Volume 17 Number 7 1989 Nucleic Acids Research

The degree of ultraviolet light damage to DNA containing iododeoxyuridine or bromodeoxyuridine is dependent on the DNA sequence

Vincent Murray and Roger F.Martin

Molecular Science Group, Peter MacCallum Cancer Institute, 481 Little Lonsdale Street, Melbourne, Victoria 3000, Australia

Received December 13, 1988; Revised and Accepted February 27, 1989

ABSTRACT

The sequence selectivity of 300 nm ultraviolet light damage to DNA containing bromodeoxyuridine or iododeoxyuridine was examined on DNA sequencing gels. This was accomplished using a system where an M13 template was employed to direct synthesis of DNA in which thymidine was fully substituted with bromodeoxyuridine or iododeoxyuridine. The sites of damage corresponded to the positions of analogue incorporation. The extent of damage varied considerably at different sites of cleavage and ranged from the undetectable to over fifteen times the limit of detection (as assessed by laser densitometer scans). Strong damage sites had the "consensus" sequence CTT while sites of no detectable damage had the "consensus" sequence GTR. Bromodeoxyuridine and iododeoxyuridine had the same sites of damage although the extent of damage varied at different sites and bromodeoxyuridine damage was slightly greater than iododeoxyuridine. DNA containing thymidine was not damaged to any detectable level in this system with 300 nm ultraviolet light. The use of three closely related DNA sequences as targets for damage confirmed that (1) the sites of analogue incorporation are the cause of ultraviolet damage; and (2) that the neigbouring DNA sequence is an important parameter in determining the extent of damage. It is proposed that the microstructure of DNA - in particular the distance between the 5-carbon of the pyrimidine base (which is attached to the halogen) and hydrogen on the 2' carbon of the 5'-deoxyribose - ultimately determines the degree of cleavage with large distances giving a small degree of damage and smaller distances a large degree of damage.

INTRODUCTION

DNA containing iododeoxyuridine (IUdR) or bromodeoxyuridine (BrUdR) is more sensitive to ultraviolet light than DNA containing thymidine (for reviews see refs. 1,2). This sensitivity has been exploited in a number of techniques for cell and molecular biology. The mechanism for this UV sensitivity appears to consist of the following: the 2' carbon of the deoxyribose on the 5'-side of the halogen-deoxyuridine is in close spatial proximity to the carbon-halogen bond; the ultraviolet light splits the carbon-halogen bond to produce a halogen radical and a uracilyl radical; the uracilyl radical reacts with the 5' deoxyribose; several reaction products result including base and deoxyribose damage, single-strand breaks and

double-strand breaks. On treatment with alkali certain base and deoxyribose damage is converted to single-strand breaks and this amounts to approximately four times the direct single-strand breaks. Double-strand breaks account for about 1-5% of the total strand breaks.

The sequence specificity of UV damage to IUdR and BrUdR containing DNA was investigated using a novel system. Single-stranded M13 templates were copied using a deoxyribosetriphosphate mix in which iododeoxyuridine (dIUTP) or bromodeoxyuridine triphosphate (dBrUTP) was completely substituted for dTTP. Thus the synthesised strand is fully substituted with IUdR or BrUdR. After UV irradiation, the damage products can be run on DNA sequencing gels and the sites of damage can be pinpointed to the nearest base pair. Since versatile M13 cloning vectors are available, the sequence specificity of any DNA sequence can be analysed. In this study a comparison was made between three closely related human alpha DNA sequences (3-6).

MATERIALS AND METHODS

i) Materials

All chemicals used were of analytical grade. Klenow polymerase I and [alpha ^{32}P] dATP (3000 Ci/mmole) were obtained from BRESA, Australia. dATP, dGTP, dCTP and dTTP were from Boehringer-Mannheim and dIUTP and dBrUTP from Pharmacia.

ii) Recombinant M13 clones

Details concerning the construction of M13 clones alpha 82, alpha 22 and alpha 32 can be found in Murray & Martin (6,7). Briefly they are 340 bp inserts derived from the human tandem repeat called alpha RI-DNA which is present in about 50,000 copies per haploid genome and the "consensus" sequence has been fully determined (3,4,5). The three clones are closely related in DNA sequence and the base substitutions are shown in Table I relative to the alpha 32 DNA sequence.

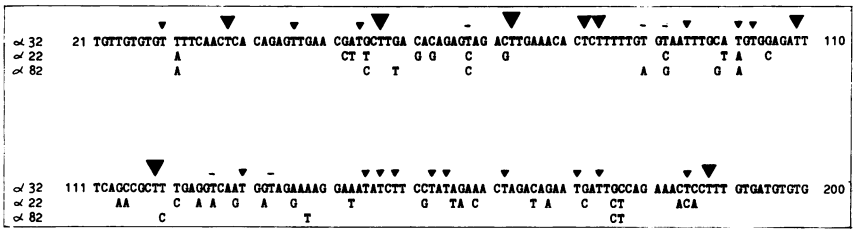
iii) Synthesis of DNA containing dIUTP or dBrUTP

The synthesis of DNA fully substituted with either iodinated or brominated uracil was accomplished using a modification of the Sanger dideoxy sequencing procedure. In a reaction volume of 10 µl, there was 0.5ng 15 bp primer which had been annealed to 500 ng of single-stranded M13 clone, 7mM $MgCl_2$, 7mM tris-HCl, pH 7.5, 50mM NaCl, 10µM dGTP, 10µM dCTP, 2 BRESA units of Klenow polymerase I, 1µCi of [alpha-^{32}P] dATP. After incubation at room temperature for 10 min , 1µl of 500µM dATP and 1µl of either 1mM dIUTP or 1mM dBrUTP was added and incubated at 37°C for 20 min . The reaction was stopped by the addition of 1µl of 250mM EDTA, pH 8.0.

iv) UV irradiation

Three µl aliquots of the DNA synthesis reaction were then irradiated for 10 minutes in Eppendorf tubes positioned at the surface of a UVB transilluminator containing four BLE-IT158 lamps (Spectronics Corporation) with a wavelength peak at about 300 nm. Measurements with a Spectoline DM-

TABLE 1



The DNA sequences of clones alpha 82, alpha 32 and alpha 22 are described in Murray and Martin (1987). Large triangles indicate strong damage sites, medium triangles - medium damage, small triangles - weak damage, dashes indicate no detectable damage in clone alpha 32. The sequence is shown 5' to 3'.

300N ultraviolet meter indicated that the fluence was 30,600 J/m^2, without correction for attentuation by the walls of the Eppendorf tubes).

v) DNA sequencing reactions

DNA sequencing reactions were run on the same gel as the irradiated samples to determine the sites of damage. The dideoxytriphosphate sequencing reactions were carried out as described by Sanger and co-workers (8,9). In other sequencing reactions 100μM dIUTP or 100μM dBrUTP was substituted for dTTP.

vi) Densitometer analysis

The autoradiographs of sequencing gels were scanned by a Zeineh soft laser densitometer (Biomed Instruments) and digitally analysed by Zeineh software (Biomed Instruments). Peak heights were determined after subtraction of control values from unirradiated samples.

RESULTS

i) Synthesis of DNA substituted with BrUdR or IUdR

The synthesis reaction, which is a modification of the Sanger dideoxytriphosphate DNA sequencing procedure, consists of two stages. In the first stage the DNA is labelled in the presence of only 3 dNTPs at the 5' end of the molecule (immediately after the primer) with [alpha ^{32}-P] dATP for about 10 bp - the three dNTPs used dATP, dCTP and dGTP contain impurities which are able to substitute for dTTP in the reaction for a small number of bps. The second stage consists of a chase period in which the addition of dIUTP or dBrUTP results in the full complement of 4 dNTPs and extensive synthesis of DNA. In this second stage the synthesised DNA is fully substituted with dIUTP or dBrUTP instead of dTTP. Routinely the extent of synthesis is greater than 3000 bp as assessed by comparison with labelled pBR322 (see figure 1).

Analysis of the labelled DNA on DNA sequencing gels show that the incidence of pauses in DNA 30-500 bps in length produces a significant in-

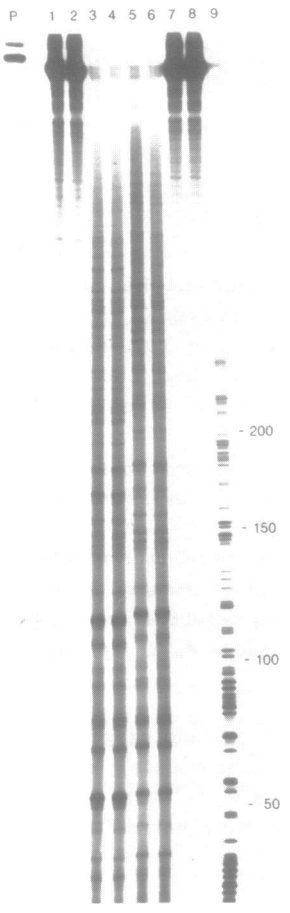


Figure 1 Autoradiograph of a DNA sequencing gel showing UV damage to IUdR- and BrUdR - containing DNA

The DNA in the various tracks are derived from reactions as described in the materials and methods. Clone alpha 32 was used. The variations are as follows:- lanes 1-4 contain BrUdR, lanes 5-8 - IUdR; lanes 3-6 are treated with UV light while all other lanes are not; lane 9 is the dideoxy T sequencing track; lane P is EcoRI-cut pBR322 which has been 3'-end labelled.

track background but this is sufficiently low to enable detection of damage. Other systems, for example - labelled restriction fragments of plasmids - would have a lower background but do not enable experiments of the sort described in this paper to be carried out.

ii) **UV damage to DNA containing IUdR**

The DNA in lanes 5,6,7 and 8 of figure 1 are aliquots from the same DNA

synthesis reaction. This DNA, which contains IUdR, is either irradiated with UV light at 300 nm (lanes 5 and 6) or not irradiated (lanes 7 and 8). The irradiated samples are extensively damaged. However, there is wide variation in the intensity of damage at individual sites (see section iv for densitometer analysis). The bands are generally "fuzzy" and not as sharp as sequencing reactions on the same gel. The position of the hotspots is related to the sites of incorporation of IUdR which is incorporated where T is present in normal DNA. The hotspots appear to migrate approximately 2 bps below dideoxy-T bands (lane 9). Where there are runs of Ts in the sequence (eg the five Ts at bps 84-88), the site of damage appears to be mainly at the T nearest the 5' end - towards the bottom of the gel.

An important control is the UV irradiation of DNA containing the normal T with no IUdR present. As shown in lanes 3 and 4 in figure 2, no detectable damage is observed when the normal T was present.

iii) Effect of Dithiothreitol

The presence of 10mM dithiothreitol (which is routinely present in dideoxytriphosphate sequencing reactions as a reducing agent) was found to reduce the degree of UV damage in iodo-substituted DNA by about three-fold. A dithiothreitol concentration curve revealed that below 0.02 mM, no dithiothreitol inhibition was apparent. The absence of dithiothreitol did not affect the extent of DNA synthesis and therefore, in the experiments described in this paper, no dithiothreitol was present.

iv) Densitometer analysis

By the use of a laser scanning densitometer, the peak heights of the damage sites were quantified. Peak heights were used because of greater reproducibility when a high background is present. This high background means that the smaller densitometer values are less accurate than the larger values. For this reason all samples were run in duplicate and densitometer values are an average of these duplicates. In table 2 it can be seen that the densitometer values for IUdR vary over a large range - the limit of detection is 0.07 relative to bp 119 and therefore by comparison with the largest peak height at bp 73 - more than a 15 fold range of damage is detected. For BrUdR the range is more than 17-fold.

Densitometer analysis allows the damage sites to be assigned into various categories of damage. Relative to bp 119, strong damage sites have values greater than 0.8, medium between 0.5 and 0.8, weak 0.07 to 0.5 and no detectable damage - less than 0.07. With IUdR for the region bp 29 to 190 in clone alpha 32 there are 3 strong sites, 5 medium, 14 weak and 5 with no significant damage.

v) UV damage to DNA containing BrUdR

In figure 1 a comparison is depicted of UV damage to DNA containing BrUdR (lanes 1-4) and IUdR (lanes 5-8). The DNA damage is similar for the two types of DNA. The damage products with BrUdR migrate slightly faster than with IUdR. Every damage site which is present with IUdR, is also present with BrUdR although differences in intensity occur.

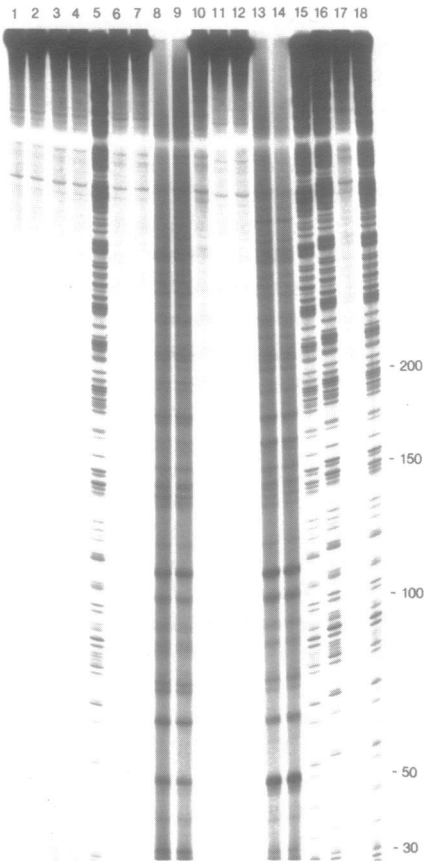


Figure 2 **Autoradiograph of a DNA sequencing gel showing UV treatment of IUdR-, BrUdR- and T- containing DNA**

The DNA reactions are described in the Materials and Methods section. Clone alpha 32 was used. The variations are as follows:-
lanes 1-5 and 16 contain T, lanes 6-10 and 17 - IUdR, and lanes 11-15 and 18 - BrUdR; lanes 3,4,8,9,13,14 are treated with UV light while all other lanes are not; lanes 5 and 16 are the normal dideoxy T DNA sequencing track; in lanes 10 and 17 IUdR replaces T in the dideoxy T reaction, while in lanes 15 and 18 BrUdR replaces T in the dideoxy T reaction.

A densitometer analysis reveals (table 2) that the ratios of the peak heights with BrUdR and IUdR ranges from 0.51 to 1.77. The average for the 23 sites examined is 1.29 which apparently reflects the greater reactivity of BrUdR - most of the ratios are greater than 1.0.

Table 2 Densitometer-derived intensity of UV damage to DNA containing IUdR or BrUdR using clone alpha 32

base pair	IUdR intensity relative to bp119	BrUdR to IUdR ratio
30	0.35	1.63
38	0.54	0.51
46	0.22	1.0
54	0.18	1.29
57	0.96	1.65
68	-	-
73	1.06	0.79
82	0.78	0.76
84	0.65	1.20
90	-	-
92	-	-
95	0.43	1.26
101	0.30	1.69
103	0.18	1.29
109	0.50	1.59
119	1.0	1.31
126	-	-
130	0.31	1.59
133	-	-
145	0.18	1.28
147	0.31	1.00
149	0.37	1.55
153/155	0.48	1.08
162	0.39	1.57
171	0.33	1.33
174	0.41	1.77
185	0.37	1.05
188	0.50	1.41

Dashes indicate that no damage is detectable at these sites.

The mobility of the BrUdR damage sites is slightly faster than with IUdR. This could occur for two reasons. First, DNA containing BrUdR could migrate at a different rate in the sequencing gel. Second, UV damage could result in a different reaction product with BrUdR which migrates with a different mobility. An experiment was designed to distinguish between these two possibilities. The dideoxy T sequencing reactions were performed with the following alterations - the reactions contained dideoxy TTP, dGTP, dCTP and dATP with dBrUTP, dIUTP or the normal dTTP. Figure 2 shows the mobility of the products from these reactions. In lanes 15 and 18, the presence of

Table 3 DNA sequences present at strong and medium damage sites, and at sites with no detectable damage in clone alpha 32.

			bp
IUdR	Strong	GCTTG	57
		ACTTG	73
		GCTTT	119
	Medium	ACTCA	38
		ACTCT	82
		TCTTT	84
		GATTT	109
		CCTTT	188
	consensus	RCTTG	
		T	
BrUdR	Strong	GCTTG	57
		GCTTT	119
	Medium	ACTTG	73
		TCTTT	84
		GATTT	109
		GATTG	174
		CCTTT	188
	consensus	RCTTG	
		T	
IUdR and BrUdR	no damage	AGTAG	68
		TGTGA	90
		TGTAA	92
		GGTCA	126
		GGTAG	133
	consensus	NGTRR	

The bases are shown 5' to 3'. R represents the purines adenine and guanine. N is any base. The letter T is shown for convenience where it is substituted with IUdR or BrUdR. The third of the nucleotides is presumed to correspond to the site of initiation of cleavage in the halogenated analogue.

BrUdR causes the bands to migrate faster than with IUdR or T which both migrate at approximately the same position. (For both IUdR and BrUdR, cleavage product is always approximately 2bps faster than the corresponding T band in the dideoxy T track). Thus, the first explanation is correct.

vi) Comparison of DNA sequences at damage sites

The neighbouring sequences at the damage sites were examined to

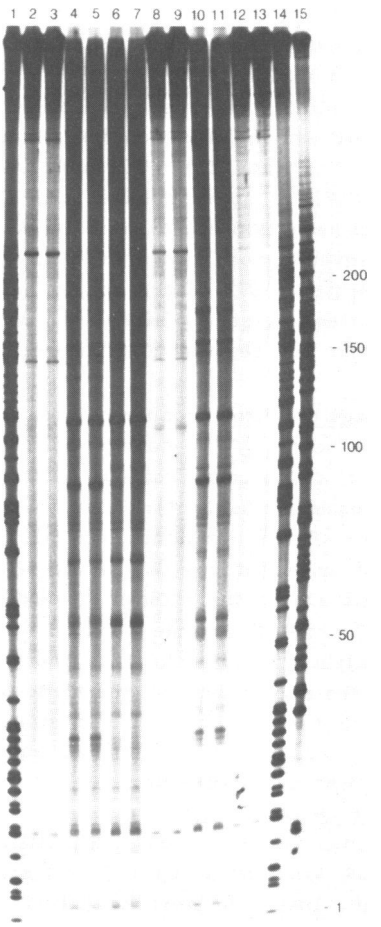


Figure 3 Autoradiograph of a DNA sequencing gel showing UV damage to IUdR-containing DNA in three cloned DNA sequences
The DNA reactions are described in the Materials and Methods section and the variations are as follows:-
lanes 1-5, 14,15 is clone alpha 82, lanes 6-9 - clone alpha 32 and lanes 10-13 - clone alpha 22; lanes 4-7,10,11 are treated with UV light while all other lanes are not; lanes 1 and 14 are the dideoxy T and lane 15 is the dideoxy A sequencing tracks.

determine whether similar sequences were present at sites with a high degree of cleavage or conversely those with a low degree of cleavage. In table 3, a listing is made of those sequences present at strong and medium damage sites; and also where no damage is detectable. (For ease of discussion T is used

where it is substituted with IUdR or BrUdR and when a DNA sequence is quoted, the third nucleotide corresponds to the presumed site of initiation of cleavage in the halogenated analogue).

For both IUdR and BrUdR the same "consensus" sequence -RCTTG/T- is present at the sites of strong and medium damage. The sites of no significant damage are the same for IUdR and BrUdR and have the same "consensus" sequence -NGTRR. At each base pair in the two "consensus" sequences, the designated base is present in at least 70% of the cases.

The "consensus" sequence does not always give rise to the expected degree of damage. The sequence CTT occurs on six occasions in the examined sequence and for IUdR, three are strong sites, two medium and one weak site. For the five occasions that the sequence GTR occurs, four have no detectable damage and one is weak.

vii) The sites of UV damage in three closely related DNA sequences

The sequence specificity of UV damage to IUdR-containing DNA was examined in three alpha RI-DNA clones from bps 29 to 150. These clones are closely related in DNA sequence (table 1). Relative to clone alpha 32, clone alpha 82 has a low number of base substitutions (10 out of 121) while alpha 22 has a larger number of substitutions (21 out of 121).

This type of analysis allows the effects of random base substitutions to be examined. Figure 3 shows UV damage to IUdR - containing DNA for the three clones alpha 82, alpha 32 and alpha 22. There is a wide variation in the degree of damage at the same sites for the three clones. The densitometer analysis is summarised in table 4 for all the sites of damage - the base substitutions are also indicated. Note that capital letters indicate potentially significant base substitutions.

Sites where a G,C or A is replaced by T

There are six occasions where a T base substitution occurs relative to clone alpha 32. The T substitution at bp 138 in clone alpha 82 results in a strong damage site (nine times the limit of detection) while no damage is detectable for clones alpha 32 and alpha 22.

The T(53) substitution in clone alpha 22 moves the site of damage one base pair towards the bottom of the gel as expected.

The T(100) a(101) alterations in clone alpha 22 effectively move the single T one base to the 5' end. As expected the site of damage also moves one base towards the bottom of the gel.

The substitution of T(143) in clone alpha 22 does not produce any detectable damage. This could be because the 2 neighbouring bases are purines.

The changes t(55) in clone alpha 22 and t(59) in alpha 82 are similar in that they extend the 3' end of a run of Ts. There is no change in the position of damage.

Sites where a T is replaced by A,G or C

There are twelve occasions where an A,G or C base substitution replaces the T in clone alpha 32. The a(31) substitution has a similar effect in both

Table 4 Densitometer-derived intensity of UV damage to DNA containing IUdR for three clones relative to clone alpha 32

base pair	alpha 82 D	alpha 82 BS	alpha 32 D	alpha 22 D	alpha 22 BS
30	2.42		1.0	3.29	
		a(31)			a(31)
32	*(1.0)		-	*(1.24)	
38	1.33		1.0	1.08	
46	0.68		1.0	1.0	
					c(52)
53	-		-	*(0.34)	T(53)
54	1.12		1.0	-	
		c(55)			t(55)
57	0.47		1.0	0.74	
		t(59)			g(61), g(63)
68	-	C(68)	-	-	c(68)
					g(72)
73	0.79		1.0	0.53	
82	0.84		1.0	0.77	
84	0.84		1.0	0.73	
90	-	A(90)	-	-	
92	-	G(92)	-	-	C(92)
95	1.48		1.0	1.48	
		g(99)			
100	-		-	*(0.61)	T(100)
101	-	A(101)	1.0	-	a(101)
103	1.0		1.0	1.0	
					c(105)
109	0.80		1.0	0.67	
					a(114), a(115)
119	0.81		1.0	0.76	
		c(120)			c(121), a(124)
126	-			-	A(126)
					g(129)
130	0.65		1.0	-	
					a(132)
133	-		-	-	
					g(136)
138	*(0.65)	T(138)	-	-	
143	-		-	-	T(143)

bBS - base substitution relative to the alpha 32 DNA sequence. A higher case letter (A,G,C or T) indicates that the base substitution is at the site of damage. A T indicates that a potential new site is created while A,G or C implies that a site is lost. A lower case t indicates a base substitution in a run of thymidines but not at the 5' end (where damage is detected). A lower case a,g or c is a base substitutions not at a damage site.

D - densitometer measurement relative to clone alpha 32 (which is defined at 1.0). A - indicates that no damage is detectable. Where no damage is found for clone alpha 32, a * indicates detectable damage for other clones; for detectable damage only in alpha 82 and alpha 22, the number in brackets in relative to clone alpha 82; for detectable damage only in alpha 82 or alpha 22, the number in brackets is relative to bp 119.

clones alpha 82 and alpha 22. The four Ts in clone alpha 32 are split into two halves and bp 32 becomes a site of damage.

The base substitutions A(101) in clone alpha 82 and A(126) in clone alpha 22 have the same effect. The removal of T from the sequence results in damage at this site being reduced to below the limit of detection.

The changes C(68) in both clones, A(90) in clone alpha 82, C(92) in clone alpha 22, and G(92) in clone alpha 82 are similar. Since there is no detectable damage at all of these sites for clone alpha 32, no discernible effect of these substitutions can be observed.

The substitutions c(120) in clone alpha 82 and c(121) in clone alpha 82 are discussed in the next section.

Base substitutions in neighbouring sequences

The two substitutions in clone alpha 22, g(72) and g(129), test the hypothesis of "consensus" sequences for strong or weak damage. At the strong damage site, bp 73, a strong "consensus" sequence of RCTTG occurs in clone alpha 32 but is changed to RGTTG in clone alpha 22. As predicted by the hypothesis a reduction in damage intensity occurs (by about 50%). At the weak damage site, bp 130, the sequence AATGG in clone alpha 32 is changed to AGTGA which is a "consensus" sequence for no detectable damage. Again as predicted by the hypothesis a reduction in the intensity of damage occurs (to below the limit of detection).

For the next three damage sites to be examined in this section, at bps 57, 95 and 119 (it also applies to bp 73 above), the sites of damage are all strong sites of damage and hence the densitometer values are more accurate at these positions.

In clone alpha 32 the damage site at bp 57 is the strong "consensus" sequence of GCTTG. It is changed to CCTTT in clone alpha 82 and TCTTG in clone alpha 22. The intensity of damage is reduced by 53% in clone alpha 82 and 26% in clone alpha 22. These reductions would appear to be caused by substitution of a purine at the second base 5' in the "consensus" sequence by pyrimidines in clones alpha 82 and alpha 22.

The damage site at bp 119 contains the sequence GCTTT (a strong "consensus" sequence) in clone alpha 32, GCTCT for clone alpha 82, and GCTTC for clone alpha 22. The densitometer peak height is reduced by 19% in clone alpha 82 and 24% in clone alpha 22. Thus it would appear that a C in the first base 3' is slightly less preferred than T, and C in the second base 3' instead of T also results in slightly less damage.

At the damage site at bp 95 the nearest base substitutions are 3 bases 5' to the damage site. Clone alpha 32 has T(92), clone alpha 82 - G(92), and clone alpha 22 - C(92). Compared to clone alpha 32, the intensities for clones alpha 82 and alpha 22 are both increased by 48%. Thus base substitutions are affecting damage over a range of 3 bps.

DISCUSSION

In this paper UV damage to DNA containing IUdR or BrUdR has been examined using a novel technique. Using this technique which involves single

stranded M13 and a primer, thymidine in the synthesised strand is completely substituted with IUdR or BrUdR. After UV treatment the sites of damage can be located to the exact base pair using DNA sequencing gels.

The extent of UV damage at the IUdR sites varies over a large range - from no detectable damage to strong hotspots - a range of over fifteen-fold as assessed by densitometer analysis. The extent of damage was correlated with neighbouring DNA sequences (see below). UV damage was dependent on analogue incorporation since no damage was detectable if the normal thymidine base was present instead an analogue.

There is a previous report on the sequence specificity of UV damage to DNA partially substituted (50%) with BrUdR (10). They used their system to examine repressor - DNA interactions and did not quantitatively analyse or discuss in any detail the extent of UV-induced DNA damage. However, close examination of their data reveals that they also found that damage was associated with each BrUdR site and a large variation in the extent of damage was apparent.

In this study the products of UV cleavage were found to move approximately 2 bps faster than the corresponding dideoxy-thymidine bands. If damage to the deoxyribose of the nucleoside 5' to the halogen substituted nucleoside left a 3'phosphoryl terminus, the 5' labelled fragment would be expected to have an electrophoretic mobility about 3 nucleotide positions faster than the fragment terminating in dideoxy T at the substituted site, assuming that species with 3'OH-and 3'deoxy-terminii have similar mobilities. Unfortunately the "fuzziness" of the damage site bands and high in-track background make it impossible to even speculate about the terminii left by the photolytic reaction; indeed multiple species may be present; gamma irradiation produces both 3' phosphate and 3' phosphoglycolate terminii (11).

A puzzling aspect to the sites of DNA damage occurs at a run of Ts eg bps 84-88. The main cleavage band is at the 5' end. This implies that interaction between neighbouring IUdR or BUdR bases occurs. Free radical transfer along the DNA in a 5' direction is a possible explanation for these observations.

As shown by other workers (2) the presence of a sulphydryl reagent - in this study dithiothreitol - greatly reduces the yield of UV damage by acting as a free radical scavenger. The 3-fold reduction found in this paper in not as great as 4-to 11-fold reduction others have found with cysteamine using other systems.

A comparison between the UV damage to DNA containing IUdR or BrUdR was made. The sites of damage were in exactly the same position after allowing for the different mobility of DNA containing BrUdR. However, the extent of damage differed slightly at the various damage sites. DNA containing BrUdR was more sensitive than IUdR-DNA by about 30%. The difference was probably greater because the UV absorbance maximum is 280 nm for BrUdR and 287 nm for IUdR (2) while irradiation was at 300 nm.

On comparing the DNA sequences present at strong and medium damage, a

"consensus" sequence emerged - 5' RCTTG/T 3'. Both IUdR- and BrUdR-DNA had the same "consensus" sequence although there were differences in the particular sites. Also a "consensus" sequence emerged for the sites where no IUdR or BrUdR associated damage was detectable - GTRR. These "consensus" sequences only give an indication of the likely extent of damage eg. of the six CTT sites present, 3 were strong sites, 2 medium and one weak; and of the five GTR sequences present, 4 had no detectable damage and one weak. On examination of the data of Ogata and Gilbert (10), it appears that their data is consistent with the "consensus" sequences found in this paper.

The use of three closely related sequences enabled the effect of random base substitutions to be analysed. This type of experiment is very powerful because it allows direct tests of hypotheses derived from examination of a single sequence.

The first hypothesis that was tested and confirmed was that UV damage is directly caused by incorporation of IUdR or BrUdR in place of T. New damage sites arose where none previously existed when a T base substitution replaces A,G or C at bps 53,100 and 138. Damage sites are abolished at sites where T is replaced by A,G or C at bps 101 and 126. However, the substitution T (143) did not result in a detectable damage site although this could be caused by neighbouring purine sequences.

The second hypothesis that was tested using three closely related sequences, was that the "consensus" sequences are important parameters in determining the degree of cleavage. The positive influence of a purine at the second position 5' to the damage site on the strong "consensus" sequence was confirmed at the bp 57 damage site. The negative effect on the degree of damage of a G at the first base 5' was shown at bp 73 and 130. The slightly negative effect of C instead of T at the first and second base 3' to the damage site was described at bp 119.

Factors other than the five bps at the "consensus" sequence also contribute to the degree of damage. Longer range effects of bases outside the "consensus" sequence were also found in this paper. The degree of damage at bp 95 was influenced by base substitutions at least 3 bp away.

The third concept that was examined was the observation that at runs of consecutive Ts, the damage site is mainly found at the 5'T. This observation was confirmed at base substitutions t(31), T(53), and T(59).

An explanation for the variation in the degree of UV damage to DNA containing IUdR (or BrUdR) is a difference in effective UV cross-section for iodine (or bromine) for photons of that energy. However, this explanation would possibly account for only a small proportion of the greater than fifteen-fold variation at different damage sites.

A simple hypothesis to explain the variation in the extent of UV damage to IUdR or BrUdR containing DNA and its sequence specificity is that the distance between the 5-carbon of the halogenated uracil and the hydrogen on the 2'-carbon of the 5'-deoxyribose is the most important parameter; a large distance results in a low degree of damage and a short distance a high degree

of damage. A bromine or iodine radical is also produced as well as the uracilyl radical and their proximity to the deoxyribose could contribute to the observed damage. However, currently the uracilyl radical is thought to be the most important radical species. If it is accepted that the hydrogen on the 2' deoxyribose is the target for the free radical because of its close spatial proximity to the 5-carbon of uracil (although there is no direct evidence for this assumption), then the distance between the two reacting species will obviously be an important determinant of the extent of the reaction.

The microstructure of B-DNA (the main form of DNA found in normal aqueous solutions) as determined by X-ray crystallography of a dodecamer oligonucleotide (12,13) differs from the regular Watson-Crick structure. Variations in propeller twist, helical twist, relative sliding of base pairs, roll angles between base pairs, and the conformation of the sugar-phosphate backbone can occur which alter the microstructure of DNA (for review see ref. 14). This microvariation is a result of interactions with neighbouring base pairs so that the conformational energy is minimised (particularly optimising base stacking energy).

Unfortunately in trying to examine the results in this paper in relation to the X-ray crystallographic data, the only published structure of an oligonucleotide with B-type DNA geometry is the dodecamer CGCGAATTCGCG. This sequence does not contain the dinucleotides GT or CT which are the critical sequences for this study. Ideally an X-ray crystal or NMR structure of an oligonucleotide containing the dinucleotides G.IUdR, C.IUdR, G.BrUdR or C.IUdR should be obtained.

However, analysis of the dodecamer structure revealed certain principles which enable predictions to be made about the "consensus" sequences CTT (strong) and GTR (no damage). The obvious differences between the two sequences is that one is three pyrimidines and the other is an alternating purine- pyrimidine-purine. This latter structure has been discussed in detail in the dodecamer oligonucleotide (15). Alternating purine pyrimidines present a structural problem because of steric hindrance between opposing purines and an alteration in structure occurs to reduce the steric hindrance. This alteration in structure can result in at least one of the following:- a lower propeller twist, moving the purine out of the stack by displacing one of the base pairs along its long axis, opening the roll angle between the base pairs, and decreasing the local rotation angle. Thus the two "consensus" sequences should have very different microstructures. An obvious prediction from our experimental results is that these differences in microstructure result in a larger distance between the 5-carbon of the halogenated uracil and the hydrogen on the 2'-carbon of the deoxyribose for the sequence GTR than for CTT.

The concept that microvariation in DNA structure is caused by neighbouring base sequences can also explain variations in the degree of damage as a result of base substitutions. These include base substitutions

in the "consensus" sequence at a damage site, and also the longer range effect of a base substitution more than two base pairs from the damage site. Likewise, the highly preferred cleavage of the 5' T of a run of Ts might also be explained by conformational features of consecutive halogenated pyrimidines.

In our previous studies on the sequence specificity of DNA damaging agents, long range effects of base substitutions were also found. For bleomycin (16) a twelve base pair effect was observed and with [^{125}I]-labelled Hoechst 33258 (17) a seven base pair effect. These long range effects were also thought to be due to a base substitution affecting the DNA microstructure at a damage site.

An interesting contrast to the sequence specificity of UV damage to DNA containing IUdR or BrUdR is the sequence specificity of bleomycin. The two DNA damaging agents are almost opposite in their specificities. Alternating purine pyrimidine sequences are strongly damaged and runs of pyrimidines are not significantly damaged by bleomycin (18,19,16) whereas the opposite is true for UV damage to DNA containing IUdR or BrUdR. Different extreme DNA microstructures are probably recognised in both cases.

Acknowledgements

This project was supported by the National Health and Medical Research Council and research funds from the Peter MacCallum Cancer Institute. We would like to thank Larry Wakelin for helpful discussions and critical reading of the manuscript.

Abbreviations

bp, base pair; IUdR, iododeoxyuridine; BrUdR, bromodeoxyuridine; dIUTP, iododeoxyuridine triphosphate; dBrUTP, bromodeoxyuridine triphosphate.

REFERENCES

1. Hutchinson, F. (1973) Quart. Rev. Biophys. 6, 201-246.
2. Hutchinson, F. and Kohnlein, W. (1980) Progr. Mol. Subcell. Biol. 7, 1-42.
3. Manuelidis, L. (1978) Chromosoma 66, 23-32.
4. Wu, J.C. and Manuelidis, L. (1980) J. Mol. Biol. 142, 363-386.
5. Darling, S.M., Crampton, J.M., and Williamson, R. (1982) J. Mol. Biol. 154, 51-63.
6. Murray, V. and Martin, R.F. (1987) Gene 57, 255-259.
7. Murray, V. and Martin, R.F. (1985) Gene Anal. Techn. 2, 95-99.
8. Sanger, F. Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
9. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161-178.
10. Ogata, R. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4973-4976.

11. Henner, W.D., Grunberg, S.M. and Haseltine, W.A. (1983) J. Biol. Chem. 258, 15198-15205.
12. Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2179-2183.
13. Dickerson, R.E. and Drew, H.R. (1981) J.Mol. Biol. 149, 761-786.
14. Shakked, Z. and Rabinovich, D. (1986) Prog. Biophys. Molec. Biol. 47, 159-195.
15. Calladine, C.R. (1982) J. Mol. Biol. 161, 343-352.
16. Murray, V. and Martin, R.F. (1988) J. Biol. Chem. 263, 12854-12859.
17. Murray, V. and Martin, R.F. (1988) J. Mol. Biol. 203, 63-73.
18. Mirabelli, C.K. Ting, A. Huang, C-H, Mong, S. and Crook, S.T. (1982) Can. Res. 42, 2779-2785.
19. Murray, V. and Martin, R.F. (1985) Nucleic Acids Res. 13, 1467-1481.