The Impact of an Exocyclic Cytosine Adduct on DNA Duplex Properties: Significant Thermodynamic Consequences Despite Modest Lesion-Induced Structural Alterations†

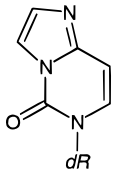
Craig A. Gelfand,‡ G. Eric Plum,‡ Arthur P. Grollman,§ Francis Johnson,§ and Kenneth J. Breslauer*,‡

Department of Chemistry, Rutgers, The State University of New Jersey, 610 Taylor Road, Piscataway, New Jersey 08854-8087, and Department of Pharmacological Sciences, State University of New York at Stony Brook, Stony Brook, New York 11794

Received May 11, 1998; Revised Manuscript Received July 17, 1998

ABSTRACT: The exocyclic base adduct 3,N^4-deoxyethenocytosine (ϵC) is a common DNA lesion that can arise from carcinogen exposure and/or as a biproduct of cellular processes. We have examined the thermal and thermodynamic impact of this lesion on DNA duplex properties, as well as the structural alterations imparted by the lesion. For these studies, we used calorimetric and spectroscopic techniques to investigate a family of 13-mer DNA duplexes of the form ($^{5'}$CGCATGNGTACGC$^{3'}$)•($^{3'}$GCGTACNCATGCG$^{5'}$), where the central N•N base pair represents the four standard Watson−Crick base pairs (corresponding to four control duplexes), and where either one of the N bases has been replaced by ϵC, yielding eight test duplexes. Studies on these 12 duplexes permit us to assess the impact of the ϵC lesion as a function of sequence context. Our spectroscopic and calorimetric data allow us to reach the following conclusions: (i) The ϵC lesion imparts a large penalty on duplex stability, with sequence context only modestly modulating the extent of this lesion-induced destabilization. This result contrasts with our recent studies of duplexes with abasic sites, where sequence context was found to be the predominant determinant of thermodynamic damage. (ii) For the ϵC-containing duplexes, sequence context effects are most often observed in the enthalpic contribution to lesion-induced duplex destabilization. However, due to compensating entropies, the free energy changes associated with this lesion-induced duplex destablization are nearly independent of sequence context. (iii) Despite significant lesion-induced changes in duplex energetics, our spectroscopic probes detect only modest lesion-induced changes in duplex structure. In fact, the overall duplex maintains a global B-form conformation, in agreement with NMR structural data. We discuss possible interpretations of the apparent disparity between the severe thermodynamic and relatively mild structural impacts of the ϵC lesion on duplex properties. We also note and discuss the implications of empirical correlations between biophysical and biological properties of lesion-containing duplexes.

The exocyclic DNA base modification 3,N^4-deoxyethenocytosine (ϵC), shown below, can result from exposure to vinyl chloride (*1*−*5*), a highly reactive carcinogen, from hepatic copper poisoning (*6*), or as a consequence of reactions between DNA and the products of lipid peroxidation (*7, 8*).



εC

Exocyclic base modifications, such as ϵC, can lead to genetic mutations due to possible alteration of coding information at the lesion site. In this regard, several polymerases show a tendency to add an A or T residue to the nascent strand across from an ϵC site (*9*−*15*), thereby resulting in either a C to T transition or a C to A transversion. Addition of G and C across from ϵC, and even U with RNA polymerase, as well as one- and two-base deletions also have been observed, with the specific outcome depending on the polymerase studied (*13, 15, 16*). In vivo mutation rates resulting from ϵC depend strongly on the organism, with a mutation frequency range of approximately 2% in bacteria (*11, 12*), to over 80% in mammalian COS cells (*12*). With regard to the repair of this lesion, glycosylase activity has been described which removes etheno-modified bases in human cells (*17, 18*).

Recent NMR studies of short DNA duplexes which contain the ϵC lesion show the duplex to have an unusual local conformation in the immediate vicinity of the damaged base, with most of the duplex remaining unperturbed in the B-form conformation (*19*−*21*). These same NMR studies reveal that a change in the identity of the base across from the ϵC lesion causes different conformations for the bases of the damaged pair, which, in turn, can result in localized bending and/or rearrangement of the neighboring base pair stacks.

† This research was supported by NIH Grants GM-23509 and GM-34469 to K.J.B. and CA-47795 to A.P.G.

* Author to whom correspondence should be addressed.
‡ Rutgers, The State University of New Jersey.
§ State University of New York at Stony Brook.

S0006-2960(98)01090-3 CCC: $15.00
Published on Web 08/20/1998

Table 1: Effect of the ϵC Lesion on the Thermal and Thermodynamic Stabilities of DNA Duplexes

sequence family	duplex	T_m (°C)[a]	ΔT_m (°C)[a,b]	ΔG(25 °C) (kcal/mol)	$\Delta\Delta G$(25 °C)[c] (kcal/mol)
-G N G- -C N C-	A·T	68.2 ± 0.5	−	19.6 ± 0.6	−
	C·G	71.4 ± 0.5	−	20.5 ± 0.5	−
	G·C	69.2 ± 0.5	−	20.4 ± 0.6	−
	T·A	68.4 ± 0.5	−	20.4 ± 1.1	−
-G ϵC G- -C N C-	ϵC·A	58.1 ± 0.5	−10.3 ± 0.7	6.8 ± 0.2	−13.5 ± 1.1
	ϵC·C	56.1 ± 0.5	−13.1 ± 0.7	6.0 ± 0.2	−14.4 ± 0.7
	ϵC·G	56.2 ± 0.5	−15.2 ± 0.7	5.1 ± 0.4	−15.3 ± 0.6
	ϵC·T	55.9 ± 0.5	−12.3 ± 0.7	7.6 ± 0.2	−12.0 ± 0.6
-G N G- -C ϵC C-	A·ϵC	57.5 ± 0.5	−10.7 ± 0.7	5.9 ± 0.2	−13.7 ± 0.6
	C·ϵC	53.3 ± 0.5	−18.1 ± 0.7	6.4 ± 0.3	−14.1 ± 0.6
	G·ϵC	57.9 ± 0.5	−11.3 ± 0.7	7.0 ± 0.2	−13.4 ± 0.6
	T·ϵC	56.7 ± 0.5	−11.7 ± 0.7	6.6 ± 0.2	−13.7 ± 1.1

[a] T_m values are for solutions containing 50 μM duplex DNA. [b] $\Delta T_m = T_m$(ϵC-containing duplex) − T_m(parent duplex). [c] $\Delta\Delta G = \Delta G$($\epsilon$C-containing duplex) − ΔG(parent duplex).

Despite these impressive structural studies, little, if anything, is known about the impact of this lesion on duplex energetics, thereby limiting our insight into the role of "energetic recognition" (*22*) of this lesion in mechanisms of mutagenesis and repair. To address this deficiency, we present here the results of our spectroscopic and calorimetric studies on the thermodynamic impact of ϵC lesions on duplex DNA. We use differential scanning calorimetry (DSC) and temperature-dependent UV absorbance spectroscopy to study the impact of the lesion, as well as the influences of the cross-strand and neighboring bases, on the thermodynamic properties of duplex DNA. Circular dichroism spectropolarimetry and intrinsic fluorescence also are used to study the global and local structures of these duplexes. By comparing the properties associated with a set of eight ϵC-containing duplexes with the corresponding undamaged parent Watson−Crick duplexes, we are able to isolate and define the thermodynamic, thermal, and structural consequences of the ϵC lesion. We present these results and discuss possible correlations between our biophysical data and the known biological, biochemical, and structural consequences of the ϵC lesion.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification. The DNA oligomers studied here have the same sequence as those used in our previous investigations of lesions (*23*−*27*), thereby facilitating comparisons. Protocols used for the synthesis and purification of the oligonucleotides previously have been described (*26*). The sequences studied here are:

strand 1: 5′-CGCATGNGTACGC-3′
strand 2: 5′-GCGTACNCATGCG-3′

where N represents the four standard nucleotide bases, as well as ϵC, the 3,N^4-deoxyethenocytosine exocyclic adduct of deoxycytosine (*28*). We refer to duplexes using the notation strand 1•strand 2, citing only the central bases of each strand. Families of duplexes are cited as ϵC•N and N•ϵC, with each family corresponding to four duplexes in which the ϵC lesion is either in strand 1 (ϵC•N) or in strand 2 (N•ϵC).

Molar extinction coefficients were determined as previously described (*27*), and were found to be (1.06 ± 0.02) × 10^5 and (1.12 ± 0.02) × 10^5 $M^{-1}\cdot cm^{-1}$ at 260 nm, 25 °C for N = ϵC in strand 1 and strand 2, respectively. The extinction coefficients of the oligomers with standard Watson−Crick bases have been previously reported (*27*). The 1:1 stoichiometry of each duplex was confirmed using the method of continuous variations, also as previously described (*27*).

All experiments, except fluorescence spectroscopy, were performed in a buffer solution of 1 M NaCl, 0.1 mM EDTA, and 10 mM sodium phosphate, at pH 7.0.

Duplex Melting. Differential scanning calorimetry (DSC) and temperature-dependent UV absorbance spectrophotometry were used to monitor the thermally induced order−disorder transition of each DNA duplex. DSC measurements were performed on a N-DSC differential scanning calorimeter (Calorimetry Sciences Corp., Provo, UT), while the UV-detected melting measurements were performed on an AVIV 14DS UV−Vis−IR spectrophotometer. Details of all the experimental protocols (*27*), as well as the data analyses (*22*, *25*, *27*, *29*, *30*), have been described previously.

Structure-Elucidating Spectroscopy. Circular dichroism spectropolarimetry (CD), performed on an AVIV 62DS spectropolarimeter, was used to evaluate the global structure of each duplex, using conditions and methods previously described (*27*). Fluorescence spectra were determined on a Perkin-Elmer LS-50 fluorometer at 25 °C using a 4 mm square quartz cuvette, with 100 μM duplex in a buffer of 100 mM NaCl, 50 mM MOPS at pH 7.0.

RESULTS AND DISCUSSION

The ϵC Lesion Induces Dramatic Thermal and Thermodynamic Destabilization of the Host DNA Duplex. Table 1 presents thermal (T_m) and thermodynamic (ΔG) stability data derived from both our UV-absorbance and DSC melting studies. Note that the presence of the ϵC lesion dramatically reduces the thermal stability of the duplex, with ΔT_m values falling between −10 and −18 °C. Further note that this thermal destabilization translates into an average lesion-induced thermodynamic destabilization relative to the parent Watson−Crick duplex ($\Delta\Delta G$) of −13.8 kcal/mol. This lesion-induced duplex destabilization represents a 60%−70% loss in *total* duplex thermodynamic stability at 25 °C, despite the modification of only 1 out of 26 bases (3.8%) [or 1 base pair out of 13 (7.7%), or 2 base pair stacks out of 12 (17%)]. This dramatic lesion-induced destabilization corresponds to a change of between 8 and 11 orders of magnitude in the

Table 2: Effect of the ϵC Lesion on the Dissociation Enthalpy Change and the Effective Molecularity of a Family of DNA Duplexes

sequence family	duplex	ΔH (kcal/mol)	$\Delta\Delta H^a$ (kcal/mol)	$\Delta H^{vH\,b}$ (kcal/"mol")	$n_{eff}{}^c$
-G N G- -C N C-	A·T	101.6 ± 2.4	−	100.8 ± 2.3	2.01 ± 0.06
	C·G	98.1 ± 1.8	−	89.8 ± 1.7	2.09 ± 0.05
	G·C	104.0 ± 1.9	−	100.0 ± 3.0	2.04 ± 0.07
	T·A	104.1 ± 4.2	−	98.2 ± 4.4	2.06 ± 0.12
-G ϵC G- -C N C-	ϵC·A	81.6 ± 0.9	−22.6 ± 4.3	93.7 ± 1.2	1.87 ± 0.03
	ϵC·C	77.2 ± 1.7	−26.7 ± 2.6	91.1 ± 1.0	1.85 ± 0.05
	ϵC·G	65.4 ± 3.5	−32.8 ± 4.0	97.6 ± 1.3	1.67 ± 0.10
	ϵC·T	97.0 ± 1.0	−4.6 ± 2.5	96.1 ± 1.6	2.01 ± 0.04
-G N G- -C ϵC C-	A·ϵC	71.4 ± 0.8	−30.2 ± 2.4	98.5 ± 1.5	1.73 ± 0.03
	C·ϵC	93.3 ± 3.0	−4.8 ± 3.6	77.8 ± 2.6	2.20 ± 0.10
	G·ϵC	82.8 ± 0.8	−20.2 ± 2.1	96.6 ± 1.4	1.87 ± 0.03
	T·ϵC	82.7 ± 0.5	−21.4 ± 4.3	94.2 ± 2.1	1.88 ± 0.05

[a] $\Delta\Delta H = \Delta H(\epsilon C\text{-containing duplex}) - \Delta H(\text{parent duplex})$. [b] From the concentration dependence of T_m determined from UV absorbance vs temperature curves. [c] n_{eff} calculated using calorimetric ΔH, and based on concentration dependence of T_m.

duplex equilibrium constant. These results indicate that although the chemical modification is localized to one base, the thermodynamic impact of the damage is far from local. Apparently, the energetic consequences of the lesion propagate well beyond the site of the structural modification.

Sequence Context Has Relatively Little Influence on the Impact of the Lesion on Duplex Stability. Inspection of the data in Table 1 reveals few, if any, significant differences in the destabilization induced by the ϵC lesion among the duplexes. In other words, the mere presence of the ϵC lesion is sufficient to impart the primary penalty, with sequence context, including both neighboring and cross-strand bases, exerting only a secondary or tertiary effect. On a quantitative level, there may be some influence on the *thermal* stability of the base identity opposite the ϵC lesion, although the differences are similar to the uncertainty in the data. To be specific, the average of all four pairings of pyrimidines across from ϵC produces a ΔT_m value of -13.8 ± 1.4 °C compared with a ΔT_m value of -11.9 ± 1.4 °C for all four pairings of purines across from ϵC. The corresponding $\Delta\Delta G$ values of -13.6 ± 1.6 kcal/mol for the four pyrimidines opposite ϵC and -14.0 ± 1.5 kcal/mol for the four purines opposite ϵC are statistically indistinguishable, consistent with the lesion-induced destabilization being independent of this feature of the identity of the base opposite the lesion.

The identity of the bases neighboring the lesion also appears to exert little or no effect on either $\Delta\Delta G$ or ΔT_m. For ϵC•N, the average $\Delta\Delta G$ value is -13.9 ± 1.6 kcal/mol and the average ΔT_m value is -12.7 ± 1.4 °C, while for N•ϵC, the average $\Delta\Delta G$ value is -13.8 ± 1.5 kcal/mol and the average ΔT_m value is -13.0 ± 1.4 °C. These average values are statistically indistinguishable, suggesting that the neighboring base identity does not have a primary influence on the damage imparted by ϵC upon the duplex, at least for the sequences investigated here. We recognize, however, that additional studies on duplexes with the ϵC lesion flanked by other neighboring bases are required to evaluate the generality of this behavior.

The Enthalpic Contribution to Duplex Stability Varies with the Sequence Context of the Lesion. In contrast to the near-independence from sequence context of the lesion-induced changes in duplex free energy and thermal stability, the corresponding enthalpic changes are far more variable. The relevant ΔH values are listed in the third column of Table 2. These enthalpy changes were derived from DSC measurements, and therefore are not dependent on an assumed melting model. The ϵC•T and C•ϵC duplexes show relatively small lesion-induced enthalpic reductions, $\Delta\Delta H$, of -4.6 ± 2.5 and -4.8 ± 3.6 kcal/mol, respectively, compared to the other six duplexes, which exhibit more dramatic $\Delta\Delta H$ values ranging between -20.2 ± 2.1 and -32.8 ± 4.0 kcal/mol. Interestingly, in the two outlying duplexes, ϵC•T and C•ϵC, a pyrimidine is placed opposite ϵC. This observation may suggest a simple steric accommodation of the base opposite the lesion. The ϵC base has an exocyclic addition to the normal C base which converts a pyrimidine, a one-ring base, into a purine-like two-ringed base. This "expanded" base might be expected to impart less distortion to the duplex when paired across from a smaller pyrimidine rather than a larger purine. However, this steric model may be too simplistic given the failure to rationalize the corresponding properties of the remaining two pyrimidine−ϵC pairs in a similar manner.

Further inspection of the data in Table 2 reveals that the differential duplex melting enthalpy, $\Delta\Delta H$, depends significantly on the identity of the bases which flank the lesion, in contrast to the near-independence of $\Delta\Delta G$ on sequence context. Comparison of the ϵC lesion paired with the same base in both sequence orientations (e.g., ϵC•A vs A•ϵC) shows a difference in $\Delta\Delta H$ between approximately 7 and 17 kcal/mol. Interestingly, the rank order of enthalpic reduction based on opposite base identity is different for the two orientations studied. For the ϵC•N family of duplexes, the $\Delta\Delta H$ order is T < A < C < G, while for the N•ϵC family the $\Delta\Delta H$ order is C < G < T < A. Nevertheless, the overall $\Delta\Delta H$ range which is encompassed by the two sequences is virtually identical. Although we cannot provide a simple explanation for the enthalpic distinctions between the two duplex families, our results do demonstrate that sequence context can influence the enthalpic contribution to duplex stability.

In the aggregate, the data discussed above reveal that even though the stability of ϵC-containing duplexes is independent of sequence context, the enthalpic impact of the ϵC lesion depends on the identity of the opposing base and the neighboring bases, with these two influences being approximately equal.

Enthalpy−Entropy Compensations Equalize the Free Energy Penalties of the ϵC Lesion, Making Them Independent of Sequence Context. As noted above, the presence of the ϵC lesion imparts a severe penalty on duplex stability. We find this large lesion-induced $\Delta\Delta G$ penalty to be essentially independent of sequence context, in contrast to the $\Delta\Delta H$ penalty. This observation is common for many types of nucleic acid damage, in which entropic differences balance enthalpic effects, thereby tempering the overall variation in free energy ($\Delta\Delta G$) induced by the damage. This compensation effect is observed even for systems with widely varying lesion-induced $\Delta\Delta H$ values. Similar enthalpy−entropy compensations have been described for a model abasic site (*27*), and for two exocyclic adducts, 1,N^2-propanodeoxyguanosine (*22*) and 8-oxodeoxyguanosine (*25*), thereby suggesting the generality of this phenomenon.

The Influence of Sequence Context Depends on the Identity of the Lesion. As described above, the impact of the ϵC lesion on duplex stability is dramatic, although relatively independent of sequence context. By contrast, recent results from this laboratory demonstrate that the thermodynamic impact of the abasic lesion strongly depends on sequence context (*27*). These contrasting results underscore the need to conduct biophysical, biochemical, and biological investigations of lesions as a function of sequence context before attempting to define the impact of the lesion.

van't Hoff Transition Enthalpies and the Duplex Melting Process. The results presented and discussed above are based upon model-independent thermodynamic parameters derived from calorimetric studies. Comparison with model-dependent van't Hoff estimates of the same parameters provides us with an opportunity to assess the influence of the lesion and sequence context on the mechanism of duplex melting. We previously have described an approach for combining spectroscopic and calorimetric data to calculate an effective molecularity, n_{eff}, which reflects the extent to which duplex melting deviates from an idealized bimolecular transition. Deviation of n_{eff} from the formal molecularity value of 2 (*25, 31*) provides some measure of the degree to which a structural perturbation, such as a lesion, alters the nature of the melting process.

Inspection of the data in Table 2 reveals the following. First, as expected, the four parent Watson−Crick duplexes all melt with effective molecularities near 2. Second, in nearly all cases, the presence of the lesion alters the nature of the duplex melting. This observation is consistent with what we have reported for other lesions (*25, 27*). Direct comparisons between the van't Hoff (ΔH_{vH}) and the corresponding calorimetric enthalpy values also reveal significant disparities for most of the lesion-containing duplexes. Conventionally, $\Delta H_{vH} < \Delta H_{cal}$ is interpreted as an indication of non-two-state transitions. Conversely, $\Delta H_{vH} > \Delta H_{cal}$ is taken to indicate multimolecular (aggregation) processes. For the data reported here, we find ΔH_{vH} generally to be larger than ΔH_{cal}. There being no independent evidence of aggregation, we interpret these data to reflect deviations from the idealized bimolecular dissociation process. This observation appears to be general for lesion-containing duplexes and, as we previously have suggested (*25*), may reflect differences in helix initiation for lesion-containing duplexes.

Circular Dichroism Shows the ϵC-Containing Duplexes To Have Similar Global Conformations. The CD spectra of the ϵC-containing duplexes are shown in Figure 1. All eight spectra show the duplexes to be primarily B-form (*32*). Slight differences exist among the eight duplexes, but no gross structural differences are apparent. CD spectra depend on base content and sequence, as well as duplex conformation. Consequently, spectral differences are expected, because each duplex has a different central base "pair", as well as two different neighboring pair stacks at the central position, both in the four control Watson−Crick duplexes and in the eight duplexes which contain the ϵC base. Our conclusion that the lesion-containing duplexes exist primarily as B-form DNA is consistent with previous NMR studies (*19*−*21*) which indicate structural alterations in the duplex only at and very near the ϵC site, with the overall duplex conformation remaining in the B-form. Thus, the dramatic energetic consequences of the lesion are not reflected in significant structural alterations of the duplex.

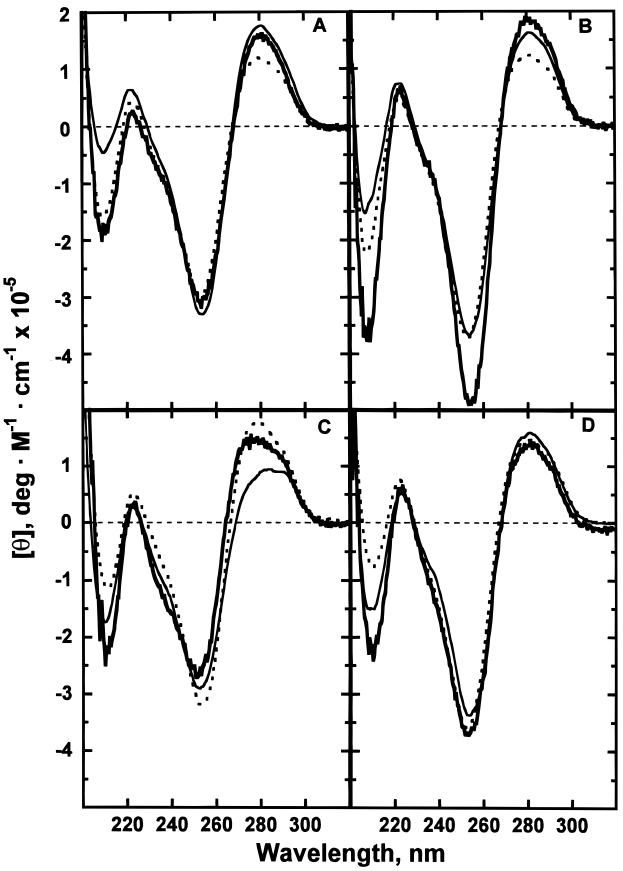


FIGURE 1: Circular dichroism spectra of DNA duplexes containing the ϵC lesions and the parent Watson−Crick duplexes. Each panel shows the spectrum for one Watson−Crick parent (thick line) and the associated ϵC-containing duplexes, with ϵC in strand 1 (ϵC•N; dotted line) or in strand 2 (N•ϵC; thin line). Panel A presents the spectra of A•T, ϵC•T, and A•ϵC, and similarly for C•G (panel B), G•C (panel C) and T•A (panel D), and their respective ϵC-containing duplexes.

Intrinsic Fluorescence Is Suggestive of Local Variations in the Conformation and/or the Microenvironment of the ϵC Base. The relative intensities and wavelengths of the maximum of ϵC intrinsic fluorescence emission for all eight ϵC-containing duplexes are summarized in Table 3. Because ϵC is the only significantly fluorescent moiety in the duplex, fluorescence spectroscopy provides us with a microscopic observable to compare with the circular dichroism and thermodynamic characterizations of the duplexes. Inspection

Table 3: Fluorescence Emission Spectral Characteristics for the ϵC-Containing Duplexes

sequence family	duplex	relative fluorescence intensity [a]	λ_{em}^{max} (nm)[a]
-G ϵC G-	ϵC·A	14.2	407
-C N C-	ϵC·C	9.7	399
	ϵC·G	22.0	400
	ϵC·T	13.8	394
-G N G-	A·ϵC	24.8	385
-C ϵC C-	C·ϵC	9.5	390
	G·ϵC	9.0	392
	T·ϵC	8.5	380

[a] Fluorescence spectra were collected with excitation at 310 nm.

of the fluorescence data in Table 3 suggests that the ϵC-containing duplexes can be segregated into two groups based on the fluorescence emission intensities and the wavelength of maximum emission (λ_{em}^{max}). Note that the N•ϵC family of duplexes exhibit emission maximum wavelengths between 380 and 392 nm, while the ϵC•N family of duplexes exhibit higher emission maximum wavelengths, between 394 and 407 nm. Thus, the λ_{em}^{max} of ϵC is sensitive to the sequence context of the lesion. The simplest explanation for this observation is that the different neighboring bases provide different microenvironments for ϵC, leading to shifts of the fluorescence emission spectrum. Further, the longer wavelength λ_{em}^{max} for the ϵC•N family is consistent with ϵC being flanked by two guanines, with the purines perhaps providing more extensive hydrophobic interactions than pyrimidines, and, accordingly, inducing longer λ_{em} for the fluorophore.

The variations in the fluorescence intensity also are of interest. We find a nearly 3-fold fluorescence enhancement for two duplexes (A•ϵC and ϵC•G) relative to the remaining duplexes. This distinction is consistent with some aspect of the microenvironment of the ϵC site being determined by the identity of the cross-strand base. In short, the fluorescence data reveal that the neighboring bases and the cross-strand base influence the microenvironment at the lesion site. Although enthalpy−entropy compensations mask the expression in ΔG of sequence-dependent effects, the ΔH data reveal sequence-dependent differences that may relate to those detected by fluorescence. Interestingly, the two duplexes that are most intensely fluorescent also are most affected enthalpically by the lesion. As there are many factors that can influence fluorescence intensity, we will not attempt further interpretation of the fluorescence data.

Comparison of Fluorescence and NMR Observables. NMR structures of three 11-mer duplexes, with sequences that are identical to the central 11 base pairs of our N•ϵC duplex family, show wide variation in the conformation of both ϵC and the "paired"/opposing base (*19*−*21*), while the remainder of the duplex exhibits normal Watson−Crick structure. It is interesting to note that the A•ϵC duplex, which exhibits enhanced fluorescence, also presents an altered solution structure in which the bases are accommodated within the helix in "a staggered conformation with each residue displaced toward the 5′-terminus and intercalated between bases on the opposite strand" (*19*). This self-intercalation model should result in an altered microenvironment for the ϵC residue relative to typical base pair stacking. This confluence of the fluorescence data and NMR structures suggests that similar localized structural variations also may be present in the ϵC•N family, and that further NMR studies with this latter sequence context for ϵC may prove useful.

Localized Structural Changes with Far-Reaching Thermodynamic Consequences. It is of interest to compare the thermodynamic and structural consequences of the ϵC lesion. Our data show that the ϵC lesion imparts a large penalty on the host duplex, destabilizing the duplex free energy by as much as 70%. Although there is variability in the enthalpic contribution to this destabilization, it is consistently offset by a compensating change in the entropy to yield a lesion-induced free energy of destabilization that is nearly independent of sequence context. In contrast to the dramatic duplex destabilization induced by the ϵC lesion, structural studies reveal that this lesion only induces local changes in DNA structure, with the bulk of the duplex retaining B-form. Clearly the apparent magnitudes of the energetic and structural perturbations induced by the ϵC lesion cannot at present be correlated in a simple manner. Several possible explanations/rationalizations are considered below.

It is possible that the ϵC lesion and associated local structural perturbations induce large changes in the associated solvent. Solvent effects can have very large influences on the thermodynamic properties of the duplex, often well in excess of changes that might be predicted from NMR-derived structures (*33*). Alternatively, subtle structural changes at the lesion site, that are outside the detection limits of a NMR structure, may propagate along the helix to induce an amplified global energetic effect. Independent of the veracity of such microscopic/macroscopic interpretations, our results underscore the importance of including thermodynamic analyses as a function of sequence context in any comprehensive study of DNA lesions.

Correlating Biophysical and Biological Data. Our biophysical studies show that the presence of the ϵC lesion dramatically decreases the thermal and thermodynamic stability of duplex DNA. Although these effects are independent of sequence context, we find the enthalpy data to exhibit a dependence on the identities of the neighboring and cross-strand bases.

Biological studies reveal that in mammalian (COS) cells, the ϵC lesion is highly mutagenic (*12*). Several reports show that bacteria are better able to cope with ϵC damage (*11*, *12*). Although we cannot now describe the basis for this discrimination, such an observation is at least consistent with the notion that repair enzymes may, in part, locate DNA lesions by detecting alterations in duplex energetics (*22*). It is also of interest to note that the tendency to pair T opposite ϵC during DNA replication in vivo (*11*, *12*, *14*) empirically correlates with the one duplex in our study, ϵC•T, which is significantly less enthalpically damaged. In the aggregate, there is a growing body of data which suggests that thermodynamics may play an important role in the overall biological outcome of ϵC and other DNA damage in both eukaryotes and prokaryotes (*25*, *31*). This possibility underscores the need for parallel biophysical and biological studies of DNA lesions.

CONCLUDING REMARKS

We have shown that the ϵC lesion dramatically decreases the thermal and thermodynamic stability of duplex DNA,

with this lesion-induced destabilization being essentially independent of sequence context. The large destabilizing effect of the lesion on the duplex is independent of the identity of the same-strand neighboring bases of the lesion and the cross-strand partner base of the lesion; however, the thermodynamic origins of the destabilization do not share that independence. This behavior contrasts with that which we have observed for the abasic lesion, where the lesion-induced thermal and thermodynamic influences are strongly modulated by sequence context. These two contrasting observations underscore the need for evaluating the influence of sequence context when probing for correlations between the biophysical impact of a lesion and its biological outcome. We also have shown that the dramatic thermodynamic impact of the ϵC lesion occurs in the absence of a corresponding dramatic change in duplex structure. Any lesion-induced structural alterations are modest and highly localized at the lesion site, with the balance of the duplex remaining in a B-like conformation. This observation is consistent with behavior we have observed with other lesions and underscores the dangers of assuming that large energetic perturbations imply significant structural alterations, and vice-a-versa. Clearly, biophysical studies on DNA lesions require parallel structural and thermodynamic studies before one can meaningfully assess the impact of a lesion on duplex properties.

ACKNOWLEDGMENT

We thank Robert Rieger for the synthesis of the ϵC-containing oligonucleotides and Dr. Jens Volker and Dr. Slawomir Mielewczyk for useful discussions and assistance.

REFERENCES

1. Laib, R. J., Gwinner, L. M., and Bolt, H. M. (1981) *Chem.-Biol. Interact. 37*, 219−231.
2. Kusmierek, J. T., and Singer, B. (1982) *Biochemistry 21*, 5717−5722.
3. Ciroussel, F., Barbin, A., Eberle, G., and Bartsh, H. (1990) *Biochem. Pharmacol. 39*, 1109−1113.
4. Matsuda, T., Yagi, T., Kawanishi, M., Matsui, S., and Takebe, H. (1995) *Carcinogenesis 16*, 2389−2394.
5. Guichard, Y., El Ghissassi, F., Nair, J., Bartsch, H., and Barbin, A. (1996) *Carcinogenesis 17*, 1553−1559.
6. Nair, J., Sone, H., Nagao, M., Barbin, A., and Bartsch, H. (1996) *J. Cancer Res. 56*, 1267−1271.
7. El Ghissassi, F. E., Barbin, A., Nair, J., and Bartsch, H. (1995) *Chem. Res. Toxicol. 8*, 278−283.
8. Nair, J., Vaca, C. E., Velic, I., Mutanen, M., Valsta, L. M., and Bartsch, H. (1997) *Cancer Epidemial Biomarkers Prev. 6*, 597−601.
9. Jacobsen, J. S., and Humayun, M. Z. (1990) *Biochemistry 29*, 496−504.
10. Palejwala, V. A., Simha, D., and Humayun, M. Z. (1991) *Biochemistry 30*, 8736−8743.
11. Basu, A. K., Wood, M. L., Niedernhofer, L. J., Ramos, L. A., & Essigmann, J. M. (1993) *Biochemistry 32*, 12793−12801.
12. Moriya, M., Zhang, W., Johnson, F., and Grollman, A. P. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11899−11903.
13. Simha, D., Yadav, D., Rzepka, R. W., Palejwala, V. A., and Humayun, M. Z. (1994) *Mutat. Res. 304*, 265−269.
14. Zhang, W., Johnson, F., Grollman, A. P., and Shibutani, S. (1995) *Chem. Res. Toxicol. 8*, 157−163.
15. Shibutani, S., Suzuki, N., Matsumoto, Y., and Grollman, A. P. (1996) *Biochemistry 35*, 14992−14998.
16. Kusmierek, J. T., and Singer, B. (1982) *Biochemistry 21*, 5723−5728.
17. Hang, B., Chenna, A., Rao, S., and Singer, B. (1996) *Carcinogenesis 17*, 155−157.
18. Hang, B., Singer, B., Margison, G. P., and Elder, R. H. (1997) *Proc. Natl. Acad. Sci. U.S.A. 84*, 12869−12874.
19. Korobka, A., Cullinan, D., Cosman, M., Grollman, A. P., Patel, D. J., Eisenberg, M., and de los Santos, C. (1996) *Biochemistry 35*, 13310−13318.
20. Cullinan, D., Korobka, A., Grollman, A. P., Patel, D. J., Eisenberg, M., and de los Santos, C. (1996) *Biochemistry 35*, 13319−13327.
21. Cullinan, D., Johnson, F., Grollman, A. P., Eisenberg, M., and de los Santos, C. (1997) *Biochemistry 36*, 11933−11943.
22. Plum, G. E., and Breslauer, K. J. (1994) *Ann. N.Y. Acad. Sci. 726*, 45−56.
23. Vesnaver, G., Chang, C.-N., Eisenberg, M., Grollman, A. P., and Breslauer, K. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3614−3618.
24. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1992) *Biochemistry 31*, 12096−12102.
25. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1995) *Biochemistry 34*, 16148−16160.
26. Gelfand, C. A., Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1996) *Biopolymers 38*, 439−445.
27. Gelfand, C. A., Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1998) *Biochemistry 37*, 7321−7327.
28. Zhang, W., Rieger, R., Iden, C., and Johnson, F. (1995) *Chem. Res. Toxicol. 8*, 148−156.
29. Marky, L. M., and Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
30. Breslauer, K. J. (1994) *Methods Mol. Biol. 26*, 347−372.
31. Pilch, D. S., Plum, G. E., and Breslauer, K. J. (1995) *Curr. Opin. Struct. Biol. 5*, 334−342.
32. Gray, D. M., Ratliff, R. L., and Vaughan, M. R. (1992) *Methods Enzymol. 211*, 389−406.
33. Chalikian, T. V., Sarvazyan, A. P., Plum, G. E., and Breslauer, K. J. (1994) *Biochemistry 33*, 2394−2401.

BI981090B