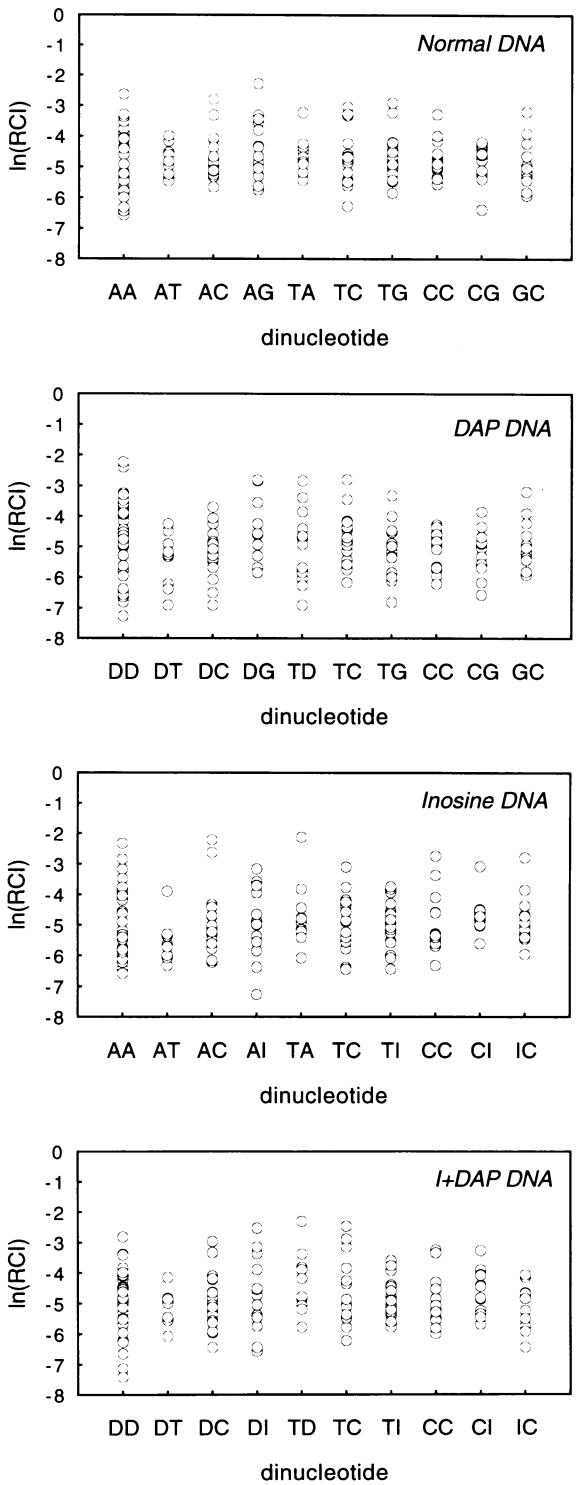


Fig. 8. Relative intensity of cleavage as a function of the cleaved dinucleotide. The natural logarithm of the cleavage intensity for 158 bonds examined in the *tyrT*(A93) DNA, comprising 70 in the Watson strand and 88 in the Crick strand, is plotted for each of the 10 unique dinucleotide steps.

according to the supplier's recommended protocol in the activity buffer provided. The primers, 5′-AATTCCGGTTACCTTTAATC and 5′-TCGGGAACCCCCACCACGGG having a 5′-OH or 5′-NH_2 terminal group, were obtained from the Laboratory of Molecular Biology, Medical Research Council, Cambridge. Checks were carried out to ensure that the primers blocked with a 5′-NH_2 group were free from contaminants and not labelled by the kinase. 1,10-Phenanthroline, 2,9-dimethyl-1,10-phenanthroline (neocuproine), 3-mercaptopropionic acid, sodium ascorbate and hydrogen peroxide were from Sigma Chemical Co. All other chemicals were analytical grade reagents, and all solutions were prepared using doubly deionized, Millipore filtered water.

Preparation, purification and labelling of DNA fragments containing natural and modified nucleotides

Plasmid pKMp27 (Drew *et al.*, 1985) was isolated from *E.coli* by a standard SDS–sodium hydroxide lysis procedure and purified by banding in CsCl–ethidium bromide gradients. Ethidium was removed by several isopropanol extractions followed by exhaustive dialysis against Tris–EDTA buffer. The purified plasmid was then precipitated and resuspended in appropriate buffer before digestion by the restriction enzymes. The 160 bp *tyrT*(A93) fragment used as a template was isolated from the plasmid by digestion with restriction enzymes *Eco*RI and *Ava*I. It is worth mentioning that this template DNA bore a 5′-phosphate due to the action of *Eco*RI and thus only the newly synthesized DNA (with normal or modified nucleotides) can be labelled by the kinase.

Polymerase chain reaction (PCR)

The protocol used to incorporate inosine and/or 2,6-diaminopurine residues into DNA is comparable to those previously used to incorporate 7-deazapurine or inosine residues with only a few minor modifications (Marchand *et al.*, 1992; Sayers and Waring, 1993). PCR reaction mixtures contained 10 ng of tyrT(A93) template, 1 μM each of the appropriate pair of primers (one with a 5′-OH and one with a 5′-NH_2 terminal group) required to allow 5′-phosphorylation of the desired strand, 250 μM of each appropriate dNTP (dTTP, dCTP plus dATP or dDTP and dGTP or dITP according to the desired DNA), and 5 units of Taq polymerase in a volume of 50 μl containing 50 mM KCl, 10 mM Tris–HCl, pH 8.3, 0.1% Triton X-100, and 1.5 mM $MgCl_2$. To prevent unwanted primer–template annealing before the cycles began, the reactions were heated to 60°C before adding the Taq polymerase (Bloch, 1991). Finally, paraffin oil was added to each reaction to prevent evaporation. After an initial denaturing step of 3 min at 94°C, 20 amplification cycles were performed, with each cycle consisting of the following segments: 94°C for 1 min, 37°C for 2 min and 72°C for 10 min. After the last cycle, the extension segment was continued for an additional 10 min at 72°C, followed by a 5 min segment at 55°C and a 5 min segment at 37°C. The purpose of these final segments was to maximize annealing of full-length product and to minimize annealing of unused primer to full-length product. The reaction mixtures were then extracted with chloroform to remove the paraffin oil, and parallel reactions were pooled. Several extractions with water-saturated *n*-butanol were performed to reduce the volume before loading the samples on to a 6% non-denaturing polyacrylamide gel. After electrophoresis for ~1 h, a thin section of the gel was stained with ethidium bromide so as to locate the band of DNA under UV light. The same band of DNA free of ethidium was excised, crushed and soaked in elution buffer (500 mM ammonium acetate, 10 mM magnesium acetate) overnight at 37°C. This suspension was filtered through a Millipore 0.22 μm filter and the DNA was precipitated with ethanol. Following washing with 70% ethanol and vacuum drying of the precipitate, the purified DNA was resuspended in the kinase buffer.

DNA labelling and purification

The purified PCR products were 5′-end labelled with [γ-^{32}P]ATP in the presence of T4 polynucleotide kinase according to a standard procedure for labelling blunt-ended DNA fragments (Maniatis *et al.*, 1982). After completion the labelled DNA was again purified by 6% polyacrylamide gel electrophoresis and extracted from the gel as described above. Finally, the labelled DNA was resuspended in the requisite buffer, i.e. in an acetate buffer, pH 6.5, for the uranyl photocleavage or in a Tris buffer, pH 7.0, for the hydroxyl radical and the DNase I cleavage reactions.

Uranyl photocleavage

Uranyl photocleavage of normal and modified DNA was performed according to previously described protocols (Nielsen *et al.*, 1988, 1990; Jeppesen and Nielsen, 1989). Briefly, the [^{32}P]DNA fragment (300–500 c.p.s.) in 100 μl 50 mM Na acetate buffer, pH 6.5, was incubated with 1 mM uranyl nitrate (diluted from a 100 mM stock solution which is stable at room temperature). The samples were irradiated for 30 min at room temperature using a Philips TL K 40W/03 fluorescent light tube emitting at 420 ± 30 nm. DNA was precipitated with 70% ethanol, 0.1 M Na acetate, pH 4.5, and analysed by gel electrophoresis in polyacrylamide/urea sequencing gels.

Table I. Phosphodiester bonds which are cleaved much better in doubly substituted (I+DAP) DNA than in normal DNA.

Nucleotide position	Strand	Sequence 5′→3′	RCI ratio[a]		
			G→I (Inosine DNA)	A→D (DAP DNA)	G→I + A→D (I+DAP NA)
		−3 +3			
79	Watson	GCG↓TCA CGC AGT	7.72	0.60	5.51
55	Crick	GTT↓GAG CAA CTC	2.90	1.21	3.65
120	Crick	GGC↓CTG CCG GAC	1.92	1.97	3.50
78	Crick	GAC↓GCG CTG CGC	2.54	0.53	3.07
105	Crick	CGG↓GAA GCC CTT	1.02	0.64	2.59
131	Crick	GCT↓TTT CGA AAA	0.45	1.41	2.57
20	Watson	GTT↓ACG CAA TGC	3.03	1.52	2.52
70	Crick	GCT↓GTA CGA CAT	1.40	1.11	1.55
93	Crick	CGC↓TTC GCG AAG	1.49	0.79	1.48
38	Watson	GCA↓ACC CGT TGG	1.79	1.13	1.46
51	Watson	TTT↓TCT AAA AGA	0.57	5.92	4.45
68	Watson	TTT↓ACA AAA TGT	0.41	5.16	3.96
128	Crick	TTT↓TTA AAA AAT	0.54	3.1	2.84
33	Watson	ATT↓ACG TAA TGC	0.66	4.56	2.66
53	Watson	TTC↓TCA AAG AGT	0.57	1.88	2.63
81	Crick	AAT↓GAC TTA CTG	1.85	1.29	2.47
27	Watson	ATG↓AAA TAC TTT	1.19	1.54	2.25

[a]Relative Cleavage Intensity ratio, comparing a given phosphodiester bond in the I+DAP DNA and the same bond in the normal DNA

EDTA·Fe^{II} cleavage

Hydroxyl radicals were generated via the reduction of hydrogen peroxide by the EDTA·Fe^{II} complex in the presence of ascorbic acid. The experimental procedure followed the protocol described by Burkhoff and Tullius (1987).

DNase I footprinting

DNase I experiments were performed essentially according to the original protocol (Low *et al.*, 1984). The digestion of the samples (6 µl) of the labelled DNA fragment dissolved in 10 mM Tris buffer, pH 7.0, containing 10 mM NaCl was initiated by the addition of 2 µl of a DNase I solution whose concentration was adjusted to yield a final enzyme concentration of ~0.01 unit/ml in the reaction mixture. The extent of digestion was limited to <30% of the starting material so as to minimize the incidence of multiple cuts in any strand ('single-hit' kinetic conditions). Optimal enzyme dilutions were established in preliminary calibration experiments. After 3 min, the digestion was stopped by freeze drying, samples were lyophilized, washed once with 50 µl of water, lyophilized again and then resuspended in 4 µl of an 80% formamide solution containing tracking dyes. Samples were heated at 90°C for 4 min and chilled in ice for 4 min before electrophoresis.

Electrophoresis and autoradiography

DNA cleavage products were resolved by polyacrylamide gel electrophoresis under denaturing conditions (0.3 mm thick, 8% acrylamide containing 8 M urea) capable of resolving DNA fragments differing in length by one nucleotide. Electrophoresis was continued until the bromophenol blue marker had run out of the gel (~2.5 h at 60 W, 1600 V in TBE buffer, BRL sequencer model S2). Gels were soaked in 10% acetic acid for 15 min, transferred to Whatman 3MM paper, dried under vacuum at 80°C, and subjected to autoradiography at −70°C with an intensifying screen. Exposure times of the X-ray films (Fuji R-X) were adjusted according to the number of counts per lane loaded on each individual gel (usually 24 h).

Quantitation by storage phosphor imaging

A Molecular Dynamics 425E PhosphorImager was used to collect data from storage screens exposed to the dried gels overnight at room temperature (Johnston *et al.*, 1990). Baseline-corrected scans were analysed by integrating all the densities between two selected boundaries using ImageQuant version 3.3 software. Each resolved band was assigned to a particular bond within the pBS fragment by comparison of its position relative to sequencing standards generated by treatment of the DNA with formic acid followed by piperidine-induced cleavage at the purine residues in DNA (G + A track).

Acknowledgements

The authors thank Dr Andrew Travers for critical reading of the manuscript. This work was supported by grants (to M.J.W.) from the Wellcome Trust, the Cancer Research Campaign and the Association for International Cancer Research; and (to P.E.N.) from The Danish National Research Foundation. C.B. thanks the INSERM and the Sir Halley Stewart Trust. We thank Drs C.Abell and J.Sagi for supplying the DAP triphosphate. The skilful assistance of Dean Gentle is gratefully acknowledged.

Table II. Bonds which are less susceptible to endonuclease cleavage in doubly substituted (I+DAP) DNA than in normal DNA

Nucleotide position	Strand	Sequence 5′→3′	RCI ratio[a]		
			G→I (Inosine DNA)	A→D (DAP DNA)	G→I + A→D (I+DAPDNA)
		−3 +3			
124	Crick	TAC↓TGG ATG ACC	0.79	0.19	0.09
64	Watson	ACA↓CTT TGT GAA	0.12	0.25	0.09
71	Watson	ACA↓GCG TGT CGC	0.54	0.15	0.09
134	Crick	TAA↓TGC ATT ACG	0.42	0.31	0.18
119	Crick	GCC↓TGC CGG ACG	0.30	0.87	0.19
87	Crick	ATA↓TCA TAT AGT	0.79	0.48	0.24
62	Crick	GTG↓TTA CAC AAT	0.41	0.47	0.33
54	Watson	TCT↓CAA AGA GTT	0.35	0.59	0.35
23	Watson	ACG↓GAT TGC CTA	0.92	0.49	0.36
101	Crick	AAG↓CGG TTC GCC	0.46	0.38	0.36
84	Watson	ATT↓TGA TAA ACT	0.30	1.00	0.37
36	Watson	ACG↓CAA TGC GTT	1.51	0.31	0.39
90	Crick	TTC↓ATA AAG TAT	0.42	1.12	0.43
59	Watson	ACG↓TAA TGC ATT	1.80	0.41	0.59

[a]Relative Cleavage Intensity ratio, comparing a given phosphodiester bond in the I+DAP DNA and the same bond in the normal DNA

References

Berman,M.L. and Landy,A. (1979) *Proc. Natl Acad. Sci. USA*, **76**, 4303–4307.

Bloch,W. (1991) *Biochemistry*, **30**, 2735–2747.

Burkhoff,A.M. and Tullius,T.D. (1987) *Cell*, **48**, 935–943.

Burkhoff,A.M. and Tullius,T.D. (1988) *Nature*, **331**, 455–457.

Diekmann,S., von Kitzing,E., McLaughlin,J., Ott,J. and Eckstein,F. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 8257–8261.

Diekmann,S., Mazzarelli,J.M., McLaughlin,J., von Kitzing,E., and Travers,A.A. (1992) *J. Mol. Biol.*, **225**, 729–738.

Drew,H.R. (1984) *J. Mol. Biol.*, **176**, 535–557.

Drew,H.R. and Travers,A.A. (1984) *Cell*, **37**, 491–502.

Drew,H.R., Weeks,J.R. and Travers,A.A. (1985). *EMBO J.*, **4**, 1025–1032.

Edwards,K.J., Brown,D.G., Spink,N., Skelly,J.V. and Neidle, S. (1992) *J. Mol. Biol.*, **226**, 1161–1173.

Heinemann,U. and Alings,C. (1989) *J. Mol. Biol.*, **210**, 369–381.

Herrera,J.E. and Chaires,J.B. (1994) *J. Mol. Biol.*, **236**, 405–411.

Jeppesen,C. and Nielsen,P.E. (1989) *Nucleic Acids Res.*, **17**, 4947–4956.

Johnston,R.F., Pickett,S.C. and Barker,D.L. (1990) *Electrophoresis*, **11**, 355–360.

Koo,H.S and Crothers,D.M. (1987) *Biochemistry*, **26**, 3745–3748.

Low,C.M.L., Drew,H.R. and Waring,M.J. (1984) *Nucleic Acids Res.*, **12**, 4865–4877.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Marchand,C., Bailly,C., McLean,M.J., Moroney,S. and Waring,M.J. (1992) *Nucleic Acids Res.*, **21**, 5601–5606.

Nelson,H.C.M., Finch,J.T., Luisi,B.F. and Klug,A. (1987) *Nature*, **330**, 221–226.

Nielsen,P.E., Jeppesen,C. and Buchardt, O. (1988) *FEBS Lett.*, **235**, 122–124.

Nielsen,P.E., Møllegaard,N.E. and Jeppesen,C. (1990) *Nucleic Acids Res.*, **18**, 3847–3851.

Nielsen,P.E., Hiort,C., Sönnichsen,S.H., Buchardt,O., Dahl,O. and Nordèn,B. (1992) *J. Am. Chem. Soc.*, **114**, 4967–4975.

Sayers,E.W. and Waring,M.J. (1993) *Biochemistry*, **32**, 9094–9107.

Sönnichsen,S.H. (1991). Thesis, University of Copenhagen.

Steitz,T.A. (1993). *Structural Studies of Protein–Nucleic Acid Interaction. The Sources of Sequence-Specific Binding*. Cambridge University Press, Cambridge.

Stubbe,J. and Kozarich,J.W. (1987) *Chem. Rev.*, **87**, 1107–1136.

Suck,D. (1994) *J. Mol. Recogn.*, **7**, 65–70.

Suck,D. and Oefner,C. (1986) *Nature*, **321**, 620–625.

Suck,D., Lahm,A. and Oefner,C. (1988) *Nature*, **332**, 465–468.

Travers,A.A. (1989) *Annu. Rev. Biochem.*, **58**, 427–452.

Travers,A.A. (1993) *DNA–Protein Interactions*. Chapman & Hall, London.

Tullius,T.D. (1988) *Nature*, **332**, 663–664.

Weston,S.A., Lahm,A. and Suck,D. (1992) *J. Mol. Biol.*, **226**, 1237–1256.

Received on December 7, 1994; revised on January 31, 1995