The Stability of Duplex DNA Containing 3,N^4-Etheno-2′-Deoxycytidine (εdC). A UV Melting and High Resolution ^{1}H NMR Study

Neil J. Gibson,[a] John A. Parkinson,[a] Thomas Barlow,[a] William P. Watson[b] and Tom Brown*[a]

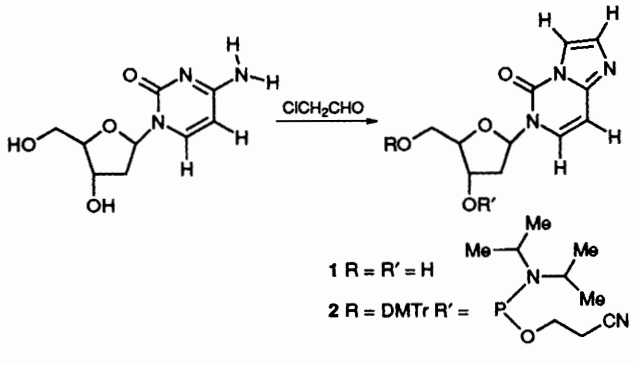
[a] *Department of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh, UK EH9 3JJ*
[b] *Shell Research Ltd., Sittingbourne Research Centre, Sittingbourne, Kent, UK ME9 8AG*

The mutagenic lesion 3,N^4-etheno-2′-deoxycytidine (εdC) greatly destabilises DNA when paired with each of the natural bases although the solution structure of a duplex containing a εC : A pair shows that the lesion does not greatly disrupt base pairing in the duplex.

Vinyl chloride, an established carcinogen in humans,[1] is epoxidised *in vivo* by cytochrome P-450 enzymes to chloroethylene oxide which rearranges to chloroacetaldehyde.[2] Both the haloxirane and the haloaldehyde alkylate DNA giving both mono and bifunctional adducts[3] including 3,N^4-etheno-2′-deoxycytidine (εdC) **1**. When single stranded DNA is treated with chloroacetaldehyde mutagenesis is almost exclusively at cytosines[4] and consists of 80% C-to-T transitions and 20% C-to-A transversions. Because the DNA polymerase is misreading the template at the εC site and there is ambiguity as to the inserted base, εC appears to form non-instructional lesions.[5] Such lesions are either devoid of template activity, as is the case with abasic sites, or have their template activity blocked by chemical adducts. When a base is incorporated opposite such a lesion it is usually found to be adenine with a lower incidence of thymine.[6] Non-instructional lesions block DNA replication and are lethal unless they are excised from the template or bypassed by the replication enzymes. According to the SOS hypothesis,[7] arrest of DNA synthesis by non-instructional lesions causes induction of the SOS regulon and the induced enzymes facilitate DNA synthesis beyond the lesion site by an unknown mechanism. However it has been observed that mutagenesis by εdC is SOS independent and highly efficient and that transfection of *Escherichia coli* with coliphage M13AB28 bearing a single εdC residue does not lead to a diminution of the survival of M13 DNA.[8] These are characteristics of a mispairing lesion. Such lesions contain template information which can be decoded by a DNA polymerase. However the information is incorrect and leads to the misincorporation of a base opposite the lesion without impairing DNA replication or cell survival.

In order to investigate whether εdC forms instructional or non-instructional lesions we have prepared the DNA duplexes d5′(GCTGεCGACG)3′/d5′(CGTCYCAGC)3′, where **Y** = A, G, T or C, which contain εdC paired with each of the four natural bases and examined their stability by UV melting. We have also examined the solution structure by high field ^{1}H NMR of one of these duplexes, d5′(GCTGεCGACG)3′/d5′(CGTCACAGC)3′, which contains an εdC : A pair.

The DNA synthesis monomer **2**, which was prepared from 2′-deoxycytidine,[9] coupled during DNA synthesis with 98.5% coupling efficiency. The crude oligonucleotide was purified by reverse phase HPLC and desalted on sephadex G10.



The thermal dissociation of duplex DNA into its constituent single strands leads to hyperchromicity. The point of maximum inflection of a plot of UV absorbency *versus* temperature is defined as the melting temperature T_{max} of the duplex. The thermodynamics of DNA duplex formation in the presence of a εdC lesion were determined according to the method of Gaffney and Jones.[10]

Table 1 shows that the incorporation of εdC into DNA leads to a large drop in the free energy of duplex formation. All four pairs to εdC are considerably less stable than the Watson–Crick C : G and A : T base pairs. These data also indicate that for the four εdC-containing duplexes, the pairing of A with εdC gives the duplex of greatest stability. However the εC : A pair is only very slightly more stable than the εC : G pair which in turn is slightly more stable than the εC : T pair while the differences in the stability of all four εC : Y pairs are not large. Consequently it would appear to be unlikely that the εdC lesion is instructional since it is observed that DNA polymerases pair εdC mainly with adenine but also with thymidine.

Analysis of the 600 MHz ^{1}H NOESY spectrum at 250 ms mixing time of the duplex d5′($G^1C^2T^3G^4εC^5G^6A^7C^8G^9$)3′/d5′($C^{10}G^{11}T^{12}C^{13}A^{14}C^{15}A^{16}G^{17}C^{18}$)3′ in 99.99% D_2O buffer containing 10 mmol dm^{-3} phosphate and 100 mmol dm^{-3} sodium chloride at 25 °C reveals that the duplex adopts the B conformation. The pattern of cross-peaks between the sugar H(1)′ resonances and the base H(6) and H(8) resonances shows that every residue adopts the anti conformation (Fig. 1).

Fig. 2 shows the 600 MHz one-dimensional ^{1}H NMR spectrum of the imino protons in 90% H_2O/10% D_2O buffer containing 10 mmol dm^{-3} phosphate and 100 mmol dm^{-3} sodium chloride at 18 °C. The H_2O signal was supressed with presaturation and the imino proton spectrum demonstrates the different rates of exchange of these protons with the solvent. The resonances were assigned from analysis of the two-dimensional 600 MHz ^{1}H NOESY spectrum at 250 ms mixing time with presaturation of H_2O. The thymidine N(3)H

Table 1 The thermodynamic parameters of hybridisation for four εdC containing duplexes and the corresponding duplexes containing Watson–Crick base-pairs 5′GCT GXG ACG3′/3′CGA CYC TGC5′

X	Y	T_{max} (1 μmol)[a]/ °C	ΔH°/ kJ mol^{-1}	ΔS°/ kJ mol^{-1} K^{-1}	ΔG_{298}/ kJ mol^{-1}	r	$\Delta\Delta G_{298}$[b]
C	G	47.2	−264.77	−0.71	−53.19	0.9965	—
A	T	40.6	−290.70	−0.81	−49.32	0.9958	3.92
εC	A	24.6	−255.03	−0.74	−34.51	0.9971	18.67
εC	G	23.9	−218.79	−0.62	−34.03	0.9985	19.17
εC	T	22.3	−270.81	−0.80	−32.41	0.9992	21.56
εC	C	19.7	−236.19	−0.69	−30.57	0.9972	22.91

[a] For each duplex the T_{max} was determined in triplicate at six concentrations between the range 3 and 34 μmol dm^{-3}. r = correlation coefficient of straight line plot of $\ln C_T$ *vs* $1/T_{max}$. T_{max} (1 μmol) = the calculated melting temperature of the duplex at a concentration of 1 μmol dm^{-3}. [b] $\Delta\Delta G_{298}$ = the decrease in the free energy of duplex formation between the duplex 5′GCT GXG ACG3′/3′CGA CYC TGC5′ indicated and duplex 1.

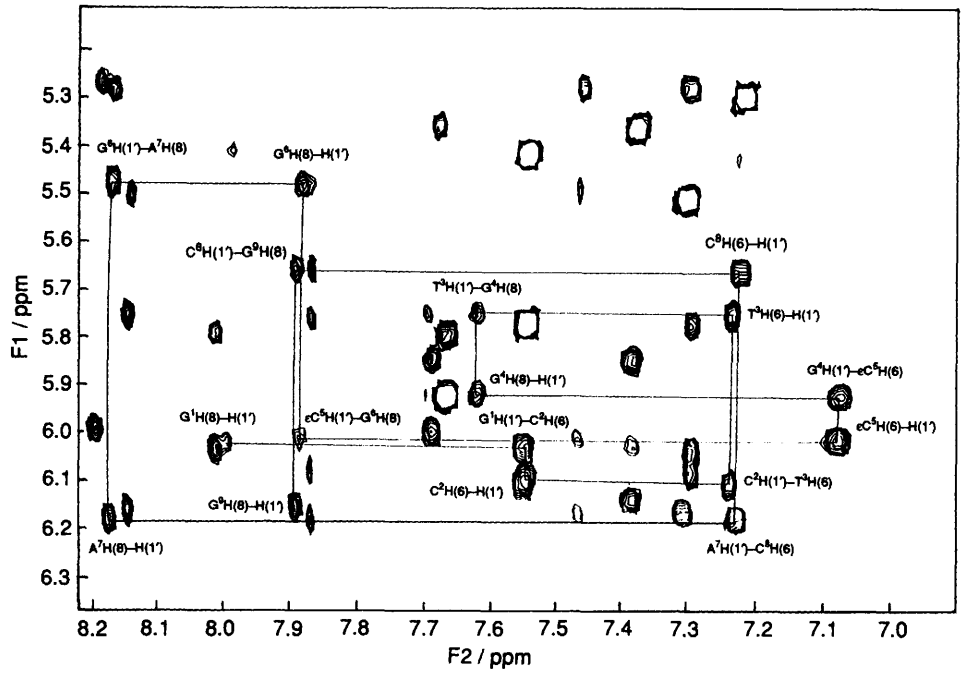


Fig. 1 Detail of the 600 MHz ^{1}H NOESY spectrum of the duplex 5′d(G^{1}C^{2}T^{3}G^{4}εC^{5}G^{6}A^{7}C^{8}G^9)3′/d5′(C^{10}G^{11}T^{12}C^{13}A^{14}G^{15}A^{16}G^{17}C^{18})3′ in 99.99% D_2O at 25 °C with 250 ms mixing time showing the connectivities between sugar H(1)′ and base H(6)/H(8) resonances and highlighting the connectivities from G^{1}H(8) to G^{9}H(1′)

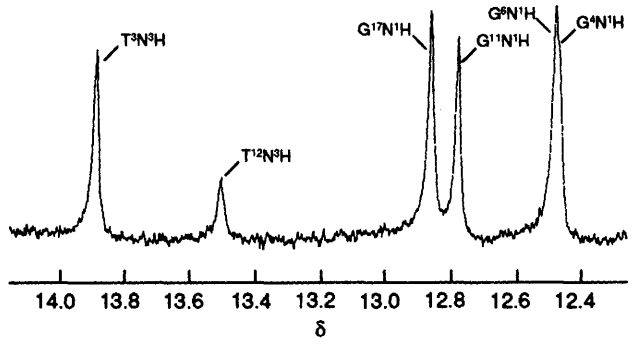


Fig. 2 Portion of the 600 MHz ^{1}H NMR spectrum of the duplex 5′d(G^{1}C^{2}T^{3}G^{4}εC^{5}G^{6}A^{7}C^{8}G^9)3′/d5′(C^{10}G^{11}T^{12}C^{13}A^{14}G^{15}A^{16}G^{17}C^{18})3′ in 90% H_2O at 18 °C showing the imino resonances

signals were assigned by observing an NOE to the corresponding cross-strand deoxyadenosine H(2) signals. The deoxyguanosine N(1)H signals were observed to give NOEs to both amino signals of the corresponding cross strand deoxycytidine, in turn these amino signals showed NOEs to the H(5) of the same base. Thus a chain of assignment could be traced from each deoxycytidine H(5) signal to the corresponding cross strand deoxyguanosine N(1)H. Of the six possible resonances potentially observable in this region (terminal imino protons are in fast exchange with the solvent) all are seen. It is noteworthy that T^{12}N(3)H, which is remote from the lesion site, is in faster exchange with the solvent than both the G^{6}N(1)H and G^{4}N(1)H which flank the lesion site and this suggests that the T^{12}:A^7 base pair is slightly melted out. The non-diminution of the G^{6}N(1)H and G^{4}N(1)H signals compared to G^{17}N(1)H and G^{11}N(1)H shows that the C:G base pairs on either side of the εC:A base pair are intact and that there is no melting of the central region of the duplex around the εC lesion. This shows that the instability of duplex DNA containing the εC lesion, as shown by UV melting, is not due to major disruption of base pairs flanking the lesion site.

We are grateful to the Royal Society of Edinburgh for a Caledonian Research Fellowship to T. Brown, to the SERC for an earmarked studentship to T. Barlow and for 600 MHz NMR time (Edinburgh University) and jointly to the SERC and to Shell Research Ltd. for a fellowship to N. J. G.

Received, 22nd June 1994; Com. 4/03801K

References

1 H. Bartsch, A. Barbin, M. J. Marion, J. Nair and Y. Guichard, *Drug Metab. Rev.*, 1994, **26**, 349.
2 F. P. Guengerich, W. M. Crawford and P. G. Watanabe, *Biochemistry*, 1979, **18**, 5177.
3 R. J. Laib, L. M. Gwinner and H. M. Bolt, *Chem. Biol. Interact.*, 1981, **37**, 219; E. Scherer, C. J. Van Der Laken, L. M. Gwinner, R. J. Laib and P. Emmelot, *Carcinogenesis*, 1981, **2**, 671.
4 J. S. Jacobsen, C. P. Perkins, J. T. Callahan, K. Sambamurti and M. Z. Humayun, *Genetics*, 1989, **121**, 213.
5 D. Simha, V. A. Palejwala and M. Z. Humayun, *Biochemistry*, 1991, **30**, 8727.
6 L. A. Loeb and B. D. Preston, *Annu. Rev. Genet.*, 1986, **20**, 201–230; S. D. Rabkin and B. S. Strauss, *J. Mol. Biol.*, 1984, **178**, 569.
7 V. A. Palejwala, R. W. Rzepka and M. Z. Humayun, *Biochemistry*, 1993, **32**, 4112.
8 V. A. Palejwala, D. Simha and M. Z. Humayun, *Biochemistry*, 1991, **30**, 8736.
9 S. C. Srivastava, S. K. Raza and R. Misra, *Nucl. Acids Res.*, 1994, **22**, 1296.
10 B. L. Gaffney and R. A. Jones, *Biochemistry*, 1989, **28**, 5881.