NMR Structures of Damaged DNA

Mark Lukin and Carlos de los Santos*

Department of Pharmacological Sciences, State University of New York at Stony Brook, School of Medicine, Stony Brook, New York 11794-8651

Received September 9, 2005

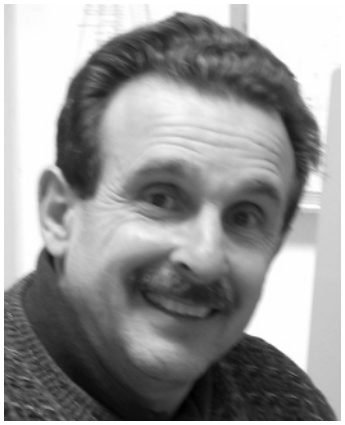
Contents

10.1021/cr0404646 CCC: $59.00
Published on Web 02/08/2006

1. Introduction

Attesting to the spectacular growth in the field of structural biology, the total number of coordinate files deposited in the Protein Data Bank was rapidly approaching 30 000 entries by the first week of March 2005.[1] While X-ray crystallography is the driving force behind this growth, NMR spectroscopy has an increasing contribution in the area of nucleic acid struture, where almost 45% of the entries have used NMR techniques. Concurrent factors contribute to the relative advantage of NMR for studying the structure of small nucleic acid molecules. On one hand, the intrinsically "polymorphic" nature of DNA and RNA molecules, with no or weakly pronounced globular structure and a largely undifferentiated molecular surface, makes it difficult to obtain good diffracting crystals.[2] On the other hand, advances in enzymatic synthesis of RNA and DNA now allow the preparation of large quantities of uniformly labeled [^{13}C-, ^{15}N-] samples,[3−5] opening the door to the development of new NMR pulse sequences that provide new insights on nucleic acid structure. In fact, the increased number of multidimensional NMR experiments now available not only provides structure independent assignment methods but also permits the observation of spin interactions, which were mostly inaccessible with unlabeled samples.[6−10] In addition, NMR spectroscopy has the capability to monitor, at the molecular level, base pair and other dynamic processes within a broad range of time scales (10^{-9}−10^{-3} s), making it a unique complement for crystallographic studies.[11−17] In this article, we will review NMR studies that have characterized the solution structure of short duplexes having damaged nucleotides that result from the action of physical or chemical insults. We will limit our survey to DNA duplex structures that bear relevance for fundamental cellular processes, including DNA repair and replication. Furthermore, we will cover NMR studies of duplexes containing artificial residues that have increased our understanding of DNA stability and protein recognition.

Mark Lukin was born in Izhevsk city, USSR. He received his M.D. in chemistry from the Samara State University. He did his Ph.D. research in the Engelhardt Institute of Molecular Biology (Moscow, Russia) under the supervision of Prof. A. Krayevsky, working on the design and synthesis of DNA polymerase inhibitors and antiviral drugs. After obtaining his Ph.D. degree in 1993, he joined the laboratory of Dr. R. Beabealashvilli at the Institute of Experimental Cardiology, Moscow, where he studied the structure and function of natural ribozymes and their artificial analogues. In 2004 he joined the de los Santos laboratory, where he is developing methods for the synthesis of ^{2}H and/or ^{13}C labeled nucleotides, needed for structural studies of damaged DNA by high-resolution NMR. His research interests include the study of the nucleic acid structure and understanding the interrelation between the primary structure of biological macromolecules, their spatial organization, dynamics, and biological functions.



Carlos de los Santos was born in 1953 in Buenos Aires, Argentina. He received his Ph.D. degree, in 1988, from the University of Buenos Aires, studying the metabolism of *Crithidia Fasciculata* using carbon-13 nuclear magnetic resonance, under the supervision of Dr. Benjamin Frydman. Soon after graduation, he came to the USA, where, supported by an International Fellowship from the CONICET (Argentina) and a Fogarty Fellowship from the NIH (USA), he was a postdoctoral fellow in the laboratory of Dr. Dinshaw Patel, at Columbia University in New York. From 1992 to 2004, he worked as an Associate Laboratory Member in the Memorial Sloan-Kettering Cancer Center in New York. In 1994, he started working as a Research Assistant Professor in the Department of Pharmacological Sciences of the State University of New York at Stony Brook, where he was later appointed Assistant Professor and Associated Professor in 2004. His continuous research interests are in the structural biology of damaged DNA as a way to understand basic molecular mechanisms of lesion recognition and repair. Presently, he is working on the structure of clustered DNA damage and exocyclic base lesions. He is a member of the American Mutagen Society and the American Chemical Society.

Despite the progress on chemical approaches for the synthesis of isotopically enriched oligonucleotides,[18] their preparation is still laborious and resource consuming. Besides, the damaged nature of DNA lesions precludes the

enzymatic preparation of labeled samples for NMR studies. Therefore, the most sophisticated multidimensional techniques have had limited application for studying the structure of damaged DNA and the large majority of studies reviewed here have followed well-established ^{1}H and, to a lesser extent, ^{31}P spectroscopic methods.[19−22] Briefly, proton chemical shifts provide information about base pair alignments and a qualitative description of lesion-induced perturbations. Analysis of NOESY spectra allows the assignment of proton signals and the computation of interproton distance bounds that are used for structural refinement. Examination of two-dimensional correlation spectra, such as COSY, DQF-COSY, and TOCSY, complete the assignments of proton signals and can measure *J*-coupling constants among sugar protons, which determine the sugar conformation. Observation of ^{31}P chemical shift alterations reflects the conformational transition in the phosphate backbone. Following the NMR characterization, computational methods generate refined structures. Manual computer-assisted modeling, distance-restrained potential energy minimization (REM), and restrained molecular dynamics (RMD) all have been used to generate three-dimensional models of damage-containing duplexes. RMD-generated models are more accurate, and in general, atomic coordinates obtained by this method are deposited in the Protein Data Bank. Throughout this review, we refer to the accession codes of specific structures.

In some instances, damage formation results in the loss of protons that are key for spectrum assignment and RMD. For example, base oxidation or formation of an abasic site removes base protons, a fact that may underdetermine the local duplex conformation at the lesion site. In these cases, special care must be taken to evaluate the final RMD structure since it may be biased by the intrinsic limitations or the implementation of molecular dynamics simulations. In addition, NOESY experiments can establish only short interproton distances, less than 5 Å, which in duplex DNA correspond to protons located in the same or sequential residues. Therefore, the use of NOE distance restraints for global shape determination leads to accumulation of experimental errors and this type of data obtained solely by these methods must be considered with caution. The use of residual dipolar coupling (RDC) restraints significantly improves the NMR accuracy of the duplex curvature determination;[23] however, the unavailability of enriched material has limited its application in the field of damaged DNA. Frequently, description of the three-dimensional structure information alone is insufficient to explain the biological properties of the damage. Measurements of chemical exchange rates, especially base pair lifetimes, have provided important information that aids understanding of local dynamics at specific DNA sites.[24−27] Whenever appropriate, we present this type of studies in the review. The experimental framework for studying spin relaxation rates and how they relate to local and global dynamics of macromolecules is well established.[28] However, since the limited number of studies performed on this area has been the topic of a recent comprehensive review,[29] we will not discuss them here.

Notwithstanding the existence of overlaps and exceptions, we have organized this review following the main pathways of damage generation and repair mechanisms. We will review first the structure of DNA lesions that result from the attack of small reactive particles, such as reactive oxygen species or free radicals, which may be generated endogenously or as a consequence of ionizing radiation. This type of damages includes DNA strand breaks, abasic sites, oxidized bases, oxidatively fragmented products, and base deamination products. Most of these lesions are omnipresent in cellular DNA and are the substrate of base excision repair (BER).

We will next discuss the structure of DNA damages that can be repaired by more than one mechanism but are grouped based on the similarity of their structure or generation mechanism. These damages include non-natural nucleotides, nucleotide alkylation products, pyrimidine photoproducts, and platinum complexes.

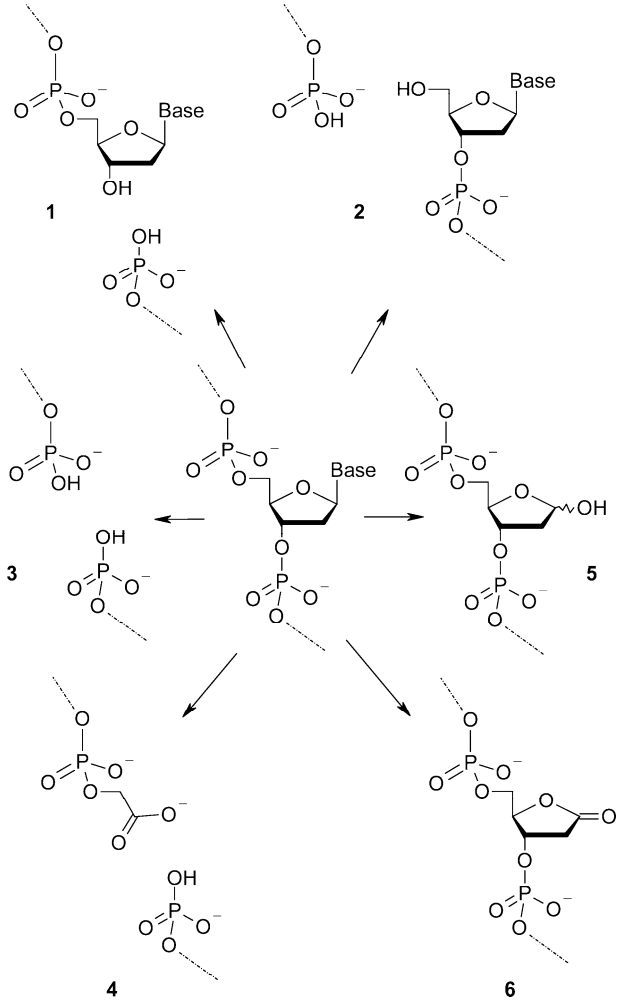
Many DNA damages result form the addition of polyaromatic hydrocarbons or heterocyclic amines to the C^8, N^6, N^2, or N^7 atoms of purine bases. They represent a chemically diverse set of ubiquitous environmental pollutants, which are present in car exhaust, tobacco smoke, pesticides, food mutagens, and fungal toxins and shared the property of being processed by nucleotide excision repair (NER). We will present all these studies in the same section.

Since they are repaired by a unique enzyme system, base pair mismatches originated during DNA replication can be regarded as lesions and many NMR studies have described the structure of many mismatch-containing duplexes.[30,31] However, we will only describe the structure of lesion-containing mismatches, which may be considered models of lesion-induced mutagenic intermediates, and leave the study of unmodified mismatches beyond the scope of this review. Also excluded from this survey are NMR studies of RNA structures, unmodified DNA duplexes, DNA triple helices, four-stranded DNA structures, parallel-stranded DNA and noncovalent intercalators, and drug binders, which, in general, fail to activate repair mechanisms. Similarly, the discussion of experimental and computational methods employed for nucleic acid structure determination is outside the scope of this article. While some excellent reviews covering the structure of a specific kind of lesions have appeared recently, more than 15 years have passed since the last comprehensive surveys on this subject.[19,20,30] Thus, this review will cover NMR structural studies of damaged DNA duplexes performed mostly during the last 15 years. Throughout this review we will follow IUPAC−IUB definitions and recommendations for DNA structural parameters.[32]

2. Abasic Sites, DNA Nicks, and Gaps

Fragmentation of the DNA sugar phosphate backbone leads to the appearance of nicks, gaps, and abasic sites (Scheme 1). DNA nicks originate by cleavage of the phosphodiester bond between adjacent nucleotides, and many enzymatic processes form them in cellular DNA. For instance, DNA endonucleases catalyze the hydrolysis of phosphodiester bonds, creating nicks **1** and **2**. Formation of a reversible single-strand break is a key step of DNA topological transitions catalyzed by type I topoisomerase. Nicks are required intermediates after DNA repair synthesis or during DNA replication of the lagging strand, before DNA ligases can seal them. One-nucleotide gaps originate after removal of a single nucleoside from double-stranded DNA. They may be formed chemically or enzymatically by the AP-lyase activity of DNA glycosylases that produce single-nucleotide gaps **3** with 3′- and 5′-phosphorylated ends. Radiomimetic antibiotics such as bleomycin can abstract the deoxyribose 4′-hydrogen, creating a free radical, which may react with molecular oxygen to form a hydroperoxy nucleotide. This compound is unstable, breaking into a single nucleotide gap **4** with 3′-phospholycolate-5′-phosphate ends.

Scheme 1. Nucleotide Fragmentation in DNA



Hydroxide radicals generated by ionizing radiation induce a similar reaction.

Many processes cause the hydrolysis of the glycosidic bond in DNA, creating an aldehydic abasic site **5** (AP sites) and a free nucleobase.[33] Even under physiological conditions, the solvolytic or oxidatively induced hydrolysis of the glycosidic bond proceeds with a small but detectable rate. In eukaryotic cells, spontaneous hydrolysis forms approximately 10^4 AP sites per cell per day.[34] Purine N^7 alkylation, by low molecular weight compounds or bulky environmental pollutants, can augment considerably the rate of glycosidic bond hydrolysis. In addition, the continuous removal of damaged bases by DNA glycosylases increases the burden of abasic sites present in the cells.[35] The main processes that generate AP sites, nicks, and gaps, as well as their repair by BER,[36] are outlined in Figure 1.

AP sites are often called apyrimidinic or apurinic sites, reflecting the identity of the base that has been lost. This designation is not a formality because, although an abasic site is simply a deoxyribose residue in DNA, its origin can be different and some of its biophysical properties depend considerably on the partner base, sometimes called the orphan base. For example, apyrimidinic sites are formed mainly via enzymatic removal of damaged bases,[37] while spontaneous as well as alkylation-induced hydrolysis of the glycosylic bond contributes solely to the formation of apurinic sites.[38]

The lactone abasic site **6**, a different type of AP lesion that is much less abundant than the aldehydic AP site, is generated as the result of radical abstraction of the 1′-hydrogen followed by base elimination. Enediyne antibiotics of the neocarcinostatin family, UV light, or γ-radiation can form lactone AP sites inside the cell.[39]

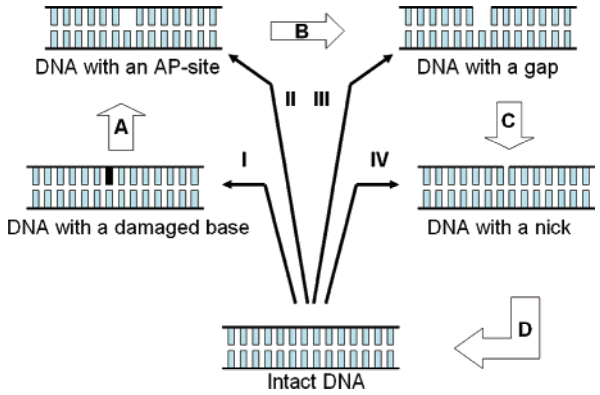


Figure 1. Cartoon representation of the BER pathway and generation of small DNA damages. Damage generation processes: I, oxidation, alkylation, or deamination of nucleobases; II, acid-catalyzed, alkylation- or oxidadion-induced, or solvolytic cleavage of the nucleotide glycosylic bond; III, oxidative fragmentation of the nucleotide moiety; IV, chemical or enzymatic incision of phosphodiester bonds. BER pathway: A, excision of damaged bases by DNA glycosylases; B, DNA strand incision AP-endonuclease or AP-lyases; C, gap-filling synthesis by repair polymerases; D, nick ligation by DNA ligases.

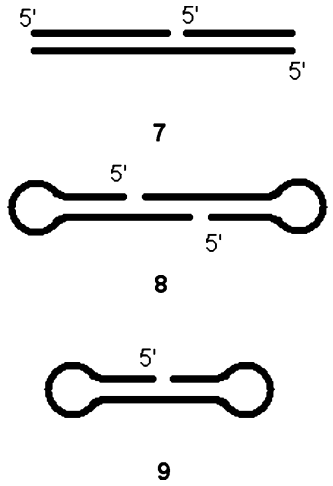


Figure 2. Models used for studying DNA with nicks or gaps.

2.1. Strand Breaks

2.1.1. Model Systems for NMR Study of Nicks and Gaps

Three types of model systems have been used to mimic DNA with nicks or gaps. The earliest system, **7**, consisted of a long oligodeoxynucleotide segment annealed to two shorter complementary fragments[40] (Figure 2). Despite its simplicity, it had two major drawbacks: the possibility of alternative sample annealing and fraying of duplex termini, which together have limited its application. These shortcomings are generally eliminated by the use of dumbbell systems. Bimolecular dumbbells **8** are composed of two identical hairpins with complementary sticky ends that, due to their symmetry, will have a reduced set of resonances.[41] The induced structural distortions of one nick on the second nick were considered negligible because of the four to six base pairs spaced between them. However, the most adequate model is that of the hexaethylene glycol-linked oligonucle-

otide dumbbell **9** reported recently.[42] The artificial linkers stabilize the duplex termini very effectively, such that the system behaves almost as an infinitely long duplex with a single damage, eliminating any concern about cooperativity between the two damages.

2.1.2. Effect of Nicks and Gaps on Duplex Structure

In contrast to expectations, early electrophoretic studies of nicked duplexes have shown that formation of a strand break does not cause a significant increase of DNA flexibility, implying that the DNA around the lesion essentially retains its ordered structure.[43] Subsequent NMR studies confirmed this suggestion. Assignment of the imino proton spectrum indicated that Watson−Crick (WC) alignments are conserved for all nonterminal base pairs of nicked duplexes.[40] In addition, key sequential NOE contacts between base and sugar protons, characteristics of a regular B-form DNA, were also present. ^{31}P NMR spectra show small chemical shift dispersion, further supporting the presence of a regular B-form helix for nicked duplex.

A direct comparison between the nicked dumbbell **8** and an unmodified duplex with the same stem sequence shows similar intensity of base and sugar NOE cross-peaks for the corresponding nucleotides, suggesting that local conformational differences between these systems are small. In addition, phosphorylation at the 5′-side of the nick does not affect its structure considerably, causing only small distance variations, estimated by the authors not to exceed more than 1 Å, between the aromatic protons of nick-flanking residues.[41] The NMR spectra of all nick mimics studied to date exhibit slightly broadened and moderately upfield-shifted imino proton signals for the nick-flanking base pairs, demonstrating partial fraying of the duplex at the nick site.[42,44−46] However, the temperature dependence of exchangeable proton spectra shows that imino proton signals of nick-flanking and internal base pairs disappear simultaneously, indicating that these dumbbells melt cooperatively.

In contrast to the case of nick-containing DNA, a conformation cannot be proposed for gap-containing DNA, where all bases are stacked without distortion of the sugar−phosphate backbone. Theoretically, four conformations are possible around the gap: a disordered state with frayed gap-flanking base pairs, a double helix with a bulged orphan nucleotide, a straight double helix with a gap, or a kinked double helix. Gel electrophoresis data predict a larger conformational mobility for duplexes with a gap than for nicked ones.[43] Solution structures of four gap-containing duplexes of different sequence context are available to date (Chart 1), two with nonphosphorylated gap ends **(10a**,**b)** and the other two with 5′-phosphorylated/3′-phosphoglycolated ends **(11a**,**b)**.

Comparison of the NMR spectra of a gap-containing duplex with those of an orphan guanine (**10a**) and those of a nicked duplex with the same sequence (**10c**) showed that guanine imino protons of the gap-flanking base pairs are WC hydrogen bonded. The amino protons of complementary cytosine residues are resolved and have a chemical shift difference of about 1 ppm, indicating that they also participate in hydrogen bond formation. In contrast to the imino proton signal of terminal base pairs that disappears at 16 °C, the imino protons of gap-flanking dG·dC pairs are quite visible at this temperature.[45] Similar observations have been done for the case of a gap-containing duplex with an orphan adenosine **10b**,[46,47] indicating that base pairs flanking the gap are more stable and ordered than the duplex ends. However, this is not the case of duplexes **11a**,**b**, mimicking the damage caused by the bleomycin antibiotic family.[48] In this case, the gap ends are phosphorylated. The imino proton signals of gap-flanking base pairs are significantly broaden, due to increased solvent exchange, and shift upfield by about 0.5−0.7 ppm. Base pairs next to those flanking the gap are affected also, showing a reduced intensity of their imino proton signals at 25 °C.[49] The different behavior could be attributed to an increase in electrostatic repulsions caused by the charged ends, although variations between duplex sequences **10** and **11** may also play a role. Despite differences in lesion-site stability, the NMR characterization of gap-containing duplexes indicates the presence of a single B-type conformation with the orphan adenine or guanine residues stacked in the helix. The presence of the gap results in the loss of essential protons, needed for accurate structural refinement, and thus, a reliable determination of duplex curvature and discrimination between straight and kinked duplex conformations is not possible. To compensate for the lack of experimental restraints and increase the accuracy of RMD, structure calculation of duplex **10a** has been done in the presence of explicit three-site transferable intermolecular potential (TIP3P) water molecules. The results suggest the presence of two about equally populated duplex conformations, one with an approximately 28° anisotropic bend at the gap site and the other with an almost straight axis.[45] Although this observation was based on a limited set of experimental distances, it is consistent with results from two-dimensional gel electrophoresis experiments that predict an anisotropic bend on gap-containing DNA.[50]

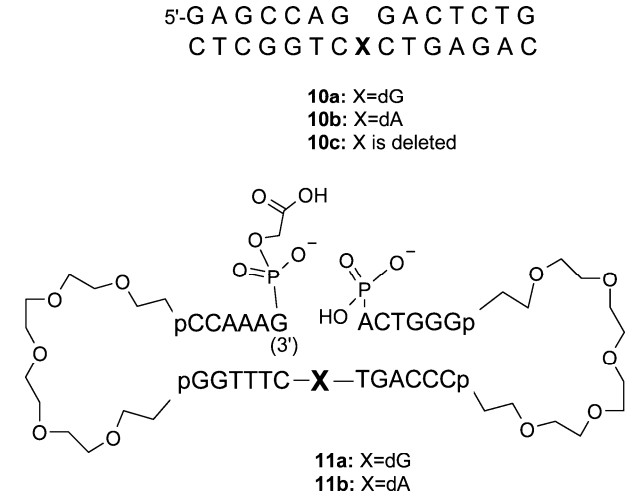
It is worth noting that gap-containing duplexes **10** and **11** have an orphan purine nucleotide. Since purine residues have greater stacking ability than pyrimidines, conclusions based on the structures of **10** and **11** cannot be automatically expanded to gapped duplexes with orphan pyrimidine residues or to those where the gap is in a different sequence context. To date, no NMR studies are available for duplexes with a pyrimidine base opposite the gap.

2.2. Abasic Sites

2.2.1. AP Site Chemistry. Methods of AP Site Generation for NMR Studies. AP Site Mimics

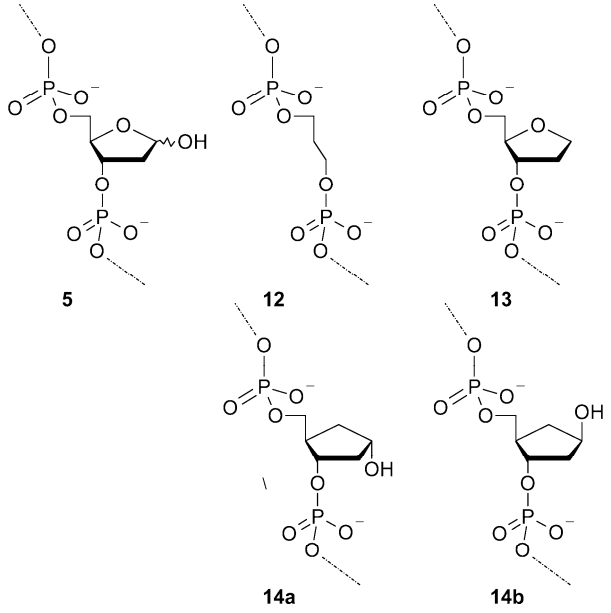
AP sites **5** are 2′-deoxyribose moieties phosphorylated at the 3′- and 5′-positions. Since these two phosphates, located

Chart 1. Gap-Containing Duplexes



at the β- or δ-position relative to the aldehyde function, are good leaving groups, the direct preparation of AP site-containing oligonucleotides for structural studies is impossible using standard phosphoramidite chemistry. Two strategies are used to circumvent this obstacle. The first approach generates the AP site postsynthetically, using an oligodoexynucleotide containing a precursor at the desired position. The abasic site precursor may be a deoxyribose moiety having a photolabile protecting group[51] or a 2′-deoxyuridine residue that generates an AP site after enzymatic treatment with uracil DNA glycosylase (UDG).[52] The second approach makes use of chemically stable analogues (Chart 2) that retain

Chart 2. AP Site 5 and Its Stable Analogues 12, 13, and 14a,b



the major structural features of natural AP sites. The simplest AP site isoster, the 1,3-propanediol linker **12**, was employed only occasionally because of its significantly greater flexibility.[53,54] The most widely used AP site analogue is the 2-hydroxymethyl-3-hydroxytetrahydrofuran residue **13** (THF),[55] which lacks the hemiacetal hydroxyl. In addition, two isomeric carbocyclic basic site analogues **14a**,**b** have been used recently in order to understand the effect of AP site anomeric configuration on duplex structure.[56]

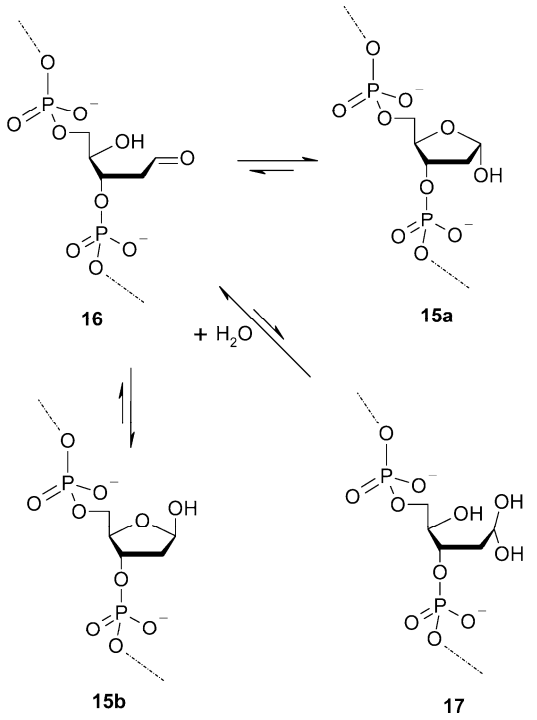
2.2.2. Chemical State of AP Sites in DNA

A deoxyribose moiety incorporated into DNA is a γ-hydroxyaldehyde and, in principle, can exist in four interconverting chemical states (Scheme 2): α or β cyclic hemiacetal (**15a**,**b**), a free hydroxyaldehyde **16**, or a geminal diol **17**, the hydrated form of the hydroxyaldehyde. To establish how DNA composition may affect the equilibrium among these forms, ^{13}C and ^{17}O NMR spectroscopy has been used to study duplexes containing a single AP site opposite dA, dT, dG, or dC.[57,58] The spectra showed that, in DNA, over 99% of the natural abasic site exists in the cyclic hemiacetal form, with less than 1% of the open hydroxyaldehyde state and no geminal diol detected. The ratio between the α- and β-anomers is about equal for all orphan residues, with the exception of dC, where only a single ^{13}C1 peak is present.[57]

2.2.3. Conformation of DNA Containing AP Sites

A thermodynamic study comparing gap- and AP site-containing duplexes showed that both lesions cause almost

Scheme 2. Interconversion of AP Site States



identical destabilization, indicating that the presence of a deoxyribose moiety does not make any meaningful addition to the stability of gap-containing samples.[54] Therefore, gap- and AP-site-containing duplexes can be expected to share many structural features, with probably a more ordered lesion site conformation for the latter lesion. Numerous studies[59,60,62−65] have mostly confirmed this expectation, showing that formation of an AP site affects the duplex conformation mainly at the lesion site and, to a lesser extent, adjacent base pairs (Table 1, Figure 3). Only when the AP site forms part of an (A·T) tract, did perturbation of the duplex structure go beyond the lesion site and neighboring base pairs.[61]

The ^{31}P NMR spectra of AP-site-containing duplexes show most phosphate resonances within the chemical shift range that corresponds to canonical B-form conformations, with only one or two signals, mapped to lesion site residues, appearing upfield. NOESY spectra display sequential cross-peaks between the base and 1′, 2′, 2″, and 3′ sugar protons of 5′-flanking residues, indicating the presence of a right-handed B-form duplex conformation. In addition, imino proton signals appear in the 12−14 ppm chemical shift range and show NOE cross-peaks to their WC partners, establishing hydrogen bonding for all canonical base pairs of the duplex. However, the duplex conformation at the lesion site is highly dependent on the type of adjacent base pairs and the orphan base.

2.2.4. Role of the Orphan Base and Sequence Context

The existing NMR data show that local duplex structure depends on the identity of the orphan base and the type of AP residue used in the study, demonstrating the fundamental role of hydrophobic and stacking interactions on lesion site conformation. Whether facing a natural or an AP site analogue, orphan purine bases, being largely hydrophobic, stack always inside the helix. The position of orphan pyrimidine residues shows more variability, but they are intrahelical when paired to the natural AP site. When the THF abasic site analogue is used, orphan dC residues adopt

Table 1. Structure of AP-Site-Containing DNA

Duplex sequence	AP site type[a]	AP site position[b]		Orphan base position[b]	PDB code	Ref
		α	β			
CGCGA**X**ACGCC GCGCTATGCGG	d-Rib	I	I	I	1A9J, 1A9I	59
CGCGA**X**ACGCC GCGCTCTGCGG	d-Rib	*	I	I	1A9H, 1A9G	59
CGCGA**X**ACGCC GCGCTATGCGG	d-Rib	E	I	I		60
CGCATT**X**TTGCG[c] GCGTAAAAACGC	d-Rib	E / Maj	I	I	1SJL, 1SJK	61
CCAAAG**X**ACTGGG GGTTTCATGACCC	d-Rib	E	E	I	1GJ0, 1GIZ, 1G5E, 1G5D	62
CGTG**X**GTGC[d] GTACTCACG	d-Rib	E ↔ I		I		51
CGTG**X**GTGC GCACGCACG	THF	I		I		63
CGTG**X**GTGC GCACTCACG	THF	E ↔ I		E ↔ I		63
CGTG**X**GTGC GCACCCACG	THF	E		E		63
CGTG**X**GTGC GCACACACG	THF	I		I		64
AGGCT**X**GACTGC[e] TCCGACCTGACG	THF	I		E / Min		73
GTAC**X**CATG CATGAGTAC	THF	I		I		55
CGCAC**X**CACGC GCGTGTGTGCG	THF	E / Min		I		65
CGCAC**X**CACGC GCGTGTGTGCG	d-Rbl	I		I		70
GTAC**X**CATG CATGAGTAC	PD			I		53
CGTAC**X**CATGC GCATGAGTACG	α-CAP	I	I	I	1XCY	56
CGTAC**X**CATGC GCATGAGTACG	β-CAP	I	I	I	1XCZ	56

[a]d-Rib, 2'-deoxyribose; THF, 2-hydroxymethyl-3-hydroxy tetrahydrofurane; d-Rbl, 2'-deoxyribonolactone; PD, 1,3-propandiol; CAP, carba AP-site. [b]I, intrahelical, E, extrahelical; NO, signal not observed for this isomer; Min, displacement toward the minor groove; Maj, displacement toward the major groove. [c]Duplex with A•T tract. [d]No separate conformation was obtained for the α and β anomers. [e]Topoisomerase II recognition site. *α-AP site isomer reportedly absent in this duplex.

extrahelical conformations, whereas an orphan T base stacks inside the helix when it is flanked by purine bases, or tends to be solvent exposed when it is flanked by cytosine residues that have weak stacking ability.

2.2.5. Effect of the Anomeric Hydroxyl

Since natural AP sites have low chemical stability, a large number of structural studies have been done using the chemically stable THF analogue. Duplex DNA containing THF abasic sites is a good substrate for the action of DNA endonucleases,[66] suggesting that it is an adequate AP site analogue. However, the lack of the anomeric hydroxyl group increases the flexibility of the THF ring, a fact that could give rise to conformational differences between natural and THF AP sites. Unfortunately, since most structural studies of natural AP-site- or THF-containing duplexes solved to date have used different sequence contexts, a direct comparison of the influence of stereoelectronic effects and anomeric OH hydrogen bonding on duplex conformation is not possible.

Another way to estimate the effect that the hemiacetal hydroxyl group has on AP site conformation is to compare the structure of α and β AP site anomers, which exchange slowly in duplex DNA. The NMR structures available to date, of duplexes with dA, dC, or T placed opposite a natural AP site, suggest that the deoxyribose position depends on the type of orphan base, the configuration of the deoxyribose acetal carbon, and lesion-flanking bases. When dA is the orphan base, both isomers of natural AP sites may adopt an external conformation or, in a different sequence context, only the β-anomer is inside the duplex (see Table 1). In the latter case, the authors proposed that a water-mediated hydrogen bond, formed between the orphan dA residue and the hemiacetal hydroxyl of β-AP site **15b**, keeps the deoxyribose moiety inside the helix, whereas formation of a water bridge is impossible for the α-AP anomer **15a**, which then adopts an extruded conformation.[59] On the basis of these observations and the X-ray structure of the human ApeI endonuclease/THF-containing duplex complex, it was proposed that the enzyme only incises the α-anomer of the lesion.[67] Pyrimidine bases are too small for the formation of such a bond that, according to the authors, could increase the probability of the extrahelical AP site orientation.

A striking finding of natural abasic sites has been the absence of the α-AP anomer **15a** in a duplex having a dC orphan residue,[59] an observation that is in full agreement with an earlier ^{13}C NMR study performed by the same group.[57] Possibly, this feature reflects the effect of the orphan cytosine residue, the most basic aglycon in DNA, on the deoxyribose

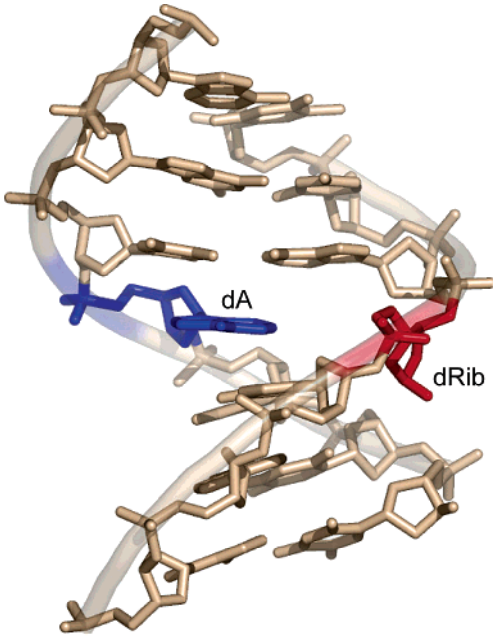


Figure 3. Structure of a DNA duplex containing a centrally located apyrimidinic site. The figure shows the 2′-deoxyribose moiety (in red) displaced out of the helix and the orphan adenine (in blue) intrahelical without major distortions of the double helix. Hydrogen atoms are not shown. Prepared from PDB entry 1G5D[62] using PyMol.[555]

mutarotation. A unique role of cytosine residues has been shown in the phosphate bond scission catalyzed by HDV ribozyme[68] and, presumably, in the opening of the malonodialdehyde exocyclic guanine adduct[69] described in section 5.3.

To better understand the effect that C1-deoxyribose configuration has on the conformation of the AP site in DNA, our group introduced nonmutarotating α- and β-carbocyclic analogues **14a**,**b** of the lesion. These AP analogues lacks *gauche* effects along the C3−C4 and C4−C5 bonds and the *anomeric* effect, which possibly drive the conformation of the sugar. But, in contrast to the case of the natural abasic site, there is no exchange between the α- and β-forms of carbocyclic AP sites, facilitating the independent observation of structural differences at the lesion site. The NMR characterization of DNA duplexes containing an α- or β-carba-AP site residue opposite dA indicates a very similar duplex conformation with both isomers of the carba-AP site inside the helix. In addition, the RMD models show that both the α- and β-hydroxyl groups participate in hydrogen bonding.[56] It is worth noting that human ApeI endonuclease incises both **14a** and **14b** AP site analogues with similar efficiency *in vitro*, indicating that C1-deoxyribose configuration is not important for enzymatic recognition.[56]

Inside the cell, natural AP sites can exist also as a deoxyribonolactone **6** that possesses higher rigidity due to the presence of an sp^2-carbon. The NMR characterization of a duplex containing a deoxyribonolactone abasic site opposite T shows a well-ordered right-handed helical structure with both the AP site and its partner T residue inside the helix.[70]

2.2.6. AP Sites in Unusual DNA Contexts

Some DNA sequences are preferential binding sites for proteins that play a vital role in the regulation of gene expression, DNA replication, and related processes. Formation of abasic sites within these sequences can significantly change their conformation, thus impairing regulatory functions. Two such motifs containing a single AP site have been studied to date: a human topoisomerase II binding site and an $d(A\cdot T)_5$ tract. Some unusual features have been observed in both cases.

DNA topoisomerases resolve topological problems associated with DNA replication and transcription. Type II topoisomerases act by creating reversible, bistrand nicks four nucleotides apart and passing one strand through the gap existent in the DNA−toposiomerase complex before sealing both nicks.[71] The presence of an AP site in the topoisomerase DNA binding sequence augments the enzymatic rate of DNA cleavage by nearly 20-fold, increasing the number of double-strand breaks far above the acceptable level and, therefore, jeopardizing genome integrity.[72] NMR spectroscopy has solved the structure of a duplex having an apurinic site located between the two cleavage sites of topoisomerase II, at their +2 positions.[73] Analysis of the NOESY spectra establishes that WC alignments are present for all canonical base pairs of the duplex, but standard base−sugar proton interactions at the lesion site are missing on both the damaged and undamaged strands. The RMD model shows the orphan cytosine residue to be extrahelical, in the minor groove of the duplex, and the THF residue to be inside the helix. As a result of partial collapse of the gap opposite the AP site, the duplex bends by 10°, reducing the distance between the topoisomerase cleavage sites from 15 Å in the undamaged duplex to 11 Å. Such steric effects could be responsible for shifting the equilibrium between the closed and cleaved forms of DNA in the topoisomerase active site.

Phased $d(A\cdot T)_n$ tracts bend the double-stranded helix, facilitating the cellular organization of DNA. In addition, $d(A\cdot T)_n$ tracts are often present near replication origins,[74] promoter regions,[75] and other consensus sequences, suggesting their importance in gene regulation processes (see ref 76 and references therein). The NMR characterization of a duplex having a natural AP site positioned at the center of a $d(A\cdot T)_n$ tract profoundly affects the double-helical conformation. The structures of both the α- and β-anomer-containing duplexes differ considerably from each other and from that of the undamaged $d(A\cdot T)_n$ tract control. The duplex containing the β-deoxyribose has a straight conformation, but the shape of its α-counterpart is more compressed (Figure 4). In contrast to the case of other AP-site-containing duplexes, the structure of these duplexes is highly perturbed, with disruption occurring as far as four base pairs away from the lesion site.[61] While a complete understanding of such differences remains to be established, it is likely that the reduced thermal and thermodynamic stability of duplexes containing $d(A\cdot T)_n$ tracts is a major contributing factor.

2.3. AP Site Deletion–Mutation Models and Clustered AP Sites

Abasic sites are noncoding lesions that can result in either mutation during DNA replication or cell death. Induction of the SOS response in *E. coli* facilitates synthesis past the AP, which proceeds predominantly by incorporating dAMP. In contrast, there is not a clear pattern of single-base mutations in eukaryotic cells,[78] and a small number of one-base deletions (−1Δ) are also observed.[79] While all the studies mentioned above can represent structural models for base substitution mutations, duplex samples with a different number of residues in each strand are needed to describe the conformation of −1Δ intermediates. NMR studies have characterized the conformations of two heteroduplexes containing an AP site in a −1Δ context, reaching basically the same conclusions.[80,81] Data analysis showed that both

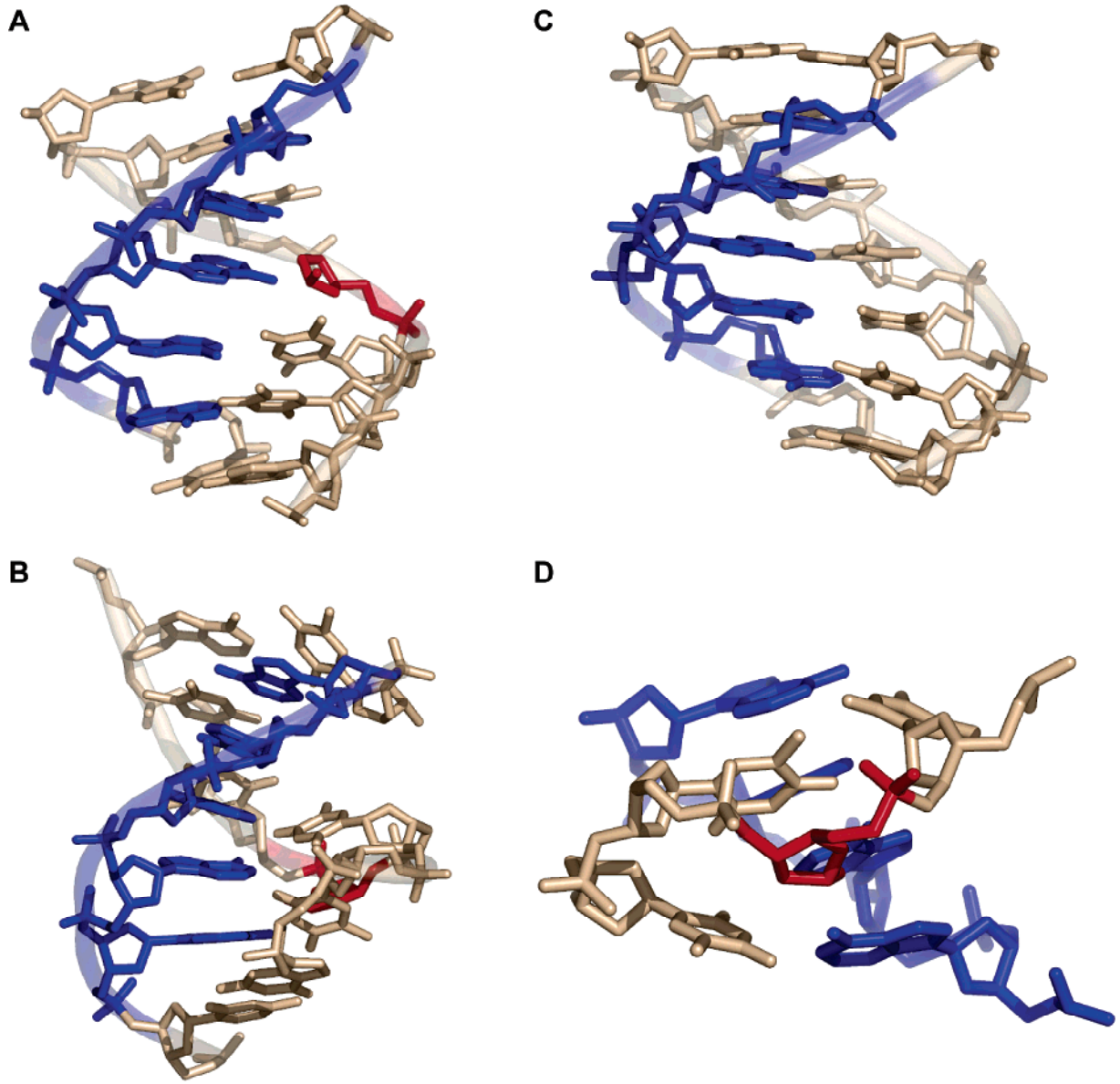


Figure 4. Structure of DNA containing a natural AP site opposite the central adenine of a dA_5 tract.[77] The figure displays a central seven-nucleotide duplex segment with the AP residue in red and the dA_5 tract in blue. Panel A, the β-AP isomer; panel B, the α-AP isomer; panel C, unmodified duplex. Coordinates are taken from PDB entries 1SJL, 1SJK, and 1FZX, respectively. Panel D depicts the central three-nucleotide segment of the duplex shown in panel B but with the damaged strand prominent. Prepared using PyMol.[555]

heteroduplexes adopt regular right-handed helical conformations with WC base pair alignments throughout and the AP site residues pushed out of the helix without adopting a defined conformation. The refined structures show the lesion-flanking base pairs properly stacked in the helix and the AP site exposed to solvent in both models.

Abasic sites formed by spontaneous depurination or by the enzymatic removal of damaged bases are stochastically scattered all through the genome. Similarly, oxidized bases and AP sites produced by chemical agents also follow a random distribution in cellular DNA. In contrast, ionizing radiation can generate several hydroxyl radicals in a single energy deposition event, which may lead to the appearance of clustered DNA lesions or multiply damaged sites (MDS), loosely defined as any combination of strand breaks, oxidized bases, and abasic sites within a helical turn.[82] Even low doses of ionizing radiation can produce clustered damages in cellular DNA, and remarkably, only 20% of them are double-strand breaks.[83] The biological effects of double-strand breaks have been widely studied, but other MDS have received attention only recently. Although the same enzyme system deals with singly damaged sites and MDS, lesion clustering generally affects protein recognition and activity.[84−87] It has been established that human ApeI and exonuclease III can readily generate double-strand breaks when two AP sites are in a 3′-direction to each other, indicated as the (+) direction, but their activity is significantly reduced when the lesions are in the opposite (−) orientation.

To identify the molecular mechanism of endonuclease inhibition, our laboratory performed the structural characterization of bistrand AP site duplexes **18a** (+1) and **18b** (−1), which differ only in the relative orientation of THF lesions within the cluster (Chart 3).[88]

The NMR spectra showed that both MDS duplexes adopt very similar right-handed helical conformations with WC alignments across all canonical base pairs. At the MDS site, duplex **18a** displays sequential NOE connectivities between base and sugar protons of residues flanking the abasic sites, suggesting that THF residues are outside the helix and lesion-flanking bases are close in space. In contrast, the absence of

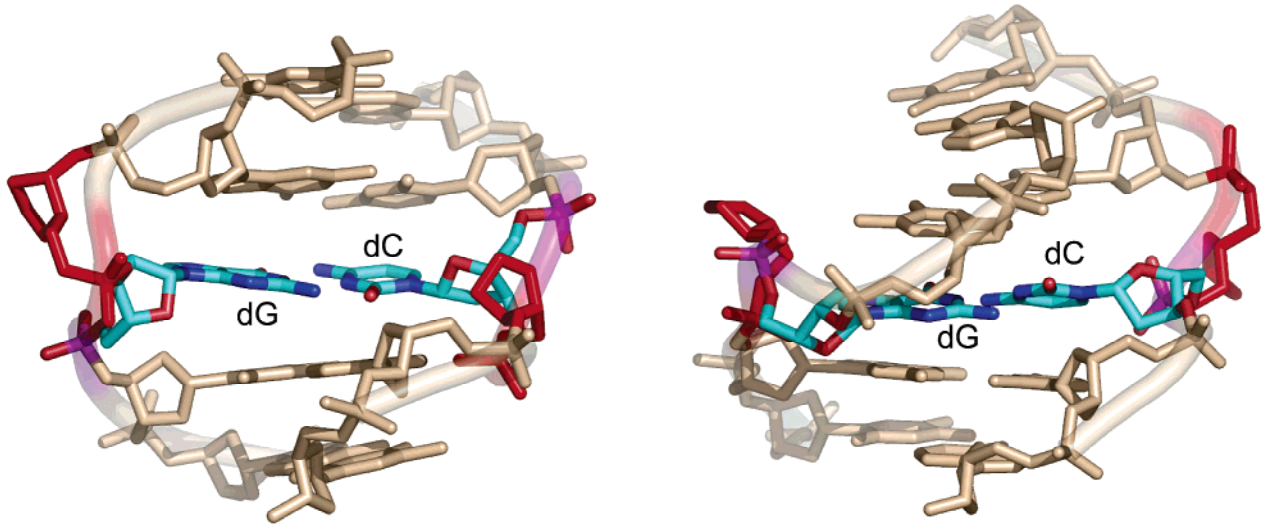


Figure 5. Structure of DNA containing clustered AP sites.[89] The figure shows the central segment of duplexes **18c** and **18d** with the extrahelical tetrahydrofuran moieties in red and the orphan dG and dC residues forming regular WC base pairs. Prepared using PyMol.[555]

Chart 3. Duplex Sequences Containing Clustered AP Sites

C G C A T G **F** G T A C G C
G C G T A C A **F** A T G C G
18a

C G C A T G **F** G T A C G C
G C G T A **F** A C A T G C G
18b

C G C A T G **F** G T A C G C
G C G T A C C **F** A T G C G
18c
F = 13

C G C A T G **F** G T A C G C
G C G T A **F** C C A T G C G
18d

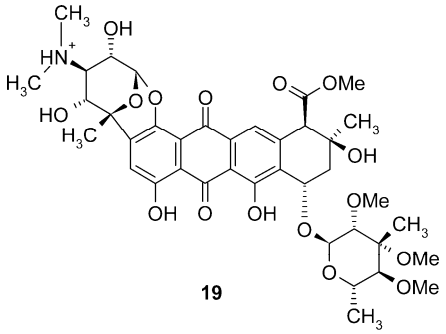
such NOE interactions in duplex **18b** indicates that both AP sites are intrahelical, acting as spacers. In agreement with the experimental distances, the RMD structure of duplex **18a** (PDB code 1HT7) shows the extrahelical THF residues, positioned at the edged of a widened major groove with the orphan bases forming a propeller-twisted alignment that lacks hydrogen bonds. In the case of duplex **18b** (PDB code 1HT4), the RMD structure is less perturbed with both AP sites aligning along the phosphodiester backbone and the orphan bases forming an almost coplanar dG·dA mismatch that also lacks hydrogen bonds. These structures lead to the proposal that the (−1) MDS duplex **18b**, with the AP residues less exposed to solvent and aligned along a more regular DNA backbone, would impose a bigger challenge for endonuclease recognition than the more distorted (+1) MDS duplex **18a**, with extrahelical abasic site residues.

To further understand the role played by the interaction between neighboring orphan bases on the conformation of bistrand AP lesions, our group also solved the structure of related duplexes **18c**,**d** having the orphan dA replaced by dC.[89] The NMR spectra indicated that the THF residues adopt extrahelical conformations in both MDS duplexes, with the subsequent formation by the orphan dG and dC residues of well-shaped WC base pairs. The RMD structure of **18c** shows the AP site residues at the edge of the major groove of the duplex while, in the duplex **18d**, they appear predominantly aligned along the DNA backbone with one of them at the edge of the minor groove (Figure 5). Importantly, the analysis of human ApeI incision reactions, using longer MDS duplexes with the NMR sequence at their center, showed that it cuts the (+1) abasic sites more efficiently,[89] suggesting that extrahelical AP sites displaced into the major groove of the helix are, as a general rule, easy targets for endonuclease.

2.4. DNA Nicks and AP Sites as Scavengers for Polycyclic Molecules

DNA molecules possess a high propensity to trap hydrophobic planar polycyclic moieties by intercalating them between adjacent base pairs, resulting in an increase in helix rise and a considerable distortion of the sugar−phosphate backbone. Opening of a hydrophobic binding pocket in duplex DNA is considerably facilitated by some structural motifs with more flexible backbones, such as DNA with bulged nucleotides that demonstrate high affinity toward intercalators.[90,91] To evaluate whether DNA containing a strand break also has increased affinity toward duplex intercalators, an NMR study has characterized the interaction of the antibiotic nogalamycin (**19**, Chart 4) with nick-

Chart 4. Nogalamycin



19

containing duplexes.[92] The NMR titration of two dumbbell (double-hairpin) duplexes having a strand break at the center of the stem region, either on a d(TpG) step **20a** or its complementary d(ApC) step **20b**, shows that nogalamycin forms a single 1:1 complex in both cases (Chart 5). Complex

Chart 5. Nicked Dumbbell Duplexes Used for the Nogalamycin Binding Study

20a **20b**

formation has a significant stabilizing effect, increasing the T_m of both nick-containing hairpins by approximately 5 °C.

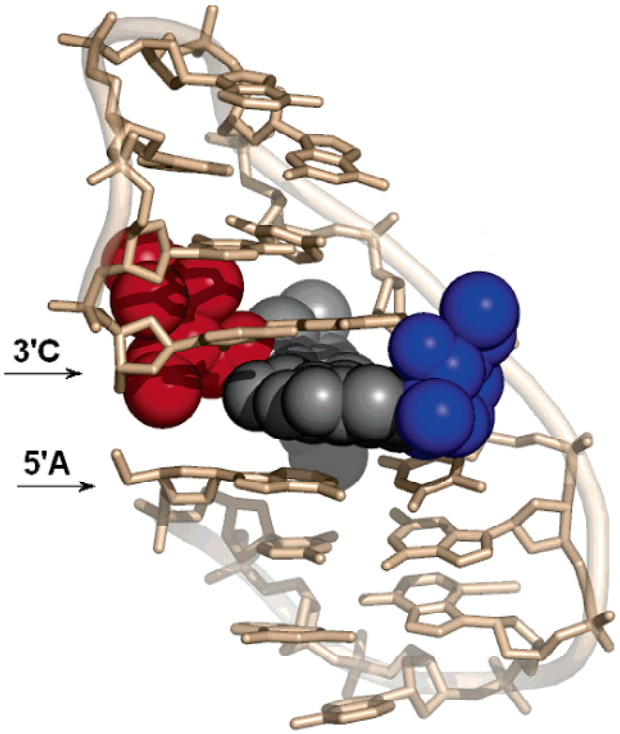


Figure 6. Structure of the complex formed between the nicked dumbbell duplex **20b** and nogalamycin **19**. The picture displays the nogalose moiety in red and the aminoglucose in blue. Prepared from PDB entry 1L0R[92] using PyMol.[555]

Analysis of NOE contacts between nogalamycin and DNA protons indicates that the antibiotic intercalates at the nick site in both samples, with the nogalose and bicyclo aminoglucose moieties positioned in the minor and major grooves of the helix, respectively. The RMD structures show that in both duplexes nogalamycin adopts a similar orientation relative to the dG-T strand (Figure 6, PDB code 1L0R), which is almost identical to that found in an intact duplex having a dGpT intercalation site.[93,94]

The increased affinity of nicked DNA for intercalators may have more important consequences than just the formation of tight noncovalent complexes. Radiomimetic antibiotics of the bleomycin family can bind duplex DNA and activate molecular oxygen, inducing the abstraction of the 4′-proton from a sugar moiety on one strand. The subsequent nucleotide fragmentation gives rise to a phosphoglycolated gap **4** along with 10−20% double-strand breaks, which account for the antibiotic cytotoxicity. The high level of double-strand breaks produced by bleomycin could hardly be explained by two statistically independent strand incision events on DNA. It was proposed that after the first cleavage event bleomycin does not leave the duplex molecule but instead rearranges and forms a stable complex with the gap-containing DNA, allowing for cleavage of the second strand after reactivation of the drug.[48]

Bleomycins need an iron ion as cofactor for the oxygen activation. Since Fe^{2+}-bleomycins are highly active and cleave DNA very rapidly, they have limited application for structural studies. Co^{2+}-bleomycins, that fail to activate molecular oxygen, have been used instead for studying the conformation of bleomycin−DNA complexes. The NMR titration of Co^{2+}-bleomycin B2 (**21**, Chart 6) with the DNA dumbbell **11b**, a mimic of the first oxidation product, demonstrates formation of a stable 1:1 complex. The solution structure of this complex (PDB codes 1G5L and 1GJ2) displays the bis-thiazole moiety of bleomycin intercalated into the gap, in parallel intercalator mode, with the guanidinium fragment forming an ion pair with the phosphoglycolate moiety at the 3′-side of the gap. The metal binding center of **21** shows sequence specific interactions with minor groove atoms of the duplex and, therefore, is located close to the intact strand near a segment opposite to the gap. As a result, the metal Co^{2+} ion is located near the H4′ of T19, in the right position for starting the second incision reaction[95] (Figure 7).

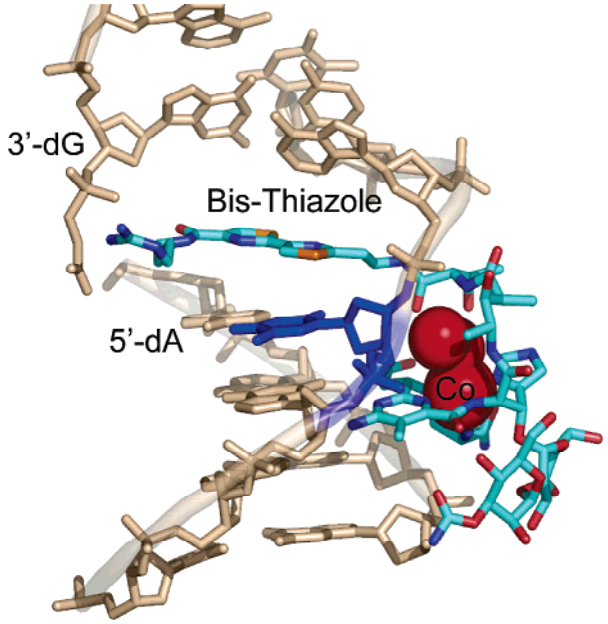
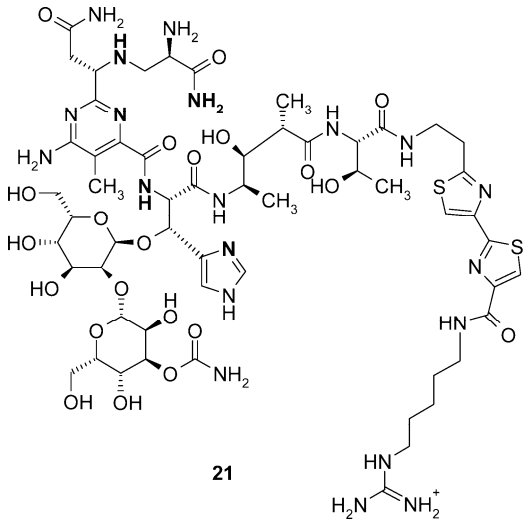


Figure 7. Structure of the complex formed between Co^{2+}-bleomycin B2 **21** and phosphoglycolated duplex **11b**. The Co^{2+} ion and a bound H_2O_2 molecule are shown in red. Prepared from PDB entry 1G5L[95] using PyMol.[555]

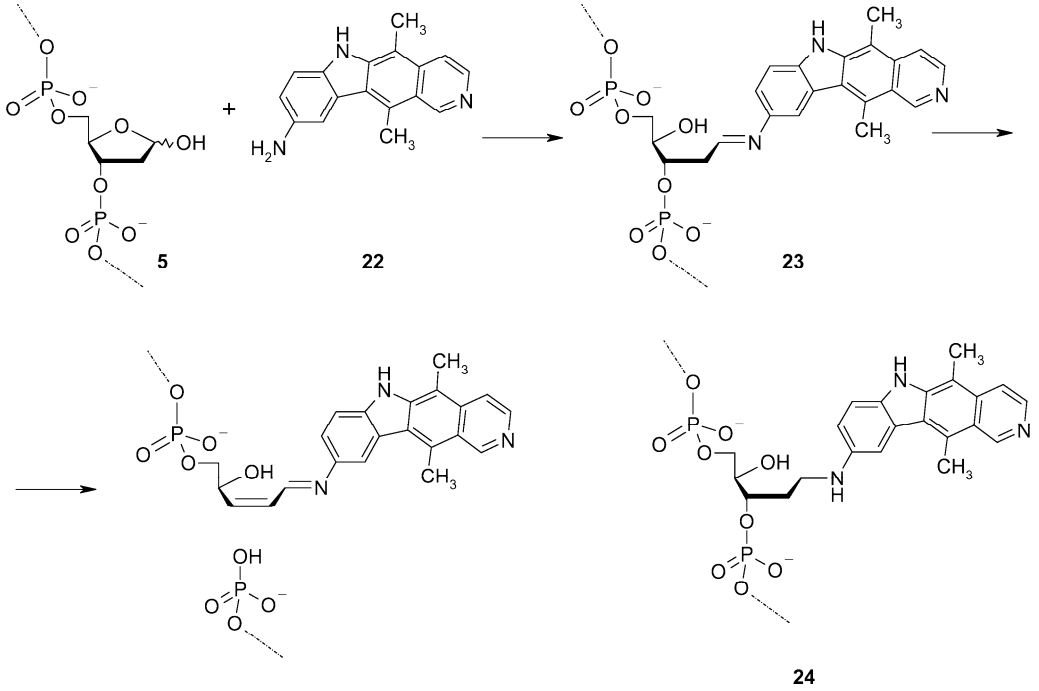
In contrast to nicked or intact DNA that needs to adjust to the presence of an intercalator, AP sites already have a preformed hydrophobic pocket for binding polycyclic compounds. 9-Aminoellipticine **22**, a member of the ellipticine family, binds to apurinic but not apyrimidinic sites in DNA, forming Schiff base conjugates with deoxyribose **23**, which can then cause strand breaks via β-elimination of the 3′-phosphate[96] (Scheme 3). To establish the structure of the

Chart 6. Structure of Bleomycin B2[a]



[a] Atoms involved in metal ion binding are shown in bold face.

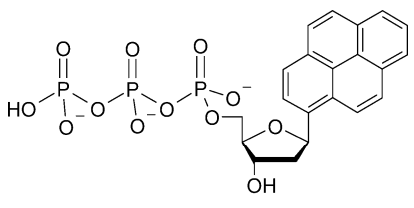
Scheme 3. Adduct Formed between 9-Aminoellipticine and Natural AP Sites



reaction intermediate, an NMR study has characterized the reduced Schiff intermediate **24** formed between **22** and an AP-site-containing duplex with an orphan thymine base.[51] In contrast to the AP-site-containing control duplex that exists in two interconverting states, with the deoxyribose moiety either intrahelical or extrahelical, the conjugate **24** adopts a single conformation in solution. Analysis of NOESY spectra indicates that the 9-aminoellipticine moiety intercalates between lesion-flanking guanine bases, with its indole ring occupying the gap present at the AP site. The pyridine part of **24** faces the complementary strand, wedging between the orphan thymine and its 5′-flanking cytosine residues and causing a partial displacement of the orphan residue into the major groove. The tight packing of **24** inside the helix is consistent with the higher thermal stability of the aminoellipticine-containing duplex, as compared to the parent AP-site-containing control DNA, giving some rationale for the inhibitory effect of **22** on AP endonuclease activity.

The early recognition that DNA abasic sites may cause reading errors of polymerase-mediated sequencing reactions, led to the discovery that a nucleoside 5′-triphosphate with a pyrene aglycon **25** (Chart 7) is preferentially incorporated

Chart 7. Pyrene 2′-Deoxynucleoside Triphosphate



opposite to AP sites by DNA polymerases.[97,98] To characterize the conformation of Pyr·AP pairs, which have no possibility for hydrogen bond formation, our group established the structure of two double-stranded samples having a Pyr·THF pair (PDB codes 1FZL and 1FZS), one at the interior of a 13-mer duplex and another at a primer−template junction.[99] The NMR spectra indicated that both molecules adopt right-handed conformations with WC alignments for all canonical base pairs. Both residues of the Pyr·THF pair are inside the helix, with the aromatic moiety intercalating between neighboring purine bases in the duplex or stacking over the 5′-flanking base pair in the template−primer. In the latter case, residues of the template single-strand region fold back over the Pyr·THF pair, protecting the hydrophobic ring from solvent exposure. Interestingly, the NOESY spectra of both samples showed exchange cross-peaks for the dA·T pair at the 5′-side of the Pyr residue, suggesting that the presence of the highly hydrophobic moiety partially alters the dynamics of local stacking or the pairing of more polar nucleobases.

3. Oxidative Lesions

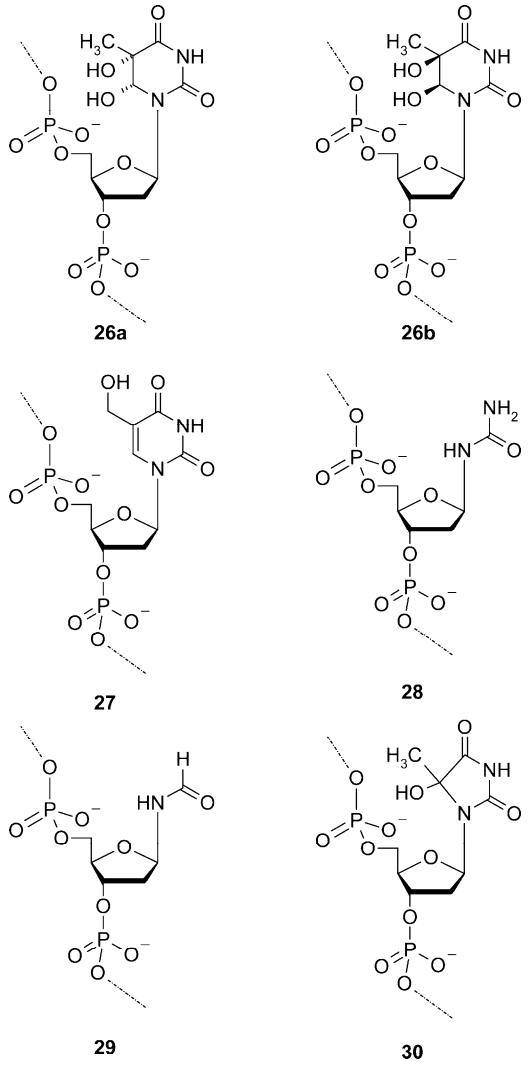
The process of mitochondria respiration or the interaction of ionizing radiation with water may produce reactive oxygen species, such as hydroxyl radicals or singlet oxygen, which react with DNA bases generating oxidative damage. Nitric oxide, a mediator of many cell processes, forms peroxynitrite, which has similar reactivity to that of hydroxyl radicals toward nucleobase oxidation.[100] DNA glycosylases continuously remove oxidized bases from nuclear DNA, reducing the burden of oxidative lesions in the cell, and while the exact number of lesions is very difficult to determine, it reaches a steady-state level of about 0.3−4.2 oxoguanine residues per 10^6 dG nucleotides in normal human cells.[101] The purine C8 position and the C5−C6 double bond of pyrimidine bases are the most frequent sites of damage.[102] We review below NMR structural studies of duplexes containing 8- or 2-oxopurine, oxopyrimidine, thymine glycol, base fragmentation products, or oxidized methyl thymine lesions in DNA.

3.1. Pyrimidine Oxidation Damages

Under aerobic conditions, hydroxyl radical attack of thymine initiates a series of conversions yielding damaged sites with five different aglycones,[103] namely isomeric thy-

mine glycols **26a**,**b**, hydroxymethyluracil **27**, urea **28**, formamide **29**, and 5- and 5-hydroxy-5-methylhydantoine **30** (Chart 8). With the exception of the last lesion, all the others

Chart 8. Thymine Oxidation Products



have been studied by NMR spectroscopy.

3.1.1. Thymine Glycol

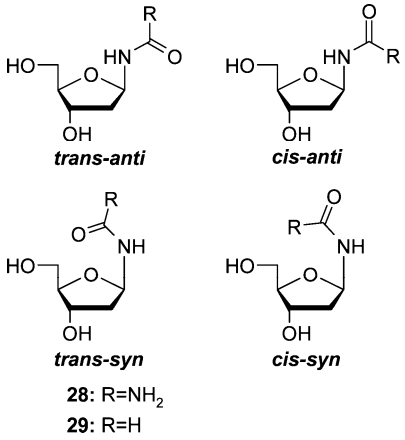
The appearance of a glycol at the thymine 5,6 position causes the loss of base aromaticity, affecting its planarity and hydrogen bonding ability and increasing the bulkiness and hydrophilicity of the base. Thus, it should come as no surprise that thymine glycol (Tg) generally blocks DNA replication[104−106] and induces base substitution mutations.[107] To evaluate the structural impact of Tg in DNA, NMR studies have characterized undecamer duplexes containing a single thymine glycol lesion in a 5′-d(...G Tg C...)[108] or 5′-d(...A Tg A...)[109] sequence context. In these studies the Tg lesion was generated postsynthetically, via permanganate oxidation of single-stranded oligonucleotides, which produces mainly *cis* glycols. Thus, only the 5*R*,6*S* Tg **26a**, the predominant form, and, to a lesser extent, the 5*S*,6*R* Tg **26b** diastereomer, are present in the damaged duplex. Furthermore, *trans* isomers are less stable and gradually epimerize into the corresponding *cis* isomers.[110] Therefore, *trans* Tg isomers are also absent on the samples.

The ^{31}P, imino proton, and NOESY spectra showed that a Tg lesion causes moderate disruption of the right-handed helical structure, which remains a B-form helix with normal WC alignments for all unmodified base pairs. At the lesion site, Tg produces significant local perturbations, which manifest by the interruption of sequential NOE connectivities on both strands of the duplex and by the loss of hydrogen bonding across the lesion-containing base pair. The RMD structure obtained for one of these samples[109] shows Tg displaced toward the major groove of the duplex, increasing significantly its exposure to the solvent. The authors concluded that the extent of structure perturbations caused by a single Tg in duplex DNA is comparable to that produced by a single AP site.

3.1.2. Pyrimidine Fragmentation Derivatives

Urea (Ur) and formamide (Frm) nucleotide moieties in DNA are fragmentation products of lesions initially formed by reaction of active oxygen species with thymine and, to a lesser extent, cytosine.[103] In addition, it has been reported recently that further oxidation of 8-oxopurines also yields some Ur moieties in DNA.[111] Like any amide, these residues may exist in two virtually not-interconverting conformations, *cis* or *trans*, with a 2:3 ratio.[112,113] In addition, like any nucleotide, they can have a *syn* or *anti* orientation around the glycosidic bond, with each orientation being nearly equally probable (Chart 9).

Chart 9. Conformation of Urea and Formamide Glycosides

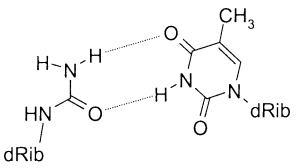


The presence of these lesions in cellular DNA is likely to have important biological consequences. Although the Ur and Frm moieties retain some hydrogen bonding potential, *in vitro* studies show that both types of damage significantly impede DNA replication. An Ur residue present in the template strand blocks DNA polymerase activity one nucleotide before the lesion or, under conditions of decreased specificity, right at the damage site.[105] The Klenow fragment of DNA Pol I can bypass a Frm lesion by incorporating dGMP, whereas Taq DNA polymerase introduces one-base deletion mutations.[114]

To explain these biological effects, several studies have characterized the conformations of duplexes having pyrimidine fragmentation products. The NMR spectra of a decamer duplex containing an Ur opposite thymine showed that, while the overall duplex structure retains the major features of B-form DNA, large changes are observed at the damage site, which were strongly dependent on the conformation of the Ur moiety.[113] The exchangeable proton signals of the Ur residue, as well as its partner and 3′-flanking nucleotides, were significantly broad, precluding observation of NOESY cross-peaks at the lesion site. Analysis of one-dimensional

NOE difference and nonexchangeable proton NOESY spectra indicates that Ur and its partner thymine residue adoptintrahelical conformations. Simultaneously, the presence on the same NOESY spectrum of cross-peaks between lesion-flanking base pairs establishes their spatial proximity, suggesting that both lesion site residues are extruded, with the subsequent collapse of the residual gap. On the basis of these contradictory observations, the authors concluded that the lesion site exists in two conformational states that depend on the isomeric form of the Ur moiety. The *trans*−*anti* isomer would be present in the intrahelical conformation where it forms the hydrogen bonded pair **31**, with the complementary thymine residue, outlined in Chart 10, whereas the *cis*

Chart 10. Proposed Ur·T Pair Alignment



31

Ur isomers are unable to pair with thymine, leading to the extrusion of both residues from the helix.

In contrast to urea derivatives, Frm-containing DNA exhibits well-resolved NMR spectra, allowing the complete structural characterization of duplexes with *cis* or *trans* lesion isomers. Analysis of the NMR data indicates a small dependence of duplex structure on lesion isomerism, when a Frm residue is opposite to dA[115] or dG.[116,117] In the Frm·dA duplex, both lesion isomers can adopt intrahelical or extrahelical conformations, whereas its partner adenine is always inside the helix. In the intrahelical conformation, the *cis*-Frm residue is *syn* around the χ angle, forming two hydrogen bonds with adenine, while the *trans* isomer remains *anti*, forming only one. Overall, the behavior of Frm lesions and their effects on duplex conformation closely resemble those of apyrimidinic sites.

The preferred incorporation of dGMP opposite the Frm lesion **29** during DNA replication suggests the existence of a specific interaction between these two residues, and the NMR characterization of the Frm·dG duplex demonstrates a well-ordered structure in solution. All the NOE connectivities that characterize B-form DNA are present through both strands of the duplex containing the *trans*-Frm isomer. The formyl proton of the Frm moiety displays normal inter- and intranucleotide NOE cross-peaks, demonstrating that it adopts an intrahelical duplex orientation. In the *cis*-Frm case, its formyl proton lacks of NOE connectivity to the H1′ of the 3′-flanking nucleotide, but other cross-peaks present on the spectra demonstrate that the damaged residue is inside the helix. The RMD structures of both duplexes show the Frm carbonyl oxygen hydrogen bonded to the imino proton of the partner guanine base, a fact that is proposed to account for stabilization of the intrahelical conformation and the miscoding properties of the lesion. It is worth noting that analysis of ROESY spectra reveals the existence of conformational exchange in the case of the *trans*-Frm isomer. The structure of the minor *trans*-Frm species has not been characterized in detail, but it most likely contains extrahelical residues.[116]

The significant number of one-base deletion mutations observed during Taq polymerase replication of Frm-containing DNA[114] has urged the NMR characterization of duplexes containing frame-shift mutation intermediates. Analysis of the NOESY spectra shows small distortions of a canonical B-form helix with lesion-flanking base pairs stacking on top of each other, indicating that the Frm residue has been extruded from the duplex. The RMD data indicate that a large number of conformations are available for the extruded Frm residue. The refined solution structure (PDB no. 1OSR) demonstrates a close analogy between duplexes containing an extrahelical formamide or an AP site residue.[118]

3.1.3. 5-Hydroxymethyluracil

The 5-methyl group of thymine may be the subject of oxidation, leading to the formation of 5-hydroxymethyl-2′-deoxyuridine **27** (dU^{HM}) in DNA.[119,120] Generally speaking, this change on the methyl group has a little effect on base pair properties,[121] a fact that has been used for the enzymatic incorporation of diverse nonradioactive labels into DNA.[122,123] In addition, dU^{HM} completely replaces thymidine in the SPO1 bacteriophage genome, demonstrating its ability to serve as a carrier of hereditary information,[124] and has a very low mutation frequency in murine cells.[125] However, the existence of 5-hydroxymethyluracil−DNA glycosylase in mammalian cells, which is responsible for the excision of dU^{HM} from single- and double-stranded DNA,[126] would suggest that its presence may have deleterious effects.

NMR spectroscopy has examined the alignment of dU^{HM}·dA or dU^{HM}·dG base pairs located at the center of heptamer duplexes.[127] The NMR spectra indicate the presence of canonical B-form duplex conformations with formation of a regular dU^{HM}·dA WC alignment, or a wobble dU^{HM}·dG pair that is essentially identical to that of the dG·T mismatch. Analysis of NOESY spectra shows that the hydroxymethyl (HM) group of dU^{HM} does not rotate freely but points away from dU^{HM} H^6, adopting a conformation that depends on the identity of the opposite purine. In the dU^{HM}·dA pair the HM hydroxyl is located on the 5′-side of the base plane, whereas it is positioned on 3′-side in the dU^{HM}·dG-pair-containing duplex. Molecular modeling of the dU^{HM}·dA duplex predicts the formation of a hydrogen bond between the HM hydroxyl and N^7 of the 5′-flanking guanine or O^4 of the dU^{HM} residue. In contrast, modeling of the dU^{HM}·dG duplex predicts the formation of only the intranucleotide hydrogen bond. While the fast solvent exchange rate precluded observation of these hydrogen bonds, the authors proposed that they might explain the different behaviors of dG·T and dU^{HM}·dG base pairs at elevated pH. In contrast to the imino protons of dG·T, dG·dU, or dU^{HM}·dA pairs that do not dissociate until the alkaline DNA denaturation starts, proton chemical shifts suggest that the pK_a value of the dU^{HM} imino group on the dU^{HM}·dG pair is as low as 9.7, allowing the transition to a WC-type pair upon ionization. Importantly, as indicated by the authors, the presence of small amounts of the ionized dU^{HM}·dG pair at physiological pH can account for the slight mutagenic effect of dU^{HM}·dG.[127]

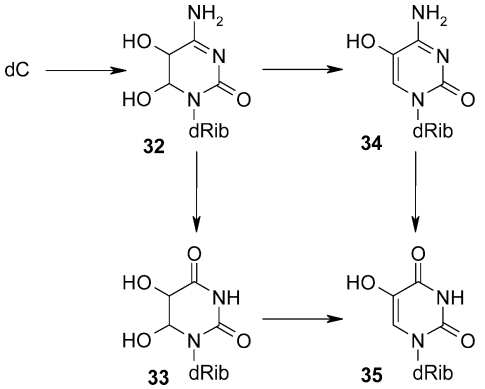
Formation of internucleotide hydrogen bonds between the dU^{HM} hydroxyl and N^7 of the 5′-flanking adenine residue has also been reported on duplexes containing a $d(A{\cdot}U^{HM})_2$ step.[128,129] The RMD structures show that simultaneous formation of these hydrogen bonds on both strands of the duplex results in an increase of base twist and roll angles, reducing stacking between lesion site bases and increasing

local DNA flexibility. The authors concluded that these conformational features may facilitate the recognition of dU^{HM}-containing sequences by transcription factors, such as the case of TF-1 of SPO1 phage.[130,131]

3.1.4. 5-Hydroxyuracil

This lesion appears generally as a consequence of oxidation of the cytosine 5,6 double bond to 5,6-cytosine glycol **32** that is unstable and undergoes hydrolytic deamination or dehydration, leading to the formation of 5,6-dihydroxyuracil **33** and 5-hydroxycytosine **34** lesions, correspondingly (Scheme 4). Levels of these lesions present in rat tissues were found

Scheme 4. Conversion of Cytosine Oxidation Products



to be comparable to that of 8-oxodG,[132] one of the most abundant oxidative lesions. Since both **33** and **34** are able to slowly convert to the hydroxyuracil (dU^{OH}) derivative **35**, considerable levels of the latter lesion are normally detected in tissues. dU^{OH} lesions are highly mutagenic, causing, predominantly, dG·dC → dA·T transitions.[133] However, the possibility of base pair formation with nucleotides other than dA has been predicted by *ab initio* 6-31G** computations. Proton and phosphorus NMR spectroscopy have characterized four self-complementary duplexes containing a single dU^{OH} lesion paired to dA, dC, dG, or T. The limited set of spectroscopic data obtained for these samples suggest that all duplexes adopt a nondistorted B_I type conformation. The shape and chemical shift of the imino proton signals indicate that the damaged base is hydrogen bonded with dA, dG, or T partners and forms a less stable pair when it is opposite to dC.[134] No additional information on dU^{OH} lesions is available to date.

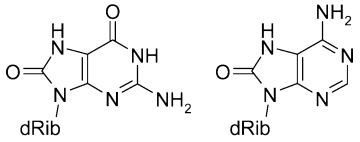
3.2. Purine Oxidation Products

Oxidative attack to purine bases causes hydroxylation at the C8 or C2 position without affecting base aromaticity. Changes in hydrogen-bonding potential and orientation of the base dipole are the main consequences of purine oxidation and only alter stacking and base pair alignment at the lesion site. Since none of the NMR studies reviewed below has observed a detectable distortion on the global shape of DNA or base pair planarity, we will limit our discussion to the lesion site structure and its alignment.

3.2.1. 8-Oxopurines

The C8 position of purine nucleotides is the most susceptible to free radical oxidation. Singlet oxygen forms a tricyclic intermediate that gives rise to 8-hydroxypurine **36** (8-oxodG) along with some other minor byproducts (Chart

Chart 11. 8-Oxopurine Lesions



11).[135] In contrast to thymine glycol and pyrimidine fragmentation products, 8-oxodG does not block polymerase activity but is highly mutagenic, incorporating dC and significant amounts of dA[136] at a ratio that is strongly dependent on the type of polymerase involved in DNA replication.[137] The product of adenine C^8 oxidation, 8-oxo-7,8-dihydroadenine **37** (8-oxodA), is less abundant and at least 1 order of magnitude less mutagenic than 8-oxodG.[138,139] However, a subsequent study reported that eukaryotic DNA polymerases α and β could incorporate dG and dG or dA, respectively, opposite 8-oxodA,[140] suggesting the existence of hydrogen-bonded alignments that stabilize the mutagenic base pair intermediates.

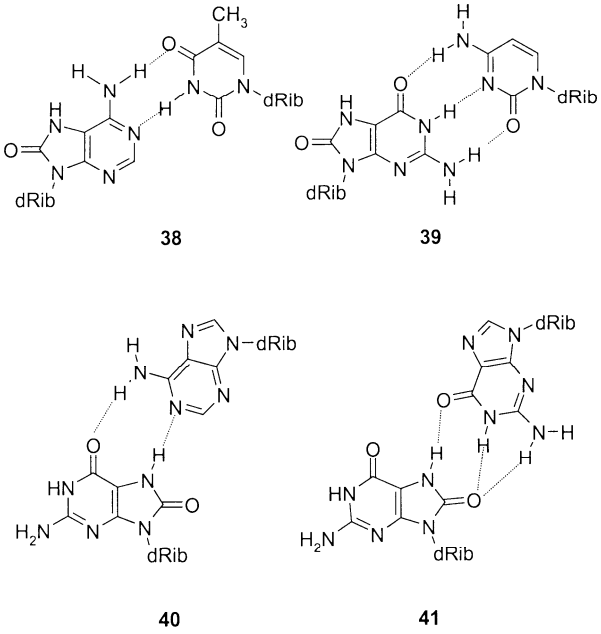
Compared to unmodified purine nucleotides, the 8-oxo lesions have two distinctive features, increased conformational variability and altered hydrogen-bonding properties. The presence of a bulkier and more electronegative 8-oxo group destabilizes the *anti* conformation, making the *syn* orientation almost equally probable. Thus, 8-oxopurine nucleotides can present their Watson−Crick or Hoogsten edge to the partner base, increasing the number of possible base pairs that it can form in DNA,[141] and notably, more than one alignment is theoretically possible with a given base. Furthermore, although ^{15}N NMR spectroscopy showed that only the amido tautomer of the lesion is found on free nucleosides,[142] the duplex environment could potentially shift this equilibrium in cellular DNA.

Solution NMR studies have studied the alignment of 8-oxopurine-containing base pairs on duplexes containing dC,[143] dA,[144] or dG[145] (PDB code 1N2W) opposite 8-oxodG, as well as of the duplex with an 8-oxodA·T pair[138,146] (PDB code 1FJB). Since the H8 proton is absent in 8-oxopurines, the nucleotide conformation around the glycosidic bond cannot be assessed as usual, by evaluating the intensity of the intraresidue NOE cross-peak between the H8 and H1′ protons. As a result of the studies discussed below, the conformation of 8-oxopurine around the χ angle was inferred mainly from NOE interactions observed for the exchangeable protons.

Observation of numerous NOE interactions between 8-oxodA-H2 and imino, as well as amino, protons of residues located at and flanking the lesion site establish that the 8-oxodA forms 8-oxodA·T WC alignments **38** (Chart 12).[138,146] The *anti* conformation of 8-oxodA places the H^7 proton on the major groove of the duplex, where it resonates around 10 ppm and shows NOE interactions with other major groove protons of the sequence. Similarly, the 8-oxodG lesion is in an *anti* conformation forming regular WC alignments **39** in a self-complementary dodecamer containing two 8-oxodG·dC base pairs[143] (Chart 12). The H^7 proton of 8-oxodG also resonates at about 10 ppm but, in this case, shows no NOE interactions, a fact consistent with its major groove localization.

When 8-oxodG is opposite to dA, it adopts a *syn* glycosidic torsion angle and forms a mutagenic Hoogsten type alignment **40** with the mismatched partner dA.[144] In the *syn*

Chart 12. Base Pair Alignments of 8-oxodA·T 38 and 8-oxodG Opposite dC, dA, and dG 39−41



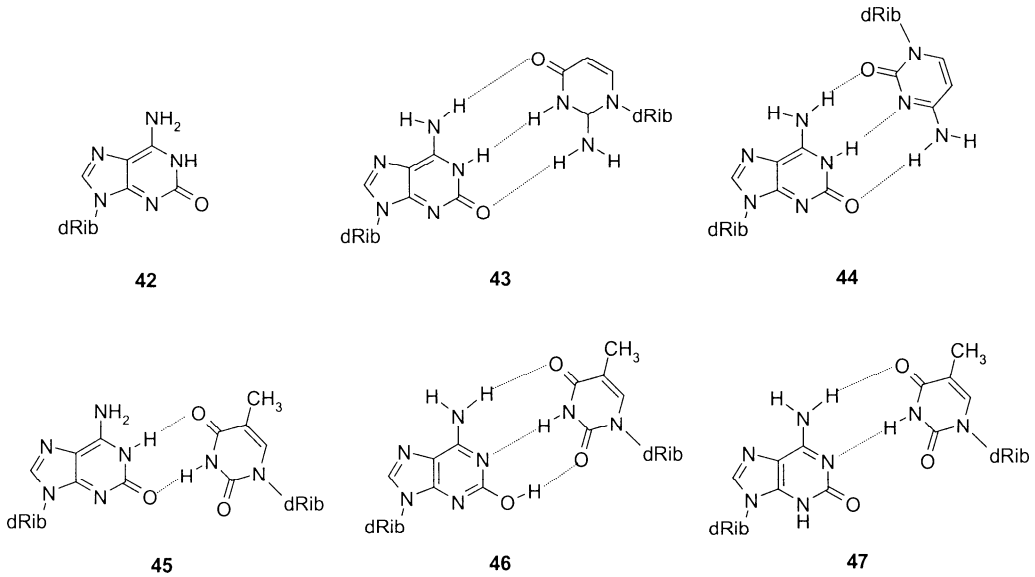
conformation, the H1 imino and $2NH_2$ amino protons of 8-oxodG are on the major groove and experience a significant upfield shift, indicating that they are not hydrogen bonded. Conversely, the H7 imino proton shifts downfield by approximately 2.6 ppm and displays numerous NOE contacts with imino and amino protons of adjacent base pairs, indicating that it is hydrogen bonded to the lesion partner dA residue. A distance-restrained energy-minimized model shows that the 8-oxodG(*syn*)·dA Hoogsten pair causes no significant distortions to the duplex conformation. Although 8-oxodG was found predominately in the 6,8-dioxo form, the presence of additional H5−H6 NOE cross-peaks for all cytosine residues of the sample leads the authors to conclude that about 15% of the lesion is present in a minor tautomeric form. Interestingly, no additional signals were observed for protons belonging to lesion site residues or flanking base pairs. Unfortunately, the small population of the alternative form precluded its characterization further.[144]

The NMR characterization of a self-complementary duplex containing two 8-oxodG·dG pairs showed essentially the same base alignment at the lesion site as the 8-oxodG·dA duplex[145] (Chart 12). However, in this case, the RMD structure (PDB code 1N2W) demonstrated that the usual Hoogsten geometry required a conformational adjustment in order to form a hydrogen-bonded 8-oxodG·dG alignment **41**. 8-OxodG moves slightly toward the major groove, and its partner guanine does in the opposite direction, which brings the imidazolone ring of 8-oxodG close to the WC edge of dG, allowing the formation of one normal hydrogen bond and one bifurcated three-center hydrogen bond. This adjustment at the lesion site affects the stacking pattern of 8-oxodG, which is now stacked preferentially with its 3′-flanking neighbor. Nevertheless, the T_m of the 8-oxodG·dG duplex is higher than that of the 8-oxodG·dA one, indicating that the latter pair causes a larger reduction of duplex thermal stability. At first sight, this observation seems to contradict the mutagenic specificity of DNA polymerases, which do not incorporate dG opposite the lesion. Therefore, keeping in mind that replicative DNA polymerases use both hydrogen bonding and geometry matching as the basis for incoming dNTP recognition,[147,148] it must be concluded that the perturbed Hoogsten geometry **41** of the 8-oxodG·dG pair is unacceptable for these polymerases.

3.2.2. 2-Oxoadenine (Isoguanine)

Isoguanine **42** (isodG, Chart 13) is a rare adenine oxidation product that has severely reduced hydrogen-bonding capabilities with normal nucleobases in DNA.[102] Since the WC edges of isoguanine and guanine are mirror images, the former cannot form WC type base pairs with dC or T. IsodG can adopt a stable hydrogen-bonded alignment **43** only when it is paired to the non-natural isocytidine base or after mirror image problems are resolved. For example, a regular WC type isodG·dC base pair **44**, or the analogous isodC·dG, readily occurs in parallel DNA duplexes[149] (PDB code 1BE5), a situation that is not known to occur *in vivo*. Usually, oxidation of the adenine C2 produces only a single isodG moiety opposite T in the context of an antiparallel double-stranded DNA. In principle, an isodG·T pair can form a wobble alignment **45** that, in comparison with the dG·T mismatch, should exhibit reversed base displacements, with

Chart 13. Isoguanine 42 and Its Alignment with IsodC 43, dC 44, and T 45−47



isodG shifted toward the major groove and T shifted in the opposite direction.

Alternatively, isodG may transform into the 2-enol or N^3-protonated tautomers forming isodG·T alignments **46** and **47**, respectively, with regular WC geometry. The fact that isodG partially exists as the 2-hydroxy isomer in a low polar environment[150] suggests that such an alignment may actually form in duplex DNA.

A combined NMR and X-ray study has characterized the structure of the self-complementary dodecamer, d(CGC[iG]-AATTTGCG), containing two isodG·T base pairs.[151] At low temperature, the presence of six partially resolved signals, accounting for all 13 imino protons of the sequence, indicates that the symmetry of the duplex is retained. In addition, they resonate in the 11−14 ppm chemical shift range, suggesting that all base pairs of the duplex are hydrogen bonded. Importantly, a strong NOE cross-peak is observed between isodG H1 and T H3 protons, indicating that a wobble isodG·T alignment is present at 2 °C. Raising the sample temperature causes the appearance of a second form that amounts to 40% of the duplex population at 40 °C. Since the imino proton signals of internal base pairs did not change at high temperature, the authors ruled out the putative formation of a hairpin and concluded the second form of the duplex has the enol tautomer of isodG forming a WC alignment. Sustaining this hypothesis is the observation that the X-ray structure of an identical duplex crystallized in a complex with Hoechst 33342 shows both the wobble **45** and WC isodG·T **46** (or **47)** base pairs present on the same duplex molecule.[151] In addition, the observation that AMV reverse transcriptase incorporates T opposite isodG during *in vitro* reverse transcription of DNA further supports the existence of the minor enol tautomer of isodG.[152]

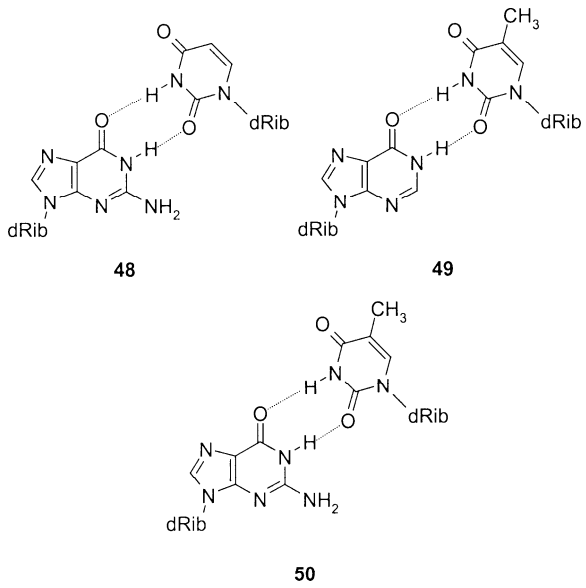
4. Nucleotide Deamination Damage

The exocyclic amino groups of nucleobases may undergo spontaneous deamination, which in the case of cytosine leads to the emergence of uracil residues in DNA.[153] Similarly, conversions of 2′-deoxyadenosine to inosine (I) and of 2′-deoxy-5-methylcytidine to thymine also occur, albeit with less frequency. Left unrepaired, the dG·dU **48**, T·dI **49**, and dG·T **50** mismatches formed in this process can cause dG·dC → dA·T, T·dA → dC·dG, and dG·dC^{Me} → dA·T transition mutations, respectively, with profound implications for genome integrity (Chart 14).

To evaluate the structural consequences of deamination products, an early NMR study has three heptamer duplexes containing a central dG·dU, T·dI, or dG·T mismatch.[154] The NMR data showed that these duplexes are regular right-handed helices, with deoxyribose residues in a C2′-endo conformation and canonical WC base pair alignments outside the mismatch sites. NOESY spectra of the exchangeable protons displayed strong cross-peaks between the imino protons of mismatched bases, establishing the presence of a wobble pair in all three cases. Although three-dimensional models of the wobble alignments were not reported, the available data indicate they are similar to that described earlier for dG·T mismatches,[155−157] with the pyrimidine and purine bases displaced toward the major and minor grooves of the duplex, respectively.

The most frequent and deleterious deamination event in cellular DNA is the conversion of cytosine to uracil, generating dG·dU mismatches. DNA uracil DNA gycosylase (UDG) is a member of a large superfamily of enzymes that

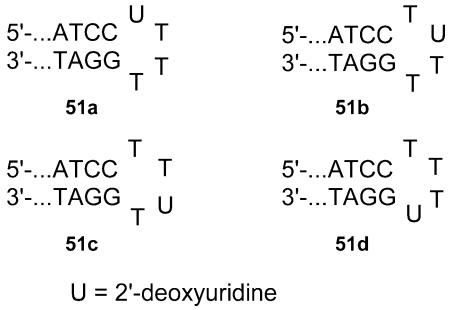
Chart 14. Mismatches Formed as a Result of Base Desamination



can remove uracil bases from single- and double-stranded DNA. UDG uses a mechanism that efficiently excludes thymine bases form its catalytic pocket in a sequence independent manner.[158] However, it has been established that *E. coli* UDG poorly excises uracil when a dU residue is located at any of the first three positions of a tetraloop region. In addition, higher K_m and lower V_{max} values for these substrates indicate that both enzyme affinity and its catalytic rate are affected by changes in conformation caused by the loop formation.[159]

To understand the modulation of UDG activity by local DNA conformation, a series of NMR studies have determined the structure of dU-containing hairpins (Chart 15), observing

Chart 15. dU-Containing Hairpin Duplexes



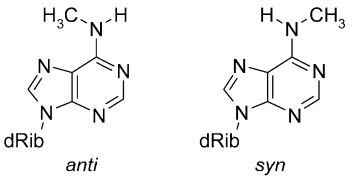
similar stem structures but with marked conformational differences at the loop region.[160−162] The RMD structure of the U4 hairpin (**52d**) (PDB code 1QE7) shows the uracil base stacked between the 3′- and 5′-flanking residues and B_I-DNA values for the sugar−phosphate backbone angles at the loop region.[160] Conversely, in the refined structure of the U2 (**52b**) hairpin (PDB code 1DGO) where UDG uracil incision is retarded the most, dU and its 3′-flanking residue adopt *syn* conformations around the glycosidic torsion angle and the α, γ, and ζ backbone dihedral angles of the loop assume the unusual *trans* conformation.[161] RMD structures of the U1 (**52a**) and U3 (**52c**) (PDB codes 1II1 and 1IDX, respectively), that are incised by UDG somewhat more efficiently than the U2 hairpin, show all loop residues in *anti* coformations and a *trans* conformation for backbone

dihedral angles of the dU residue.[162] On the basis of these observations and of X-ray data available for *E. coli* UDG,[163] the authors proposed that a stretched out backbone conformation of the dU nucleotide reduces the affinity of the scissile residue for UDG.

5. Alkylation Damage

Besides oxidative and nucleophilic attacks, nucleic acid bases can be damaged by carbon electrophiles that preferentially add to ring and exocyclic nitrogens and, to a lesser extend, to base peripheral oxygens, giving rise to alkyl or Shiff base derivatives. In living cells, three major mechanisms lead to the appearance of alkylated bases in DNA. The first mechanism involves the sequence specific enzyme-catalyzed methylation of DNA bases, which occurs in most prokaryotic and eukaryotic organisms as a control for various cellular processes.[164] Although only the *anti* rotamer of N^6-Me-adenine or N^4-Me-cytosine (defined relative to the imino nitrogen position) can form canonical WC alignments, these lesions usually have minor effects on DNA conformation. The NMR characterization of an oligodeoxynucleotide containing the d(G A T C) sequence motif hemimethylated at adenine-N^6 showed that the duplex has all residues in the *anti* conformation around the χ torsion without significant distortions of the B-type DNA conformation.[165] Rotamer interconversion proceeds slowly in the NMR time scale, and both the *anti* **52a** (PDB code 1UAB) and syn **52b** (PDB code 1OQ2) conformers (Chart 16) are embedded in the stem,

Chart 16. *Syn* and *Anti* Isomers of N^6-Methyladenine



52a **52b**

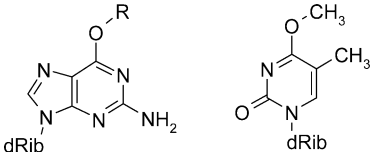
causing a small adjustment of the alignment of the methylated adenine. Similarly, N^6-methylation at a d(TpA) step of a 16-mer duplex does not cause considerable changes of the global DNA conformation but only a subtle variation in DNA dynamics.[166] Since N^6-adenine or N^4-cytosine methylation is a fundamental epigenetic mechanism, these products will not be considered DNA damages, and we will leave the discussion of such modifications beyond the scope of this review.

The second mechanism of DNA alkylation involves the direct attack of exogenous low molecular weight compounds, such as alkyl halogenides, sulfonates, nitroso compounds, and other common pollutants, widely present in chemical waste and tobacco smoke. Alternatively, intracellular oxidation transforms nonreactive unsaturated organic pollutants into highly reactive alkylating agents, which may react with nucleobases causing DNA damage. In the third mechanism, alkyl lesions may appear in the cell as a result of reactions between nucleobases and endogenously generated compounds, such as acrolein and malondialdehyde. Alkylating agents that react with DNA bases can be mono- or bifunctional, and in the latter case, formation of either a cyclic adduct or an interstrand cross-link is possible. All these types of damage, namely, monoalkylated bases, cyclic adducts, and interstrand cross-links, are discussed below.

5.1. Products of Alkylation by Monofunctional Agents

It has been established several decades ago that nitrosamine agents, such as *N*-methyl-*N*-nitrosourea and *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine, are able to both induce cell transformation and suppress tumor growth. The chemical basis for these phenomena is the same, namely, the ability of nitrosamines to alkylate almost all oxygens and endocyclic nitrogens on DNA. The positions most frequently attacked are the N^7 of guanine and backbone phosphate oxygens, which in combination account for about 85% of all alkylation products.[167] However, NMR characterization of a dodecamer duplex containing a single N^7-methylguanine residue shows that the alkyl damage has almost no effect on DNA conformation.[168] Conversely, alkylation of guanine or thymine oxygens has severe structural consequences because the imino nitrogen of O^6-alkyl guanine **53** or O^4-alkyl thymine **54** lesions (Chart 17) changes its character from a hydrogen

Chart 17. *O*-Alkylated Guanosine and Thymidine Residues



53a : R = methyl
53b : R = ethyl

54

bond donor to an acceptor, preventing the formation of canonical WC base pair alignments.[169] As a result, the carcinogenicity of a given alkylating agent does not correlate directly with the overall number of alkylation products but rather with the fraction of these two minor alkyl lesions.[170]

5.1.1. O^6-Alkyl Guanine

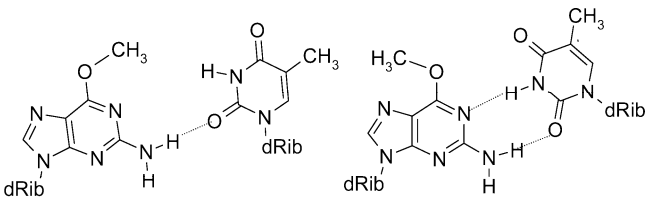
O^6-Alkyl guanine lesions considerably destabilize DNA. A UV-melting study has determined the stability of modified self-complementary Dickerson dodecamers d(C G **N** G A A T T C **M** C G), where **M** is O^6-MedG (**53a**) and **N** is any of the four natural deoxynucleotides, observing lower T_m values than those of unmodified duplexes.[171] Depending on the base pair type, T_m of modified duplexes decreases in the following order: dG·dC > dG·T > O^6-MedG·dC > O^6-MedG·dA > O^6-MedG·dG > O^6-MedG·T. However, bacterial DNA polymerases incorporate only dCMP and TMP opposite O^6-MedG during *in vitro* primer extension studies, with TTP being severalfolds a better substrate,[172] and the lesion causes only dG·dC → dA·T transition mutations *in vivo*.[173] In other words, formation of the least stable base pair appears to be highly preferable during DNA replication. To understand the structural bases of this phenomenon, a comparative NMR study has characterized a series of modified Dickerson dodecamers having a O^6-MedG·T,[174] O^6-EtdG·T[175] (O^6-EtdG, **53b**), O^6-MedG·dC,[176] O^6-EtdG·dC,[177] dG·T, or dG·dC base pair at the 10th position, with the latter two as controls for wobble and WC base pair alignments.

A common feature in all these duplexes is the observation of characteristic inter-nucleotide NOE cross-peaks that determine the *syn* (relative to guanine-N^1) conformation of the alkyl group, a fact most probably due to stereoelectronic effects similar to the *gauche* effect on aliphatic compounds. However, the small dispersion of ^{31}P NMR signals and the similarities in NOESY spectra put the dG·dC-, O^6-MedG·

T-, and O^6-EtdG·T-containing duplexes in one structural group, which has the regular B-type conformation and the normal WC type geometry at the lesion-containing base pair. In contrast, the O^6-MedG·dC- and O^6-EtdG·dC-containing duplexes have spectra similar to those of the dG·T duplex, establishing a second group with wobble geometry at the lesion site.

Because of the significant steric repulsion with thymine-O4 or cytosine-NH_2, the *syn* conformation of the O^6-alkyl group does not favor the WC base pair geometry with either base. As a result, the distance between O^6-alkdG-N1 and the thymine imino proton is too long and formation of a hydrogen bond does not take place. This is consistent with the observation that the imino proton signal of the O^6-MedG·T pair has a lower chemical shift value and a faster solvent exchange rate than other hydrogen-bonded imino protons of the duplex. Thus, only one hydrogen bond stabilizes the O^6-MedG·T base pair **55**, which is open slightly toward a major groove. It is worth noting, however, that this alignment is in contradiction with two subsequent X-ray studies, which establish an *anti* orientation of the O^6-alkyl group and formation of two hydrogen bonds across the O^6-MedG·T pair **56** in the crystalline state (Chart 18).[178]

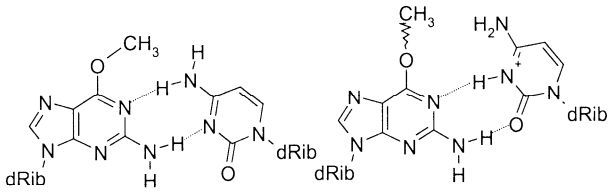
Chart 18. Solution 55 and Crystalline State 56 Alignments of the O^6-MedG·T Pair



55 **56**

In contrast to the WC alignment, the wobble geometry of the O^6-alkdG·dC base pair **57** (Chart 19) does not cause steric

Chart 19. O^6-MedG·dC Base Pair Alignments



57 **58**

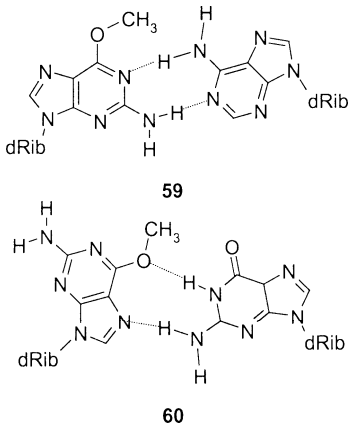
clashes when the alkyl group assumes a *syn* orientation. In accordance with this, the cytosine amino protons appear as a single downfield-shifted peak, implying that the amino group is rotating fast on the NMR time scale and forming a weak hydrogen bond. This base pair alignment changes when cytosine is protonated under acidic conditions. The effect of cytosine protonation on the behavior of O^6-alkdG·dC pairs was studied using a modified Dickerson dodecamer bearing a ^{15}N-label at the N^1 or N^2 position of O^6-methylguanine. The pH dependence of ^{15}N chemical shifts displayed a sigmoidal shape with an effective pK_a value close to 5, which was associated with protonation of the lesion partner cytosine. On the basis of the disappearance of the sigmoidal pH dependence at high temperatures and the amplitude of chemical shift changes upon protonation, the authors concluded that lowering the pH induces transition of the wobble O^6-alkdG·dC pair **57** to a protonated WC alignment **58**.[179] However, there was no experimental data presented to describe the conformation of the O^6 methyl group in this case, and hence, it is not clear how steric problems in the base pair **58** were resolved.

The increased basicity of cytosine in the O^6-alkdG·dC pair implies that a significant fraction of the protonated O^6-alkdG·dC$^+$ pair **58** is present at physiological pH and, therefore, may account for dC incorporation during DNA replication. This conclusion is partially corroborated by X-ray results demonstrating the WC type geometry for this pair in crystalline Z-DNA.[180]

The stability of O^6-MedG·dA and O^6-MedG·dG pairs is between those of O^6-MedG·dC and O^6-MedG·T pairs, suggesting that some ordered alignment takes place there. According to an early NMR study,[181] both bases of the O^6-MedG·dA pair assume *anti* conformations around their χ torsion angle and are stacked inside the duplex. At the same time, significant chemical shift changes of sugar protons at and adjacent to the lesion site suggest that the local geometry is perturbed considerably, a common characteristic of purine−purine mispairs. On the basis of these facts, the authors proposed a tentative mispair alignment **59** where the adenine base is displaced toward the minor groove and forms two H-bonds with the O^6-MedG residue.

The O^6-MedG·dG duplex shows a strong NOE cross-peak between the H8 and H1′ protons of the methylated nucleotide. Further, NOESY cross-peaks are absent between O^6-MedG-H8 and aromatic protons of adjacent bases that are usually seen between nucleotides with the χ angle in the *anti* range. On the other hand, the lesion partner dG residue shows normal spectral properties. On the basis of these observations, the authors conclude that O^6-MedG adopts a *syn* conformation around the χ torsion and forms a Hoogsteen O^6-MedG·dG base pair a alignment **60**, stabilized by two hydrogen bonds[181] (Chart 20).

Chart 20. O^6-MedG·dA 59 and O^6-MedG·dG 60 Base Pair Alignments



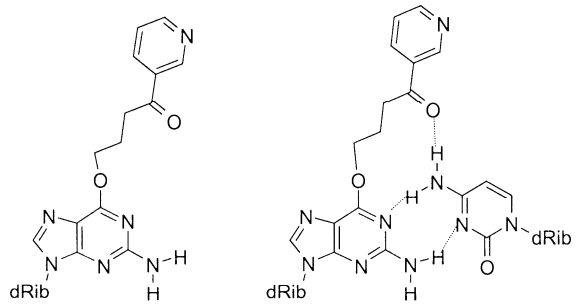
59

60

The above-mentioned studies only made qualitative descriptions of base pair alignments and hydrogen bonding at the lesion site without reporting the refined structural model of O^6-alkyl guanine lesion-containing DNA. A refined solution structure (PDB code 1PYJ) of an 11-mer duplex containing O^6-[4-oxo-4(3-pyridyl)-butyl]guanine ([POB]dG) **61**, a product of guanine-O^6 alkylation by the tobacco specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) paired to dC has been reported recently.[182] Analysis of the NMR spectra demonstrates that this duplex shares major features with the previously studied O^6-methyl and

-ethyl derivatives. All unmodified nucleotides exhibit sharp, well-resolved resonances consistent with a predominant B-type helical conformation. The O^6-[POB]dG residue exhibits conformational heterogeneity, which broadens the base and sugar proton signals, complicating the assessment of its conformation by standard NOE methods. However, the ^{13}C chemical shift value of the 1′-carbon of O^6-[POB]-dG suggests that its glycosidic bond is in the *anti* conformation and internucleotide NOE cross-peaks establish that the damaged residue stacks in the helix. The O^6-[POB]dG·dC base pair adopts a wobble alignment (Chart 21), similar to

Chart 21. O^6-[POB]dG and Its Alignment with dC



61

that reported previously for the O^6-MedG·dC pair **57**. NOE interactions between the aliphatic chain of POB and lesion-flanking bases indicate the POB group is coplanar with the modified guanine. In addition, its pyridine ring shows no NOE cross-peaks at short or moderate mixing times and, thus, is exposed to solvent. The conformation of the alkyl chain around the C6−O6 bond is intermediate between *syn* and *anti* values, suggesting a comparatively low rotational barrier. The RMD structure predicts the formation of an additional hydrogen bond between POB and the cytosine amino group of the lesion-containing base pair (Figure 8).

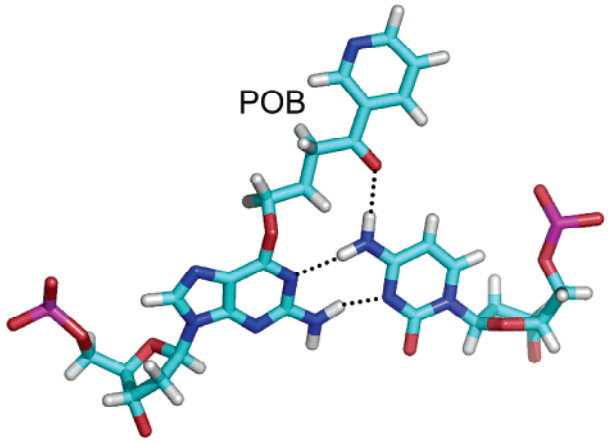


Figure 8. Alignment of the O^6-[POB]dG·dC pair with atoms are colored by type. Prepared from PDB entry 1PYJ[182] using PyMol.[555]

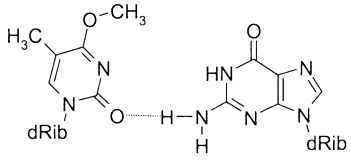
Despite this putative additional interaction, the presence of the POB lesion considerably decreases the duplex thermal stability. Overall, the structural characteristics of this duplex are consistent with previous observations made for O^6-methyl and O^6-ethyl derivatives.

Comparison of the base pair alignments described for O^6-alkyl guanine shows that, while any normal nucleotide is able to pair with O^6-MedG, only the O^6-MedG·T pair **56** retains a normal WC geometry, aiding understanding of the mutagenic profile of O^6-alkyl lesions. It has been noted that free energy differences between matched and mismatched base pairs, between 0.2 and 4.0 kcal/mol,[183] are insufficient to explain nucleotide incorporation. It was proposed later that not only the energy of the base pair but also the correct geometry of the alignment for the incoming nucleotide is a key selection criterion during DNA replication.[184,185] In the case of O^6-MedG, only O^6-MedG·T and O^6-alkdG·dC^+ (**58**) pairs meet this criterion, explaining incorporation of T and dC opposite O^6-MedG, as well as the absence of other mismatches during DNA replication.

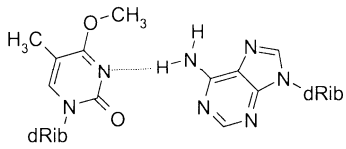
5.1.2. O^4-Alkyl Thymine

During DNA replication, O^4-MeT also induces dA·T → dG·dC transition mutations, and as in the O^6-alkylguanine case, wobble and WC base-pairing modes have been considered to explain this behavior. It has been observed that a duplex containing an O^4-MeT·dG base pair shares many key structural features with the O^6-MedG·T duplex.[186] Nucleotides at the alkyl lesion-containing base pair are in the *anti* conformation around the χ torsion angle with their WC edges facing each other. The conformation of the O^4-alkyl group is *syn* relative to the imino nitrogen. The chemical shift of the guanine imino proton in the O^4-MeT·dG base pair **62** (Chart 22) is upfield from the value observed

Chart 22. O^4-MeT Pairing with Purine Bases



62



63

in the normal dG·dC and wobble dG·T base pairs, ruling out the possibility of a wobble alignment across the O^4-MeT·dG base pair. Instead, the authors suggest the formation of a WC-like O^4-MeT·dG pair stabilized by the formation of a single hydrogen bond.

In contrast, the NMR evidence indicates that O^4-MeT pairs to adenine, forming a wobble alignment geometry (**63**) with a single hydrogen bond between the O^4-MeT imino nitrogen and a proton from the adenine amino group.[187]

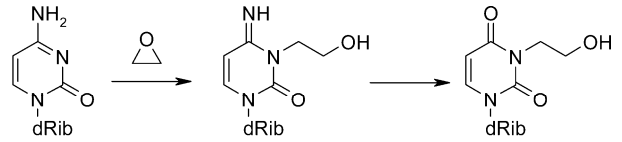
5.1.3. Bases Alkylated by Small Aliphatic Epoxides

Many low molecular weight aliphatic epoxides are widely used as industrial chemical intermediates. The smallest member of this class, ethylene oxide, is also use for sterilization of food and surgical products. Furthermore, small epoxides may be produced endogenously by the enzymatic oxidation of corresponding alkenes, which are common environmental pollutants present in car exhaust and tobacco smoke. All these epoxides may react with soft nucleophile centers present in cellular DNA, such as the endocyclic base nitrogens, giving rise, preferentially, to β-hydroxyethyl

derivatives of cytosine-N3, adenine-N1, and guanine-N7. In many instances, the base-pairing properties of alkylated residues are severely compromised, explaining why some epoxides, or their alkene precursors, are established carcinogenic agents in animals and are probably involved in human cancer etiology.

Reaction of ethylene oxide with DNA gives rise to, among other products, an N^3-hydroxyethyl derivative **65** that undergoes rapid deamination, producing 3-hydroxyethyl-2′-deoxyuridine (HEdU) **66** (Scheme 5),[188] which cannot form

Scheme 5. Reaction between Ethylene Oxide and Cytidine

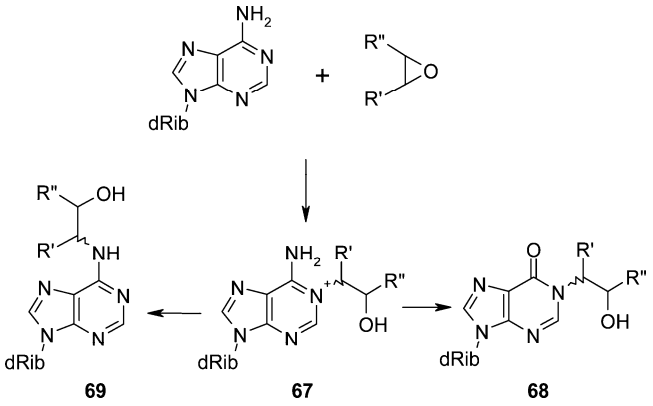


a WC base pair alignment. *In vitro* primer extension studies using A-family polymerases demonstrate that the presence of HEdU in the template strand blocks DNA replication.

However, in experiments with a mutated Klenow fragment or T7 DNA polymerases lacking proofreading activity, HEdU induces dG·dC → dA·T transition mutations and, to a lesser extent, dG·dC → T·dA transversion mutations.[189] These results indicate a misinstructive rather than noninstructive behavior of the lesion, suggesting that some specific interactions between HEdU and an incoming adenine or thymine base may take place during DNA replication. The NMR characterization of a DNA duplex containing HEdU **66** opposite adenine partially clarifies the nature of these interactions.[190] The NMR spectra show that all nonterminal nucleotides adopt a B-type helical conformation, with both bases of the HEdU·dA base pair in the *anti* range around the χ torsion angle. ^{31}P NMR chemical shifts have low dispersion, indicating a uniform B_I-type conformation of the sugar−phosphate backbone, and NOESY spectra indicate that all base pairs of the duplex, including the lesion-containing HEdU·dA pair, retain WC orientations. The RMD structure (PDB code 1KXS) demonstrates, however, that the bulky hyrdohyethyl (HE) group increases the C1′−C1′ distance on the lesion-containing pair from an average value of 10.8 Å, for a normal dA·T pair, to about 13 Å, forming a small bulge at the lesion site. Although variations of this distance were small during the RMD simulations, the HEdU·dA base pair and, especially, the HE group fluctuate among several conformations. The authors observe that on ∼20% of these states the distance between the OH proton of HE and adenine N1 favors the formation of a hydrogen bond. For other conformations, a hydrogen bond between HEdU O4 and the adenine amino group is possible. While formation of these hydrogen bonds does not contradict the NMR data, signals of the HE alcoholic and adenine amino protons were not observed in the spectra.[190] In addition, the putative formation of these hydrogen bonds would suggest stabilization of the HEdU·dA pair, explaining dA incorporation during DNA replication, but is insufficient to account for the observed selectivity of the lesion. Additional structural data about base pair alignments of HEdU and other nucleobases are needed in order to understand the misinstructional behavior of HEdU.

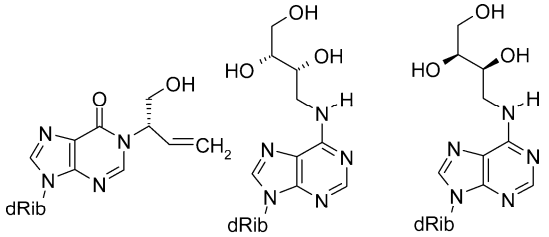
Inside the cell, cytochrome P450 dependent monooxygenases may oxidize butadiene (BD),[191] producing a series of mono- and diepoxides that attack at adenine N1 a, creating

Scheme 6. Products of Adenine and Epoxide Interactions



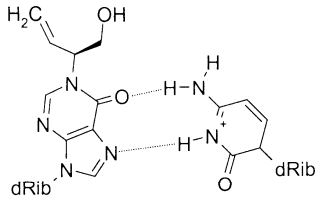
a second type of epoxy-derived adducts (Scheme 6). Reaction of BD epoxide with adenine N^1 gives rise to an unstable intermediate **67**, which may undergo either deamination to the N^1-alkylated deoxyinosine **68**[192] or Dimroth rearrangement to the N^6-alkyladenine derivative **69**[193] (Scheme 6). NMR spectroscopy and RMD simulations have solved the solution structures of these DNA lesions located at the second codon-61 position of the human *N*-ras protooncogene. The NMR studies have determined the structure of duplexes containing the 1,2-epoxy-3-butene-N^1-deoxyinosine adduct (BDOdI) (**70**) or the (2*R*,3*R*) (**71**) or (2*S*,3*S*) (**72**) conformations of the 1,2-dihydroxy-3,4-epoxybutene-N^6-adenine adduct, named *R*-BDTdA and *S*-BDTdA, respectively (Chart 23).

Chart 23. 1,2-Epoxy-3-butene Purine Lesions



In contrast to nonalkylated inosine, BDOdI is incapable of forming WC base pairs with dC, due to the presence of a bulky group at the WC edge. Nevertheless, this adduct is strongly mutagenic in *E. coli*, inducing dCTP incorporation.[194] The NMR spectra of an 11-mer duplex containing the most mutagenic isomer of BDOdI, the *S* enantiomer, paired to T show that the global DNA conformation is almost unperturbed by the presence of the lesion.[195] However, to avoid steric clashes, *S*-BDOdI assumes a *syn* conformation around its glycosidic bond, placing the alkyl chain in the major groove of the duplex. The RMD structure (PDB code 1U6N) displays the hypoxantine base displaced toward the major groove without interacting with the thymine residue in the complementary strand. The gap left by *S*-BDOdI is presumably filled by water, increasing the solvent exchange rate of the thymine imino proton and explaining its absence in the spectra. On the basis of these observations the authors propose that during DNA replication a protonated cytosine would pair to the *syn* BDOdI residue **73** (Chart 24), forming an alignment similar to that observed earlier for 1,N^2-propano-dG[196] (see section 5.3) and explaining the preferential dCTP incorporation opposite to the lesion. Structural

Chart 24. Proposed Alignment of the *S*-BDOdI·dC Pair



73

studies of DNA containing a BDOdI·dC base pair are needed to verify this hypothesis.

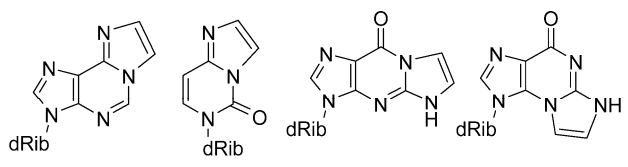
In contrast to BDOdI **70**, *R*-BDTdA **71** and *S*-BDTdA **72** adducts are much less mutagenic. They do not block DNA replication considerably and mutation levels are smaller than 0.25%.[197] The NMR characterization of DNA containing either an *R*-BDTdA·T[198] or an *S*-BDTdA·T[199] (PDB code 1U6C) pair shows normal WC alignment at the lesion-containing base pair, with the side chain of BDT embedded in the major groove in both cases. However, the type of mutations induced by BDTdA depends on the configuration of BDT,[200] implying that some specific interaction of the lesion side chain and DNA takes place. Since *S*-BDTdA induces dA·T → dC·dG transversion mutations, a DNA duplex with an *S*-BDTdA·dG mismatch has been studied.[201] The lesion-containing mismatch shares common properties with its nonalkylated counterpart, which has both bases oriented with their WC edges facing each other and the guanine residue slid toward the minor groove.[202] The authors propose the formation of a hydrogen bond between the β-hydroxyl group and dG, which could partially explain the mutagenic profile of the lesion, but the corresponding proton was not detected in the spectra, a fact probably due to fast exchange with solvent.

5.2. Alkylation Damages Caused by Bifunctional Agents

5.2.1. Acetaldehyde Derivatives

The need of fluorescent probes for studying nucleic acid structure and dynamics led to the introduction of ethenobases, more than 30 years ago. Interest in the biological properties of ethenobases followed the realization that the P540 system activates chemicals of widespread industrial use, such as vinyl chloride, chloroacetaldehyde, or urethane, to bifunctional alkylating agents, which then can react with DNA bases forming 1,N^6-etheno-dA **74** (ϵdA), 3,N^4-etheno-dC **75** (ϵdC), 1,N^2-etheno-dG **76** (1,N^2-ϵdG), and N^2,3-etheno-dG (N^2,3-ϵdG) **77** (Chart 25).[203] Subsequent epidemiological

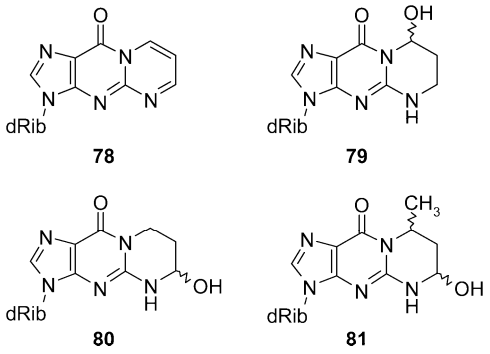
Chart 25. Etheno Nucleotide Lesions



74 75 76 77

studies linked these chemicals to the higher frequency of liver tumors observed in workers of paint factories, leading to strong regulatory policies. Promptly, the list of exocyclic lesions expanded to include the dG adduct formed by malondialdehyde **78** (M_1dG), 6-hydroxy-1,N^2-propano-dG **79** (γ-(OH)-PdG), 8-hydroxy-1,N^2-propano-dG **80** (α-(OH)-PdG), and 6-methyl-8-hydroxy-1,N^2-propano-dG **81** (Chart 26), which result from the reaction of malondialdehyde,

Chart 26. Six-Membered Exocyclic Guanine Adducts

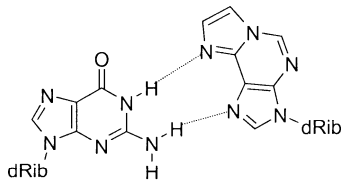


78 79 80 81

acrolein, and crotonaldehyde with DNA. The development of ultrasensitive detection methods led to the discovery of background levels of exocyclic adducts in DNA samples recovered from unexposed humans or from laboratory rats fed with a chemical-free diet. The search for the mechanism of formation of these endogenous adducts identified several lipid peroxidation products, including *trans*-4-OH-2-nonenal, acrolein, and molondialdehyde, as the most likely culprits. Furthermore, the cellular level of these lesions increases with stress, inflammatory processes, and metal store diseases, making the study of correlations between exocyclic adducts and downstream biological effects, such as mutagenesis, cell transformation, and cancer, the subject of active research.[204]

The first etheno base structurally characterized in DNA was ϵdA **74** paired to T or dG residues at the center of a 9-mer duplex.[205,206] Sequential interactions observed on the NOESY spectra recorded at 15 °C indicate that both duplexes adopt right-handed helical conformations with an increase of nucleotide dynamics at and around the damaged base, mostly seen in the ϵdA·dG duplex. All undamaged residues adopt an *anti* conformation around their glycosylic torsion angle and canonical WC base pair alignments. In the case of the ϵdA·T duplex, ϵdA is also *anti* with no spectroscopic evidence of hydrogen bonding across the ϵdA·T base pair. An energy-minimized model of the duplex structure derived using NOE distance restraints shows that the bases of the ϵdA·T pair are not coplanar. The ϵdA adduct is perpendicular to the helical axis and stacks properly between flanking bases, but its partner T residue is inclined toward its 5′-side without forming any hydrogen bonds. In the case of the ϵdA·dG duplex, ϵdA adopts a *syn* conformation, resulting in a ϵdA(*syn*)·dG(*anti*) alignment **82** that is stabilized by formation of two hydrogen bonds, with the striking feature that both hydrogen donors are located on the same molecule (Chart 27). The ϵdA(*syn*)·dG(*anti*) alignment proposed in the early

Chart 27. ϵdA(*syn*)·dG(*anti*) Pair Alignment



82

NMR study was later confirmed by X-ray crystallography methods.[207]

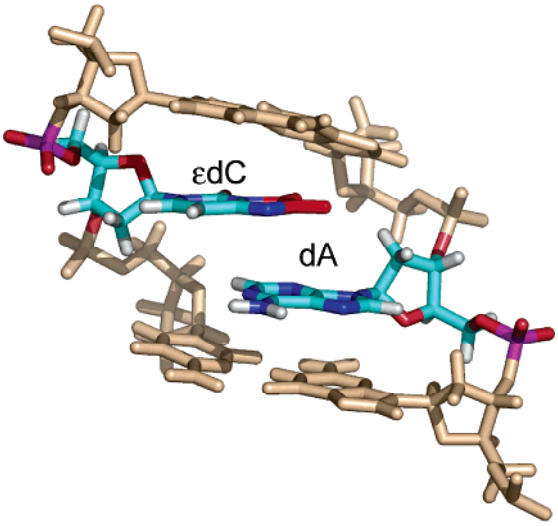


Figure 9. Highly staggered ϵdC·dA pair as seen from the major groove. The etheno fragment is colored red. Prepared from PDB entry 1B6Y[208] using PyMol.[555]

A second etheno lesion that has been structurally characterized by NMR spectroscopy and restrained molecular dynamics was ϵdC **75**. Our laboratory has reported the solution structures of 11-mer duplexes having a centrally located ϵdC residue paired to dA, T, dC, or dG.[208−211] The NMR spectra indicate that all four duplexes are regular right-handed helices with residues having an *anti* conformation around their χ torsion angle and base pairs stabilized by WC hydrogen bonds. Proton signals of nucleotides outside the lesion site have very similar chemical shifts, narrowing structural differences among these duplexes to the lesion-containing base pair and its interaction with flanking residues. The lesion partner determines the structural adjustments needed to accommodate the lesion inside the helix, and in all four cases a defined average structure is present in solution without evidence of alternative conformations at 25 °C. The most unusual structure is that of the ϵdC·dA duplex where both residues, ϵdC and dA, have χ torsions in the *anti* conformation but a highly staggered base pair alignment (4.8 Å), with residues displaced toward their 5′-side, avoiding any steric clash that would result from the χ *anti* conformation. The aromatic planes of ϵdC and dA are almost perpendicular to the helical axis, facilitating stacking interactions at the lesion site, but there are no hydrogen bonds across the lesion-containing base pair (Figure 9).

The other ϵdC duplexes accommodate the lesion in a similar fashion by forming a sheared base pair alignment where ϵdC and its partner base are displaced toward the major and minor grooves, respectively. The ϵdC residue adopts a *syn* conformation, forming a strong ϵdC(N4)−T(H3) hydrogen bond, and the duplex is slightly kinked at the lesion site. On the contrary, the helix remains more or less straight and all residues are in the *anti* conformation in the ϵdC·dG and ϵdC·dC duplexes, allowing the formation of hydrogen bonds across the lesion-containing base pair, namely ϵdC(O2)−dG(N1H) and ϵdC(O2)−dC(N4H) (Figure 10).

The structure of these duplexes provided a partial explanation for the observation that bacterial MUG and mammalian hTDG, which remove uracil or thymine residues when they are present in dG·dU or dG·T mismatches, can cleave ϵdC residues paired to any base in double-stranded DNA.[212] As in the case of the ϵdC duplexes, the dG·T or dG·dU mismatch alignment is highly sheared with the T residue

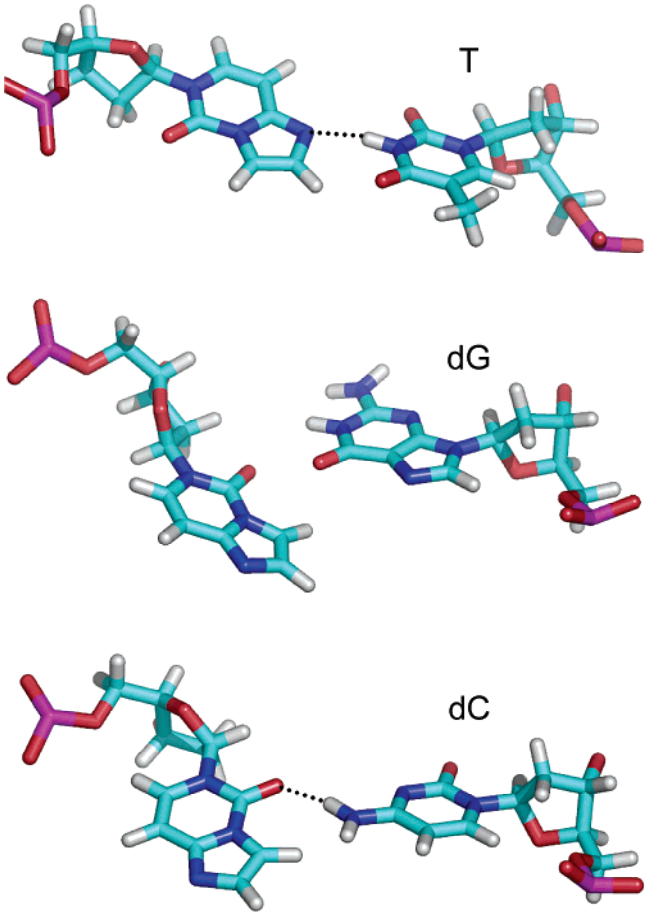


Figure 10. Alignment of the ϵdC·T, ϵdC·dG, and ϵdC·dC base pairs. Putative hydrogen bonds are shown as dotted lines. Prepared from PDB entries 1B5K, 1B6X, and 1B60, respectively,[209−211] using PyMol.[555]

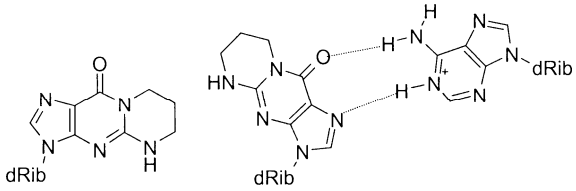
displaced toward the major groove of the helix.[157] When the repair protein binds, its substrate intercalates, from the minor groove side of the duplex, a three-amino-acids wedge inside the DNA helix and extrudes the scissile base into its catalytic pocket.[213] It is conceivable that a major-groove-positioned ϵdC or T residue does not block or even favor the extrusion process and formation of a productive DNA−protein complex. A scissile residue located in the minor groove of the helix would interfere with the extrusion process blocking the enzyme activity. In addition, while the staggered ϵdC·dA alignment does not support this hypothesis, it does not refute it either. It is worth noting that lesion-flanking base pairs were different for the enzymatic and structural studies, opening the possibility that sequence dependent effects may be hindering the observation of a staggered ϵdC·dA alignment. Modeling studies show that, in principle, ϵdC·dA is able to adopt a sheared alignment that could be stabilized by a ϵdC(O2)−dA(N6H) hydrogen bond. Additional structures of the ϵdC·dA duplex, as well as the structural characterization of the 1,N^2-ϵdG adduct, which is also repaired by MUG, are needed to further evaluate the structural basis for the recognition of ϵdC by these proteins.

5.2.2. Acrolein and Malonaldehyde Derivatives

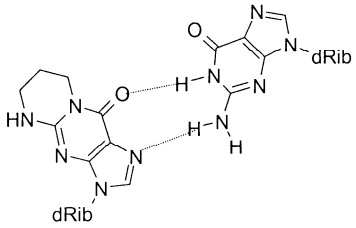
To overcome problems associated with the synthesis and purification of M_1dG (**78**) and acrolein adducts (**79**, **80**), many biological and structural studies have used 1,N^2-propano-dG **83** (PdG), as a chemically stable lesion analogue. NMR spectroscopy and restrained molecular dynamics have

determined the solution structure of duplexes containing PdG opposite dG, dA, or dC, abasic sites, and two-base deletion in a loop region of a hairpin sample. Early studies of 9-mer duplexes having a centrally located PdG residue paired to dA or dG establish that duplexes are right-handed helices with unmodified nucleotides in the *anti* conformation around the χ torsion angle and WC hydrogen bonds present on all canonical base pairs, including those flanking the lesion.[214−217] At the lesion site, the PdG·dA-containing duplex displays a pH dependent conformational exchange, with broad NMR signals at neutral pH and a pK_a value around 7.6. Characterization of the duplex structure at pH 5.8 indicates a *syn* conformation for the χ torsion angle of PdG with concomitant protonation of its partner dA residue. The refined structure shows a PdG(*syn*)·dA$^+$(*anti*) base pair **84** properly stacked between flanking base pairs and stabilized further by formation of two hydrogen bonds (Chart 28). The PdG·dG

Chart 28. Propano-dG 83 and Its Alignment with dA 84 and dG 85



83 **84**

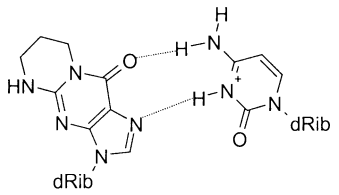


85

duplex shows similar NMR parameters for all residues, including those located at the lesion site, but in this case, formation of a PdG(*syn*)·dG(*anti*) base pair alignment **85** is pH independent.[214,216] The initial NMR characterization of the PdG·dA duplex at basic pH indicates that all residues, including those at the lesion site, adopt the *anti* conformation around the χ torsion angle. The PdG(*anti*)·dA(*anti*) alignment shows PdG inclined in the 3′-direction and partially intercalated between its partner dA and the 5′-neighboring residue. The partner dA residue adjusts in an equivalent manner, appearing inclined toward its 3′-side and intercalated between PdG and its 5′-neighbor residue. No hydrogen bonds stabilize the alignment at basic pH, suggesting that a pH-driven balance between hydrophobic and hydrogen-bonding contributions to duplex stability determines the PdG·dA base pair alignment in solution.[215] Further refinement of the PdG·dA duplex structure at high pH keeps intact the essential elements of the PdG·dA alignment described above but shows the adduct largely displaced in the major groove of the helix. Since the major groove location of PdG allows rotation around the χ torsion angle, the authors used this structure to propose a path for the exchange between the acidic and basic conformations of the PdG·dA duplex in solution.[217]

The NMR characterization of a PdG residue paired to dC has been performed under acidic conditions in two different sequence contexts, specifically at the center of a d(CG)$_3$ segment and in the 5′-d(...T X T...) context.[218,196] In the latter sequence context (PDB codes 1LAS and 1LAQ), PdG reduces the thermal stability of the duplex by 14 °C. NMR experiments performed at 20 °C show that the sample has very similar properties to those reported for the PdG·dA and PdG·dG duplexes, with residues having χ torsion angles in the *anti* conformation and canonical base pairs stabilized by formation of WC hydrogen bonds. At the lesion site, PdG adopts a *syn* conformation around the χ torsion angle, forming a PdG(*syn*)·dC$^+$(*anti*) **86** base pair alignment stabilized by two hydrogen bonds[196] (Chart 29). In the d(CG)$_3$

Chart 29. PdG(*syn*)·dC$^+$(*anti*) Pair Alignment



86

stretch, PdG causes a larger thermal stability reduction with a 24 °C drop in the duplex melting temperature.

The exchangeable proton spectra indicate that at least two, or even more, cytosine residues are protonated at the working pH value of 5.8. As seen with other PdG duplexes, the lesion also adopts a *syn* conformation around the χ torsion angle and forms a PdG(*syn*)·dC$^+$(*anti*) base pair **86**. The presence of additional NOE interactions leads to the conclusion that the χ torsion angle of the 3′-side lesion-neighboring dG residue exchanges between *syn* and *anti* conformations, generating multiple duplex structures in solution.[218] On the basis of these observations, the authors proposed a correlation between the presence of sequential Hoogsteen base pair alignments and the high frequency of two-base deletion mutations measured in the d(CG)$_3$ context. It is worth noting that in this work the adduct was only three base pairs apart from the duplex end, a fact that could have promoted partial dissociation of terminal base pairs with simultaneous protonation of cytosine residues and formation of alternative structures. The unusually high reduction of T_m suggests that end-sequence effects may have played a determining role in the thermal stability of this duplex. Hence, to understand the biological implications of these observations, it should be of interest to study a DNA duplex with a PdG-containing d(CG)$_3$ segment farther away from the end.

Additional studies of PdG-containing duplexes have placed the lesion opposite an abasic site (AP),[219] as the 5′-residue on a two-base bulge,[220,221] or on the loop region of a hairpin sequence[222] (PDB code 1LAE). At 10 °C and pH 7.3, the PdG·AP-containing duplex adopts a right-handed helical structure with all residues including the PdG lesion in the *anti* conformation around the χ torsion angle and WC base pair alignments throughout the duplex. At the lesion site, the PdG adduct is inside the helix with its exocyclic ring partially filling the cavity left by the absence of a partner base and properly stacking between flanking base pairs. The AP residue is also intrahelical and faces the PdG moiety. This structure and the structure of the PdG·dA duplex at basic pH discussed above were the first reports supporting the concept that hydrophobic interactions may stabilize a helical duplex structure without the need for hydrogen bonds.[219] Stacking of an intrahelical PdG adduct with neighboring base pairs without formation of hydrogen bonds

is observed also in the structure of a double-stranded oligonucleotide that has a single PdG lesion as the 5′-residue of a targeted XpC bulge. At 25 °C, the duplex is kinked at the bulge site by 20−35°, separating two regular helical domains with base pairs stabilized by canonical WC alignments. Direct NMR evidence places the PdG residue inside the helix in the *anti* conformation around χ and stacking with the guanine moiety of the 5′-flanking dC·dG pair. The dC residue at the 3′-side of the bulge is extruded toward the major groove of the helix, where it stacks with a 3′-neighboring guanine[220,221] (Figure 11). An additional example of

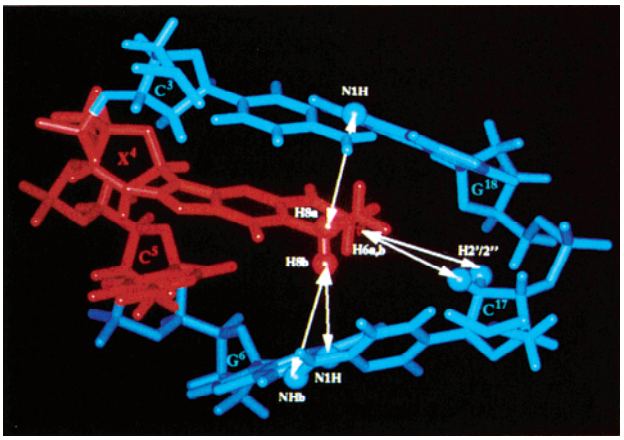


Figure 11. Lesion site structure of duplex DNA having a targeted 5′(PdG dC) bulge. This lesion site view shows the major groove prominent with bulged residues in red. Reproduced with permission from ref 221. Copyright 1995 American Chemical Society.

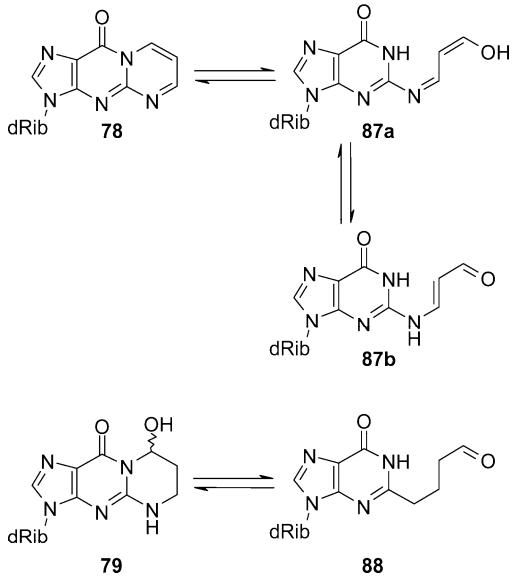
an oligodeoxynucleotide structure having a PdG adduct stabilized without formation of hydrogen bonds is that of a stem-loop sample where the adduct was at the center of a three-residue loop region. In this case the lesion stacks over the final dC·dG pair of the stem region with one side of the PdG ring almost completely exposed to solvent. The structure and the thermodynamic stability of this duplex were very similar to those of a control hairpin where a dG residue substituted for the PdG lesion.[222]

For several years the structure and the biological properties of DNA-containing PdG residues served as models for the lesions produced by malondialdehyde and acrolein, two endogenous alkylating agents formed during lipid peroxidation and oxidation damage.[204] However, the validity of this structural analogue was partially called into question after M_1dG and γ-acrolein adducts **78**−**80** were efficiently introduced into DNA samples and the NMR characterization of these duplexes revealed unexpected results.[69,223−225] When a duplex sample has an M_1dG·dC pair, the exocyclic adduct spontaneously and quantitatively converts to N^2-(3-oxo-1-propenyl)-dG (**87a**), a ring-opened conformation that reverses to M_1dG upon thermal denaturation (Scheme 7). The authors postulated that the lesion partner dC residue plays an active role during rearrangement to a ring-opened conformation by participating in the transient formation of a Shiff base intermediate. Interestingly, the NMR data suggested that ring-opened states of M_1dG **87** do not form WC hydrogen bonds with dC.[69] The refined duplex structure shows the N^2-(3-oxo-1-propenyl)-dG lesion **87** in the *anti* range around χ with the propenyl chain located in the minor groove causing only local perturbations on the duplex structure. The lesion partner dC residue protrudes slightly into the major groove of the helix, a displacement that was considered sufficient to hinder the formation of WC hydrogen bonds at the lesion-containing base pair.[223a] However, during the recent characterization of mutagenic properties of M_1dG, the same research group identified NMR evidence supporting the formation of a WC alignment at the M_1G·dC base pair,[223b] an alignment that had been reported previously for the γ-(OH)-PdG **78** adduct in duplex DNA.[225]

In contrast to the open conformation seen in duplex DNA, M_1dG is chemically stable and remains ring-closed when it is the 5′-residue of a two-nucleotide bulge[224] (PDB code 1K29). Outside the bulge site, the duplex takes on a regular helical structure with base pairs stabilized by WC alignments. Both residues of the bulge are in the *anti* conformation around the χ angle and appear inside the helix, where they lack hydrogen-bonding interactions. Globally, the duplex structure is similar to that of the duplex containing a two-nucleotide bulge, with PdG as the 5′-bulged residue in the same sequence context.[221] However, a notable difference between these duplexes is that the aromatic moieties of the bulged residues are parallel and properly stacked on top of each other only in the case of the M_1dG lesion.

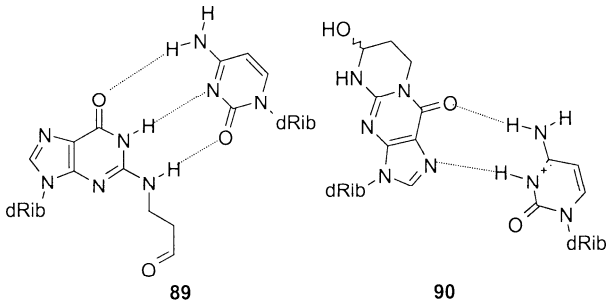
Chemical rearrangement of an exocyclic ring also occurs in the case of γ-(OH)-PdG **78**, the main lesion formed by acrolein (Scheme 7).[225] When paired to dC, duplex annealing triggers the opening of the 1,N^2-propano ring of γ-(OH)-PdG to the N^2-γ-oxopropyl form (**88**), which places the γ-oxopropyl chain in the minor groove of the helix and facilitates the formation of a regular WC pair **89** with the opposing cytosine. The ring-opened conformation of γ-(OH)-PdG causes minimal perturbation of the helical structure, and standard WC alignments are present throughout the duplex sequence. Despite the similarities in the chemical rearrangements experienced by γ-(OH)-PdG and M_1dG adducts, these lesions have different biological properties, where only the M_1dG lesion is mutagenic in bacteria. The insertion of dCMP during DNA replication is likely to trigger the chemical rearrangement of both lesions to their ring-opened forms, facilitating translesion DNA synthesis. It is possible, though, that the highly flexible propenal chain of γ-(OH)-PdG **88** can be more easily accommodated in the catalytic pocket of the DNA polymerase than the more rigid conjugated system of M_1dG **87a**, explaining their different mutagenic activities.

Scheme 7. Conversion of 78 and 79 to Ring-Opened Forms



Recently, our laboratory has established the solution structure of a related duplex containing the α-acrolein adduct **80** paired to dC at the center of an 11-mer duplex.[226] The duplex is stable at room temperature, has all unmodified residues in the *anti* range around the χ angle, and canonical WC alignments on all normal base pairs. At the lesion site, α-(OH)-PdG has its χ torsion angle in the *syn* range, is coplanar with its base pair partner, and stacks mainly with the 5′-neighboring residue. Clear NMR evidence indicates protonation of the lesion-partner dC, even at pH 6.9, with formation of an α-(OH)-PdG(*syn*)·dC$^+$(*anti*) base pair **90** that is stabilized by at least one hydrogen bond (Chart 30). In

Chart 30. γ-(OH)-PdG·dC 89 and α-(OH)-PdG·dC 90 Base Pair Alignments



89 **90**

summary, several precepts emerge from the structures of duplexes having exocyclic lesions. First, rotation around the χ torsion angle of the adduct to the *syn* conformational range with formation of an adduct(*syn*)·partner(*anti*) base pair alignment is often the lowest energy structure of duplexes having exocyclic lesions. Second, exocyclic lesions can be stabilized inside the helical DNA structure without participation of the damage residue in hydrogen bond formation. This principle applies in general to all bulky aromatic systems, including polycyclic aromatic hydrocarbon lesions and pyrene residues. Finally, the wobble base pair alignment adopted by ϵdC lesions explains the processing of this lesion and dG·T mismatches by the same DNA glycosylase.

5.3. Interstrand Cross-links

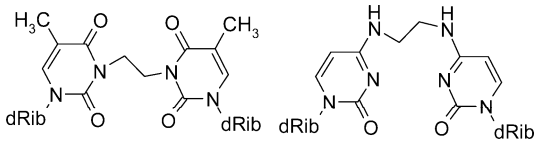
Bifunctional alkylating or acylating agents are capable of producing interstrand cross-links (ICLs). DNA cross-links are formed by exotic compounds that are generally absent in the cell, including nitrogen mustards, diepoxybutane, 2,5-bis(1-aziridinyl)-1,4-benzoquinone, and cisplatin. The DNA major repair pathways, BER, NER, and MMR, remove lesions from one strand of the duplex without affecting the complementary strand, which serves as the template for subsequent gap-filling synthesis. An ICL that covalently binds both strands of a duplex requires a more elaborate repair mechanism and, thus, represents a major challenge for the cell. Further, ICLs also are an impassable block for DNA replication and may casue cell death, a property that explains the use of many cross-linking agents as anticancer drugs.[227]

To understand putative mechanisms involved in recognition and repair of ICL, the spatial structure of DNA fragments containing such lesions must be known. However, site specific incorporation of relevant ICLs into DNA is a difficult task, because such a synthesis is not compatible with solid-phase phosphoramidite chemistry. To circumvent this problem, DNA duplexes containing artificial linkers have been studied as models for the lesions occurring *in vivo*.

5.3.1. Cross-linked Bis-pyrimidines

An original approach has allowed the preparation of self-complementary d(C G A A A **X** T T T C G)$_2$ undecamers containing an N^3T-ethyl-N^3T (**91**, Chart 31)[228] or N^4dC-ethyl-

Chart 31. Ethylene Interstrand Cross-links



91 **92**

N^4dC (**92**)[229] fragment at the middle. The structural characterization of duplexes having the **91** (PDB code 1S37) or **92** (PDB code 1N4B) ICL at the central step shows that the linker location is different in each case. In the **91**-containing duplex, the linker is located between the WC edges of the cross-linked thymines and, therefore, remains inside the helix, whereas on duplex **92** it is exposed to the major groove. The geometry of both cross-linked mismatches is close to that of the canonical WC pair, with a face-to-face base orientation and a C1′−C1′ distance of about 10.2 Å, very close to the 10.7 Å value of normal WC pairs. As a result, both duplexes form a compact stem with a regular B-type conformation.[230,231] Despite the flexibility of the ethyl linker, the tetrahedral geometry of its carbons causes, in both cases, the loss of coplanarity of the cross-linked bases, although the magnitude of this effect is different. In the case of the thymine ICL **91** duplex, the cross-linked bases show only a moderate propeller twist that has negligible consequences on adjacent base pairs and the rest of the DNA (Figure 12).[230]

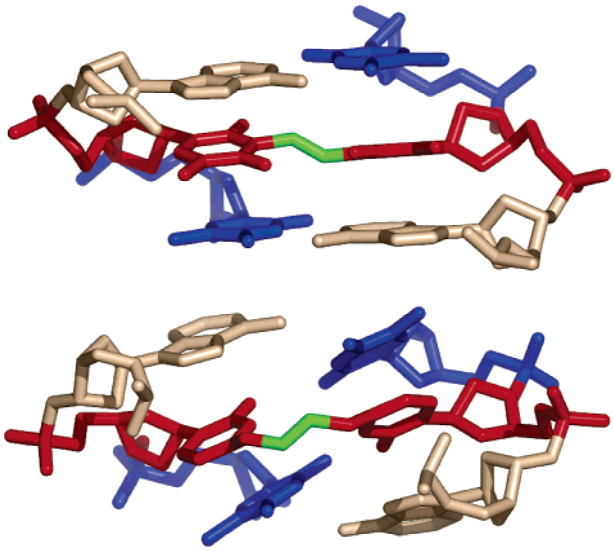


Figure 12. Lesion site structure of duplexes containing ICL **91** (top panel) and **92** (bottom panel). The pictures display central three-nucleotide segments with cross-linked residues in red, the ethylene tether in green, and flanking thymine and adenine residues in blue and wheat, respectively. Hydrogen atoms are not shown. Prepared from PDB entries 1S37[230] and 1N4B[231] using PyMol.[555]

In contrast, cytosine ICL **92** significantly alters the structure of its surroundings.[231] Thymine imino protons of ICL-flanking dA·T base pairs appear slightly upfield and show strong NOE cross-peaks to the hydrogen-bonded and exposed amino protons of their partner adenine and to the methylene protons of the linker. Strikingly, they do not exhibit cross-peaks to the adenine-H2 proton, which is a normal NOE

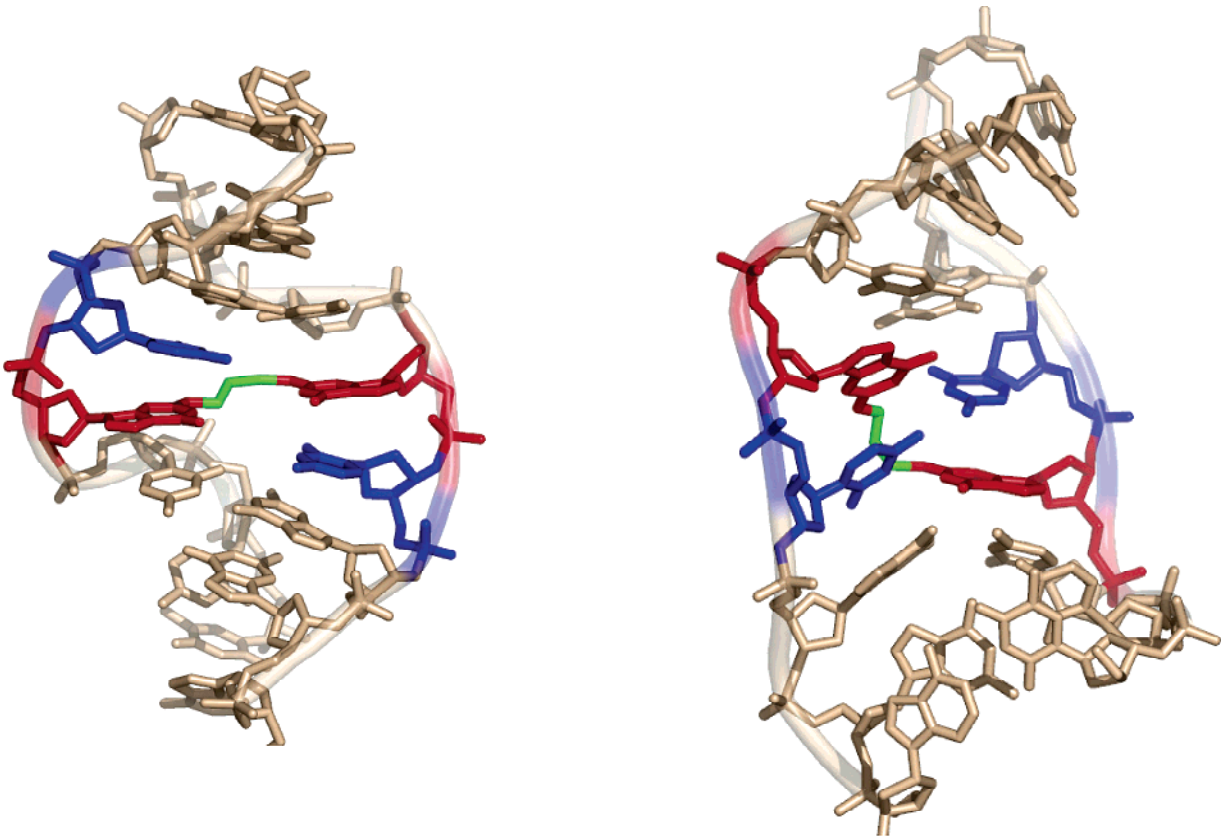


Figure 13. Structures of self-complementary duplexes **95** (left panel) and **96** (right panel) containing the trimethylene linker **94**. The pictures show cross-linked guanine residues in red, opposing cytosine residues in blue, and the linker in green. Hydrogen atoms are not shown. Prepared from PDB entries 1HZ2[233] and 1LUH[234] using PyMol.[555]

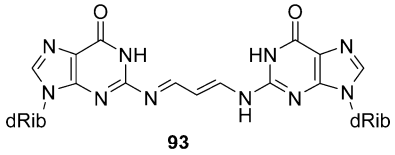
interaction present across dA·T base pairs. The RMD structure of this duplex shows the cytosine ICL fragment significantly propeller twisted, causing the partial rupture of the alignment of adjacent dA·T base pairs, which has only one hydrogen bond between thymine 4-oxo and the adenine amino group (Figure 12). The lesion-flanking dA·T pairs are also highly propeller twisted and sheared, showing the adenine and thymine residues are shifted toward the minor and major grooves of the helix, respectively. These perturbations place flanking thymine imino protons far from the adenine-N3 and close to the cytosine-O2 of the ICL, an arrangement that is consistent with the NMR observations. In addition, the helical twist at the d(ApC) step increases, almost doubling the value observed in canonical B-DNA. The local helix diameter at the ICL site, measured as the distance between phosphate groups of complementary nucleotides, increases by 2.5 Å, with a concomitant decrease in the width of both the major and minor grooves.[231]

5.3.2. Guanine ICLs

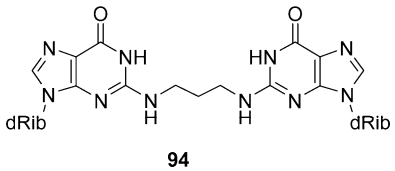
In the proper sequence context, some bi-electrophiles discussed in the exocyclic lesions section are theoretically capable of forming ICLs. It has been reported that malonodialdehyde can form Schiff base type guanine ICL **93** (Chart 32), only in oligonucleotides containing d(CpG) steps.[232] Since the linker **93** has limited stability and exists in equilibrium with its open form, its structural characterization is problematic. As a result, the factors responsible for sequence specific formation of ICL **93** remain to be established. To perform NMR studies, the guanine trimethylene linker **94**, a chemically stable mimic of the malonodialdehyde ICL, has been incorporated at the central d(CpG) or d(GpC) step of self-complementary d(A G G C G* C C T)$_2$ **95** and d(T C C G* C G G A)$_2$ **96** octamers, where G* indicates cross-linked residues.[233,234]

According to NMR data, cross-linked guanine bases adopt an *anti* conformation around the χ torsion angle with the alkyl tether substituting the amino hydrogen normally involved in WC hydrogen bonding. The tether conformation and the relative position of guanine bases across the ICL are completely different in duplexes **95** and **96**, leading to quite distinct overall conformations. In duplex **95**, the linker adopts a W-type conformation, having *anti*-periplanar C−C and C−N bonds, and produces an almost perfectly planar dG·dG mismatch at the lesion site with a longer C1′−C1′ distance. Cytosine residues, which were paired in the parent duplex to the now cross-linked guanines, lack base pair partners and form a wedge between the guanine ICL and adjacent dG·dC pairs (Figure 13). The two wedges are symmetrically situated relative to the ICL **94** and mutually compensate their inherent bend so that the main global

Chart 32. Three-Carbon Guanine ICL



93



94

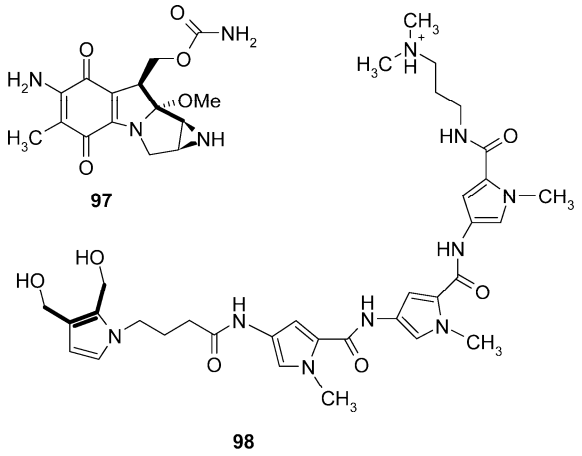
distortion of the helix is only a sharp decrease of the twist angle at the lesion site, with values of about 10° (PDB code 1HZ2).[233]

In duplex **96**, the linker adopts a nonplanar skewed U-type conformation, with *gauche* C−C and C−N bonds, allowing central cytosine residues to remain opposite to their partner cross-linked guanines (Figure 13). However, the linker causes steric restrictions that induce a large propeller twist in the lesion-containing pairs and a prominent buckle in the ICL-flanking base pairs. Furthermore, the four central base pairs of the duplex show almost zero helical twist, flattening the major and minor grooves in the vicinity of the lesion site[234] (PDB code 1LUH).

In contrast to the trimethylene linker, the malonodialdehyde linker is a chain of conjugated double bonds, which impose additional limitations on its conformation and require it to be planar. As shown by the **95** and **96** duplex structures, a planar tether conformation is present only in the case of the d(CpG*) cross-link, explaining the specificity of malonodialdehyde ICL formation at d(CpG) steps.

It is interesting to compare the structure of the guanine trimethylene ICL with those of four-carbon cross-links formed by mitomycin C **97**[235] or the 2,3-dihydroxymethylpyrrole−distamycin conjugate **98** (Chart 33).[236] These

Chart 33. Mitomycin C 97 and the 2,3-Bis(hydroxymethyl)pyrrol−Distamycin Conjugate 98[a]

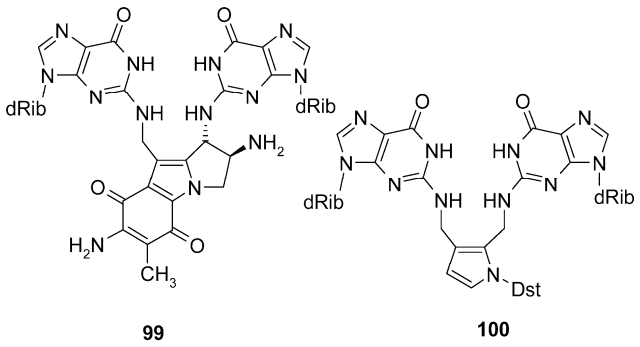


97 **98**

[a] Shown in bold face is the four-carbon ICL-forming segment.

compounds have modified alkyl groups at positions 2 and 3 of a pyrrole nucleus, which are capable of forming covalent bonds with the exocyclic amino group of guanine, producing isosteric ICL **99** and **100** (Chart 34). The bis-hydroxymethyl

Chart 34. ICL Formed by Mitomycin C and the 2,3-Bis(hydroxymethyl)pyrrol−Distamycin Conjugate

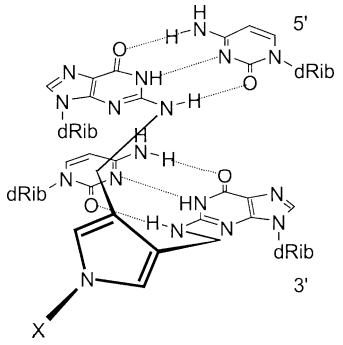


99 **100**

Dst = distamycin residue

pyrrole component of **98** is lost after formation of the ICL **100** but is needed for the delivery of ICL-forming groups to the DNA. The NMR characterization of duplexes having cross-links **97** or **98** at d(CpG) steps shows smaller perturbation of the DNA structure than that for the trimethylene linker. In both duplexes, the NMR spectra indicate the amino protons of the cross-linked guanines are hydrogen bonded to the 2-oxo groups of partner cytosine residues and are in slow exchange with solvent, indicating that, in contrast to the case of the trimethylene linker, solvent-exposed guanine amino protons were substituted by the **97** and **98** linkers. As a result, two dG·dC base pairs are formed at the cross-link sites of both duplexes, as represented in Chart 35,

Chart 35. Schematic Representation of Base Pair Alignments and the Linker Conformation of ICL 99 and 100



without disrupting the formation of regular WC alignments in the other base pairs of these duplexes. In the case of the mitomycin C ICL, NMR-based energy minimization produced two models that differed only in the pucker of the nonplanar pyrrol ring of the linker.[235] In both models, the bulky aromatic fragment is in the minor groove of the helix, partially exposed to solvent and without causing steric clashes. In the case of the ICL **100**, an energy-minimized model shows the bulky conjugate **98** in the minor groove of the duplex, where the shape of the distamycin moiety nicely fits the curvature of the double helix.[236] ICL **100** induces small perturbations of the duplex conformation, including a buckle in opposite directions of the cross-linked dG·dC base pairs that brings the guanine moieties close to each other and decreases the helical twist at the lesion site.

In summary, structural comparison of duplexes having a saturated trimethylene ICL with those containing a four-carbon partially unsaturated skeleton of ICL **99** or **100** indicates that the latter moiety is an optimized linker that does not disrupt d(CpG) steps.

6. Pyrimidine Photo-cross-link Products

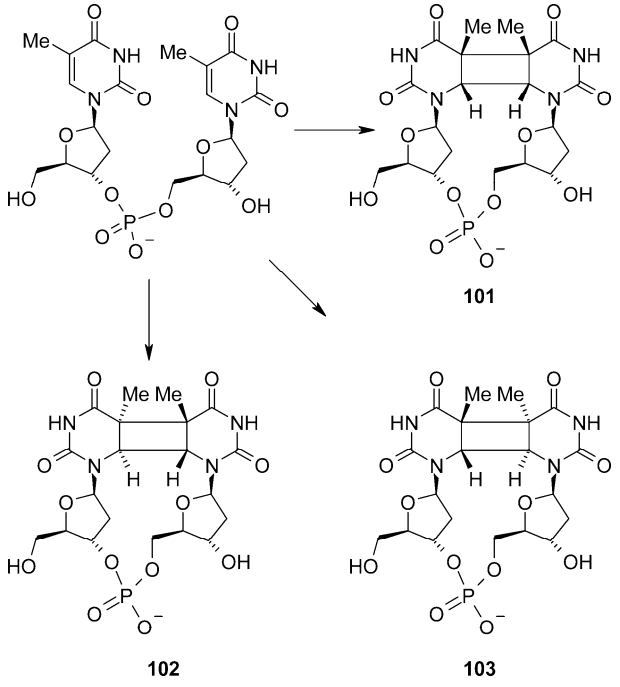
6.1. Type of UV Lesions and Reaction Mechanisms and Stereochemistry of Products

When two pyrimidine bases are properly positioned, excitation of electronic transitions in one of them may cause formation of cross-links. Depending on UV radiation wavelength, DNA conformation, and direction of the attack, different chemical reactions may take place in DNA, forming several addition products, including cyclobutane pyrimidine dimers (CPDs), 6-4 pyrimidine−pyrimidone products (6-4PP), and 6-4 Dewar products. They are briefly discussed below.

6.1.1. [2+2] Cycloaddition

CPDs, the [2+2] cycloaddition products of two pyrimidine bases, are the predominant lesions formed by UV irradiation of DNA. Formation of CPDs is more rapid upon irradiation with UVC light, $\lambda < 290$ nm, but they remain the major photoproducts of UVB or UVA radiation, that has longer wavelength or sunlight irradiation.[237] As all [2+2] photoinduced olefin cycloaddition reactions, formation of CPD proceeds as a *syn* cycloaddition and, depending on the relative orientation of pyrimidine bases, three different stereoisomeric products are possible. In right-handed duplexes, where bases are in the *anti* conformation around the χ angle and stack with each other, formation of the *cis-syn* CPD **101** is heavily favored, with less than 2% of *trans* isomers as minor products. When pyrimidine bases have random orientation, as is the case in molten DNA, all possible CPD stereoisomers are formed, including the *trans-syn-I* **102** (major) and the *trans-syn-II* **103** (minor) lesions[238] (Scheme 8). To form a CPD, pyrimidine moieties must be spatially

Scheme 8. Structure of *cis-syn* 101, *trans-syn-I* 102, and *trans-syn-II* 103 Cyclobutane Pyrimidine Dimers



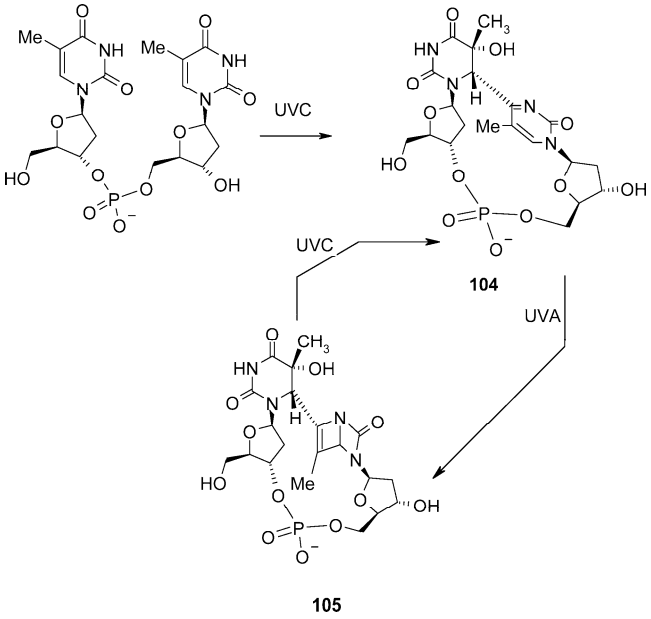
close but not necessarily contiguous in the DNA sequence. Indeed, formation of a CPD has been reported to occur between thymine bases separated by a bulged nucleotide.[239] CPD formation occurs predominantly between thymine bases, but cytosine may also participate and minor quantities of dC-T, T-dC, and dC-dC CPDs are present after DNA irradiation.[240,241] When cytosine takes part in a CPD, the loss of aromaticity increases its deamination rate, giving rise to a dU-T, T-dU, or dU-dU CPD. In addition, fluorouracil can also form CPDs as well as an additional lesion that results from opening of the cyclobutane ring.[242]

6.1.2. 6,4-Cycloaddition and Photo-rearrangement of the Pyrimidine (6-4) Pyrimidinone Adduct

The second type of photoinduced lesion involves the formation of a covalent bond between the 3′-nucleotide C4 and the 5′-nucleotide C6, giving rise to a pyrimidine (6-4) pyrimidinone (6-4PP) adduct **104**. The 5-hydroxypyrimidine moiety of 6-4PP always has a *5R,6S* configuration[243] and, as in the CPD case, the wavelength of UVC light is the most efficient for generating these lesions.

The resulting photoproduct **104** has an absorption maximum at 324 nm, so subsequent irradiation with UVA or sunlight induces its rearrangement into the Dewar isomer **105**, which retains the configuration of the chiral atoms. The (6-4)Dewar photoadduct is chemically stable but upon UVC irradiation rearranges back to the 6-4PP (Scheme 9).[244] As

Scheme 9. Formation of the 6-4PP 104 Lesion and Its Rearrangement to the 6-4Dewar 105 Dimer



a result, UV irradiation of DNA activates all three processes depicted in Schemes 8 and 9, giving rise to the formation of *cis-syn* CPD, 6-4PP, and 6-4Dewar lesions in a 17:3:1 ratio.[237]

Photodimerization, the process that leads to the appearance of pyrimidine photoadducts in living cells, is also used for the preparation of dimer-containing oligonucleotides for structural studies. While solid-phase synthesis approaches are available for the site specific incorporation of pyrimidine photoproducts in oligonucleotides,[245,246] the majority of NMR studies have used samples prepared using photodimerization reactions. This approach involves UVB or UVC light irradiation of a single-strand oligonucleotide containing a couple of adjacent thymine residues, followed by HPLC purification and subsequent annealing to the complementary strand. Since only the pair of targeted thymine residues can be present in the damage-containing strand, the structural information of pyrimidine photoproducts available to date relates mainly to dimers flanked by purine residues.

6.2. Conformation of DNA-Containing Pyrimidine Photo-cross-links

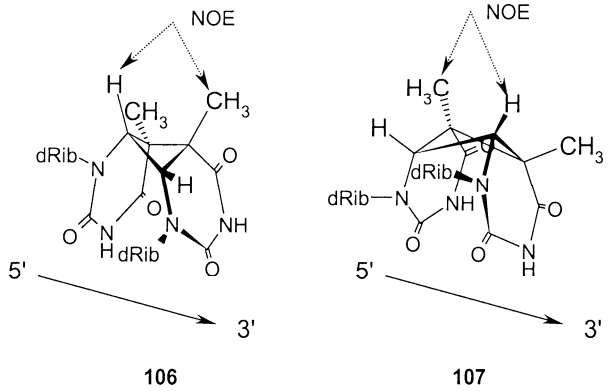
Early NMR and X-ray studies of dinucleotides containing pyrimidine dimers established that the relative orientation of damaged residues is different from that observed in unmodifed DNA. In addition, all the dimers have the main centers for hydrogen bond formation modified, a fact that should severely affect the stability of damaged duplexes. Nevertheless, the NMR characterization of duplexes containing pyrimidine dimers available to date indicates that the

structural changes caused by these lesions are confined to a four-nucleotide segment covering the photoadduct and flanking base pairs. These structural changes are discussed below.

6.2.1. Cis-syn Cyclobutane Pyrimidine Dimers

Since the cyclobutane ring of *cis-syn* CPD is not planar, dihydrothymine residues can adopt, in principle, a right-handed **106** or left-handed **107** twist (Chart 36). In addition,

Chart 36. Right-Handed 106 or Left-Handed 107 Twist of *cis-syn* CPD[a]



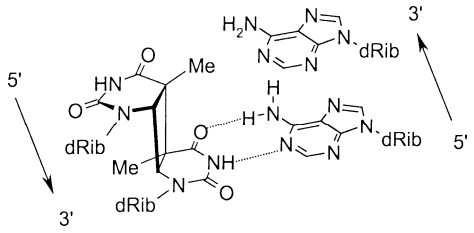
[a] Double arrows show the strongest NOE interaction in each case.

the cyclobutane ring pulls the pyrimidine bases together and, as a result, they cannot have the parallel orientation observed between nucleobases of unmodified DNA duplexes. Early X-ray and solution-state NMR studies established that the cyanoethyl ester of the *cis-syn* CPD dinucleotide adopts a left-handed conformation (**107**), displaying a −29° base twist, as opposed to the right-handed 36° value observed in B-form DNA.[247−249] Thus, the presence of this *cis-syn* CPD conformation in double-stranded DNA should disturb its structure considerably. Nevertheless, the NMR studies of lesion-containing duplexes[250−252] revealed that formation of a *cis-syn* CPD causes only small distortions in B-form DNA, an observation that is fully consistent with an early computer simulation prediction made by Kollman and co-workers.[253] NOE interactions between the photoadduct H6 and CH_3 methyl groups (Chart 36) established that, in duplex DNA, the *cis-syn* CPD changes the cyclobutane conformation from a left-handed twist **107**, observed in the isolated dimer, to a right-handed twist **106** form[251] (Figure 14). Similarly, the recent NMR characterization of a uracyl CPD-containing duplex also indicates a right-handed twist between the damage bases.[254] Transition to a right-handed twist partially alleviates the structural perturbations caused by CPD formation, and all imino protons of the duplex, including those at the lesion site, are hydrogen bonded.[255] However, NMR and UV melting studies suggest that the presence of these base pairs slightly reduces the thermal stability of duplex DNA.

6.2.2. Trans-syn Cyclobutane Pyrimidine Dimers

Until recently, a detailed three-dimensional model of DNA-containing *trans*-pyrimidine dimers was not available. Since at least one base is constrained in an unusual conformation, as described schematically in Chart 37, this lesion would be expected to induce a larger distortion in the DNA structure than the *cis-syn* CPD. NMR spectroscopy was used to characterize the conformation of a decamer duplex containing a thymine *trans-syn-I* CPD.[251] Analysis of chemical shift perturbations and the pattern observed in the NOESY

Chart 37. Base Pair Formed by *trans-syn-I* CPD 102



spectra indicated that distortions of the duplex structure are localized to the 5′-moiety of the *trans-syn-I* CPD and neighboring bases. The strong intensity of the base-anomeric proton NOE cross-peak suggests that the glycosidic bond of the lesion 5′-thymine is in a *syn* conformation, so that its WC edge is facing the major groove of the helix and only the 3′-component of the dimer **102** forms a hydrogen-bonded base pair with the opposing adenine residue (Chart 37). In agreement with this alignment, the T_m of the *trans-syn-I* CPD-containing decamer is 20° lower than that of a related duplex containing the *cis-syn* CPD. In contrast to the case of the *cis-syn* CPD, the duplex spectra of the *trans-syn-I* lesion showed no unusually shifted ^{31}P or imino proton signal, suggesting the absence of major distortions in the conformation of the sugar−phosphate backbone. However, since the imino proton signals of the duplex were not assigned, the hydrogen-bonding status of the lesion 3′-thymine could not be experimentally determined.[251]

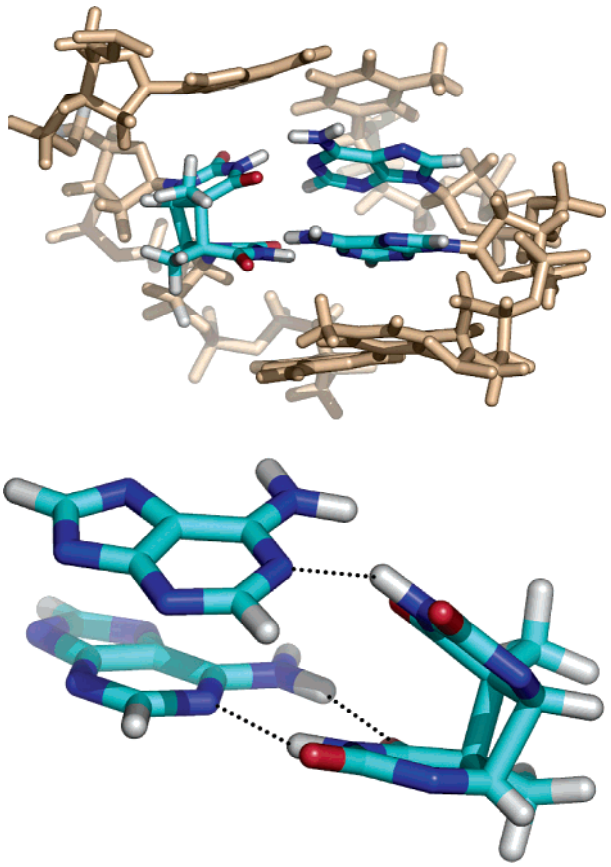


Figure 14. Structure of a *cis*−*syn* CPD **101** containing duplex. The picture displays the central four-nucleotide segment with the major groove prominent (top panel) and the **101**·d(ApA) base pair alignments as seen from the minor groove (bottom panel, only base atoms are shown.). The *cis*−*syn* CPD adopts a right-handed conformation and the base pair formed between the 5′-thymine and the complementary adenine is highly propeller twisted. Prepared from PDB entry 1TTD[265] using PyMol.[555]

6.2.3. (6-4) Pyrimidine–pyrimidinone and Dewar Adducts

The formation of 6-4PP **104** and 6-4Dewar **105** dimers causes dramatic changes in the hydrogen bond potential of the cross-linked bases. One WC hydrogen-bonding atom of the lesion 3′-pyrimidine, the O^4 of a 3′-thymine or the N^4 of a 3′-cytosine, is lost and the N^3 position becomes sterically hindered. The 5′-pyrimidine residue loses aromaticity and acquires an additional hydrophilic group. In addition, the two rings become almost perpendicular, making impossible their simultaneous base pairing and stacking inside the duplex. Despite their close topology, the geometries of 6-4PP and 6-4Dewar lesions are different. The rearrangement of the flat aromatic pyrimidinone ring to the extended and folded Dewar type bicyclic system changes the relative ring orientation within the lesion, affecting the duplex conformation at the damaged site. The NMR characterization of a 6-4PP dimer-containing duplex showed that the 5′-residue of the lesion remains essentially unperturbed.[256] Despite being considerably propeller twisted, the 5,6-dihydro-5-hydroxythymine forms a hydrogen bond with its partner adenine and stacks partially with an adjacent purine residue. No hydrogen bond is present between the 3′-component of 6-4PP and the opposing adenine, and according to the RMD model, the 5-OH group of the 5,6-dihydro-5-hydroxythymine is hydrogen bonded to the N3 of this base.

The conformation of a DNA duplex containing the 6-4Dewar adduct **105** is dramatically different.[257] In contrast to the 6-4PP, the most perturbed part of the 6-4Dewar lesion is the 5,6-dihydro-5-hydroxythymine base, which does not show NOE cross-peaks with its 5′-flanking dA residue. Instead of these interactions, the Dewar bicycle displays NOE peaks with its 3′-flanking base. There are no hydrogen bonds between 5,6-dihydro-5-hydroxythymine and its partner adenine residue, which, according to an RMD structure, shifts toward the Dewar moiety of the photodimer, supposedly forming a single hydrogen bond between its amino group and the Dewar bicycle N3. Despite the absence of experimentally observed hydrogen bonds between 6-4Dewar bases and opposing adenine residues, this lesion produces smaller distortions in the global DNA conformation than the 6-4PP, causing only a helical bend of about 21°, as opposed to the 44° value produced by the latter.[258]

6.3. Base Pairing of Pyrimidine Photoproducts

While the three major pyrimidine dimers, *cis-syn* CPD, 6-4PP, and 6-4Dewar, cause mutations in bacteria, *cis-syn* CPD is the least mutagenic of them, inducing about 5% of TT → TA and 1% of TT → TC base substitution mutations. A thermodynamics study showed that the presence of a *cis-syn* CPD adduct at the center of a dodecamer duplex reduces its stability by about 1.5 kcal/mol and incorporation of a T residue complementary to the lesion 3′-pyrimidine causes a slight additional stability penalty of just 0.7 kcal/mol,[259] suggesting that the *cis-syn* CPD participates in interstrand hydrogen bonding in both cases. In agreement with these data, analysis of NOESY spectra reveals that the thymine imino protons of the *cis-syn* CPD are hydrogen bonded, although they resonate somewhat upfield from the usual chemical shift range (Figure 14). While it was initially proposed that this upfield shift reflected the presence of weak hydrogen bonding, the altered diamagnetism of 5,6-saturated thymine rings and stacking interactions with flanking bases can explain these shifts satisfactorily. The NMR structures of mutagenic duplex intermediates containing dG or T paired to the 3′-thymine of the *cis-syn* CPD closely resemble those of the matched duplexes without any further distortion.[260,261] The presence of two well-resolved imino proton signals that show a strong NOE cross-peak between them indicates the formation of a classical dG·T* wobble alignment between the dG residue and the lesion 3′-thymine, which is displaced toward a minor groove.[260] Similarly, a T·T* wobble alignment stabilized by two hydrogen bonds is present when thymine is opposing the lesion 3′-thymine, which in this case is displaced toward the major groove of the duplex.[261] In contrast, a *cis-syn* CPD-containing duplex having to two dG residues paired to the damaged thymines adopts an extremely distorted conformation with highly propeller-twisted dG·T* wobble pairs and a sharp kink at the 5′-site of the dimer.[260]

Contrary to the *cis-syn* CPD, 6-4PP lesions are highly mutagenic in bacteria, inducing incorporation of dG opposite the lesion 3′-pyrimidone ring.[262] A thermodynamic study showed that a centrally located 6-4PP paired to d(ApA) causes an about 6 kcal/mol stability reduction of a dodecamer duplex,[259] a value close to that observed for duplexes having a four-nucleotide internal loop or a single abasic site.[54] 2′-Deoxyguanosine substitution of the dA residue opposite the lesion pyrimidone ring increases the thermodynamic stability of the 6-4PP-containing dodecamer by 0.5 kcal/mol and that of a 6-4Dewar lesion-containing duplex by 0.3 kcal/mol.[259] Interestingly, the latter lesion causes much lower mutation levels,[262] comparable to those predicted by the A-rule for noncoding damages. Therefore, the mutagenic behavior of 6-4PP and 6-4Dewar photoproducts cannot be explained in terms of free energy differences that the presence of these lesions causes in duplex DNA.

The NMR characterization of DNA duplexes containing a 6-4PP or the 6-4 Dewar photoproduct and adenine or guanine paired to the 3′-side of the lesions gave some rationale for these observations.[256−258,263,264] The experimental data show that an adenine residue positioned opposite the 3′-part of 6-4PP or 6-4 Dewar dimers is unable to form hydrogen bonds, whereas its replacement by guanine leads to the appearance of hydrogen bonds, causing significant alteration of the global DNA shape. However, the consequences of it are opposite for 6-4PP or 6-4 Dewar adducts. In the 6-4PP·d(GpA)-containing duplex, the pyrimidinone O2 group of the lesion forms hydrogen bonds with the imino and amino groups of its partner guanine and the WC edge of the dimer 5-hydroxypyrimidine moiety is hydrogen bonded to the opposing adenine base. As a result, when 6-4PP pairs to d(GpA), distortions of the DNA backbone and global duplex shape are smaller than those when d(ApA) is opposing the lesion[257] (Figure 15). In the (6-4)Dewar·d(GpA) duplex case, the lesion partner guanine also forms two hydrogen bonds; however, in this case, one of them is formed between its amino group and the Dewar bicycle O2 groups, whereas the second hydrogen bond involves the guanine imino proton and the 5-hydroxypyrimidine 2-oxo group. As a result, the latter moiety becomes almost orthogonal to its opposing adenine, causing a sharp kink in the helix and significant local distortions[264] (Figure 15), which are greater than those observed for a duplex having d(ApA) opposed to the lesion. These effects on global duplex conformation would oppose the increase of stability produced by hydrogen bonding, explaining in part the low frequency of T → C transitions observed for (6-4)Dewar adducts. In addition, the geometry of the 3′-Dewar bicycle·dA alignment

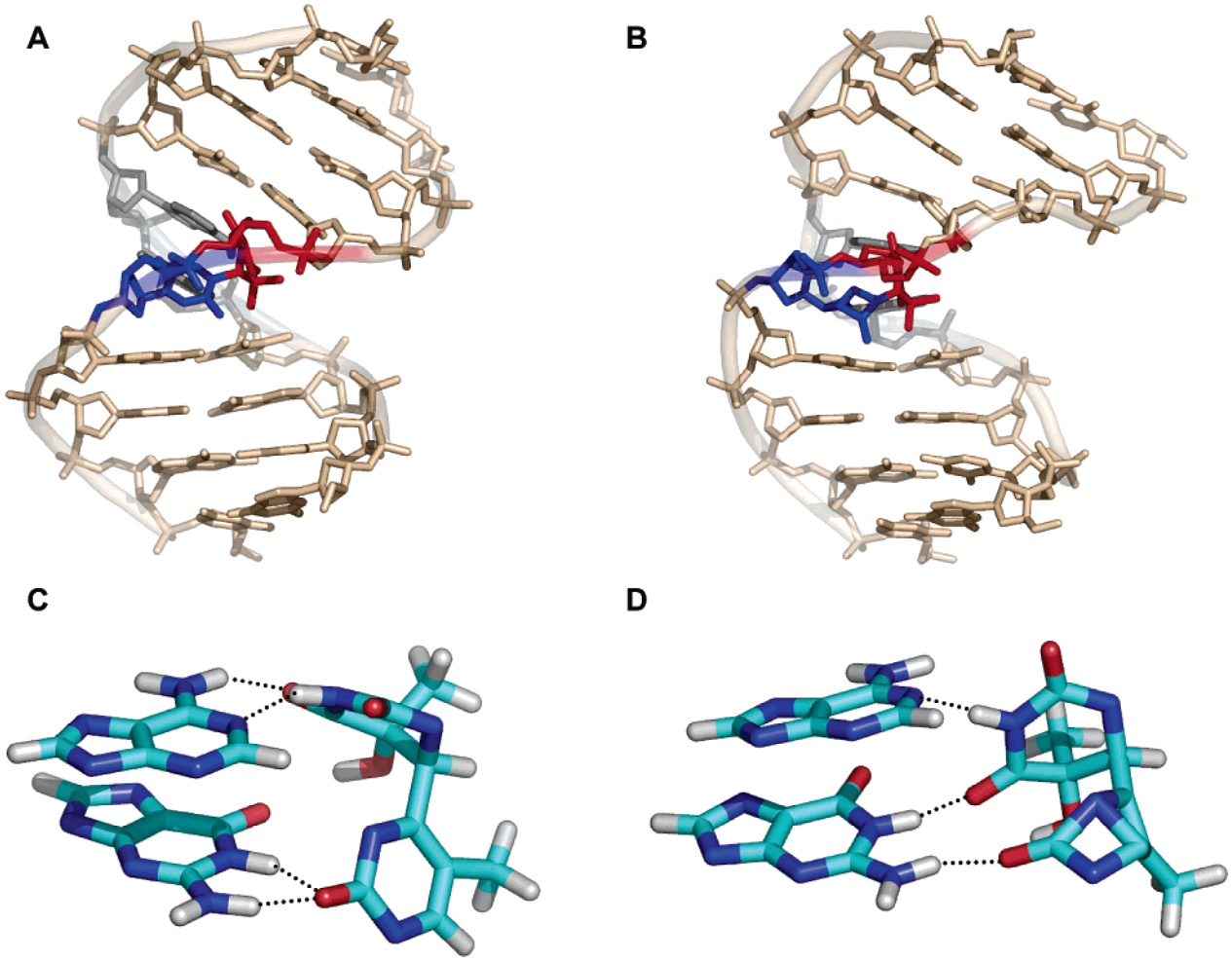


Figure 15. Global conformations of duplexes containing 6-4PP (**104**) or 6-4Dewar (**105**) dimers paired to a d(ApG) segment (panels A and B, hydrogen atoms not shown). Lesion site conformation of these duplexes (panels C and D, only base atoms are shown). Putative hydrogen bonds are shown with gray lines. Prepared from PDB entries 1CFL[263] and 1QKG[264] using PyMol.[555]

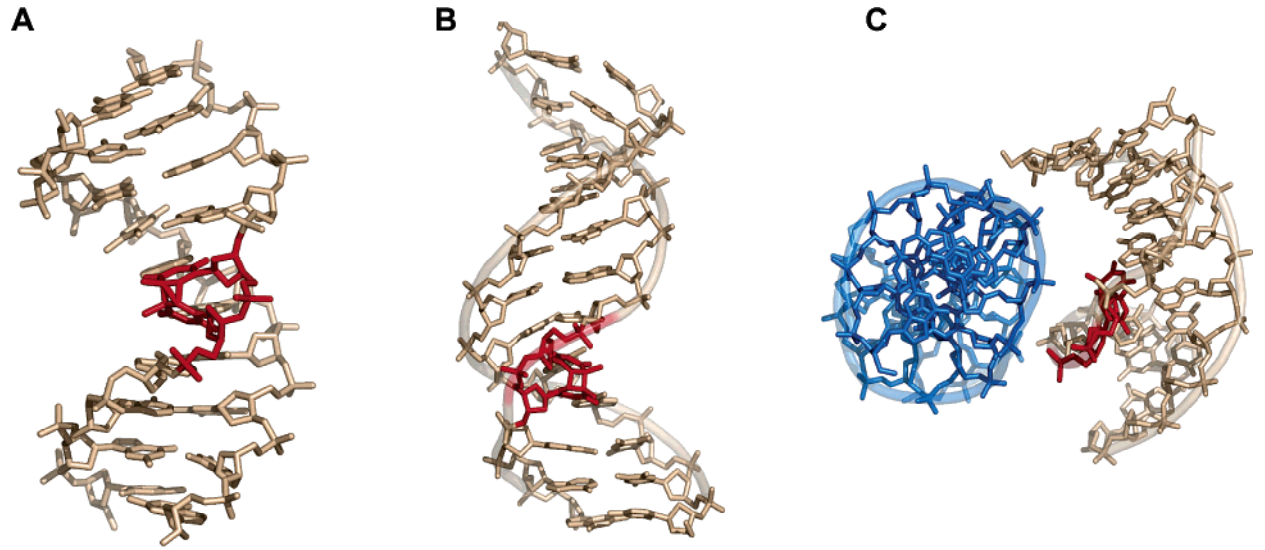


Figure 16. X-ray (A) and NMR (B) structures of duplexes containing a CPD lesion (shown in red). Panel C depicts packing interactions in the crystal structure with one duplex embedded in the major groove of another. Hydrogen atoms are not shown. Prepared from PDB entries 1N4E[271] and 1TTD[265] using PyMol.[555]

is highly regular, favoring the incorporation of a properly stacked dA residue, despite the lack of hydrogen bonding the base pair.

Table 2 summarizes the structural properties of pyrimidine photodimers.

6.4. *Cis-syn* CPDs and Changes in Global DNA Conformation

There are conflicting reports regarding the extent of *cis-syn* CPD-induced changes in the global DNA conformation when d(ApA), d(GpA), or d(TpA) residues are opposing the

Table 2. Structural Properties of Pyrimidine Photoproducts

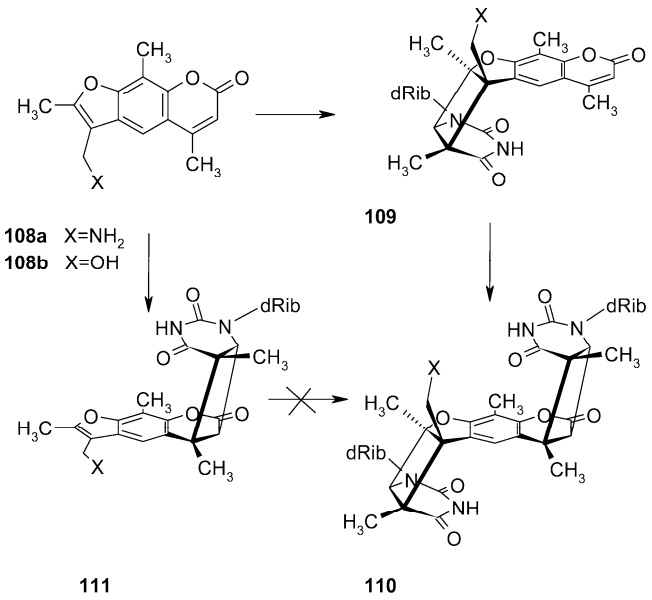
Lesion Type and Sequence[a]	Lesion Site Properties[b]		Helix Shape[c]			
	5'-side	3'-side	Bend (°)	Unwind (°)	PDB	Ref.
CPD						
5'-(..AA **T^T** AA..) (..TT **A-A** TT..)-5'	WC (2) Bck, Prp,	WC (2)	7		1TTD	265
CPD						
5'-(..AA **T^T** AA..) (..TT **A-A** TT..)-5'	WC (2) Bck, Prp	WC (2)	30		1N4E[d]	271
CPD						
5'-(..CA **T^T** AC..) (..GT **A-A** TG..)-5'	WC (2) Prp	WC (2)	~9	15		252
CPD						
5'-(..CA **T^T** AC..) (..GT **A-G** TG..)-5'	WC (2) Prp	Wbl (2) Prp, Bck	~10		1PIB	260
CPD						
5'-(..CA **T^T** AC..) (..GT **G-G** TG..)-5'	Wbl (2) Prp	Wbl (2) Prp (24°) Bck (44°)	~44		1SNH	260
CPD						
5'-(..CA **T^T** AC..) (..GT **A-T** TG..)-5'	WC (2) Prp (45°)	Wbl (2) She, Stg	12		1QL5	261
PP						
5'-(..CA **T^T** AC..) (..GT **A-A** TG..)-5'	WC (2)	NR Prp	44	32		256
PP						
5'-(..CA **T^T** AC..) (..GT **A-G** TG..)-5'	WC (2)	NR (2) Prp	27	2	1CFL	263
DW						
5'-(..CA **T^T** AC..) (..GT **A-A** TG..)-5'	NR (1) Prp	NR	21	16		257 258
DW						
5'-(..CA **T^T** AC..) (..GT **A-G** TG..)-5'	NR (1) Prp (57°)	NR (2)	43	39	1QKG	264

Lesion type indicated by ^. [a]*CPD, cis-syn* cyclobutane pyrimidine dimer; *PP,* (6-4) pyrimidino-pyrimidinone adduct; *DW,* (6-4) Dewar adduct. [b]WC, Watson-Crick; Wb, wobble; NR, non-regular. Between parentheses is the number of hydrogen bonds. [c]Bck, buckle; Prp; propeller twist; Stg, stagger; She, shear. [c]All duplexes were bent toward a major groove. [d]X-ray structure.

lesion. On one hand, the NMR solution studies show minimal distortions in the global shape of DNA[252,265] (Figure 16b), including a change in the backbone phosphate conformation at the 3′-side of the lesion that reduces the interstrand phosphate distance across the minor groove and causes a small helix bend toward the major groove and a slight unwind of the duplex at the lesion site (Table 2). These observations are fully consistent with gel electrophoresis results of duplexes having phased *cis-syn* CPD lesions[266] and several free MD structures.[267,268] On the other hand, a number of experimental and theoretical studies, including free MD simulations,[269] ligase circularization of *cis-syn* CPDs containing linear DNA,[270] and the X-ray structure of a lesion-containing dodecamer[271] (Figure 16a), predict up to a 40° bend of the helix toward the major groove of the duplex after adduct formation. Importantly, it was proposed that the bend induced by the *cis-syn* CPD is responsible for damage recognition by photolyase and DNA endonucleaseV repair enzymes.[271]

In the case of *cis-syn* CPD lesions, bend differences between X-ray and NMR structures cannot be simply ascribed to sequence specific structural variations, because duplexes used in these studies have almost identical sequences (Table 2). Two factors are usually considered to cause global shape discrepancies between X-ray and NMR DNA structures, namely, the effect of packing forces in the crystal and the inherent limitation of short-range NOE interactions to predict accurately duplex curvature. Analysis of packing interactions in the CPD-containing dodecamer reveals that one duplex molecule is embedded in the major groove of another (Figure 16c), suggesting that the bend observed in the structure could result from crystal packing forces. Similar conclusions were reached also in the case of DNA molecules containing A-tracts where crystal packing influences the extent of bending.[272−274] However, it was later shown that packing forces, being generally weak, are not able to bend a canonical B-DNA helix[275] and can only shift, instead, an already existent conformational equilibrium in duplexes that have increased flexibility. Furthermore, using an essential dynamics approach,[276] it has been found that thermal fluctuations in a *cis-syn* CPD-containing duplex have a significantly larger global bend component than those in the parent duplex and that the d(ApA) segment opposite the lesion contributes significantly to these motions, serving as a hinge.[277] Therefore, assuming that the *cis-syn* CPD lesion decreases the local DNA rigidity without affecting its average conformation in solution, it is conceivable that the bend observed in the crystal structure represents one conformational state of the duplex in solution. Accordingly, the inconsistencies between the X-ray and DNA ligation results, on one side, and the NMR and PAGE data, on the other, would disappear. Further experimental confirmation of these findings would suggest that not only the shape but also the flexibility of the lesion-containing DNA molecule could serve as a distinctive feature for *cis-syn* CPD recognition.

Scheme 10. Formation of Psoralene Mono 109 and Cross-linked 110 Adducts



6.5. Psoralen Adducts and Interstrand Crosslinks

Formation of a [2+2] cycloaddition product in DNA is not restricted to two pyrimidine moieties only. Psoralenes **108**, a family of furocoumarine intercalators with specific affinity for d(TpA)$_2$ steps, can also produce *syn-cis* cyclobutane adducts with thymine upon irradiation of the intercalated complex with UVA light.[278] Since both the furane and pyrone rings of psoralene **108** participate in the cycloaddition reaction independently, three photoproducts are possible, namely the furane monoadduct **109** (MAf), the pyrone monoadduct **111 (**Mpy), and the interstrand cross-link **110** (XL) (Scheme 10). Initially, the furane cyclobutane lesion **109** is the predominant product, but with further UVA irradiation of DNA, a second cycloaddition reaction occurs between the psoralene pyrone ring and the thymine residue on the opposite strand, generating the cross-link product **110**. Since the psoralene furane ring loses aromaticity after formation of the cyclobutane monoadduct **109**, UV irradiation of DNA with 50 nm shorter wavelength light is optimal for generating the XL product. As a result, formation of psoralene cross-links can be stopped after the first cycloaddition reaction, provided that 400 nm monochromatic light is used for irradiation. In contrast, when the psoralene pyrone ring reacts first and forms the cyclobutane lesion **111**, the remaining benzofurane moiety no longer absorbs UVA light and interstrand cross-links are not formed upon further irradiation.[279] Psoralen monoadducts and XL adducts and related compounds are cytotoxic for proliferating cells, explaining the therapeutic use of these drugs for local treatment of psoriasis.

NMR spectroscopy was used to characterize the structure of self-complementary octamer duplexes having a 4′-aminomethyl-4,5′,8-trimethylpsoralen **108a** (AMT)[280,281] or a 4′-hydroxymethyl-4,5′,8-trimethylpsoralen **108b** (HMT)[282,283] located at a central d(TpA) dinucleotide step. In the AMT case, only the XL adduct has been characterized, whereas RMD structures have been obtained for the XL and MAf lesions of HMT (PDB codes 203D and 204D, correspondingly).

MAfter formation of a thymine psoralen monoadduct, the cyclobutane moiety shares many structural features with the *syn-cis* CPD, including the sp^2 to sp^3 transition on the hybridization of cross-linked atoms. As a result, a parallel orientation of the psoralen and thymine moieties is no longer possible, a fact that should cause greater structural perturbation of DNA than that seen with classical noncovalent intercalators. The NMR spectra of the MAf-containing duplex suggest that the psoralene residue stacks properly

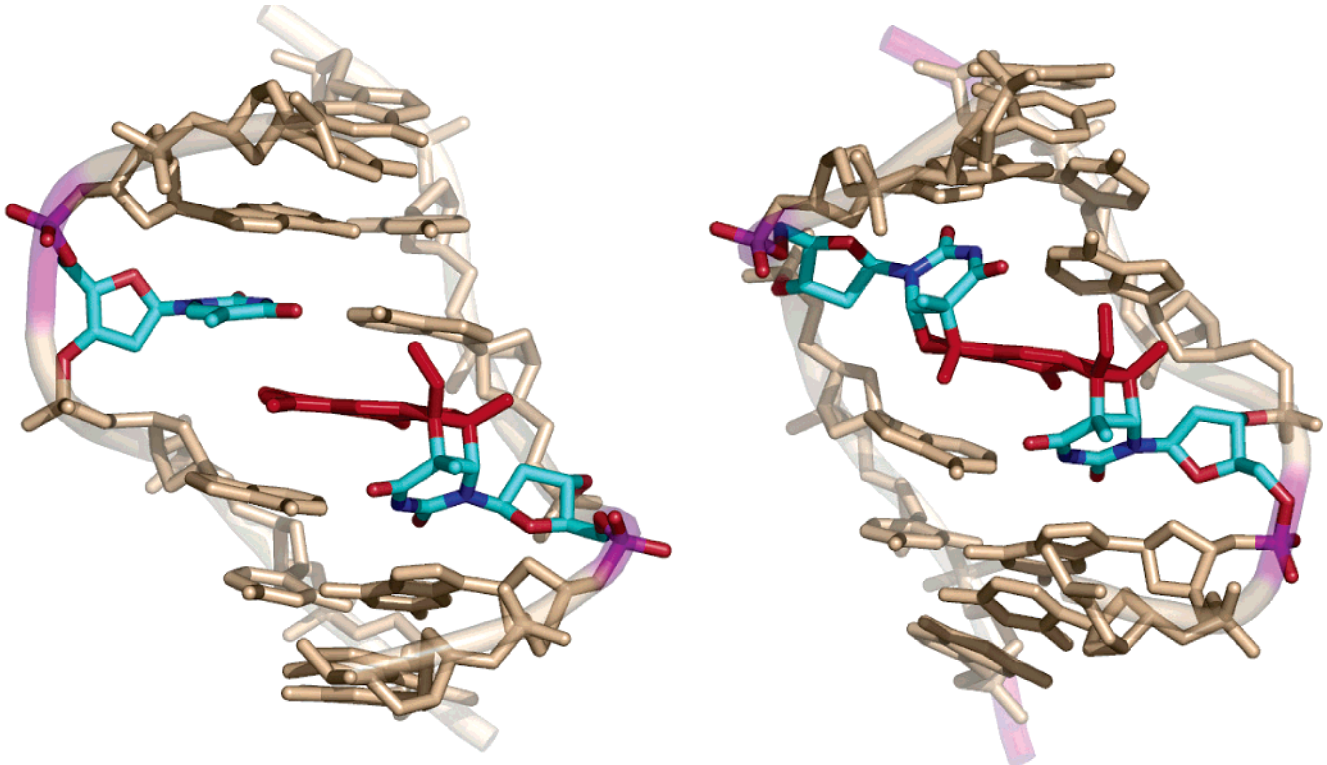


Figure 17. Structure of DNA duplexes containing a Maf **109** (left panel) or an XL **110** (right panel) psoralene adduct (colored red). Hydrogen atoms are not displayed. Prepared from PDB entries 203D[283] and 204D[283] using PyMol.[555]

inside the helix, adopting the usual parallel intercalating mode, and causes almost no perturbation to the 3′-flanking dA·T pair. In contrast, the dA·T base pair containing the adducted thymine is significantly buckled, allowing the adenine base to maintain stacking interactions with both psoralen and a flanking cytosine residue (Figure 17). Although the buckled dA·T alignment does not favor proper hydrogen bonding, analysis of the exchangeable proton spectra indicates the adducted thymine H3 proton has a slow solvent exchange rate suggesting the formation of a strong hydrogen bond. Even at 50 °C, when the duplex melting process is quite advanced, this imino proton exhibits the sharpest signal in the spectrum.[283]

When the MAf monoadduct forms the second cyclobutane ring originating the XL product, the conformation at the furane side of the lesion undergoes minimal additional changes. However, the DNA structure near the pyrone cyclobutane ring of the psoralen XL is dramatically perturbed (Figure 17). A large buckle of the dA·T pair makes the plane of the bases almost orthogonal, affecting the following two base pairs of the duplex. As a result, the imino proton signals of these base pairs start widening at 20 °C and disappear completely from the spectrum around 40−50 °C. In contrast, the imino proton signal of the furan-side dA·T pair is clearly visible at these temperatures.[283] These observations are consistent with early results demonstrating higher thermal and thermodynamic stability in duplexes containing a MAf monoadduct than in those with an Mpy lesion.[284]

The NMR structures of MAf and XL adducts of HMT in DNA show that these lesions cause 34° and 25° unwinding of the helix, respectively, agreeing with an early circular-plasmid titration study that established values similar to the unwinding produced by ethidium bromide.[285] Although a combined NMR and computational study predicted that the XL adduct of AMT would produce an up to 56° kink on the duplex,[280] a more recent NMR study demonstrated that the MAf monoadduct causes only an 8° kink in the duplex and the XL lesion does not bend the DNA at all.[282]

7. Cross-links Induced by Platinum Compounds

7.1. Formation of Adducts and Their Structure

By the end of the sixties, it was found that *E. coli* could not divide under the effects of an electric field. Subsequent studies attributed this phenomenon to the presence of *cis*-diamminedichloroplatinum **112** (Chart 38), a well-known inorganic compound that appeared in the media as a product of the reaction between the platinum electrodes and ammonium chloride present in the buffer. Commercially known as cisplatin, **112** is cytotoxic, possesses a pronounced antitumor activity in mammals, and, along with some of its close analogues, is one of the most effective and widely used anticancer agents (for a review, see ref 286). While many additional platinum compounds have been investigated to improve further the anticancer properties of cisplatin, the more promising candidates, such as carboplatin **113** or oxaliplatin **114**, still retain the essential structural features originally identified in **112**, namely a square platinum(II) complex with two ammine ligands oriented *cis* and two leaving groups. In contrast, platinum(II) complexes having leaving groups in a *trans* orientation, such as transplatin **115**, possess no or very little biological activity. The primary target of cisplatin and related compounds is cellular DNA, where, after the nucleophilic hydroxyl substitution of its chloride anions, they bind to the N7 atoms of two nearby purines, giving rise to the formation of cross-linked adducts (Chart 39).

Chart 38. Relevant Diammino Platinum(II) Compounds

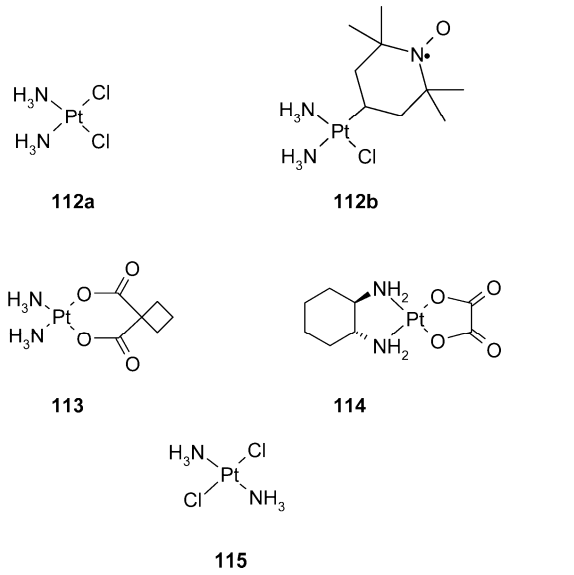
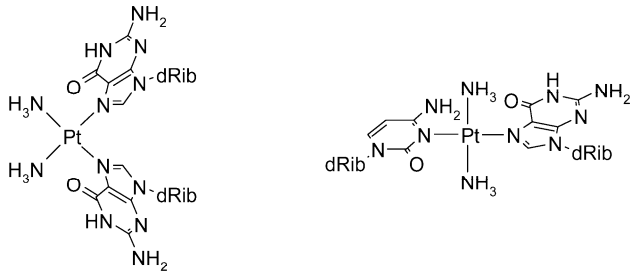


Chart 39. Prominent Cisplatin 116 and Transplatin 117 Cross-link Adducts



Cisplatin reacts preferentially with adjacent purine bases, and the *cis*-Pt-d(GpG) intrastrand cross-links account for about 60% of the adducts detected in DNA. Other important lesions are about 25% of Pt-d(ApG) cross-links and less than 8% of a cross-linked product between nonadjacent guanine bases, Pt-d(GpXpG), and an interstrand *cis*-Pt(GG) adduct.[287,288] Formation of *cis*-Pt lesions has multiple consequences for the normal cell function. Pt adducts partially block DNA replication and induce several types of mutations. Cisplatin inhibits transcription and telomere growth, and it increases the affinity of some transcription factors and high-mobility-group (HMG) proteins for DNA.[286] It is believed that the therapeutic effects of cisplatin result from the combination of all these properties.

Despite being biologically inactivity, transplatin and related compounds also form DNA adducts, although, due to steric reasons, cross-links between adjacent purine bases cannot form. In double-stranded DNA, transplatin forms adducts across dG·dC pairs, giving rise to the formation of ICL **117**,[289] whereas, in single strand DNA samples, cross-linked adducts occur between bases separated by at least one nucleotide.[290]

7.2. Geometry of Major Pt Lesions

Since platinum(II) complexes have a square geometry, the N^7-Pt-N^7 angle of the adduct should be 90°, unless the cross-linked purine bases imposed severe steric constraints that deform its geometry. Early X-ray diffraction studies confirmed that in di- and trinucleotides having *cis*-Pt lesions this angle is close to 90°.[291] In these structures, guanosine

Table 3. Structural Properties of *cis*-Pt-N^7,N^7-d(GpG) Intrastrand Cross-links

Duplex Sequence	Lesion Properties					Helix Shape			
	N^7-Pt-N^7 angle	Pt out of plane[a] (5'-G*/ 3'-G*)	Guanine planes dihedral	N-type sugar	WC Pairing	Bend	Unwind	PDB	Ref
	(°)	(Å)	(°)			(°)	(°)		
CCTCAGGCCTCC[b] GGAGTCCGGAGG	89.8	1.32 / 0.35	35	5'	3', 5'[d]	31		1PGC	294
CCTCTGGTCTCC[b,c] GGAGACCAGAGG	98.4	1.3 / 0.8	25	5', 3'		30		1IHH	295
CTCCGGCCT GAGGCCGGA	88.7	0.89 / 0.03		5'	3', 5'[e]	49.9		1KSB	296
CCTGGTCC GGACCAGG	87.5	0.85 / 0.56	42	5'	3'	58	21	1AU5	297
CCTCTGGTCTCC GGAGACCAGAGG	90	0.75 / 0.83	49	5'	3'	78		1A84	298
CCTCTGGTCTCC[c] GGAGACCAGAGG		1.2 / 0.8	26	5'		40			299
ATACATGGTACATA TATGTACCATGTAT	69	0.88 / 0.3	44	5'	3', 5'[c]	52	26		300
CTCTCGGTCTC GAGAGCCAGAG		0.5 / 0.6	50	5'	3'	80	25		301
CTCAGCCTC GAGTCGGAG					3'	55			302

[a]Displacement of the Pt atom out of cross-linked guanines plane. [b]Oxaliplatin complex. [c]X-ray data. [d]Observed at low temperature. [e]Very weak signal due to fast water exchange.

residues adopt an *anti* conformation around their χ angle and are oriented "head-to-head", with the WC edge of the bases pointing toward the same side. The N^7-Pt-N^7 angle is in the 76−87° range, making stacking of the cross-linked guanine bases impossible. Such a geometry is hardly consistent with a B-type DNA conformation, suggesting that the formation of a Pt cross-link in double-stranded duplexes must result in a major perturbation of the DNA structure or, alternatively, in a change in the adduct geometry. Early NMR studies established that the ^{195}Pt chemical shift values observed for a complex of *cis*-Pt and two 2′-deoxyguanosine 5′-monophosphates, (−2455 ppm for the *cis*-Pt-d(GpG) cross-link in single-stranded oligonucleotides and −2450 ppm[292] in double stranded duplexes)[293] are all very close. Since ^{195}Pt chemical shifts are very sensitive to the complex geometry, these data imply that platinum cross-links in DNA share common features with low molecular weight cisplatin adducts.

All NMR studies of DNA containing *cis*-Pt-d(GpG) cross-links reported up to date[294−302] (Table 3) reveal that this is mostly the case. Analysis of the NOESY spectra show that duplexes having a *cis*-Pt-d(GpG) adopt regular right-handed helical conformations, with residues in the *anti* conformation around the glycosidic bond and canonical WC alignments for all unmodified base pairs. At the lesion site, a strong NOE cross-peak between the H8 protons of cross-linked guanine bases is always present, establishing a "head-to-head" orientation of the *cis*-Pt-d(GpG) adduct in the duplex. In addition, the observation of strong intraresidue base−H3′ NOE cross-peaks, which are weak for undamaged duplexes, indicates that the sugar of the lesion 5′-guanosine (and sometimes those of both damaged residues) adopts a 3′-endo conformation, typical of A-form DNA.

The RMD models of duplexes containing a *cis*-Pt-d(GpG) lesion show that the adduct mostly retains a flat square geometry, with all angles of the platinum(II) complex 90°, but the Pt atom is slightly displaced from the guanine planes, decreasing in this way the roll angle between the bases. The dihedral angle (Table 3) formed by the plane of the platinated guanines is smaller than that observed in the dinucleotide *cis*-Pt-d(GpG) complex, but nevertheless, the bases remained largely unstacked (Figure 18). Since platinated bases are no longer parallel, their ability to simultaneously form WC base pairs with partner cytosine residues is considerably impaired. The experimental data consistently demonstrate that the

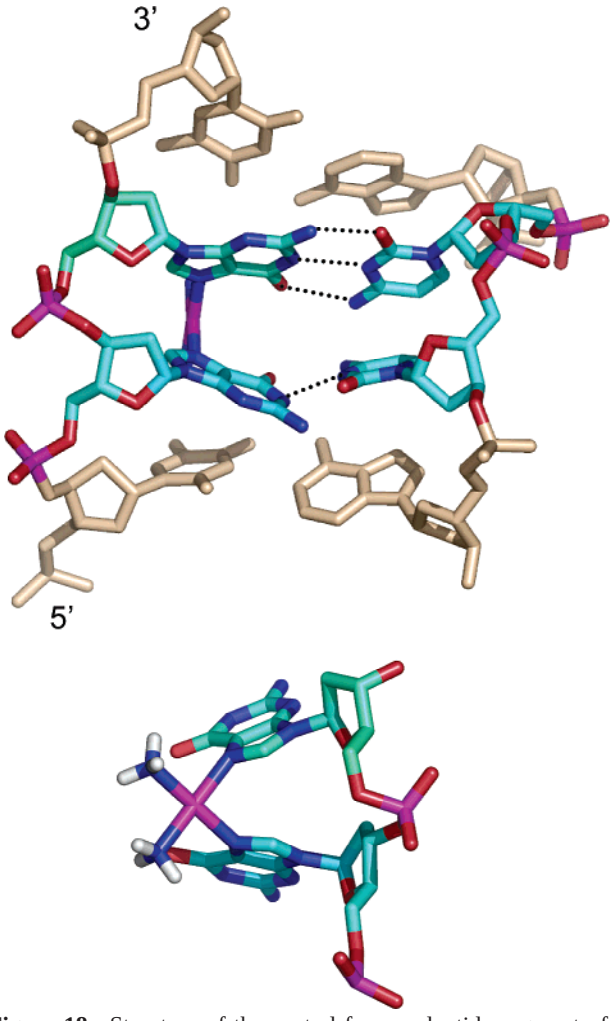


Figure 18. Structure of the central four-nucleotide segment of a *cis*-Pt-d(GpG)-containing duplex (top panel). Hydrogen atoms are not shown. Close-up view of the cross-linked d(GpG) fragment **116**(bottom panel). Prepared from PDB entry 1A84[298] using PyMol.[555]

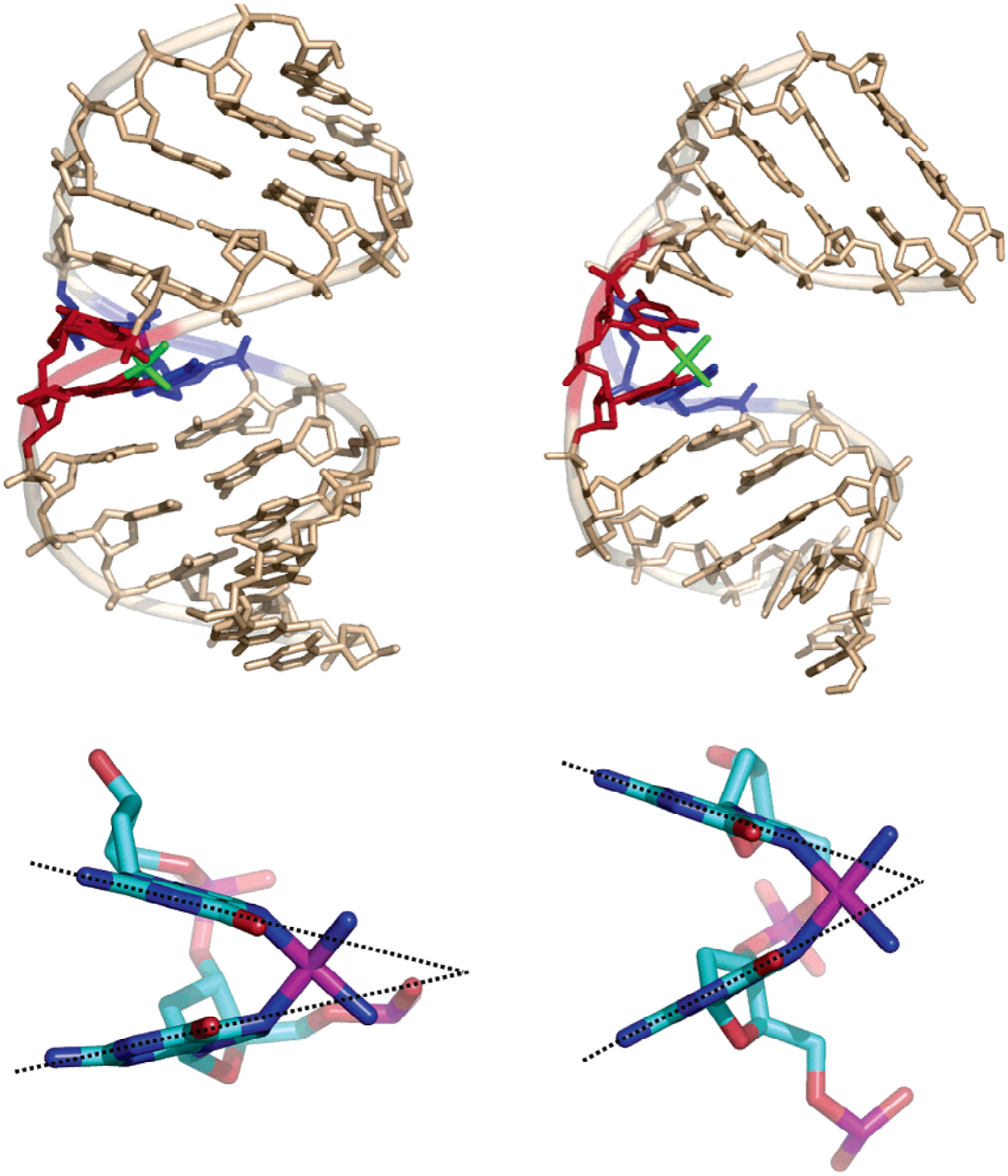


Figure 19. X-ray (top left panel) and NMR (top right panel) structures of duplexes containing a *cis*-Pt-d(GpG) lesion in identical sequence context. The pictures display the damaged d(GpG) fragment in red, the Pt cross-link in green, and the lesion partner cytosine residues in blue. Close-up view of the cross-linked d(GpG) segments in the X-ray (bottom left panel) and NMR (bottom right panel) structures, showing differences in the crystalline and solution states. Hydrogen atoms are not shown. Prepared from PDB entries 1AIO[554] and 1A84[298] using PyMol.[555]

lesion 3′-guanine base forms a stable WC base pair in duplex DNA. In contrast, evidence for participation of the 5′-platinated guanine in WC base pair alignments is limited and observed only at low temperatures.

7.3. *cis*-Pt–d(GpG) Cross-links and Global DNA Shape

Structures of DNA having cisplatin intrastrand cross-links available to date, obtained from NMR or X-ray data (Table 3), share essential common features, including the widening of the minor groove near the lesion site, transition of some nucleotides at or near the lesion from a 2′-endo to a 3′-endo conformation, and a considerable DNA bend toward the major groove, in the 30−80° range. The latter observation is consistent with a bend value of about 40° determined using PAGE mobility assays of DNA containing phased *cis*-Pt-d(GpG) adducts.[303] However, despite these similarities, some quantitative differences occur between structures obtained in solution and the crystal state. A comparison of NMR[294,298] and X-ray[295,299] structures obtained for duplexes having close or identical sequences has established that in solution the DNA is more bent and the square planar geometry of the platinated d(GpG) segment is less perturbed. In the crystal the situation is the reverse (Figure 19). It is worth noting that a partial transition to an A-form DNA is present in the X-ray structure of duplexes containing a *cis*-Pt-d(GpG) lesion, which probably results from the combined effect of adduct-induced structural perturbations and crystal packing forces.

To better establish the extent of DNA bend induced by *cis*-Pt-d(GpG) in solution, NMR spectroscopy was used to characterize the structure of an 11-mer duplex having a cisplatin analogue bearing a paramagnetic TEMPO moiety **112b**. The presence of the paramagnetic group causes a decrease of cross-peak intensity in the NMR spectrum, which follows a $1/r^6$ distance dependency. The large gyromagnetic ratio of the electron allowed the determination of distance bounds up to 20 Å apart.[301,304] The RMD computations using a combined set of NOE and paramagnetic distances produced a duplex conformation very close to that obtained using NOE distances alone. However, distances between the paramagnetic label and protons at duplex termini were considerably smaller than those computed for a straight or moderately bent DNA. The RMD structure revealed a pronounced 80° duplex bend, being the largest value ever obtained for platinated DNA by NMR spectroscopy. In addition, these results are the first direct measurement of long-range distance in the platinated duplex and suggest a highly kinked shape of cisplatin cross-linked DNA in solution.

A second question concerning the shape of *cis*-Pt-d(GpG) lesion-containing DNA is the need for separating cisplatin-induced effects from sequence specific variations of DNA structure. Preparation of platinated duplexes generally involves the introduction of *cis*-Pt-d(GpG) cross-links in single-stranded oligodeoxynucleotides, followed by their purification and subsequent annealing to the complementary strand. Thus, the *cis*-Pt-d(GpG) adduct is always present in a pyrimidine rich context and, in addition, most duplexes characterized up to date are A-T deficient. To compensate for these limitations, an NMR study has characterized the structure of a 14-mer duplex having a *cis*-Pt-d(GpG) cross-link in an adenine rich context.[300] The RMD structure showed that the duplex has essentially the same conformational features identified earlier, including a 52° bend toward the major groove that, in contrast to the cases of previous structures, was localized at the 3′-side of the lesion. Clearly, the diversity of bend values observed in *cis*-Pt-d(GpG)-containing duplexes suggests that formation of the intrastrand cross-link may actually increase DNA flexibility at the lesion site.

7.4. Platinum Interstrand Cross Links

While *cis*-Pt-N^7,N^7-d(GG) interstrand cross-links are minor lesions produced by cisplatin, they may have crucial biological effects because the repair of ICLs in the cells is less efficient. Two NMR studies have characterized the structure of DNA having a *cis*-Pt-d(GG) interstrand cross-link located at d(GpC) steps of the duplex.[305,306] Although in these studies the lesion is present in different sequence contexts, the structures of both duplexes are very similar and close to the X-ray structure published a couple of years later[307] (PDB code 1A2E). In B-form DNA, the distance between guanine N7 atoms at d(GpC) steps is about 7.5 Å; therefore, a significant structural rearrangement is necessary for the formation of a *cis*-Pt-d(GG) interstrand cross-link, which greatly affects the duplex conformation at the lesion site. Each damaged guanine intercalates in the opposite strand between the other platinated base and its 5′-neighboring residue. The damaged guanine bases stack over each other in face-to-tail fashion, explaining the observation of a weak NOE cross-peak between their H8 protons. The platinated residues are overturned form their normal position in B-form DNA in a way that places their N7 atoms facing the minor groove of the duplex, allowing the presence of a square *cis*-Pt-d(GG) complex geometry. The lesion partner cytosine residues are extruded from the helix into a position completely exposed to solvent, facilitating the inversion of the damaged bases at the lesion site. In addition, the helix is left-handed, displaying a net unwinding value of 87° for the three-step segment at the lesion site, and bent toward the minor groove of the duplex by about 40°.[305] (Figure 20).

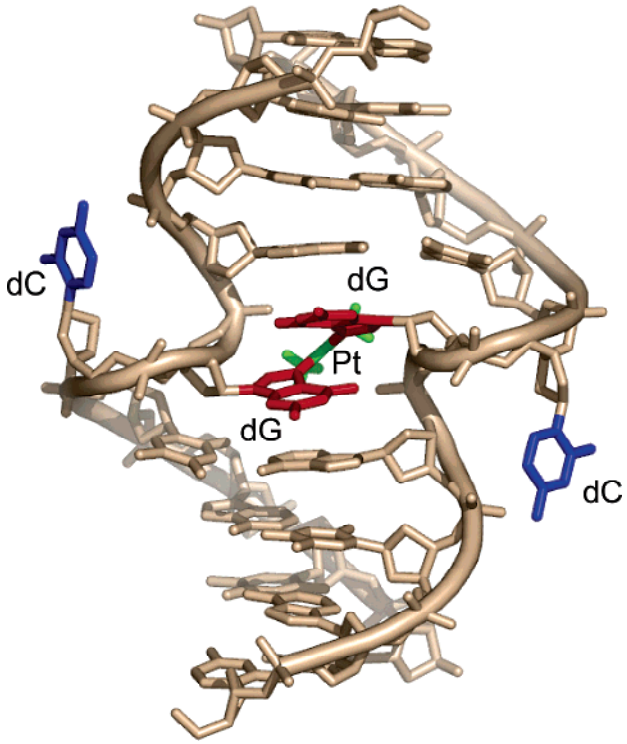
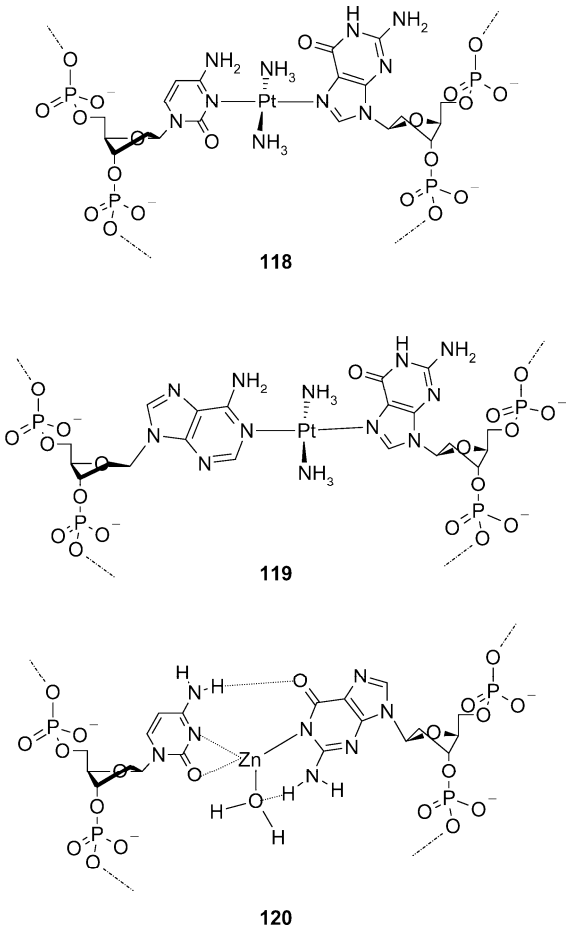


Figure 20. Structure of a DNA duplex containing a *cis*-Pt-d(GpG) interstrand cross-link. The picture shows the d(GpG) fragment in red, the Pt cross-link in green, and the complementary cytosine residues in blue, adopting an extrahelical conformation. Hydrogen atoms are not shown. Prepared from PDB entry 1DDP[305] using PyMol.[555]

In contrast to *cis*-Pt cross-links, *trans*-Pt forms ICLs across dG·dC base pairs **117**, producing smaller distortions of the double helix. The NMR characterization of a dodecamer duplex having a transplatin ICL shows that the major structural perturbation is the *anti* to *syn* transition of the cross-linked guanine **118** (Chart 40).[308] The geometry of the cross-linked fragment closely resembles that of a Hoogsten base pair but with an increased distance separating the platinated dG and dC bases, which may explain the fact that while lesion-flanking base pairs adopt proper WC alignments, their imino protons have increased solvent exchange rates. The energy-minimized model shows that these pairs remain stacked inside the helix, but to avoid steric clashes with the ammine groups of the platinum complex, there is an increase of the helical rise at the lesion site. In addition, there are compensatory variations of base pair twists that result in an average value of 36° for the whole duplex.[308] A similar alignment was observed in an 12-/11-mer duplex containing a trans-Pt d(G·A) ICL.[309] The sequence used for this study also contained a dG·T mismatch one base pair apart from the *trans*-Pt ICL that was apparently inconsequential for the structure adopted by the platinated mismatch. The duplex structure retained a regulator B-form conformation with the *trans*-Pt d(G·A) ICL assuming a pseudo-Hoogsten alignment

Chart 40. Alignment of (dG-dC) 118 and (dG-dA) 119 *trans*-Pt ICL and Proposed M-DNA 120 Base Pair

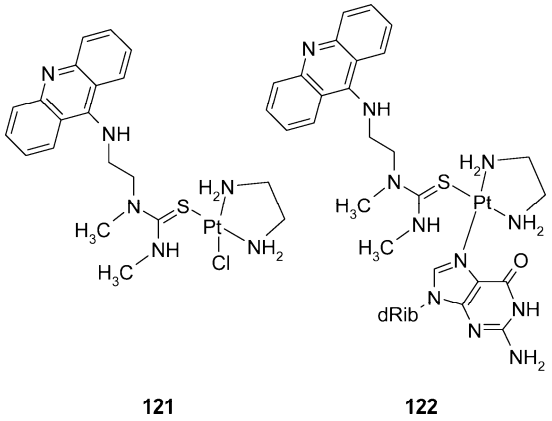


119 and the dG·T mismatch adopting a wobble geometry. It is worth noting that these observations are in contrast to the base pair alignments in M-DNA **120**, which also contains a metal ion per base pair but chelated between the two WC edges of the complementary amino acid.[310]

7.5. Other Platinum Adducts

Early studies showed a marked difference in the antitumor activity of mono- and bidentate platinum lesions, leading to the conclusion that only cross-linked adducts were biologically significant. However, the need to reduce cisplatin nephrotoxicity and widen its anticancer spectrum resulted in the screening of thousands of platinum(II) and platinum(IV) chelates and, among them, a number of hybrid compounds containing intercalating moieties (for a review, see ref 311). One of these compounds with promising antigrowth activity in many tumor cell lines, PT-ACRAMTU (**121**), contains an acridine moiety tethered to a *cis*-Pt complex with a single leaving group. This compound forms DNA adducts **122** (Chart 41) having a single bond with guanine or adenine N7 and intercalates its acridine moiety at the 5′-side of the platinated residue, resulting in very tight complexes. In addition, the high affinity of acridine for d(CpG), d(TpA), and d(GpA) steps drives the formation of PT-ACRAMTU adducts to those sites, expanding the affinity of cisplatin that is mostly restricted to d(GpG) steps.[312] A recent NMR study has characterized the solution structure (PDB code 1XRW) of an octamer duplex containing a PT-ACRAMTU adduct.[313] The NMR spectra indicate the presence of a right-handed

Chart 41. PT-ACRAMTU 121 and the PT-ACRAMTU-N^7-dG Monoadduct 122



helical conformation with residues in the *anti* range around the χ torsion angle and WC base pair alignments throughout the duplex, including at the damaged-containing dG·dC pair. At the lesion site, the platinum complex retains a square planar geometry with the acridine moiety directed toward the 5′-side of the damage strand, facilitating its intercalation inside the duplex (Figure 21). The helix is partially unwound

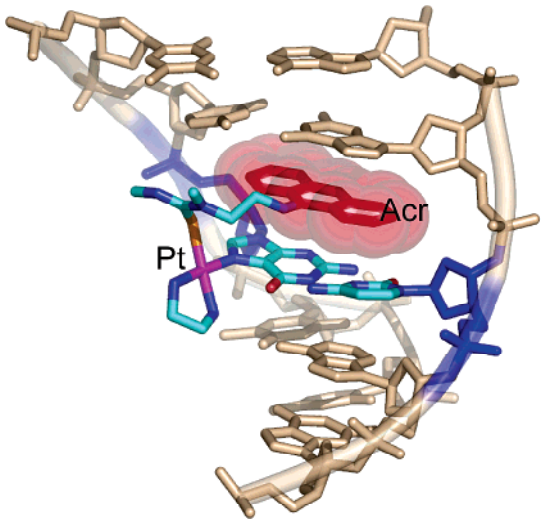


Figure 21. Structure of the central five-base-pair segment of a DNA duplex containing the Pt-ACRAMTU adduct **122**. The picture shows the intercalated acridine moiety in red and the lesion-containing base pair in blue. Hydrogen atoms are not shown. Prepared from PDB entry 1XRW[313] using PyMol.[555]

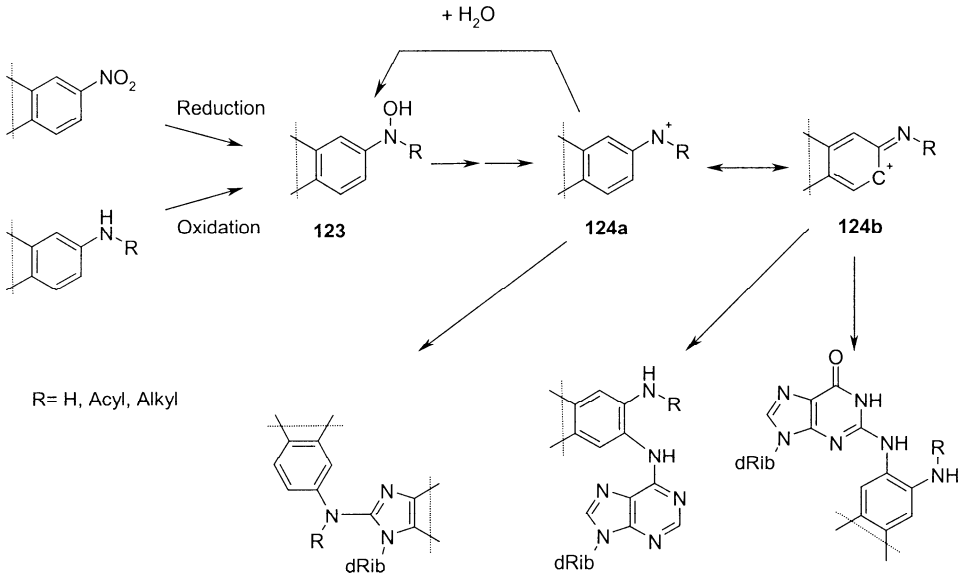
at the lesion site, but the pronounced bend that characterizes the structure of *cis*-Pt interstrand cross-links is completely absent. Interestingly, formation of the intercalative acridne platinum complex augments significantly the duplex thermal stability, increasing 13 °C its T_m value. On the basis of its intercalation manner and the effect on duplex stability, the PT-ACRAMTU lesion is structurally closer to the N^7-alkylation−intercalation compounds discussed in the next chapter than to *cis*-Pt cross-link adducts.

8. Bulky DNA Lesions

8.1. Endogenous Activation of Polyaromatic Compounds and Their Reaction with DNA

A large number of polycyclic aromatic hydrocarbons (PAHs), their nitro derivatives (NO_2-PAHs), biotoxins, and

Scheme 11. Endogenous Activation of Arylamines and Nitroaromatic Compounds

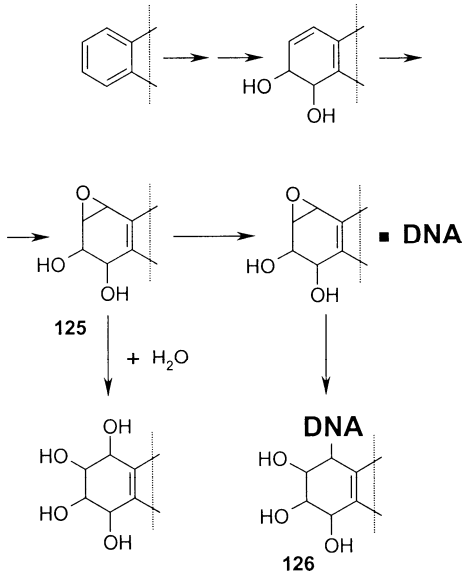


food mutagens cause the formation of bulky DNA lesions *in vivo*. Most of these lesions disrupt cellular replication and transcription of DNA, producing a high level of mutations or effectively halting DNA and RNA synthesis. The nucleotide excision repair (NER) pathway repairs bulky lesions. This pathway is able to recognize a large number of structurally diverse bulky adducts. NER involves multiple protein complexes and several steps, which include lesion recognition, excision of a 24−32-base segment surrounding the damage site, repair DNA synthesis, and ligase reaction. The last two steps of this process are relatively clear by now,[314] but the factors leading to initial damage recognition still remain to be completely understood. Indeed, understanding correlations between the structure of bulky lesions and their biological properties has been the subject of vigorous research for the last four decades. A comprehensive review specifically dedicated to the structure of DNA containing PAH lesions was published about eight years ago.[315]

The majority of bulky lesions result from the electrophilic addition of an activated polyaromatic compound predominantly to the amino or N^7 group of purine bases or to the C^8 position of guanine (for a review, see refs 316 and 317). Addition at cytidine N^4 adducts can also occur, albeit only small amounts of these lesions are formed.[318] Since polyaromatic compounds are essentially nonreactive *per se*, their cellular activation is a prerequisite for DNA damage generation (Schemes 11 and 12). In living organisms, a reductive pathway may transform nitroaromatic compounds to hydroxylamine derivatives **123**. Similarly, metabolic oxidation can convert food mutagens and other aromatic amines to corresponding hydroxylamines **123**. Upon further activation by enzymatic esterification, **123** forms cationic species **124**, which can react as nitrenium ions **124a**, leading to the appearance of purine C^8 adducts, or carbenium ions **124b**, generating aminoaryl derivatives of guanine or adenine (Scheme 11) (for a review, see refs 319 and 320).

Activation of unsaturated compounds, such as polycyclic aromatic hydrocarbons or aflatoxins, depends on their oxidation by the cytochrome P450 monooxygenase system (CYP) to the corresponding epoxides. Similar to the case of butadiene *in vivo* epoxidation (section 5.1.3), oxidation of the terminal ring of PAH and subsequent partial hydrolysis yields some amount of diol epoxide **125** (Scheme 12).[321]

Scheme 12. Metabolic Activation of PAHs and Their Addition to DNA

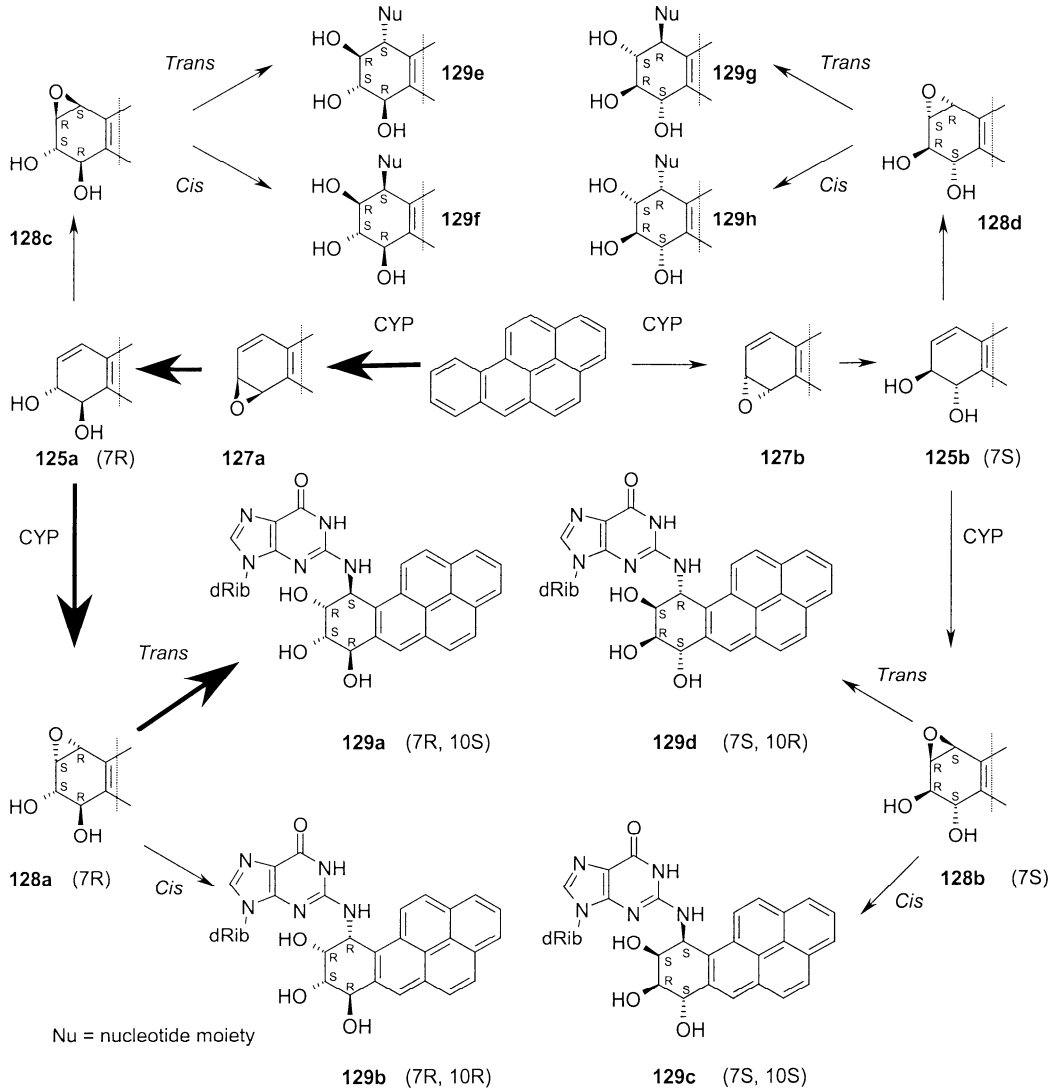


Most bulky moieties are either good helix intercalators or groove binders, and the formation of noncovalent complexes with DNA increases considerably the production of covalent DNA adducts, with the common structure **126**.[322]

The ability of PAH diol epoxides (DEs) **125** to form harmful DNA adducts is extremely dependent on the position of the epoxy group. Epoxidation of the terminal ring adjacent to a "bay" or "fjord" region (Chart 42) produces highly mutagenic and tumorogenic species.[323] In most cases, the base nitrogen reacts with the epoxide position adjacent to an aromatic PAH ring, to the aflatoxin alkoxy group, or to any other carbocation stabilizing group present in the molecule.

To present here the large number of lesion-containing duplex structures that have been solved, we will group them according to the type of damaged nucleotide and the position of the bulky group. Furthermore, in order to facilitate structural comparisons, we will discuss here the structure of other base modifications that, while less studied than PAH adducts, meet the broad criteria of bulky DNA lesions.

Scheme 13. Isomeric Benzo[*a*]pyrene Diol Epoxides and the N^2-Guanine Adducts They Form[a]



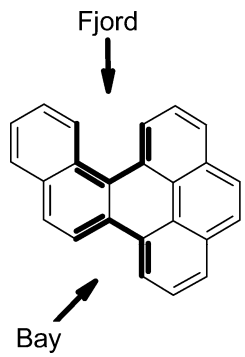
[a] Thick arrows show the major metabolic pathway.

8.2. N^2-Deoxyguanosine Lesions

8.2.1. Benzo[a]pyrene (B[a]P). Lesion Stereochemistry and Structural Organization.

DNA lesions formed by benzo[*a*]pyrene (B[*a*]P), one of the most potent and well-known PAH carcinogens, were among the first PAH lesions to be fully characterized by NMR spectroscopy and computational methods. These studies have particular relevance since general rules correlating the stereochemistry of BPDE adducts **126** and the damaged conformations they induce appear to apply to other bulky lesions.

Chart 42. "Bay" and "Fjord" Regions of PAH



The stereochemistry of B[*a*]P DE adduct formation, which also applies to the vast majority of other PAH DE adducts, is described in Scheme 13. Four chiral centers are present in DE adducts **126**; consequently, 16 diastereomers are theoretically possible. Due to the stereoselectivity of the CYP system, each epoxidation step produces about 95% of epoxides **127a** and **128a**, respectively, making DE **128a** the predominant metabolically activated form of B[*a*]P in the cell. In addition, **128a** is the most carcinogenic isomer of all four DEs **128a**−**d**. As for most epoxides, nucleotide *trans* addition is preferred for the *R*,*S*,*S*,*R*-**128a** and *S*,*R*,*R*,*S*-**128b** DEs, whereas opening of their enantiomers **128c**,**d** proceeds via *cis* addition (see ref 324 and references therein). The reaction of BPDE with DNA yields a complex mixture of products, including a major fraction of adducts that form at the N^2 position of guanine, as well as N^6-dA and N^7-dG lesions. The biological activity of each of these adducts and sequence dependent effects are only partially known.

The majority of NMR studies of PAH lesions have been done on adducts sharing the stereochemistry of lesions **129a**−**d**; and therefore, it is possible to describe their stereochemistry using only two independent parameters. In

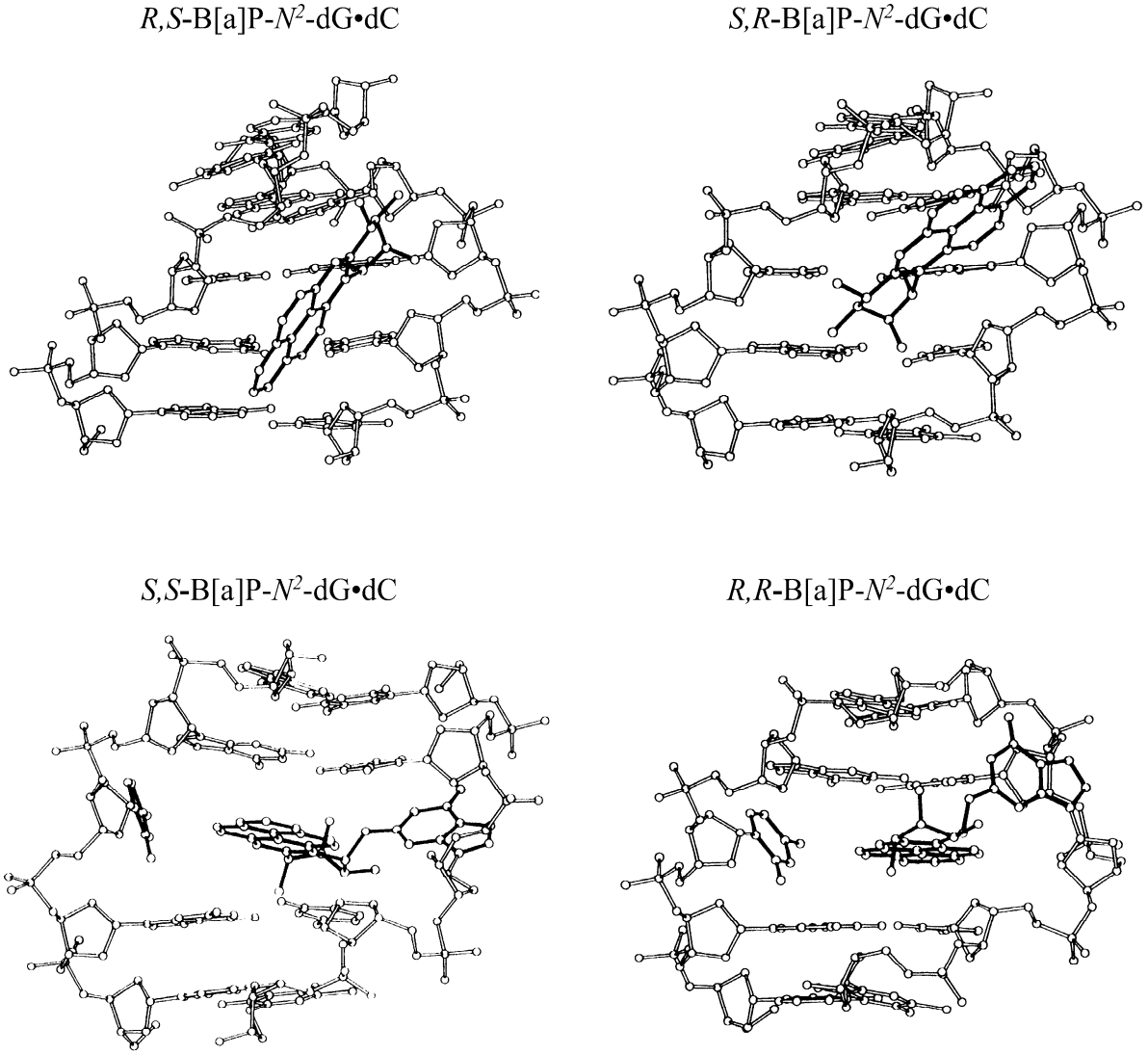


Figure 22. Structures of duplexes having B[a]P-N^2-dG lesions derived from isomeric B[a]P-DE **128a** and **128b**. All pictures show the central five-base-pair segment with the minor groove prominent. Adapted with permission from ref 315. Copyright 1997 American Chemical Society.

the present section we will use a two-letter notation with the first and second letters indicating the configurations of benzylic carbons connected to the hydroxyl and the base nitrogen, respectively.

NMR studies have established that duplexes containing isomeric B[*a*]P triol adducts **129a**−**d** share several common features. The NMR characterization of an 11-mer duplex having a centrally located **129a** residue opposite dC, *R,S*-B[*a*]P-N^2-dG·dC, flanked by C·G base pairs, reveals the existence of a stable right-handed helix, with residues in the *anti* range around the χ torsion angle and WC alignments throughout the duplex, including at the lesion-containing base pair. Observation of specific NOE interactions between the adduct and minor groove duplex protons, as well as the large shielding experienced by H1′ protons of residues located at the 5′-side of the lesion, places the pyrene moiety in the minor groove of the helix, where it is directed toward the 5′-end of the lesion-containing strand (Figure 22).[325]

Characterization of a related duplex having the *S,R*-B[*a*]P-N^2-dG lesion **129d** indicates the formation of a similar double-helical structure in solution with WC base pair alignments throughout the duplex. As in the case of the adduct **129a**, specific NOE interactions between adduct and minor groove duplex protons locate the aromatic moiety in the minor groove of the helix. However, significant shielding occurs for the H1′ proton of the lesion partner dC and its 5′-neigboring residue, demonstrating that in this case the pyrene ring is directed toward the 3′-end of the lesion-containing strand (Figure 22).[326] This structural motif, named the external minor groove conformation, is frequently adopted by lesion-containing duplexes to accommodate bulky DNA adducts. The different orientation of the B[*a*]P moiety in the minor groove observed in these early reports illustrates a precept that was later extended to other PAH lesions: the absolute configuration of the lesion linkage carbon determines adduct directionality, with the *S* isomer pointing to the 5′-end of the damaged strand and the *R* isomer toward the opposite side.

It is worth noting that analogous NMR features, including the formation of a right-handed helix with WC base pair alignments throughout the duplex, had been reported previously in duplexes having guanine N^2 lesions of pyrrole[1,4]-benzodiazepine antitumor antibiotics.[327−330] Interestingly, correlation between the minor groove orientation of the bulky substituent and the configuration of the lesion linkage carbon was also noted[328] (see below).

Ensuing NMR studies of duplexes containing the *R*,*R*-B[*a*]P-N^2-dG **129b**[331] or *S*,*S*-B[*a*]P-N^2-dG **129c** isomer[332] opposite dC in the same sequence context lead to quite different conclusions. In these duplexes, the aromatic pyrene moiety intercalates inside the helix, displacing the lesion-partner dC residue from the base stack into the major groove of the duplex and blocking the formation of a lesion-containing base pair (Figure 22). Pyrene intercalation displaces also the damaged bases, which now appear in the minor (in the **129b** case) or major (in the **129c** case) groove of the helix. As a result, each isomer intercalates from a different side of the helix, affecting the "groove look" of these structures, a characteristic that has been put forward to explain differences in their biological properties. This structural motif, called base-displaced intercalation, is another conformation that frequently accommodates bulky adducts in DNA duplexes.

The possibility that the duplex structure may be exchanging between an external, minor-groove and an intercalated base-displaced conformation was entertained early in these studies. The NOESY spectra of the *R*,*R*-B[*a*]P-N^2-dG·dC and *S*,*S*-B[*a*]P-N^2-dG·dC duplexes show the presence of exchange cross-peaks at low temperature,[331,332] indicating the existence of a second duplex conformation that, in the case of the *R*,*R*-B[*a*]P-N^2-dG·dC isomer **129b**, has the pyrene moiety not intercalated in the helix with hydrogen bonding across the lesion-containing base pair. Furthermore, while the duplexes having *R*,*S*-B[*a*]P-N^2-dG·dC or *S*,*R*-B[*a*]P-N^2-dG·dC flanked by G·C base pairs adopt a single structure in solution, an 11-mer duplex containing a *R*,*S*-B[*a*]P-N^2-dG·dC pair in the (5′-...T X C...)·(5′-...G C A...) sequence context shows two different conformations. The main conformation, with the aromatic moiety external in the minor groove of the helix and pointing toward the 5′-end of the modified strand, exchanges with a minor form, which hypothetically has the pyrene ring intercalated between flanking base pairs.[333] The effect of base context on duplex conformation has been epitomized recently by a report showing that a 5-methyl-dC substitution of the 5′-lesion-flanking dC residue in the (5′-...C X C...)·(5′-...G C A...) sequence changes the conformation of the *S*,*R*-B[*a*]P-N^2-dG lesion from the external, minor-groove alignment[326] to a base-displaced intercalated structure (PDB code 1Y9H).[334] In this case, pyrene intercalation displaces the lesion-partner residue into the major groove and the damaged dG residue in the minor groove of the helix, resembling the arrangement previously seen with the *R*,*R*-B[*a*]P-N^2-dG isomer on the nonmethylated duplex.[331] Interestingly, the *R*,*S*-B[*a*]P-N^2-dG (**129a**), which also adopts the external, minor-groove alignment, does not change to an intercalated structure upon methylation of the 5′-lesion-flanking dC, showing, once again, a correlation between adduct configuration and duplex structure. It was hypothesized that favorable hydrophobic interactions between the cytosine methyl group and the pyrene moiety, which are only present with the *S*,*R*-B[*a*]P-N^2-dG isomer, were the driving force behind this structural change.

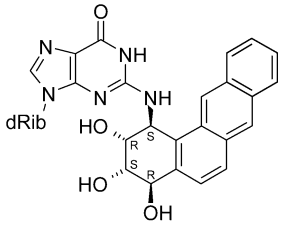
Patel and co-workers have studied the conformation of *R*,*S*-B[*a*]P-N^2-dG, *S*,*R*-B[*a*]P-N^2-dG (PDB code 1AXL), and *R*,*R*-B[*a*]P-N^2-dG adducts in the context of one-base deletion mutation (1Δ) intermediates where, basically, the lesion-partner dC residues were removed from the unmodified strand of their original 11-mer duplex sequence.[335−337] In all three cases, analysis of the NMR spectra showed that the presence of the bulky adduct stabilized the 11-mer/10-mer deletion duplexes, which have the pyrene moiety inside the helix with formation of 10 canonical WC base pairs. The most notorious difference between these systems is the groove location of the damaged residue, which in turn determines the site from where the intercalation takes place. In the case of the *R*,*S*-B[*a*]P-N^2-dG **129a** 1Δ model, the damaged residue locates in the major groove of the helix with the intercalated pyrene moiety pointing toward the minor groove. Conversely, in the *S*,*R*-B[*a*]P-N^2-dG **129d** and *R*,*R*-B[*a*]P-N^2-dG **129b** 1Δ samples, damaged dG residues are located in the minor groove of the helix, from where the aromatic moiety intercalates. Interestingly, while all the lesion-containing 1Δ models are more stable than the unmodified 1Δ control, the *R*,*S*-B[*a*]P-N^2-dG 1Δ model has 9 and 19 °C lower T_m values than the *S*,*R*-B[*a*]P-N^2-dG and the *R*,*R*-B[*a*]P-N^2-dG 1Δ duplexes, respectively, suggesting an energy penalty when pyrene intercalation proceeds from the major groove of the helix.[338] In addition, the order of thermal stability correlates with the fact that 1Δ mutations introduced by human Pol *k* during translesion DNA synthesis are not detected in the presence of the *R*,*S*-B[*a*]P-N^2-dG lesion but are highest for *R*,*R*-B[*a*]P-N^2-dG.[339]

Two manuscripts have characterized the conformation of putative replication intermediates with a *R*,*S*-B[*a*]P-N^2-dG **129a** at a primer−template junction, mimicking the DNA structure when the damaged residue is the next base to be copied[340] and after insertion of dCMP opposite the lesion (PDB code 1AXO).[341] The duplex segment of both samples adopts a regular B-form conformation with residues in the *anti* conformation around their χ torsion angle and WC base pair alignments. The single-stranded region of the pre-dCMP incorporation sample shows a high degree of conformational freedom without adopting a defined structure. The χ torsion angle of the *R*,*S*-B[*a*]P-N^2-dG residue is in the *syn* range, and the purine moiety appears on the major groove, allowing stacking of the pyrene ring on top of the duplex segment. After dCMP incorporation, the single-stranded region is more structured, *R*,*S*-B[*a*]P-N^2-dG forms a WC pair with dC, and the pyrene ring appears in the minor groove of the nascent duplex. The X-ray structure of a thermophilic DNA polymerase in complex with damaged DNA containing the *R*,*S*-B[*a*]P-N^2-dG in the template strand has been reported recently.[342] Despite differences induced by the presence of bacterial enzyme, the DNA structure of the insertion complex solved by X-ray crystallography (PDB code 1XC9) shares some important features with that seen by NMR spectroscopy. Namely, the *R*,*S*-B[*a*]P-N^2-dG residue is *anti*, forming a canonical WC base pair with the incoming dC residue, and its pyrene moiety is located external in the nascent minor groove of the helix.

8.2.2. Benzo[a]anthracene (B[a]A)

The NMR characterization of N^2-dG adducts of different polycyclic aromatic hydrocarbons has complemented our understanding of correlations between duplex structure and adduct topology. Benzo[*a*]anthracene has three linearly arranged aromatic rings that create an elongated hydrophobic moiety with a bay region at one end of the molecule. The NMR characterization of an 11-mer duplex having a centrally located *R*,*S*-B[*a*]A-N^2-dG triol adduct **130** (Chart 43) paired to dC establishes the presence of a predominant right-handed helical structure with residues in the *anti* conformation around the χ torsion angle and WC alignments throughout the duplex, including at the lesion-containing base pair.[343] The polyaromatic moiety in **130** resides in the minor groove

Chart 43. *R,S*-((−)-*trans-anti*)-B[a]A-N^2-dG Adduct



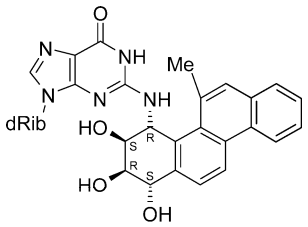
130

of the helix, where it is directed toward the 5′-end of the damaged strand, adopting an alignment similar to that previously seen with the *R,S*-B[*a*]P-N^2-dG lesion **129a**,[325,333] with which it shares the *S* configuration at the benzylic linkage carbon of the adduct.

8.2.3. 5-Methylchrysene (MCh)

The solution structure of an 11-mer duplex containing an *S,R*-MCh-N^2-dG•dC base pair in the (5′-...C_5 X_6 C_7...)•(5′-...G_{16} C_{17} G_{18}...) sequence context has been established.[344] In relation to B[*a*]P, MCh lacks one aromatic ring and has a methyl substitution at the edge of its bay region, converting it into the more crowded fjord region (Chart 44). Analysis

Chart 44. *S,R*-((−)-*trans-anti*)-5-*MCh*-N^2-dG Adduct



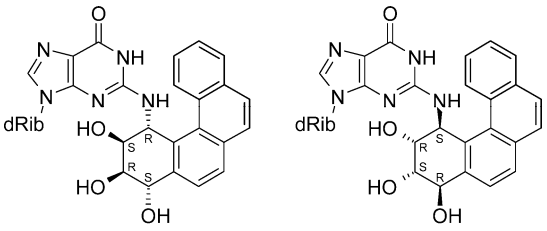
131

of the NMR spectra reveals the existence of a regular right-handed helical structure stabilized by formation of WC base pairs all through the duplex, including at the lesion containing the *S,R*-MCh-N^2-dG•dC pair. The MCh moiety is located in the minor groove of the helix, stacking mostly over the complementary strand with its aromatic ring system directed toward the 3′-end of the damaged strand. The methyl group of MCh is at the floor of the minor groove, where it wedges between the X_6•C17 and C_7•G_{16} base pairs, bending the helix about 47°. In a broad sense, the structure resembles that of the *S,R*-B[*a*]Pd-N^2-G•dC duplex,[326] but in this case, the lesion-containing base pair is buckled and the polyaromatic MCh moiety is more exposed to solvent. Interestingly, the related duplex having the *R,S*-MCh-N^2-dG lesion placed in the same sequence context shows broad NMR signals and is structurally heterogeneous. Its conformation has not been refined yet.[315]

8.2.4. Benzo[c]phenanthrene (B[c]Ph)

The NMR characterization of duplexes having the *S,R*-B[*c*]Ph-N^2-dG (**132a**) or *R,S*-B[*c*]Ph-N^2-dG (**132b**) lesion paired to dC at their centers illustrates further the dramatic effect that adduct topology can have on duplex conformation.[345] In this case, guanine bases bind to a benzylic carbon of benzo[*c*]phenathrene that, as in the *S,R*-MCh-N^2-dG lesion, is next to the more crowded fjord region of the aromatic ring system (Chart 45), imposing additional constraints to the double-stranded structure. NMR spectra of the *R,S*-B[*c*]-Ph-N^2-dG•dC-containing duplex show a predominant con-

Chart 45. *S,R*-((+)-*trans-anti*)-B[c]Ph-N^2-dG 132a and *R,S*-((−)*trans-anti*)-B[c]Ph-N^2-dG 132b Adducts



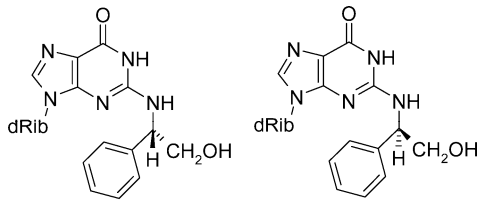
132a 132b

formation with all residues in the *anti* range around the χ torsion angle and WC base pair alignments throughout, including across the lesion-containing base pair (PDB code 1HWV). Intrinsic steric clashes induce a considerable twist at the fjord region of the B[*c*]Ph moiety, allowing its intercalation inside the helix at the 5′-side of the lesion with partial exposure of the nonaromatic ring in the minor groove. B[*c*]Ph intercalation increases the base pair rise at this step of the duplex and induces buckle of the lesion-containing and 5′-flanking base pairs without disrupting their hydrogen bonding (Figure 23). The duplex having the *S,R*-B[*c*]Ph-N^2-dG isomer shows a similar B-type conformation with WC base pair alignments all the way through (PDB code 1HX4). The B[*c*]Ph moiety intercalates in this case at the 3′-side of the lesion, increasing the distance between base pairs and inducing a small buckle on the lesion-containing pair. As in the *R,S*-B[*c*]Ph case, intercalation causes some loss of planarity on the aromatic ring system and displacement of the guanine amino nitrogen from the base plane, resulting in a chair conformation for the cyclohexene ring of B[*c*]Ph. It is worth emphasizing that the conformation of *S,R*-B[*c*]-Ph-N^2-dG (**132a**) and *R,S*-B[*c*]Ph-N^2-dG (**132b**) lesion-containing duplexes is very different from that observed for B[*a*]P adducts. In the latter case, the *S,R* and *R,S* isomers do not intercalate at all and intercalation of *R,R*- and *S,S*-B[*a*]P-N^2-dG causes the complete disruption of the lesion-containing base pair.[331,332] In this regard, the conformation of the *S,R*- and *R,S*-B[*c*]Ph-N^2-dG•dC duplexes makes up a third structural motif for accommodating minor-groove-attached bulky PAH lesions in DNA.

8.2.5. Styrene Oxide (SO)

The structure of DNA duplexes having an *S* (**133a**) or *R* (**133b**) α-styrene oxide-N^2-dG adduct ((α-SO-N^2-dG) paired to dC has been described.[346,347] As compared to previous PAH lesions, α-SO-N^2-dG adducts (Chart 46) have a single

Chart 46. *S*-α-SO-N^2-dG 133a and *R*-α-SO-N^2-dG 133b Lesions



133a 133b

benzyl group attached to the guanine amino nitrogen, reducing minor-groove crowding at the lesion site. The NMR-refined models establish the presence of right-handed helical structures with all residues in the *anti* range around the χ torsion angle and WC base pair alignments throughout the

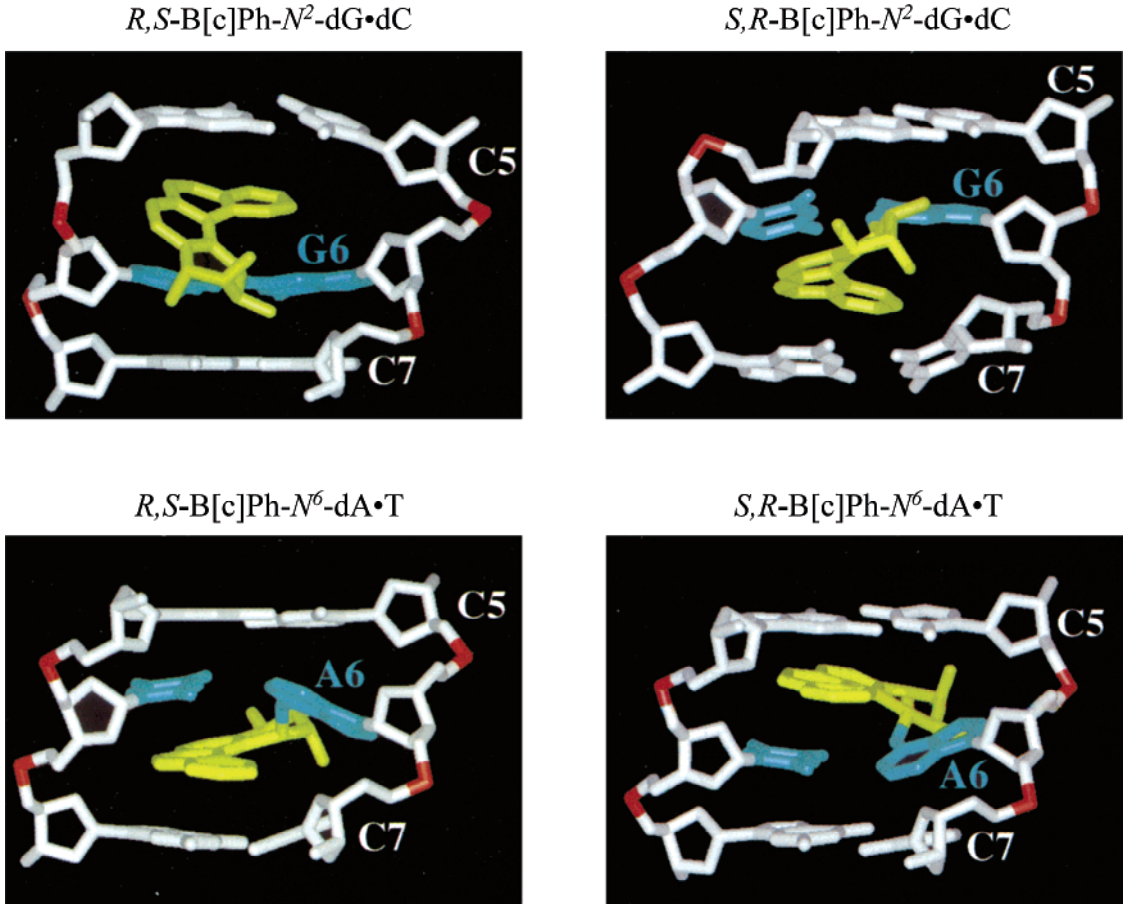


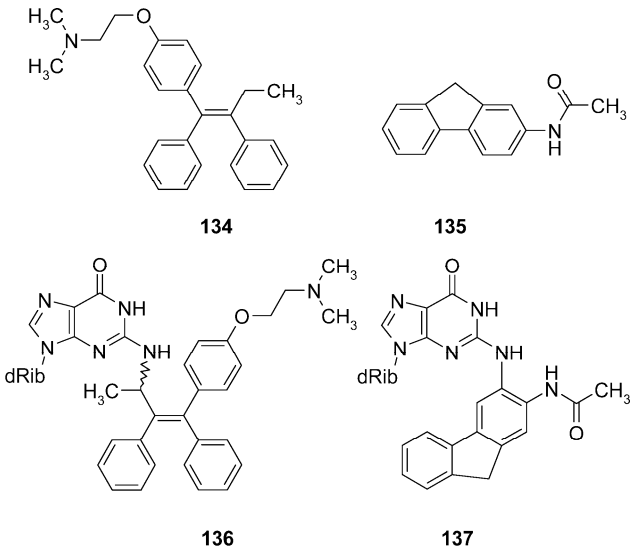
Figure 23. Structure of duplexes having isomeric B[c]Ph-N^2-dG lesions **132a**,**b** (top panel) and B[c]Ph-N^6-dA lesions **147a**,**b** (bottom panel). All pictures show the central three-base-pair segment with the minor groove prominent. Adapted with permission from ref 345. Copyright 2001 Elsevier.

duplex, including at the lesion containing base pair. In both duplexes, the benzyl ring is external in the minor groove of the helix, where it can rotate rapidly in the NMR time scale. The *R* isomer is oriented toward the 3′-end of the lesion-containing strand, and the *S* isomer is oriented in the opposite direction, following the directionality initially observed for the *R*,*S*- and *S*,*R*-B[*a*]P-N^2-dG adducts.[325,326] Interestingly, the presence of the *R*-α-SO-N^2-dG adduct induces a larger perturbation of the duplex structure at the lesion site than the *S* isomer, and when the sequence context changed from (5′-...G X T...) to (5′-...A X G...), a refined structure could not be obtained.[347]

8.2.6. Other N^2-Guanine Lesions

Additional studies dealing with the structure of duplexes containing other bulky N^2-dG lesions are the reports describing adducts of *S*-*cis*-tamoxifen (*S*-*cis*-TAM, **134**) and 2-acetylaminofluorene (2-AAF, **135**) (Chart 47). NMR spectra of an 11-mer duplex containing an *S*-*cis*-TAM-N^2-dG adduct (**136**) opposite dC indicate the presence of a right-handed helical structure with all residues in the *anti* conformation around the χ torsion angle and WC base pair alignments throughout. The tamoxifen molecule resides in a widened minor groove, which has almost doubled its width, with the substituted phenoxy ring directed toward the 5′-end of the modified strand and the two phenyl rings directed in the opposite direction. Similar to the α-SO-N^2-dG case, the spectra show a single averaged signal for the *ortho* and *meta* protons of each phenyl ring, indicating that all three aromatic moieties are in fast rotation.[348] On the basis of these observations, the authors concluded that the presence of an

Chart 47. *cis*-Tamoxifen 134, 2-Acetylamino Fluorene 135, and Their N^2-dG Adducts



S-*cis*-TAM moiety in the minor groove of the duplex might compromise the fidelity of DNA polymerases.

In the 2-acetylaminofluorene adduct case, 2-AAF-N^2-dG **137**, our laboratory just concluded the NMR characterization of an 11-mer duplex having a 2-AAF-N^2-dG•dC pair at its center. The NOESY spectra establish the presence of a regular right-handed helix with all residues in the *anti* conformation around the χ torsion angle and WC base pair alignments throughout the duplex.[349] The 2-AAF moiety

Table 4. Data for Duplexes Having N^2-dG Adducts

Double-Stranded Duplexes Having Bulky N^2-dG Lesions

lesion name[a]	adduct topology[b]	duplex length	sequence context[c]	structural motif	PDB entry	ref
R,*S*-B[*a*]P **129a**	bay	11-mer	5′-(...C G* C...)	minor groove, 5′-side oriented		325
R,*S*-B[*a*]P **129a**	bay	11-mer	5′-(...T G*C...)	minor groove, 5′-side oriented (2nd form: B[*a*]P stacked)		333
R,*S*-B[*a*]P **129a**	bay	11-mer	5′-(...5^{Me}C G* C...)	minor groove, 5′-side oriented		334
S,*R*-B[*a*]P **129d**	bay	11-mer	5′-(...C G* C...)	minor groove, 3′-side oriented		326
S,*R*-B[*a*]P **129d**	bay	11-mer	5′-(...5^{Me}C G* C...)	intercalated, dG*: minor groove displaced	1Y9H	334
R,*R*-B[*a*]P **129b**	bay	11-mer	5′-(...C G*C...)	intercalated, dG*: minor groove displaced		331
S,*S*-B[*a*]P **129c**	bay	11-mer	5′-(...C G*C...)	intercalated, dG*: major groove displaced		332
R,*S*-B[*a*]A **130**	bay	11-mer	5′-(...G G* T...)	minor groove, 5′-side oriented		343
S,*R*-5-MeC **131**	bay	11-mer	5′-(...C G* C...)	minor groove, 3′-side oriented		344
S,*R*-B[*c*]Ph **132a**	fjord	11-mer	5′-(...C G*C...)	B[*c*]Ph 3′-side intercalated dG*: WC-paired	1HX4	345
R,*S*-B[*c*]Ph **132b**	fjord	11-mer	5′-(...C G*C...)	B[*c*]Ph 5′-side intercalated dG*: WC-paired	1HWV	345
11*S*-anthramycin **138a**		6-mer	5′-(...T G* C...)	minor groove, 5′-side oriented		327, 329
S-α-SO **133a**		11-mer	5′-(...G G* T...)	minor groove, 5′-side oriented		346
		11-mer	5′-(...A G* G...)	minor groove, 5′-side oriented		347
R-α-SO **133b**		11-mer	5′-(...G G* T...)	minor groove, 3′-side oriented		346
		11-mer	5′-(...A G* G...)	minor groove, 3′-side oriented		347
S-*cis*-TAM **134**		11-mer	5′-(...C G* C...)	minor groove, 5′-side oriented		348
2-AAF **135**		11-mer	5′-(...C G* C...)	minor groove, 3′-side oriented		349
R-mitomycin-C **139**		9-mer	5′-(...C G* T...)	minor groove, 3′-side oriented	199D	352
11*S*-tomaymycin **138b**		6-mer	5′-(...T G* C...)	minor groove		328

One-Base Deletion Mutation Intermediates Having Bulky N^2-dG Lesions

lesion name[a]	adduct topology[b]	duplex length	sequence context[d]	structural motif	PDB entry	ref
R,*S*-B[*a*]P **129a**	bay	11-mer	5′-(...C G*C...)	intercalated, dG*: major groove displaced		335
S,*R*-B[*a*]P **129d**	bay	11-mer	5′-(...C G*C...)	intercalated, dG*: minor groove displaced	1AXL	336
R,*R*-B[*a*]P **129b**	bay	11-mer	5′-(...C G*C...)	intercalated, dG*: minor groove displaced		337

[a] Lesion linkage carbon is underlined. B[*a*]P, benzo[*a*]pyrene; B[*a*]A, benzo[*a*]anthracene; 5-MeC, 5-methylchrysene; B[*c*]Ph, benzo[*c*]phenanthrene; α-SO, α-styrene oxide; *cis*-TAM, *cis*-tamoxifen; 2-AAF, 2-acetylaminofluorene. [b] Indicates PAH topology and configuration of benzylic carbons. [c] Lesions always paired to C. [d] No residue paired to the lesion. Na = not applicable.

resides in the minor groove of the helix, where it is directed toward the 3′-end of the lesion-containing strand, as was the case of the *S*,*R*-B[*a*]P-N^2-dG paired to dC (Figure 24).

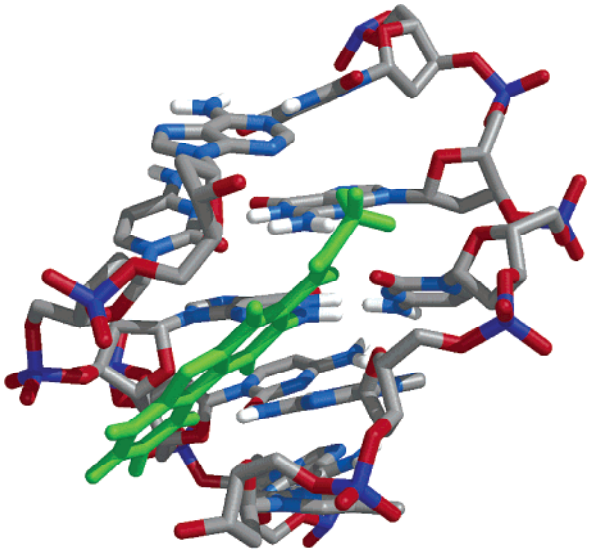
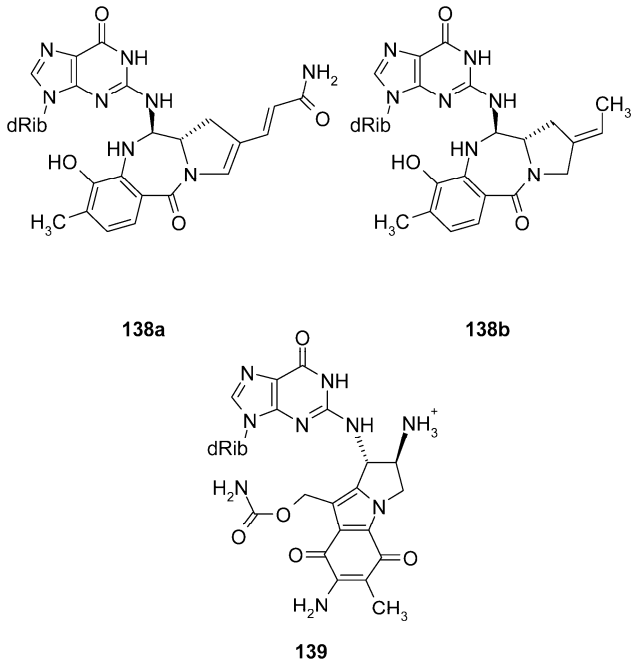


Figure 24. Structure of duplex DNA having a 2-AAF-N^2-dG lesion **137** paired to dC. This view shows the central region of the sample with the minor groove prominent.

Remarkably, the lesion considerably increases duplex stability, giving a 6 °C higher T_m and 1.80 kcal/mol less Gibb's free energy than the unmodified counterpart. Interestingly, this 2-AAF adduct has been detected in rat liver almost one year after discontinuation of carcinogen feeding,[350] suggesting that it may resist repair and supporting the hypothesis that adduct-induced stability prevents lesion recognition by the NER system.[351]

Chart 48. Anthramycin 138a, Tomaymycin 138b, and Mitomycin C 139 N^2-dG Adducts



Furthermore, some *Streptomyces* species produce antibiotics of broad use in cancer chemotherapy that react with DNA forming covalent dG lesions. NMR studies have characterized the structures of duplexes having an anthramycin-N^2-dG,[327,329] a tomaymycin-N^2-dG,[328,330] or a mitomycin C-N^2-dG monoadduct[352] (PDB code 199D) paired to dC (Chart 48). In all cases, the duplex adopts a regular B-form conformation, stabilized by formation of WC alignments

throughout and the bulky antibiotic moiety external in the minor groove of the helix. In the tomaymycin-N^2-dG adduct, two species were formed during preparation of the duplex sample, which corresponded to the *S* and *R* configurations at the lesion linkage carbon. While the specific minor groove orientation of each isomer could not be established, the antibiotic moiety showed a different orientation in each case.[328,330] Only one adduct with the *S* configuration at the lesion linkage carbon was present in the anthramycin-N^2-dG·dC duplex, and in this case, the alkyl side chain of the pyrrole ring was directed toward the 5′-end of the modified strand.[327,329] Similarly, in the case of the mitomycin C-N^2-dG monoadduct, only one product with *R* configuration at the linkage carbon formed during the preparation of the sample and the duplex structure showed that the mitomycin moiety is pointing toward the 5′-end of the modified strand.[352]

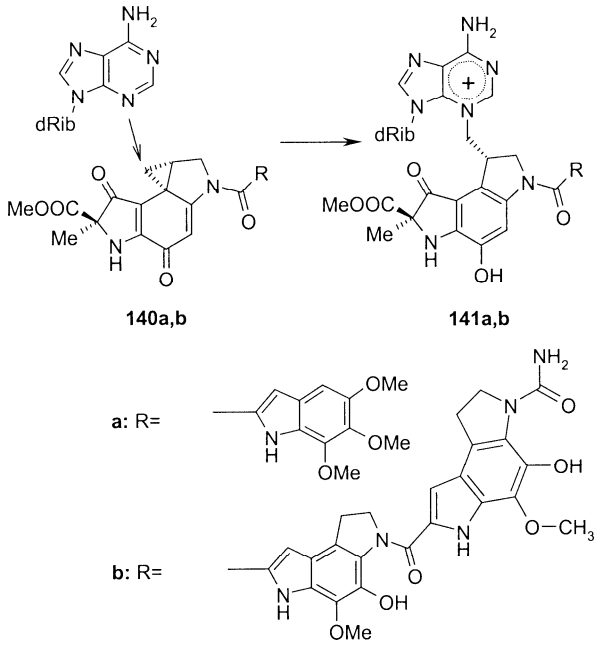
In summary, three different structural motifs have been described to date for bulky N^2-dG adducts in duplex DNA: (i) external conformation, in the minor groove of the duplex, (ii) intercalated with groove displacement of the damaged base, and (iii) intercalated without displacement of the damaged base. Concurrent factors determine the local duplex conformation at the damaged site including the absolute configuration of the adduct linkage carbon, the adduct topology, and the lesion sequence context. The existence of these structural motifs may help to explain the diverse biological responses triggered by bulky N^2-dG lesions. Table 4 compiles NMR studies performed on duplexes having N^2-dG adducts.

8.3. N^3-Adenine Lesions

Streptomyces species also produce antitumor compounds, such as duocarmycin (**140a**) and CC1065 (**140b**) antibiotics that react with the endocyclic adenine N^3 in double-stranded DNA. These highly potent minor groove binders consist of a pyrroloindole moiety ("ring A") linked by amide bonds to two additional pyrroloindole units, in the case of CC1065, or to an indol ring, in the case of duocarmycins. The cyclopropane moiety of ring A is part of a semiquinone-type system, which makes it prone to nucleophilic opening. After fast noncovalent minor-groove binding with 5′-d(A/T A A A*) or 5′-d(A/T T T A*) sequence specificity, **140a**,**b** slowly forms a covalent bond with the N^3 of adenine denoted with an asterisk,[353,354] yielding CC1065-N^3-dA **141b** or duocarmycin A-N^3-dA **141a** lesions (Scheme 14). The NMR characterization of duplex DNA containing a CC1065-N^3-dA lesion **141b** ^{15}N-labeled at the 6-amino group showed two doublets that originated from the damaged adenine amino protons, resonating in the 9.0−9.2 ppm range. Additionally, the imino proton of the lesion-partner thymine residue moves drastically upfield and appears as a sharp signal around 11.5 ppm. These observations led to the conclusion that the predominant tautomeric form of the lesion has the amino group doubly protonated, with the resulting positive charge delocalized over the entire adenine moiety.[355] Further studies of duplexes containing a CC1065-N^3-dA lesion **141b**[356,357] or the related CPI-CDPI$_2$-N^3-dA adduct[358] indicated that all residues are *anti* around the χ torsion angle and form WC alignments, including at the lesion-containing base pair that is minimally distorted. The bulky substituent locates in the minor groove of the duplex, directing its longer part toward the 5′-end of the damaged strand and forming roughly a 45° angle with the helical axis (Figure 25).

A subsequent study of a sample having a DuoA-N^3-dA·T base pair showed essentially identical NMR parameters,

Scheme 14. Duocarmycin 140a, CC1065 140b, and Their N^3-dA Adducts



indicating the presence of a positively charged adenine base.[359] The damaged duplex adopts a regular right-handed duplex conformation, with residues in the *anti* conformation around the glycosidic angle and WC base pair alignments. The refined three-dimensional structure (PDB code 107D) shows the lesion in the minor groove with the same orientation previously defined for the CC1065 adduct. In closing, all duplexes having adenine N^3 lesions characterized to date showed the presence of a minimally perturbed B-form conformation, with the adenine ring positively charged and

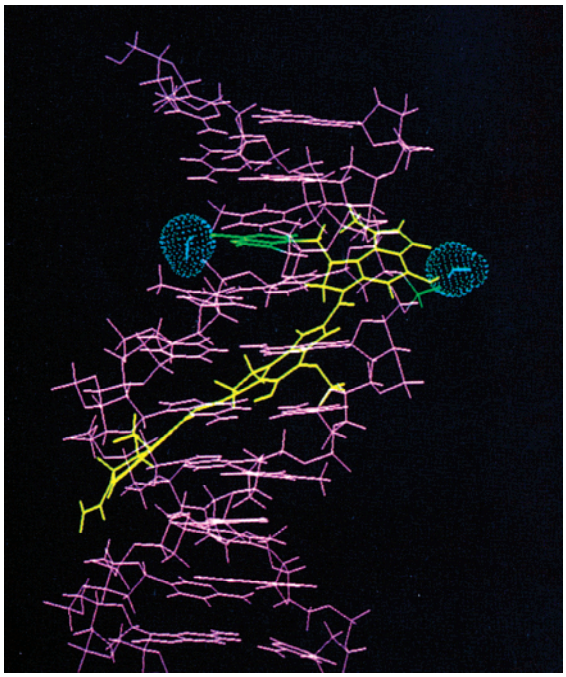


Figure 25. Central region of a dodecamer duplex having a CC1065-N^3-dA lesion **140b** paired to T, shown with the minor groove prominent. The picture depicts the van der Waals surface of two water molecules proposed to mediate the affinity of CC1065 (colored yellow) for AAA and TTA sequences. Adapted from ref 357. Copyright 1991 American Chemical Society.

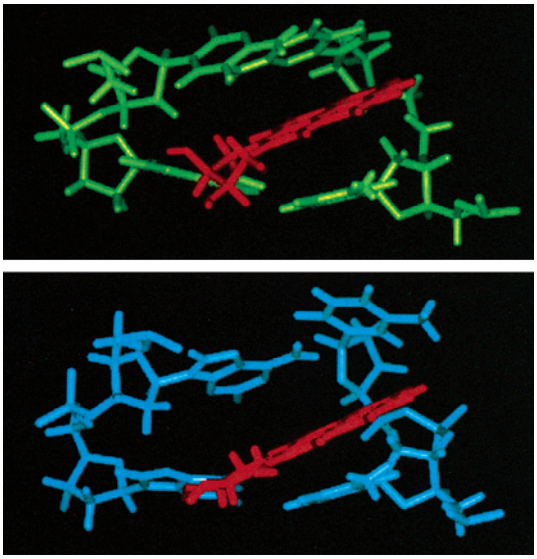


Figure 26. Intercalation of the *S*,*R*-B[a]P-N^6-dA adduct **142a** (in red) in duplex DNA. The top panel shows a quasi-intercalated conformation seen in the 5′(...C A*A...) sequence context. The bottom panel depicts the fully intercalated structure seen in the 5′-(...A A* G...) sequence context. Both views display the major groove prominent. Adapted with permission from ref 361. Copyright 1998 American Chemical Society.

stacked inside the helix and the bulky ring system external in the minor groove of the duplex.

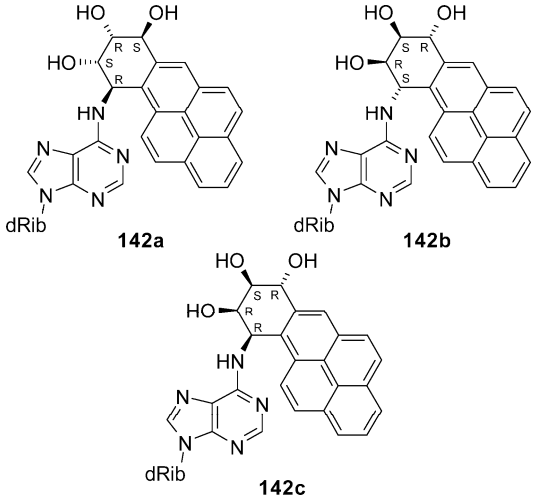
8.4. N^6-Deoxyadenosine Lesions

8.4.1. Benzo[a]pyrene (B[a]P)

Besides guanine, a second major target of B[*a*]P DE (**128a**−**d**) attack is the 6-amino group of dA. This reaction follows the same stereochemical rules indicated for dG lesions and produces a set of diastereomeric B[*a*]P-N^6-dA adducts similar to those shown in Scheme 13. The major-groove location of the adenine amino group suggests that the most likely location of a bulky aromatic moiety would be external on the major groove of the duplex or intercalated in the base pair stack. NMR spectroscopy, in combination with restrained molecular dynamics, has determined the structure of DNA duplexes having an *S*,*R*-B[*a*]P-N^6-dA residue **142a** (Chart 49) paired to T in two different sequence contexts.[360−362] In the (5′-...C X A...) sequence context, the data establish the presence of a right-handed helix with all residues in the *anti* conformation around their χ torsion angles and WC alignments all through the duplex. The pyrene moiety intercalates only partially at the 5′-side of the damaged residue, stacking largely with the 5′ lesion-flanking base pair, a geometry called "quasi-intercalated" by the authors (Figure 26). To accommodate the aromatic moiety, the lesion-containing base pair and the 5′-lesion-flanking base pair buckle in opposite directions and the helix raised at this step doubles the value that is seen in canonical B-form duplexes. Base H5 and H6 protons of lesion-flanking dC and cyclohexene H8 and H9 protons of B[*a*]P show broad NMR signals indicating increased flexibility at the 5′-side of the lesion, presumably associated with an exchange between two-half-chair conformations of the B[*a*]P cyclohexene ring.[360] This structural family and its variations is named the 5′-intercalated, WC aligned conformation.

Additional studies have established the structure of duplexes when residues flanking the *S*,*R*-B[*a*]P-N^6-dA lesion

Chart 49. *S*,*R*-((−)-*trans*-*anti*)-B[a]P-N^6-dA 142a, *R*,*S*-((+)-*trans*-*anti*)-B[a]P-N^6-dA 142b, and *R*,*R*-((+)-*cis*-*anti*)-B[a]P-N^6-dA 142c Adducts



are changed from (5′-...C A* A...) to (5′-...A A* G...).[361,362] The global structural properties of these duplexes are within the 5′-intercalated family, with canonical WC base pair alignments throughout the duplex and B[*a*]P intercalation at the 5′-side of the lesion, causing a buckle at the lesion-containing and 5′-flanking base pairs. The pyrene moiety intercalates from the major groove of the helix but, in this case, crosses the base pair stack and appears on the floor of the minor groove (Figure 26). In addition, change of the duplex sequence increases conformational flexibility at the intercalation site, a fact that resulted in the observation of two different exchange processes. In one study, computer simulation of deoxyribose J-coupling constants suggests that the sugar pucker of the 5′-lesion-flanking dA is best described as about equal populations of fast exchanging 2′-endo and 3′-endo conformers.[361] Refined models with the sugar pucker fixed on one or the other conformation show that the 2′-endo to 3′-endo exchange moves the dA residue more toward the major groove. In the other study, the presence of exchange cross-peaks in the NOESY spectra in combination with *ab initio* calculations of chemical shift perturbations led to the conclusion that the χ torsion angle of the damaged residue is in slow exchange between a prominent (>95%) *anti* and a minor (<5%) *syn* conformation.[362] The structure of the minor *syn* conformer was not refined. While both processes can be occurring simultaneously, the *anti* ↔ *syn* exchange involves a larger number of atoms undergoing longer excursions, a fact that, qualitatively, is consistent with the slow exchange character of this process. Interestingly, the *R*,*S*-B[*a*]P-N^6-dA adduct **142b** incorporated in the same sequence was too dynamic and the solution structure of this duplex could not be established.

The NMR characterization of an 11-mer duplex having a centrally located *R*,*R*-B[*a*]P-N^6-dA adduct **142c** paired to T in the (5′-...C X C...) sequence context establishes the presence of a regular right-handed helix with all residues in the *anti* conformation around the χ torsion angle and WC base pair alignments throughout the duplex.[363] The refined structure (PDB code 1AXV) shows that the aromatic moiety intercalates at the 5′-side of the damaged residue, perturbing somehow the WC geometry of the lesion-containing base pair that appears propeller twisted and buckled and the helical twist at the lesion site that is reduced (Figure 27). As in the case of a *S*,*R*-B[*a*]P-N^6-dA·T pair positioned in the (5′-...C

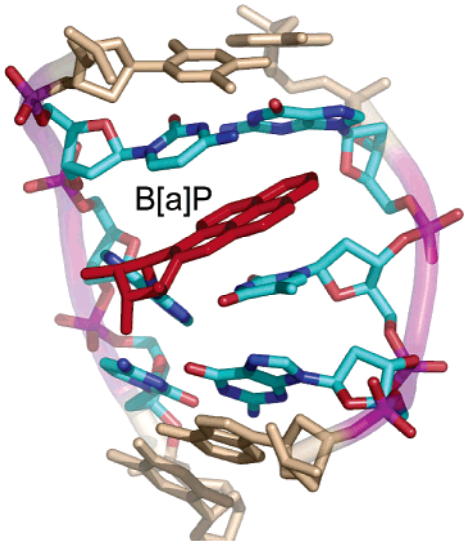
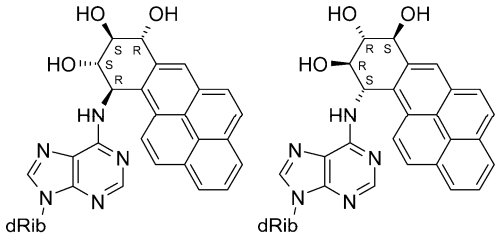


Figure 27. Structure of DNA duplex having an *R*,*R*-B[a]P-N^6-dA lesion **142c**, shown with the major groove prominent and the intercalated B[a]P moiety in red. The picture displays the central five-base-pair segment. Prepared from PDB entry 1AXV[363] using PyMol.[555]

X A...) context,[360] adduct intercalation is only partial by favoring stacking of B[*a*]P with the 5′-lesion-flanking base pair. Interestingly, the NMR characterization of a related benzo[*a*]pyrene-N^6-dA lesion, that lacks the 7- and 8-hydroxyl groups but shares the configuration of chiral atoms with the *R*,*R*-B[*a*]P-N^6-dA adduct **142c**, generates basically the same structure (PDB code 1N8C). However, the absence of hydroxyl groups allows deeper penetration of the aromatic moiety in the helix and improves stacking at the lesion site, suggesting that OH substitutions on the lesion cyclohexene ring modulate the extent of B[*a*]P intercalation.[364]

It is illustrative to compare the previous structures of duplexes containing lesions derived from the *anti*-B[*a*]P-DE with those obtained from the rare *syn* diasteromers **128c**,**d** of the diol epoxide. NMR studies have characterized the structures of duplexes having a 7*R*,8*S*,9*S*,10*R*(−)-*trans*-*syn*-B[*a*]P-N^6-dA **142d** or a 7*S*,8*R*,9*R*,10*S*(+)-*trans*-*syn*-B[*a*]P-N^6-dA lesion **142e** (Chart 50) (note that a four-letter notation is used to describe the configuration of these two compounds) paired to T in an identical sequence context.[365,366] As seen in other adducts with the 10*R* configuration at the benzylic linkage carbon, the *R*,*S*,*S*,*R*-B[*a*]P-N^6-dA-containing duplex

Chart 50. *R*,*R*-((−)-*trans*-*syn*)-B[a]P-N^6-dA 142d and *S*,*S*-((+)-*trans*-*syn*)-B[a]P-N^6-dA 142e Adducts



adopts a regular right-handed helical structure within the 5′-intercalated, WC-aligned family.[365] Similarly, the NMR spectra of the *S*,*R*,*R*,*S*-B[*a*]P-N^6-dA isomer indicate the presence of a right-handed helix with χ torsion angles in the *anti* range on nonmodified residues and WC alignments on all canonical base pairs. However, clear differences are present at the lesion site of the latter duplex where the damaged residue adopts a *syn* conformation and is not WC hydrogen bonded to its partner T. The refined structure (PDB code 1JDG) shows the pyrene system of the *S*,*R*,*R*,*S*-B[*a*]P-N^6-dA residue intercalated toward its 3′-side, causing an evident buckle at the lesion-containing and 3′-flanking base pairs (Figure 28). Therefore, as in the case of isomeric B[*a*]P-N^2-dG adducts, the configuration of the cyclohexene linking carbon primarily defines the orientation of the aromatic moiety at the lesion site.

Other examples of a direct relationship between lesion configuration and duplex structures are the results obtained with 9-mer mismatched duplexes having the *S*,*R*-B[*a*]P-N^6-dA (**142a**) or the *S*,*R*,*R*,*S*-N^6-dA (**142e**) adduct paired to dG.[367−369] The *S*,*R*-B[*a*]P-N^6-dA·dG duplex adopts a perturbed right-handed helical structure with WC base pair alignments outside the lesion site. The damaged dA residue is in the *anti* range with the aromatic moiety intercalated toward the 5′-side of the lesion, displacing its partner dG residue into the major groove and causing a large buckle at the 5′-lesion-flanking base pair.[367] The *S*,*R*,*R*,*S*-N^6-dA·dG-containing duplex also forms a right-handed helix stabilized by WC alignments on all canonical base pairs but with increased conformational dynamics at the lesion site. The duplex structure is in slow exchange between a main conformation where *S*,*R*,*R*,*S*-N^6-dA is in the *syn* range around the χ torsion angle and a minor form where it is in the *anti* conformation. In both cases, two hydrogen bonds are present

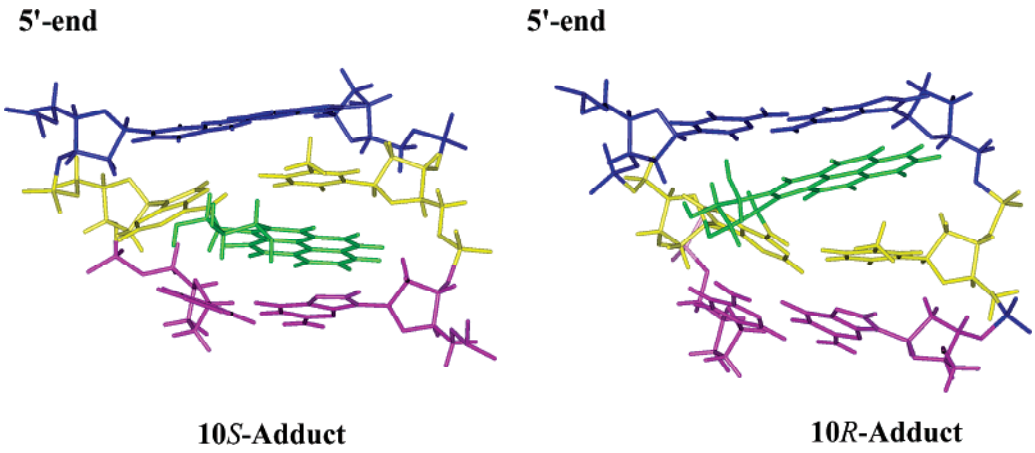


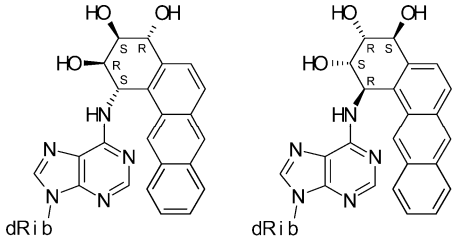
Figure 28. Intercalated structure of *S*,*S*-B[a]P-N^6-dA **142e** and *R*,*R*-B[a]P-N^6-dA **142d** adducts, derived from the rare *syn* diolepoxide, paired to T. The picture displays the central three-base-pair segment with the major groove prominent. The lesion-containing base pair is colored yellow and the pyrene moiety green. Reproduced with permission from ref 366. Copyright 2001 American Chemical Society.

at the lesion-containing base pair and the pyrene moiety intercalates at the 3′-side of the damaged residue.[368,369]

8.4.2. Benzo[a]anthracene (B[a]A)

As in the case of N^2-dG adducts, additional studies of different PAH-N^6-dA lesions complement our understanding of correlations between adduct geometry and duplex structure. Relatively recent studies have established the structure of DNA duplexes having the bay region *R,S*-B[*a*]A-N^6-dA **143a** or *S,R*-B[*a*]A-N^6-dA **143b** lesion (Chart 51) paired to

Chart 51. *R,S*-((+)-*trans-anti*)-B[a]A-N^6-dA 143a and *S,R*-((−)-*trans-anti*)-B[a]A-N^6-dA 143b Adducts



T.[370−372] Independent of the configuration at the benzylic linking carbon of the lesion, the NMR data show the presence of a right-handed helical structure with residues in the *anti* range for their χ torsion angle and WC alignments throughout duplex. The aromatic moiety intercalates inside the helix in an isomer dependent manner, appearing toward the 5′-side of the damaged dA when the lesion has the *S,R* configuration[370] or toward the 3′-side in the *R,S* isomer case.[371] With both lesion isomers, the B[*a*]A plane is almost orthogonal to the damaged adenine. In addition, base pairs flanking the intercalated aromatic moiety buckle in opposite directions, creating a cavity that allows stacking of B[*a*]A with the lesion-partner T residue. Furthermore, the *R,S* isomer **143a** stacks preferentially over the 3′-side lesion-flanking residue whereas the *S,R* isomer **143b** does it over the residue at the 5′-side of the lesion, increasing the stability of damaged duplexes (Figure 29). Perturbation of the B-form structure is more pronounced with the *R,S* isomer where the WC alignment of the lesion-containing base pair is practically lost due to its extreme buckle. Interestingly, this duplex adopts a second conformation with the damaged *R,S*-B[*a*]A-N^6-dA residue in the *syn* range around the χ torsion angle, in slow exchange with the major *anti* form.[371]

Sequence composition effects were partially investigated in the case of the *S,R*-B[*a*]A-N^6-dA isomer by changing lesion-flanking residues from (5′-...C X A...) to (5′-...A X G...).[372] As reported previously for B[*a*]P-N^6-dA adducts where similar research was done,[360,361] the global structural features and mode of PAH intercalation are basically identical in both sequence contexts.

Interestingly, changing the lesion linkage carbon on the B[*a*]A-N^6-dA lesion from the bay region of B[*a*]A to a nonbay region position has a profound effect on duplex conformation. The NMR characterization of a duplex having the *S,R*-B[*a*]A-N^6-dA lesion **144** (Chart 52) paired to T in the (5′-...C X A...) sequence context shows the aromatic moiety intercalating inside the helix at the 5′-side of the damaged residue, stacking poorly with flanking residues, and causing the loss of WC alignments at the lesion-containing and 3′-flanking base pairs.[373] On the basis of these observations, it was suggested that the increased perturbation of the

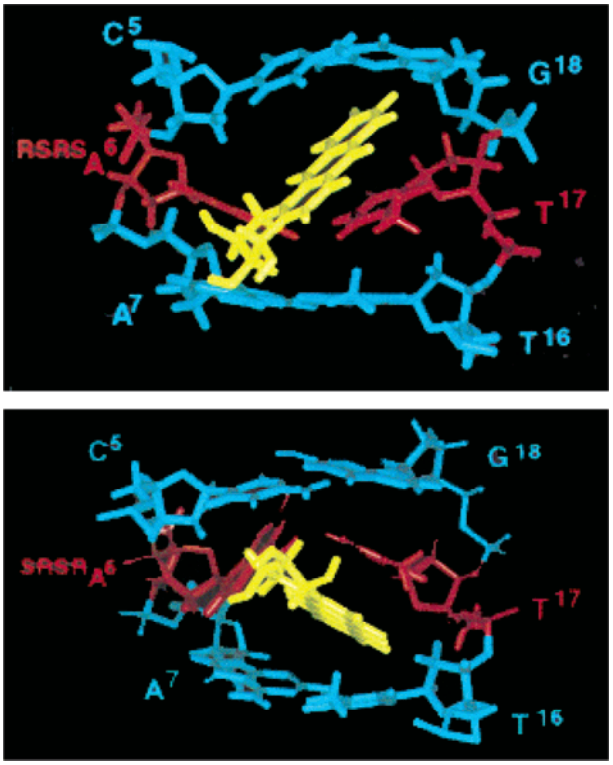
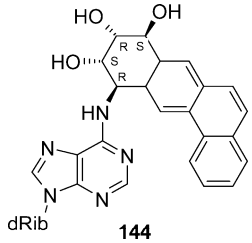


Figure 29. Intercalated structure of (top)*S,R*-B[a]A-N^6-dA **143b** and (bottom) *R,S*-B[a]A-N^6-dA **143a** adducts paired to T. The picture shows the central three-base-pair segment with the major groove prominent. Lesion-containing base pairs are colored red and the anthracene moiety yellow. Reproduced with permission from ref 371. Copyright 1999 American Chemical Society.

Chart 52. 8*S*,11*R*-(*trans-anti*)-B[a]A-N^6-dA Adduct

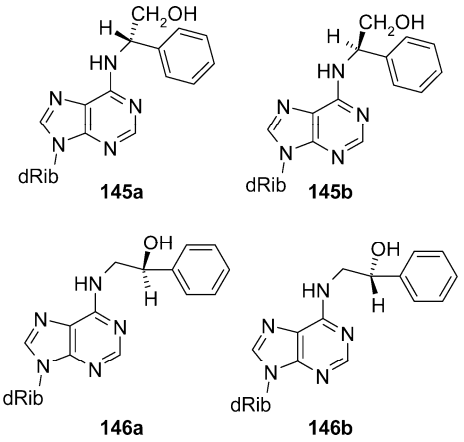


duplex structure makes this nonbay adduct a better substrate for NER, explaining its lack of mutagenicity.

8.4.3. Styrene Oxide (SO)

Early NMR studies have characterized the structure of duplexes having an *R*-α-SO-N^6-dA **145a** or *S*-α-SO-N^6-dA lesion **145b** (Chart 53) paired to T within the (5′-...C X A...) or (5′-...A X G...) sequence context.[374,375] The NMR data show that duplexes having *R*- or *S*-α-SO-N^6-dA lesions are regular right-handed helices with residues in the *anti* range around the χ torsion angle and WC base pair alignments throughout, including at the lesion-containing pair. The *R*-α-SO phenyl ring is external, in the major groove of the helix, directed toward the 5′-end of the damaged strand, and flipping slowly on the NMR time scale. Signal overlap of the lesion phenyl protons, particularly in the (5′-...C X A...) sequence context, hindered exact positioning of *R*-α-SO in the major groove of the helix and complicated the analysis of sequence dependent effects. However, three-dimensional models obtained by rMD show that *R*-α-SO is less edgewise and more perpendicular to the base pair plane in the (5′-...A X G...) context.[374] In contrast to the case of the *R* isomer,

Chart 53. α-SO-N^6-dA 145a,b and *β*-SO-N^6-dA 146a,b Adducts



the phenyl ring of the *S*-α-SO is directed toward the 3′-end of the lesion-containing strand and flipping rapidly on the NMR time scale, establishing a less crowded major groove at the 3′-side of the lesion. Two-dimensional spectra show that *S*-α-SO-N^6-dA adopts a single conformation in the (5′-...A X G...) context but equilibrates between two conformations in the (5′-...C X A...) sequence. In the latter duplex, the structure of the minor conformer could not be refined.[375]

Dramatic sequence dependent effects are observed for the structure of duplexes having the *R*- or *S*-α-SO-N^6-dA lesion mismatched to dC, in the sequence contexts indicated above.[376,377] In the (5′-...C X A...) sequence, the alignment of the *R*-α-SO-N^6-dA·dC mismatch is pH independent and the global stability of the duplex is reduced under acidic conditions. Furthermore, the *R*-α-SO-N^6-dA·dC mismatch induces major distortions in the duplex structure, hindering the refinement of a three-dimensional model. In contrast, the *S*-α-SO-N^6-dA·dC mismatch-containing duplex adopts in the same sequence context a defined right-handed helical structure with WC alignments present at all canonical base pairs. The *S*-α-SO moiety is located in the major groove of the duplex and oriented toward the 3′-end of the adduct-containing strand.[376]

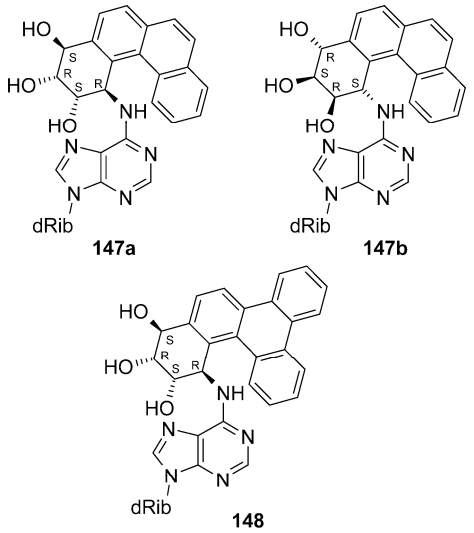
In the (5′-...A X G...) sequence context, the α-SO-N^6-dA·dC duplex is more stable at low pH, presumably due to the formation of a protonated wobble mismatch alignment. In this sequence, there is an inversion of the major grove directionality of the styrene ring whereby the *R*-α-SO points toward the 3′-end of the damaged strand and the *S*-α-SO isomer points in the opposite direction. Once again, the *R*-α-SO-N^6-dA isomer induces the largest perturbation on duplex structure, disrupting hydrogen bonding at the lesion-containing and the 5′-lesion-flanking base pairs.[377] While the biological implications of these structures have to be established, they compile a remarkable example of the type of conformational changes that may result from changing the duplex sequence.

The same group determined the structures of duplexes having an *R*-*β*-SO-N^6-dA **146a** or *S*-*β*-SO-N^6-dA lesion **146b** (Chart 53) paired to T. In these duplexes, the lesion linkage carbon is no longer asymmetric and the separation between adenine and the phenyl ring of SO increases by a CH_2 group. The NMR characterization shows that increasing the tether length of this lesion reduces adduct-induced perturbations of the duplex structure, which is now very close to that of a canonical B-form conformation. Additionally, it eliminates the influence of stereochemistry at the linkage carbon such that both the *R*- and *S*-*β*-SO-N^6-dA·T duplexes have similar conformations. The refined structures show a regular helical structure with WC alignments throughout and the *β*-SO phenyl ring in the major groove of the helix, where it is remarkably exposed to solvent.[378]

8.4.4. Benzo[c]phenanthrene B[c]Ph

As in the case of guanine lesions, studies of isomeric B[*c*]-Ph-N^6-dA lesions were the first to determine the structural consequences of having bulky fjord region adducts on the major groove of the helix.[379,380] The NMR spectra of 11-mer duplexes containing either an *S*,*R*-B[*c*]Ph-N^6-dA **147a**[379] or *R*,*S*-B[*c*]Ph-N^6-dA **147b** lesion[380] (Chart 54) paired to T

Chart 54. *S*,*R*-((+)-*trans*-*anti*)-B[c]Ph-N^6-dA 147a, *R*,*S*-((−)-*trans*-*anti*)-B[c]Ph-N^6-dA 147b, and *S*,*R*-((+)-*trans*-*anti*)-B[g]C-N^6-dA 148



in the (5′...C X C...) sequence context reveal the presence of a predominant conformation in solution, with all residues in the *anti* conformation around the *χ* angle and WC alignment throughout, including at the lesion-containing base pair. The B[*c*]Ph moiety intercalates in a configuration dependent manner, with the *S*,*R* enantiomer appearing at the 5′-side and the *R*,*S* isomer at the 3′-side of the damaged residue and causing a minor buckle of the lesion-containing base pair in both cases. In both isomers, the B[*c*]Ph cyclohexene ring adopts a distorted half-chair conformation, with the aromatic ring system showing a nonplanar propeller-like conformation and stacking partially with residues of the unmodified strand (Figure 23).

The NMR characterization of an 11-mer duplex containing the *S*,*R*-benzo[*g*]chrysene-N^6-dA adduct **148** (Chart 54) paired to T establishes a conformation that is similar to that of the *S*,*R*-B[*c*]Ph-N^6-dA·T duplex.[381] As in the *S*,*R*-B[*c*]Ph-N^6-dA lesion, the lesion linkage carbon of *S*,*R*-B[*g*]Ch-N^6-dA is next to the fjord region of the aromatic moiety. However, benzo[*g*]chrysene has an additional aromatic ring that increases the hydrophobicity of the lesion. The lesion-containing duplex adopts a regular right-handed helical conformation stabilized by the formation of WC base pairs (PDB code 1CR3). The B[*g*]Ch ring system intercalates at the 3′-side of the damaged dA, has edges on both grooves of the duplex, and stacks preferentially with the lesion-partner T residue.

It is interesting to compare the intercalative structures of the isomeric fjord region adducts **147a**,**b** and **148** with those

Table 5. Structures of Duplexes Having Bulky N^6-dA

lesion name[a]	adduct topology[b]	duplex length	sequence context[c]	structural motif	PDB entry	ref
S,*R*-B[*a*]P **142a**	bay	11-mer	5′-(...C A* A...)	B[*a*]P, 5′-side partially intercalated A*, distorted WC pair		360
S,*R*-B[*a*]P **142a**	bay	11-mer	5′-(...A A* G...)	B[*a*]P, 5′-side intercalated A*, WC pair		361
S,*R*-B[*a*]P **142a**	bay	12-mer	5′-(...A A* G...)	B[*a*]P, 5′-side intercalated A*, WC pair	1FYY	362
R,*R*-B[*a*]P **142c**	bay	11-mer	5′-(...C A* C...)	B[*a*]P, 5′-side partially intercalated A*, distorted WC pair	1AXV	363
R-B[*a*]P[d]	bay	11-mer	5′-(...C A* C...)	B[*a*]P, 5′-side intercalated A*, WC pair	1N8C	364
R,*R*-B[*a*]P[d]**142d**	bay	11-mer	5′-(...C A* C...)	B[*a*]P, 5′-side intercalated A*, WC pair		365
S,*S*-B[*a*]P[e] **142e**	bay	11-mer	5′-(...C A* C...)	B[*a*]P, 3′-side intercalated A*(*syn*), weakly hydrogen bonded	1JDG	366
R,*S*-B[*a*]A **143a**	bay	11-mer	5′-(...C A* A...)	B[*a*]A, 5′-side intercalated A*, distorted WC pair		370
S,*R*-B[*a*]A **143b**	bay	11-mer	5′-(...C A* A...)	B[*a*]A, 3′-side intercalated A*, distorted WC pair	1DL4	371
R,*S*-B[*a*]A **143a**	bay	11-mer	5′-(...A A* G...)	B[*a*]A, 5′-side intercalated A*, distorted WC pair	1I7V	372
S,*R*-B[*a*]A **143b**	bay	11-mer	5′-(...A A* G...)	B[*a*]A, 3′-side intercalated A*, distorted WC pair		372
S,*R*-B[*a*]A **144**	nonbay	11-mer	5′-(...C A* A...)	B[*a*]A, 5′-side intercalated A*, *anti* not hydrogen bonded	1DJD	373
R-α-SO **145a**		11-mer	5′-(...A A* G...)	major groove, 5′-end oriented A*, WC pair		374
R-α-SO **145a**		11-mer	5′-(...C A* A...)	major groove, 5′-end oriented A* , WC pair		374
S-α-SO **145b**		11-mer	5′-(...A A* G...)	major groove, 3′-end oriented A*, WC pair		375
S-α-SO **145b**		11-mer	5′-(...C A* A...)	major groove, 3′-end oriented A*, WC pair		375
R-*β*-SO **146a**		11-mer	5′-(...C A* A...)	major groove, no preference A*, WC pair	1K5E	378
S-*β*-SO **146b**		11-mer	5′-(...C A* A...)	major groove, no preference A*, WC pair	1K5F	378
S,*R*-B[*c*]Ph **147a**	fjord	11-mer	5′-(...C A* C...)	B[c]Ph, 5′-side intercalated A*, WC pair		379
R,*S*-B[*c*]Ph **147a**	fjord	11-mer	5′-(...C A* C...)	B[c]Ph, 3′-side intercalated A*, WC pair		380
S,*R*-B[*g*]C **148**	fjord	11-mer	5′-(...C A* C...)	B[g]C, 5′-side intercalated A*, WC pair	1CR3	381

[a] B[*a*]P, benzo[*a*]pyrene; B[*a*]A, benzo[*a*]anthracene; α-SO, α-styrene oxide; *β*-SO, *β*-styrene oxide; B[*c*]Ph, benzo[*c*]phenatrene; B[*g*]C, benzo[*g*]chrysene. [b] Indicates PAH topology and configuration at the linkage site. In the B[*a*]P case, the configuration of the next asymmetric carbon is also listed. [c] Lesions always paired to T. [d] Cyclohexene ring reduced at positions 7 and 8. [e] Derived from the *syn* diolepoxide.

Table 6. Structure Types of Mutagenic Intermediates Containing Bulky N^6-dA Adducts

lesion name[a]	adduct topology[b]	lesion partner	duplex length	sequence context	structural motif	ref
S,*R*-B[a]P **142a**	bay	dG	9-mer	5′-(...C A* C...)	B[*a*]P, 5′-side partially intercalated A*·dG, not hydrogen bonded dG displaced in the major groove	367
R,*S*-B[a]P **142b**	bay	dG	9-mer	3′-(...C A* C...)	B[*a*]P, 5′-side partially intercalated major form: A*(*syn*)·dG two H-bonds minor form: A*(*anti*)·dG two H-bonds	368 369
R-α-SO **145a**		dC	11-mer	5′-(...C A* A...)	highly distorted. SO not defined A*·dC, not hydrogen bonded	376
S-α-SO **145b**		dC	11-mer	5′-(...C A* A...)	SO, major groove, 3′-end oriented A*·dC, not hydrogen bonded	376
R-α-SO **145a**		dC	11-mer	5′-(...A A* G...)	SO, major groove, 3′-end oriented A*·dC, protonated wobble pair	377
S-α-SO **145b**		dC	11-mer	5′-(...A A* G...)	SO, major groove, 5′-end oriented A*·dC, protonated wobble pair	377

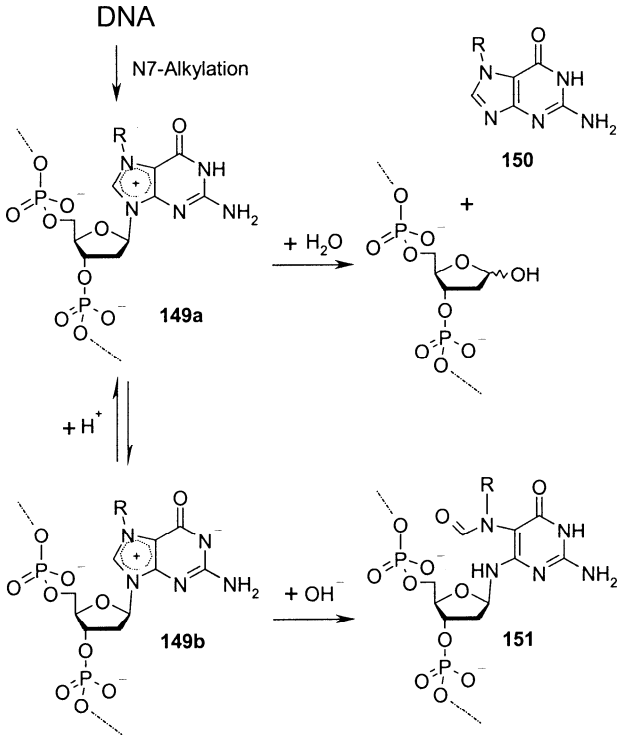
[a] B[*a*]P, benzo[*a*]pyrene; α-SO, α-styrene oxide. [b] Indicates PAH topology and configuration at the linkage site. In the B[*a*]P case, the configuration of the next asymmetric carbon is also listed.

of bay region N^6-dA lesions discussed previously.[360,369−372] In all these cases, the configuration of the benzylic linkage carbon determines the orientation of PAH intercalation in the helix. *R* isomers intercalate at the 5′-side of the damaged dA residue whereas *S* isomers do at the 3′-side. However, bay region lesions induce larger structural perturbations at the lesion site and reduce the thermal stability of the duplex, and a minor conformation is frequently present. In contrast, due to their intrinsic nonplanarity, fjord region lesions distort the helical structure to a lesser extent and PAH intercalation does not change or slightly increase the duplex thermal stability. These structural differences have been associated with the resistance of fjord region PAH adducts to nucleotide excision repair.[351] Table 5 lists structure types of bulky-N^6-dA·T duplexes, and Table 6 lists those of mismatches having bulky N^6-dA adducts.

8.5. N^7-Deoxyguanosine Lesions

Alkylation of peripheral amino groups causes only small alterations of the electronic structure of purine moieties, their hydrogen-bonding ability, and the basicity of the nucleobases. The major effects of exocyclic alkylation damages are

Scheme 15. N^7-Guanine Alkylation Damage and Its Degradation Pathway



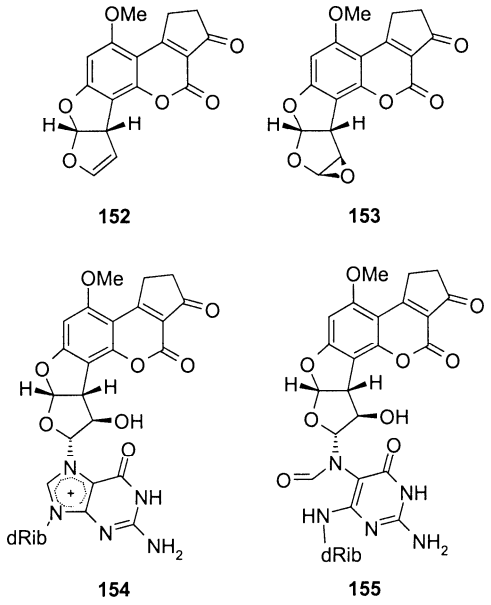
associated with the steric clashes caused by alkyl or aryl substituents or with their intercalation properties. In contrast, alkylation of the endocyclic guanine N7 nitrogen gives rise to cationic lesions **149a**, which exist in equilibrium with the more stable neutral zwitterionic form **149b** (Scheme 15). Since the pK_a of **149a** is close to 7.0,[382] similar amounts of both lesion forms are present at neutral pH. Lowering the pH shifts the equilibrium toward **149a**, which undergoes fast glycosidic bond cleavage yielding an N^7-alkylguanine product **150** and an AP site. In contrast, a pH increase facilitates a nucleophilic attack on the imidazolium ring, giving rise to alkylated FAPY lesions **151**.[383]

Probably due to an unfavorable conformation of the epoxy-containing cyclohexen ring, most PAH diol epoxides do not react with guanine N7 in DNA. However, reactive species generated from some micotoxins, including aflatoxin B_1 and related oxidatively activated alkenes, can efficiently alkylate the guanine N7 position. To date, NMR solution structures of duplexes containing several bulky N7 guanine lesions have been reported.

8.5.1. Aflatoxins B_1, Sterigmatocystin, and Pluramycin Lesions

In contrast to PAH diol epoxides that have the epoxy cyclohexenyl ring directly attached to a polyaromatic ring system, the oxidizable etheno fragment of aflatoxin and related compounds is attached to an intercalating moiety via a cyclic or linear aliphatic linker. This allows proper positioning of the electrophilic carbon of these compounds for alkylation of the guanine N7 position. Aflatoxin B_1 **152** (AFB_1), a mycotoxin produced by several species of the genus *Aspergillus*, contaminates food supplies worldwide. AFB_1 induces mutations in bacterial and mammalian cells, is an established hepatocarcinogen in rodents, and has been correlated with the onset of liver cancer in humans (review in ref 384). Cytochrome P450 enzymes metabolize AFB_1 to an *exo*-8,9-epoxide **153**, which then reacts with DNA, generating the cationic AFB_1-N^7-dG lesion **154**, with the subsequent formation of an AP site or an AFB_1-formamidopyrimidine (AFB_1-FAPY) adduct **155** (Chart 55).

Chart 55. Aflatoxin-B_1 152, Its Activated Form 153, and N^7 Guanosine Adducts 154 and 155



Early studies characterized the reaction products of AFB_1-8,9-epoxide with two short self-complementary duplexes, d(A T C G A T)$_2$ and d(A T G C A T)$_2$.[385,386] Analysis of UV melting profiles showed that formation of an AFB_1-N^7-dG lesion results in a small increase of duplex thermal stability. One-dimensional NMR spectra indicated that, in agreement with the 2:1 ABF_1:DNA stoichiometry observed during preparation of the sample, only the d(A T G C A T)$_2$ duplex maintains the self-complementary symmetry after formation of the AFB_1-N^7-dG lesion. In contrast, the number of proton signals doubles in the spectra of the d(A T C G A T)$_2$ duplex, which had shown a 1:1 stoichiometry of lesion formation. Both duplexes adopt right-handed helical conformations with small perturbations localized at the adducted guanine. Chemical shift perturbations and NOE experiments indicated that ABF_1 intercalates at the 5′-side of the damaged dG residue with its methoxy-cyclopentenone moiety located on the minor groove and the fused furan rings on the major groove of the helix. Due to low thermal stability, refined structural models of these short duplexes could not be produced. Subsequent NMR studies using longer AFB_1-N^7-dG lesion-containing duplexes, with the 6-mer sequences embedded at their centers, allowed the computation of their solution structures.[387,388] Both duplexes are right-handed helices with residues in an *anti* conformation around the χ torsion angle and WC alignments throughout, including at the AFB_1-N^7-dG·dC base pair. AFB_1 intercalates at the 5′-side of the damaged residue with its methoxy-cyclopentenone edge directed toward the minor groove of the helix. The duplex is moderately unwound at the intercalation step with a significant increase in the helical rise, which is about twice the value commonly observed on unmodified DNA. Base pairs flanking the intercalated moiety buckle in the opposite direction, forming a pocket that fits the shape of ABF_1 (Figure 30). In both sequence contexts, the structures are very similar, suggesting that duplex composition plays a

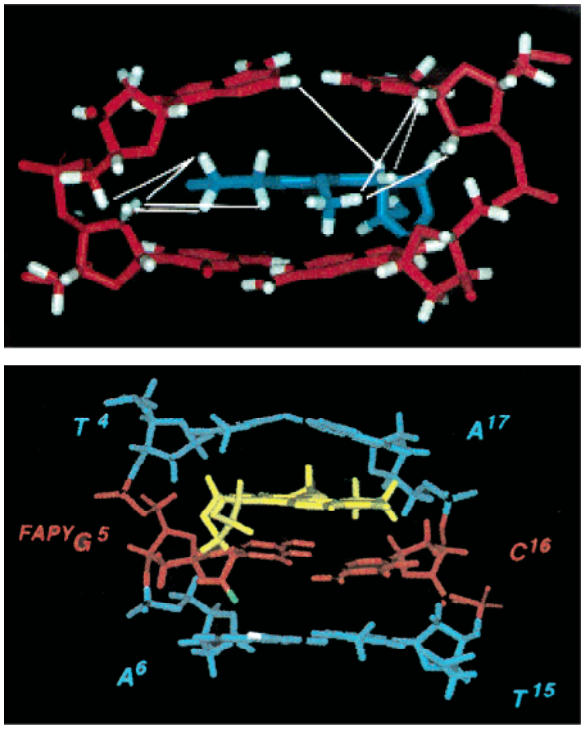


Figure 30. Intercalated structures of aflatoxin B1 lesions. The top panel shows the AFB1-N^7-dG lesion **154** paired to dC, seen with the minor groove prominent and the aflatoxin moiety in blue. Adapted with permission from ref 387. Copyright 1998 American Chemical Society. The bottom panel displays the AFB1-FAPY-dG adduct **155** paired to dC, shown with the major groove prominent and the aflatoxin moiety in yellow. Adapted with permission from ref 393. Copyright 1998 American Chemical Society.

minor role in the structure of AFB_1-containing DNA. Similarly, the presence of a dC·dA mismatch next to AFB_1-N^7-dG has only a slight influence on the structure of the AFB_1-containing duplex. The NMR characterization of a 10-mer duplex containing an AFB_1-N^7-dG·dC base pair flanked at its 5′-side by a dC·dA mismatch shows a right-handed helical structure, stabilized by WC alignment for all canonical base pairs.[389] The ABF_1 moiety intercalates following the precepts described above, including an increase in the helical rise, that almost doubles the canonical value, partial unwinding of the helix, proper stacking of AFB_1 inside the duplex, and a small buckle of flanking base pairs. In addition, the presence of ABF_1-N^7-dG stabilizes the dC·dA mismatch-containing duplex, which has a higher T_m than that of the unmodified dC·dA control sample. The mismatched dA residue is displaced toward the minor groove of the helix, forming a wobble-type dC·dA alignment, which is proposed to be exchanging between nonprotonated and adenine-N1 protonated forms (Figure 31).

Two additional studies, describing the solution structure of mutagenic duplex intermediates, further support the principle that AFB_1 intercalates at the 5′-side of the damaged residue. The NMR spectra of an ABF_1-N^7-dG·dA mismatch-containing duplex indicate the presence of a single conformation at neutral pH, with all residues in the *anti* conformation around the χ torsion angle, WC alignments at all canonical base pairs, and AFB_1 intercalation at the 5′-side of the lesion. The mismatched dA residue is displaced toward the major groove of the helix and does not participate in hydrogen bonding. Despite the absence of hydrogen bonds across the ABF_1-N^7-dG·dA mismatch, the T_m of this duplex is higher than that of the unmodified dG·dA control, indicating that an increase of duplex thermal stability is a general corollary of ABF_1 intercalation[390] (Figure 31). The other study determined the conformation of an ABF_1-N^7-dG·dC-containing duplex, with an additional dA residue positioned 5′ to the lesion partner dC (+1 insertion mutation model).[391] The NMR data establish the presence of a predominant right-handed helical conformation with WC alignments throughout the duplex, including at the lesion-containing base pair, and intercalation of ABF_1 at the 5′-side of the lesion. The extra dA residue also stacks inside the helix, with its pyrimidine ring partially inserted between the ABF_1 moiety and the damaged guanine ring. The refined model shows that this noncoplanar alignment significantly bends the helix axis (36°) and unwinds the duplex (24°), facilitating stacking interactions at the lesion site. Intercalation of ABF_1 also increases the stability of this duplex, which has a 5 °C higher T_m than the unmodified +1 insertion mutation model control.

As mentioned above, the imidazole ring of the AFB_1-N^7-dG adduct opens under basic pH conditions, forming AFB_1-FAPY **155**, which has been shown to be more toxic and persist longer inside the cell than the original lesion.[384] Thus, following the same approach implemented for obtaining sterigmatocystin-FAPY adducts,[392] base-catalyzed opening of an AFB_1-N^7-dG lesion already present in a double-stranded duplex allowed the generation of an AFB_1-FAPY-dG·dC sample for structural studies.[393] A recently formed AFB_1-

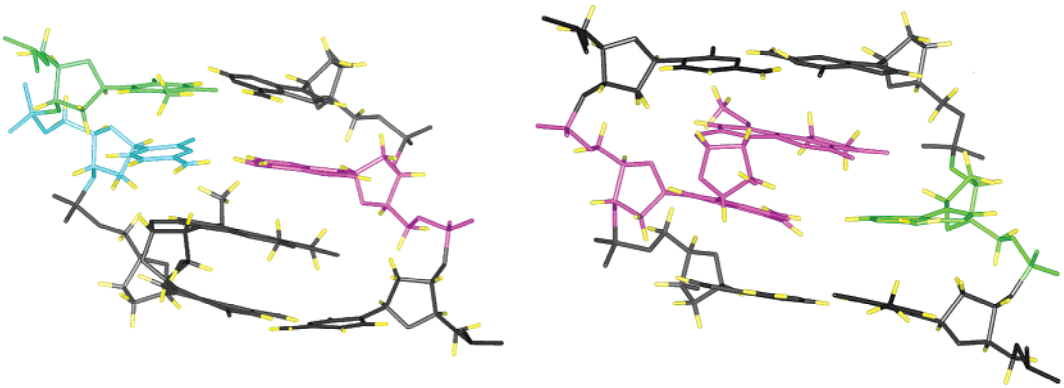
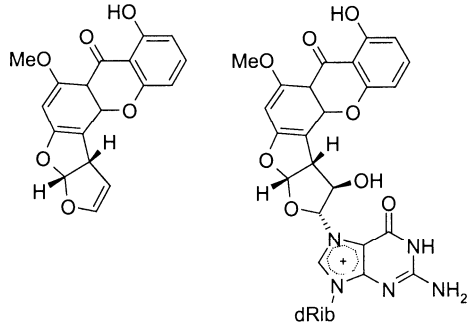


Figure 31. Structure of mismatched duplexes containing the AFB1-N^7-dG lesion **154**. The pictures show the central three-base-pair segment with the major groove prominent. The left panel displays the intercalated AFB1 moiety flanked by a 5′ dC·dA mismatch. The right panel shows the damaged residue forming an AFB1-N^7-dG·dA mismatch. Adapted with permission from refs 389 and 390. Copyright 2003 and 2002 American Chemical Society.

FAPY-dG lesion exists as a mixture of rotational isomers around the C5−N5 bond, but after several days at 4 °C, they convert to a single conformer, retaining a β configuration at the anomeric carbon. The NMR characterization of this sample establishes the presence of a regular right-handed helix with residues in the *anti* conformation around the χ torsion angle and WC base pair alignments throughout the duplex, including at the AFB_1-FAPY-dG·dC base pair. The AFB_1 moiety intercalates at the 5′-side of the lesion, very much as the AFB_1-N^7-dG did, doubling the helical rise value at this step, unwinding the duplex by around 15°, and causing a buckle in the 5′ lesion-flanking base pair (Figure 30). Remarkably, the T_m of the AFB_1-FAPY-dG·dC duplex is 15 °C higher than that of the unmodified control duplex. Analysis of the refined structure suggested that the increase in duplex thermal stability might result from better stacking of AFB_1 inside the duplex and the formation of an additional interstrand hydrogen bond between the FAPY formyl oxygen and an amino proton of a flanking dA residue. The authors hypothesized that the increased thermal stability of AFB_1-FAPY-dG might cause significant differences in the processing of AFB_1 adducts *in vivo*. Interestingly, a recent study using a host cell reactivation assay found that while both lesions, AFB_1-N^7-dG and AFB_1-FAPY-dG, are mainly repaired by NER, cells could tolerate AFB_1-N^7-dG adducts better.[394]

Sterignatocystin **156** (Stg) (Chart 56), the biosynthetic precursor of aflatoxins, is another mycotoxin produced by *Aspergillus* species that may contaminate world food supplies although to a lesser extent.

Chart 56. Sterignatocystin 156 and Its N^7-dG Adduct 157

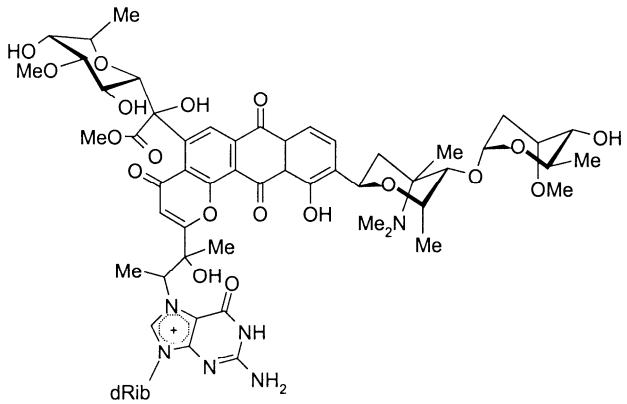


156 **157**

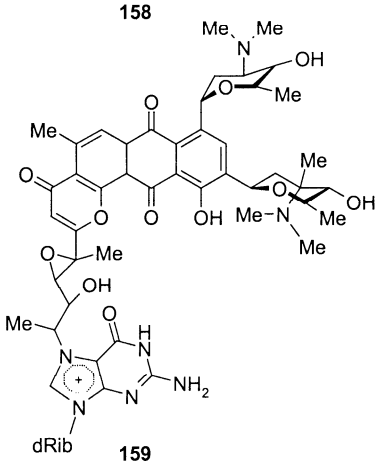
Stg-N^7-dG **157** causes mutations in bacteria[395] and induces liver cancer in rodents,[396] but its biological activity is about 10-fold lower than that of AFB_1. Interest in the structure of duplexes having the Stg-N^7-dG lesion was born in an attempt to explain biological differences between these two structurally similar lesions. The NMR characterization of the lesion-containing d(A A T G C A T T)$_2$ self-complementary duplex[397] showed that the NMR spectra remain degenerated after formation of the Stg-N^7-dG adduct, indicating the presence of two symmetrically positioned Stg moieties. Analysis of NOESY data established formation of WC alignments throughout the duplex and displayed right-handed sequential connectivities that are interrupted between the Stg-N^7-dG and 5′-lesion-flanking residues. The distance-refined model shows a regular duplex structure with Stg intercalated inside the helix at the 5′-side of the lesion but without disrupting the alignment of flanking base pairs. In contrast to the AFB_1 case discussed above, the furan linkage carbon of Stg-N^7-dG is out of the guanine plane, allowing the almost perfect stacking of Stg between flanking base pairs. On the basis of this model, the authors argued that upon alkylation the hybridization of guanine N7 changed from sp^2 to sp^3, facilitating the smooth insertion of Stg in the duplex. However, the fast solvent exchange of the Stg-N^7-dG H8 proton, similar to that observed for N^7-methylated guanosine,[398] strongly supports the presence of an aromatic imidazolium ring in **157** with an sp^2 guanine N7. Furthermore, no reasonable chemical structure can be proposed for cation **157** having guanine N7 in the sp^3 state. Although the authors did not disclose the parametrization of the damaged residue, the most probable explanation for the lack of N7 planarity could be the use of nonrelevant parameters during RMD calculations. Thus, the structure of the Stg-N^7-dG·dC-containing duplex should be reexamined further.

Streptomyces species synthesize a large number of antibiotics of the pluramycin family, which show good *in vitro* antitumor activity against a wide variety of tumor cell lines.[399] Chemically, they consist of a planar aromatic ring system, related to the anthraquinone chromophore, substituted with different sugar and epoxide functionalities. It has been established that the aromatic chromophore intercalates inside the helix, with sequence specificity determined by the saccharide substitution[400,401] and, subsequently, the epoxide group alkylates the N7 guanine position, producing a cationic lesion. NMR studies[402−406] have characterized the solution structures of duplexes having altromycin B **158** or hedamycin **159** adducts, which differ mainly by the absence of a C8 sugar substitution on the aromatic chromophore of the former (Chart 57) The NMR spectra of the self-complementary d(G A A G T A C T T C)$_2$ having an altromycin B lesion on the internal dG establish the presence of two drug molecules

Chart 57. Altromycin B-N^7-dG 158 and Hedamycin-N^7-dG 159 Adducts



158



159

per duplex.[403] As in the case of AFB_1, the H8 proton of the damaged residue exchanges fast with solvent and resonates at 9.7 ppm, downfield from the normal 7−7.8 ppm range observed for dG, indicating the presence of a cationic altromycin B-N^7-dG lesion. Sequential NOE connectivities of the duplex are interrupted at the 5′-side of the damaged residue, indicating the site of chromophore intercalation inside the helix. In addition, NOE cross-peaks between drug and duplex protons position the C10 disaccharide in the minor groove while the C5 sugar locates in the major groove of the helix. The refined structural model shows a regular right-handed helix with WC alignments throughout the duplex, including at the lesion-containing base pair. Other than an increase in the helix rise and a small buckle at the lesion-containing pair, insertion of altromycin B chromophore causes no perturbation of the duplex structure (Figure 32). Similar three-dimensional structures were

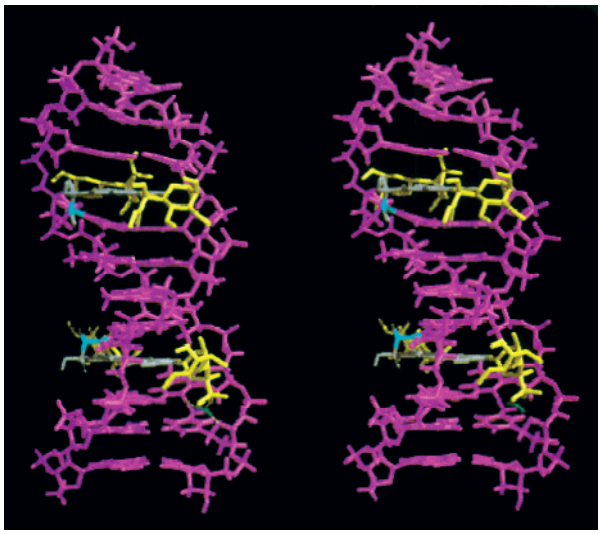


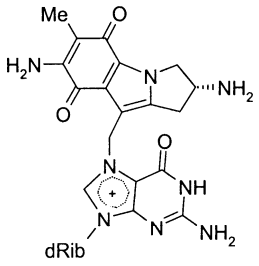
Figure 32. Stereoview of a decamer duplex containing two altromycin B-N^7-dG lesions **158**. Reproduced with permission from ref 403. Copyright 1995 American Chemical Society.

derived for the related hedamycin-N^7-guanine lesion in the 5′-(...T G*)[404] and 5′-(...C G*) sequence contexts.[405,406] In all the structures, the aromatic chromophore intercalates inside the helix at the 5-side of the damaged dG residue without causing major perturbations to the DNA structure. The amino sugars reside in the minor groove of the duplex, but their precise orientation varies among these studies.[405,406] In summary, duplexes containing pluramycin antibiotics covalently bound to guanine N7 adopt 5′-side intercalative structures similar to those of the ABF_1 family.

8.5.2. 2,7-Diaminomitosene

It has been established that treatment of mouse mammary tumor cells with mitomycin C produces several guanine adducts, including the guanine-N^7 alkylation product of 2,7-diaminomitosene (2,7-DAM), a product of mitomycin C degradation.[407] The NMR solution structure of the self-complementary d(G T G X T A T A C C A C) dodecamer, where X is a 2,7-DAM-N^7-dG lesion **160** (Chart 58), has been recently reported.[408] The NMR spectra show that the sample retains the C_2 symmetry after formation of the lesion, indicating the presence of two 2,7-DAM molecules per duplex. All residues are in the *anti* conformation around the χ angle and form WC alignments throughout the duplex, including at the lesion-containing pair. Sequential NOE

Chart 58. 2,7-Diaminomitosene-N^7-dG Adduct



160

connectivities between base and sugar H1′ protons are present without interruptions along the entire length of the duplex, including at the lesion site, suggesting that the 2,7-DAM does not intercalate inside the helix. Furthermore, several NOE cross-peaks between 2,7-DAM and duplex protons position the mitosene moiety in the major groove at the 3′-side of the damaged residue. The refined three-dimensional model (PDB code 1JO1) shows the 2,7-DAM moieties in the major groove of the duplex with both faces exposed to solvent but without perturbing the overall B-form helical structure (Figure 33). Interestingly, although 2,7-DAM-N^7-dG is a

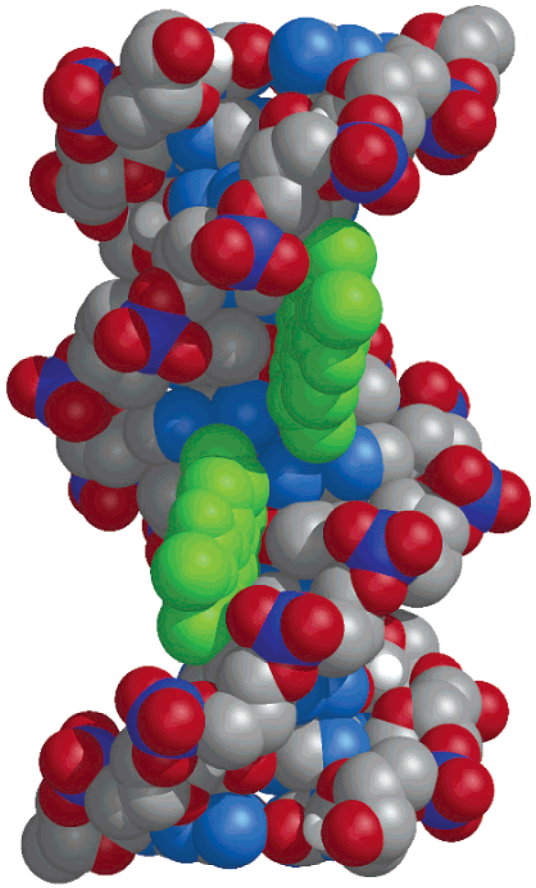


Figure 33. Structure of a dodecamer duplex having two 2,7-diaminomitosene-N^7-dG **160** lesions paired to dC. The figure shows the major groove prominent with duplex atoms colored by type and mitosene moieties in green. Prepared from PDB entry 1JO1[408] using MidasPlus.[556]

predominant lesion formed *in vivo* after mitomycin C treatment, it is 2 orders of magnitude less toxic than mitomycin C-N^2-dG adducts.[409] The lack of 2,7-DAM-N^7-dG toxicity is also in sharp contrast to the cases of other bulky N^7-dG lesions, such as AFB_1 or pluramycin antibiotics,

which can damage cells irreversibly. The authors hypothesized that the unusual major groove-aligned structure of the 2,7-DAM-N^7-dG adduct may be efficiently repaired *in vivo* or properly replicated by high-fidelity DNA polymerases, explaining its lack of toxicity. Further research is needed to address this possibility.

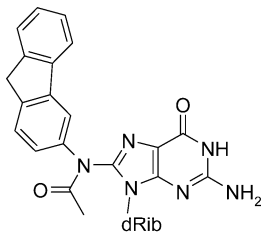
8.6. C^8-Purine Lesions

A large number of nitroaromatic compounds, heterocyclic amines produced by cooking red meat or fish, the widely studied 2-aminofluorene, and related aromatic amines, all react with purine bases producing, predominantly, bulky C8 guanine lesions. None of these compounds is active *per se*, and they must be activated prior to the formation of a covalent lesion. As discussed previously, metabolic activation of these agents follows a two-step process (Scheme 11), with similarities and differences that can partially explain their different biological activities and cell specificities. Nitroaromatic compounds, such as 1-nitropyrene or 2-nitrofluorene, undergo the enzymatic reduction of their nitro group to *N*-hydroxylamine, a reaction that in humans is catalyzed by xanthine oxidase or microsomal NADPH-cytochrome *C* reductase.[410] In the case of aromatic and heterocyclic amines, the cytochrome P450 enzyme system oxidizes their amino group to more reactive *N*-hydroxylamine derivatives.[320,411] Independent of their genesis, *N*-hydroxylamine products experience a second cellular activation step by phase II drug-metabolizing enzymes, which generate highly reactive *O*-substituted *N*-hydroxylamine intermediates that form C8 guanine lesions. Like other bulky adducts, C8 guanine lesion adducts severely affect vital cellular processes, including gene transcription and DNA replication, and their presence in chromosomal DNA is associated with the development of many types of cancer in laboratory animals and humans.[320,410−414] In fact, the structural characterization of 2-aminofluorene lesions has been the paradigm to understand at the molecular level processes of chemical mutagenesis. An extensive review specifically dedicated to the structure of DNA having aromatic amine adducts has been published.[415] To describe here the large amount of structural studies performed on duplexes containing bulky C8 adducts, we will group them by type, as we did with other bulky lesions.

8.6.1. 2-Acetylaminofluorene (2-AAF)

To date, there is just one published structure of a DNA duplex containing a 2-AAF-C^8-dG **161** lesion (Chart 59)

Chart 59. 2-AAF-C^8-dG Adduct



161

paired to dC.[416] Analysis of the NMR data shows the presence of two conformations in slow exchange on the NMR time scale. The NOESY spectra indicate that the predominant form of the duplex (about 70% population) has all unmodified residues in the *anti* conformation around the χ torsion angle with concomitant formation of WC base pairs. The χ torsion angle of 2-AAF-C^8-dG is in the *syn* range, and its partner residue remains *anti* with no evidence of hydrogen bonding across the lesion-containing base pair. A distance−energy-minimized/unrestrained molecular dynamics model of the main conformer (Figure 34) shows the lesion

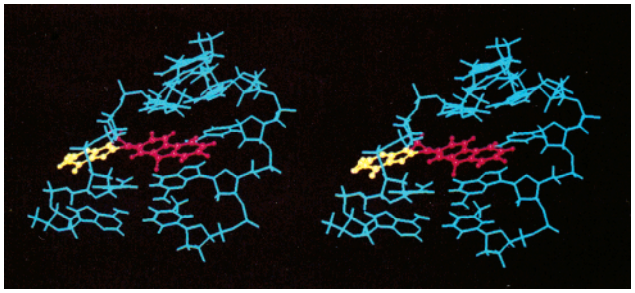


Figure 34. Stereoview of a DNA duplex having the 2-AAF-C^8-dG lesion **161** paired to dC. The picture shows a five-base-pair segment with the damaged dG residue in yellow displaced into the major groove, allowing the intercalation of the 2-AAF moiety in red. Adapted with permission from ref 416. Copyright 1993 American Chemical Society.

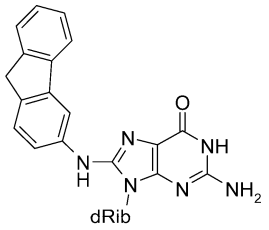
partner residue inside the helix and the damaged guanine base displaced into the major groove, in agreement with the base-displaced and insertion-denaturation models proposed more than 20 years earlier.[417−419] The 2-AAF moiety intercalates inside the duplex and stacks with the 5′-flanking base pair. The acetyl group of 2-AAF resides in the major groove of the helix with its oxygen atom *trans* to the C8 guanine position. In agreement with the conformational heterogeneity described in this paper, a relatively recent ^{19}F NMR study of a damaged duplex having a 7-fluoro-AAF-C^8-dG shows the presence of two fluorine signals in slow exchange on the NMR time scale.[420] On the basis of the analysis of solvent-induced isotope and magnetic anisotropy effects, the authors conclude that the damaged duplex adopts base-displaced structures in both conformations, differing only in the *cis* or *trans* orientation of the acetyl oxygen atom with respect to guanine C8.

Two additional NMR studies have characterized mutagenic duplex intermediates having an 2-AAF-C^8-dG lesion opposite to a targeted -1Δ or -2Δ deletion on the complementary strand.[421,422] At the working temperature, the one-base deletion 2-AAF-C^8-dG duplex shows two conformations, a fact that limited the resolution of the spectra, voiding the refinement of its three-dimensional structure. However, on the basis of proton chemical shift values and the presence (or absence) of NOE interactions, the authors concluded that the minor population of the duplex has the 2-AAF moiety intercalated inside the duplex, while it is external and points to the 5′-direction of the lesion-containing strand, in the predominant form. Similarly, the two-base deletion 2-AAF-C^8-dG duplex showed conformational heterogeneity in solution, a fact that precluded the refinement of a three-dimensional model. On the basis of proton chemical shift values and a limited NOE data set, the authors proposed that all duplex conformations have the fluorene moiety intercalated inside the helix and that, in the predominant form (about 80% population), a *syn* 2-AAF-C^8-dG residue and the 3′ lesion-flanking cytosine are extruded from the helix. Interestingly, the 2-AAF-C^8-dG lesion increases the thermal stability of both duplexes, which have a 15 °C higher T_m than the unmodified base deletion control samples, suggesting a strong fluorene-stabilizing effect that seems to be strikingly independent of its conformation.

8.6.2. 2-Aminofluorene (2-AF)

Contrary to the 2-AAF case, extensive NMR studies have characterized the structure of DNA duplexes having a 2-AF-C^8-dG lesion **162** (Chart 60) paired to dC, in different sequence contexts.[423−428] All the experimental data recorded to date show that 2-AF-C^8-dG-containing duplexes exist as a mixture of two conformations in slow exchange on the NMR time scale, with a lifetime of a few milliseconds at 30 °C. The relative population of each state depends on duplex composition, with ratios varying from 1:1 to 9:1 among the sequences studied. On the basis of the large chemical shift changes observed for the AF protons between the two forms and some specific NOE cross-peaks between duplex and AF protons, the authors of early reports concluded that the fluorene moiety is exchanging between external and intercalated conformations.[423−425] In the external conformation, all residues are in an *anti* range around the χ torsion angle with WC base pair alignments throughout the duplex. The AF moiety resides in the major groove of the helix without affecting the global B-form conformation of the duplex. The energy-minimized model of the external conformer showed that the AF moiety is in the major groove of the duplex, with its C9-methylene edge pointing to the 5′-side of the lesion-containing strand (Figure 35).[423,424] In the intercalated conformation, the fluorene moiety is inside the helix with its C9-methylene edge pointing toward the major groove of the duplex. AF intercalation takes place without disruption of lesion-flanking base pairs, which remain WC hydrogen bonded, stacking properly with the aromatic moiety, and displaces the lesion partner dC residue into the major groove of the helix. Since the absence of the 2-AF-C^8-dG H8 proton impedes a clear determination of the χ angle of the damage residue, its position in the structure is ill defined. Further, restrained molecular mechanics generated three-dimensional models that equally fit the experimental data with the 2-AF-C^8-dG residue in the *syn* or *anti* conformation, suggesting that both conformations are equally probably. On the basis of kinetic considerations, by which a bulky 2-AF-C^8-dG residue would be unlikely to rotate to a *syn* torsion angle conformation within a 10 ms lifetime, the authors concluded that the damaged residue remains *anti* in the intercalated conformation also and, thus, appears in the minor groove of the helix (Figure 35).[423,424] Two additional studies confirmed the main properties of these models and completed the characterization of the AF-intercalated structure.[427,428] On the basis of the downfield shift value of the 2-AF-C^8-dG H2′ proton (3.71 ppm), the authors concluded that in the AF-intercalated structure the damaged dG residue adopts a *syn* conformation around the χ angle and resides in the major groove of the helix (Figure 36).[427]

Chart 60. 2-AF-C^8-dG Adduct



162

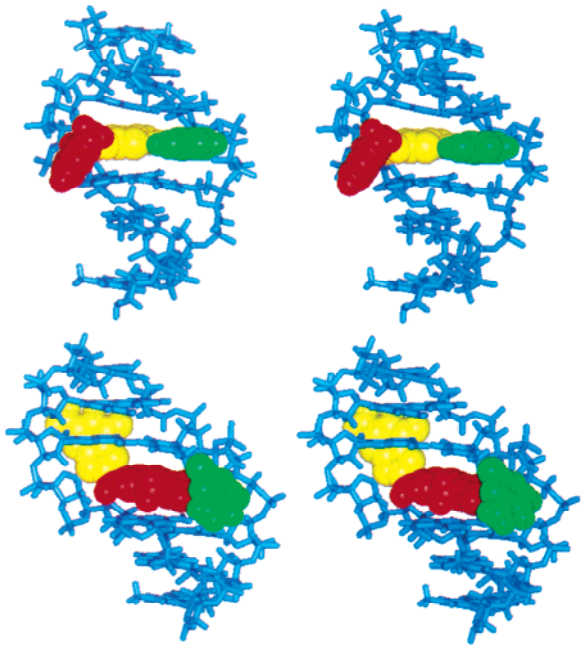


Figure 35. Stereoviews of duplex DNA containing a 2-AF-C^8-dG adduct **162** paired to dC. The top panel shows the major groove external conformation, and the bottom panel, the based-displaced intercalated structure. The picture displays the damaged dG base in yellow, its partner dC residue in green, and the 2-AF moiety in red. Reproduced with permission from ref 424. Copyright 1994 American Chemical Society.

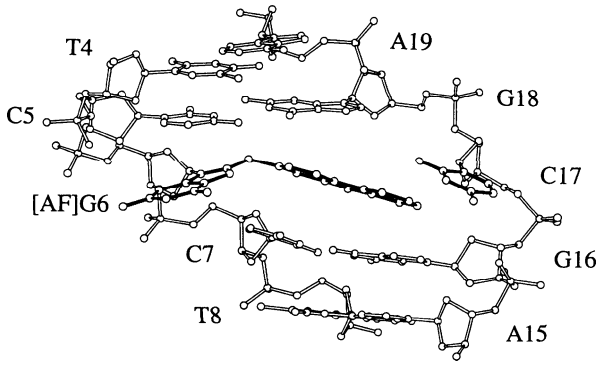


Figure 36. Base-displaced intercalated conformation of 2-AF-C^8-dG after ref 427. The picture displays the five-base-pair segment with the damaged residue located in the major groove of the duplex, after adopting a *syn* conformation around the glycosidic torsion angle. Reproduced with permission from ref 427. Copyright 1998 American Chemical Society.

NMR studies have established also the solution structure of several AF-containing samples that represent replication intermediates of the base-substitution and −1Δ and −2Δ deletion mutation processes.[429,430−433] The observation that dAMP and, to a lesser extent, dGMP are inserted opposite 2-AF-C^8-dG lesions[434,435] gave the rationale for the structure determination of 11-mer duplexes containing a centrally located 2-AF-C^8-dG·dA or 2-AF-C^8-dG·dG mismatch flanked by dC·dG base pairs.[429,430] In contrast to the case of 2-AF-C^8-dG·dC duplexes discussed previously, the NMR spectra of these mismatches show a single set of relatively sharp proton signals, suggesting the presence of a unique duplex conformation in solution. Assignment of the NOESY spectra established that all the unmodified residues adopt an *anti* conformation around the χ torsion angle and form regular WC base pair alignments. The 2-AF-C^8-dG H2′ proton resonates around 3.6 ppm, downfield from the normal H2′−H2″ range, suggesting that the χ torsion angle of 2-AF-C^8-dG is *syn* and, thus, locating the AF ring in the minor groove

Table 7. Structure of Duplexes Having Bulky C^8-dG Lesions

lesion site[a]	duplex length	sequence context	predominant structural motif	PDB entry	ref
AAF-dG·dC	9-mer	5′-(CG*C)·5′-(GCG)	dG*, *syn* major groove displaced dC, intrahelical AAF, intercalated		416
AF-dG·dC[a]	10-mer	5′-(AG*G)·5′-(CCT)	dG*·dC WC paired AF, major groove external, 5′-oriented		423, 424
AF-dG·dC[a]	10-mer	5′-(AG*G)·5′-(CCT)	dG* *anti* or *syn*, groove displaced dC, major groove displaced AF, intercalated		423, 424
AF-dG·dC	15-mer	5′-(TG*A)·5′-(TCA)	dG*·dC, WC paired AF, major groove external		425
AF-dG·dC	11-mer	5′-(CG*C)·5′-(GCG)	dG*, *syn* major groove displaced dC, major groove displaced AF, intercalated		427
AF-dG·dC	11-mer	5′-(GG*C)·5′-(GCC)	dG*·dC, WC paired AF, major groove external, 5′-oriented		428
AF-dG·dA	11-mer	5′-(CG*C)·5′-(GAG)	dG*, *syn* intrahelical dA, *anti* intrahelical AF, minor groove external		429
AF-dG·dG	11-mer	5′-(CG*C)·5′-(GGG)	dG*, *syn* intrahelical dG, *anti* intrahelical AF, minor groove external		430
AP-dG·dC	11-mer	5′-(CG*C)·5′-(GCG)	dG*, *syn* major groove displaced dC, major groove displaced AP, intercalated		439
AP-dG·dA	11-mer	5′-(CG*C)·5′-(GAG)	dG*, *syn* major groove displaced dA, major groove displaced AP, intercalated	1AXU	440
ABP-dG·dC	15-mer	5′-(TG*A)·5′-(TCA)	dG*·dC, WC paired ABP, major groove external		442
PhIP-dG·dC	11-mer	5′-(CG*C)·5′-(GCG)	dG*, *syn* major groove displaced dC, major groove displaced PhIP, intercalated	1HZ0	443

[a] AAF, 2-acetylaminofluorene; AF, 2-aminofluorene; AP, 1-aminopyrene; ABP, 4-aminobiphenyl; PhIP, phenylimidazopyridine. [b] Similar population of AF-external and AF-intercalated duplex conformations.

of the helix. Three-dimensional models derived using restrained energy minimization showed very similar structures that provided some details of the fluorene conformation at the lesion site. In agreement with the experimental distances, the AF moiety is located in the minor groove with its C4−C5 edge facing the solvent, tilted toward the 3′-side of the damaged strand, and partially protected from water exposure (Figure 37). A difference between these duplexes

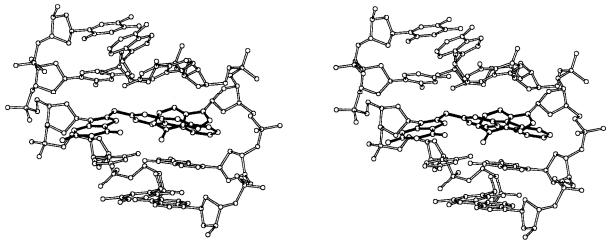


Figure 37. Stereoview of duplex DNA having a 2-AF-C^8-dG·dG mismatch. Reproduced with permission from ref 415. Copyright 1998 American Chemical Society.

is that the mismatch alignment and the relative position of AF in the minor groove of the helix are pH sensitive only in the 2-AF-C^8-dG·dA case.[429] In this duplex, the 2-AF-C^8-dG residue retains the *syn* conformation around its χ torsion angle at pH 5.9, but chemical shifts of the fluorene H5, H6, H7, and H8 protons, as well as the H1 and H8 protons of the 3′-flanking dG residue, move dramatically downfield. The authors proposed a structural model for the duplex under acidic conditions that has the lesion-partner dA protonated and hydrogen bonded with the 2-AF-C^8-dG residue in the *syn* conformation. However, the H2 and H8 proton chemical shifts of the lesion-partner dA residue showed negligible changes upon lowering the pH, challenging the proposition that it is protonated at pH 5.9. Table 7 lists structure types of duplexes containing 2-AF-C^8-dG and 2-AAF-C^8-dG lesions.

NMR spectroscopy was used to characterize the structure of a −1Δ mutation model containing a nonpartnered 2-AF-C^8-dG residue flanked by G·C base pairs.[431] Assignment of NOESY spectra indicated that the duplex adopts a right-handed helical conformation with all unmodified residues in the *anti* conformation around the χ torsion angle and canonical WC base pair alignments. The damaged residue is in the *syn* conformation and appears displaced in the major groove of the helix, facilitating the intercalation of the fluorene moiety inside the helix, where it performs rapid 180° flips and stacks preferentially with the 5′-flanking base pair. The refined structure shows that AF intercalation forms a wedge at the lesion site, with the lesion-flanking dC residues more separated than their dG partners in the complementary strand (Figure 38). Helix intercalation of the 2-AF ring is in sharp contrast to the proposed major groove 2-AAF location in the main conformer of the 2-AAF-C^8-dG −1Δ mutation duplex,[421] which, in principle, could explain their different mutagenic profiles. However, a refined structure of the latter duplex has not been produced yet and any comparison between the AF and 2-AAF −1Δ mutation models would be premature at this time.

NMR studies have described the solution conformation of two different −2Δ mutation intermediates, one having a targeted d(G*pA) deletion[432] and the other with an untargeted −2Δ deletion where the 2-AF-C^8-dG lesion was flanked by dC residues.[433] In both cases, the NMR spectra showed

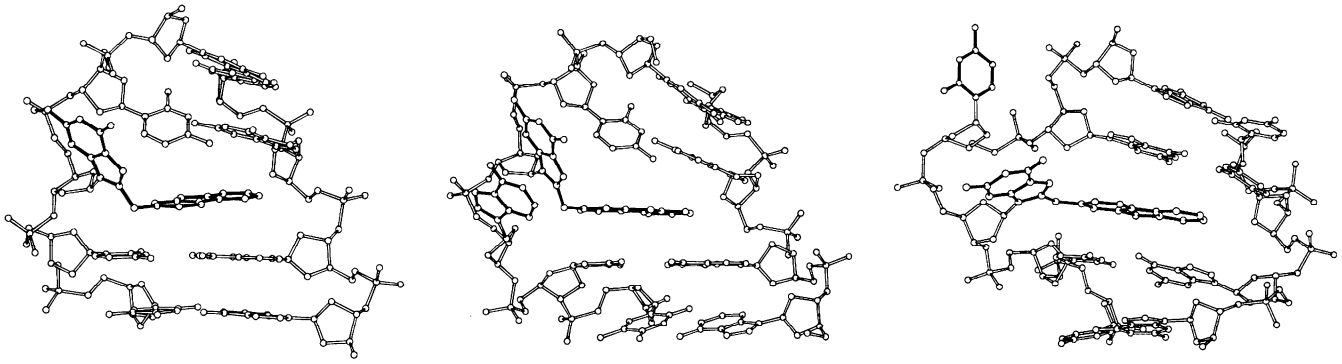


Figure 38. Structure of 2-AF-C^8-dG lesion-induced mutagenic intermediates. The pictures display extrahelical residues and the intrahelical 2-AF moiety in darker color. Left panel, −1Δ intermediate; center panel, targeted −2Δ intermediate; right panel, untargeted −2Δ intermediate. Adapted from refs 431−433. Copyright 1995 and 1997 American Chemical Society.

Table 8. Structure of One- and Two-Base Deletion Mutation Intermediates Having Bulky C^8-dG Lesions

lesion[a]	system type	sequence context	predominant structural motif	PDB entry	ref
AAF −1Δ	11/10-mer	5′-(CG*C)·5′-(G-G)	dG*, *anti* intrahelical; AAF, major groove external, 5′-oriented		421
AAF −2Δ	12/10-mer	5′-(GCG*CC)·5′(G-G-C)	dG*, *syn* major groove displaced; 3′-dC, major groove displaced; AAF, intercalated		422
AF −1Δ	11/10-mer	5′-(CG*C)·5′-(G-G)	dG*, *syn* major groove displaced; AF, intercalated		431
AF −2Δ	12/10-mer	5′-(CG*AC)·5′(G-G)	dG*, *syn* major groove displaced; 3′-dA *anti*, major groove displaced; AF, intercalated		432
AF −2Δ	12/10-mer	5′-(GCG*CC)·5′(G-G-C)	dG*, *syn* major groove displaced; 5′-dC, major groove displaced; AF, intercalated	1AX6	433

[a] AAF, 2-acetylaminofluorene; AF, 2-aminofluorene.

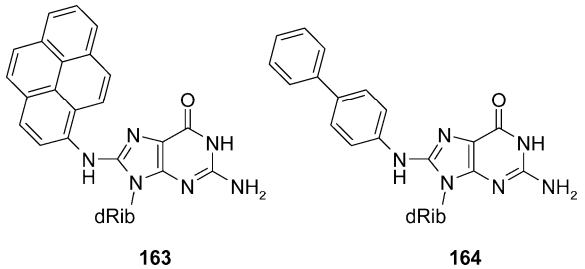
excellent resolution, allowing the assignment of most proton signals and the refinement of three-dimensional models. The duplexes are regular right-handed helices with all unmodified residues in the *anti* conformation around their χ angle and WC base pair alignments (Figure 38). In the targeted d(G*pA) deletion, the 2-AF-C^8-dG residue is *syn* and appears in the major groove of the duplex, facilitating intercalation of the AF moiety inside the helix. Its deletion partner dA, set at the 3′-side of 2-AF-C^8-dG, is displaced also into the major groove where it stacks with the damaged guanine base. An equivalent arrangement at the lesion site is present in the targeted d(G*pC) −2Δ duplex having 2-AF-C^8-dG and the 5′-lesion-flanking dC residue displaced into the major groove (Figure 38). Interestingly, a similar base-displaced intercalated structure was proposed for a 2-AAF-C^8-dG −2Δ duplex in identical sequence context but with the 3′-lesion-flanking dC residue extruded in the major groove.[422] However, the low resolution of the NMR data, which precluded the refinement of a three-dimensional structure, impedes an analysis of the differences between the 2-AF-C^8-dG −2Δ and 2-AAF-C^8-dG −2Δ duplexes. Table 8 lists the structure types of deletion mutation intermediates containing 2-AF-C^8-dG and 2-AAF-C^8-dG lesions.

NMR studies have established the conformation of a 2-AF-C^8-dG residue located at a primer template junction, before incorporation of any lesion partner[436] and after addition of a dC or dA residue.[437] In all three structures the damaged 2-AF-C^8-dG residue is displaced in the major groove with its χ torsion angle in the *syn* conformation and the fluorene moiety stacking over the last base pair of the duplex segment. This conformation shows few changes after the incorporation of a lesion-partner dC or dA residue, which appears displaced into the minor or major groove of the helix, respectively. Interestingly, in the latter case, the relative orientation of the 2-AF-C^8-dG lesion and the newly incorporated dA residue resembles, to some extent, their alignment in the 2-AF-C^8-dG·dA mismatch but with the bases far apart and displaced into the major groove. On the other hand, a recent X-ray structure of a complex formed between an active DNA polymerase fragment and a 2-AF-C^8-dG-containing primer-template demonstrates the important role played by the enzyme in the conformation adopted at the junction.[438] The structure of the preinsertion complex shows the 2-AF-C^8-dG-lesion in the *syn* conformation with the fluorene ring fully protected from solvent exposure. After dC insertion in the crystalline state, the damaged residue changes to *anti* and is WC hydrogen bonded, with the AF ring displaced toward the nascent major groove and unexposed to water. Thus, it is possible that the highly hydrophobic fluorene moiety, which is protected from solvent in the pre- and postinsertion complexes, may bias the NMR study toward the observation of an intercalative 2-AF-C^8-dG conformation.

8.6.3. 1-Aminopyrene (1-AP) and 4-Aminobiphenyl (4-ABP)

1-Nitropyrene, one of the most abundant nitroaromatic compounds found in environmental samples, may be transformed inside the cell to hydroxylaminopyrene, which, upon further activation, can react with guanine, forming 1-aminopyrene-C^8-dG (1-AP-C^8-dG) lesions **163** (Chart 61).[412]

Chart 61. 1-AP-C^8-dG 163 and 4-ABP-C^8-dG 164 Adducts



Having four fused benzene rings, 1-AP is very hydrophobic and bulkier than the fluorene moiety, facilitating the identification of correlation between lesion topology, hydrophobicity, and DNA structure. Two studies have reported the

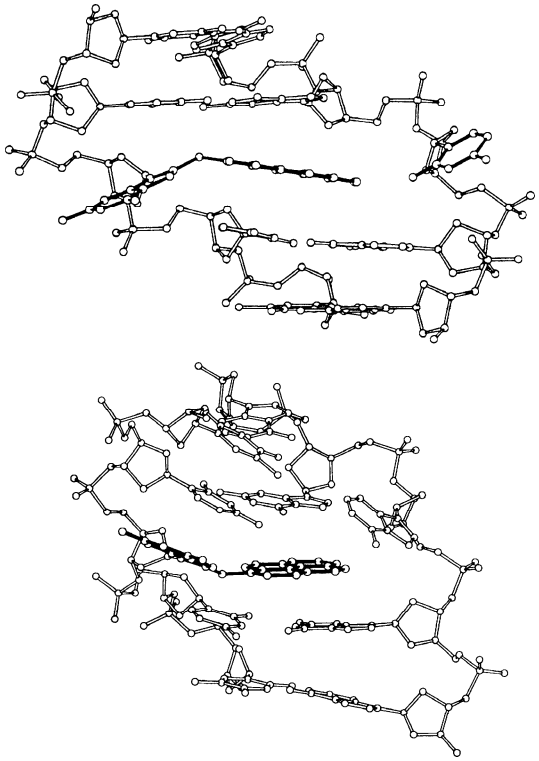


Figure 39. Structure of the central five-base-pair segment of duplexes having 1-AP-C^8-dG **163** lesions opposite dC (top panel) or dA (bottom panel). Adapted from refs 439 and 440. Copyright 1996 and 1999 American Chemical Society.

solution structures of 11-mer duplexes having a 1-AP-C^8-dG·dC or 1-AP-C^8-dG·dA base pair at their centers and flanked by C·G base pairs.[439,440] The NMR spectra showed partially resolved protons signals that were somewhat broadened at the lesion-containing and lesion-flanking base pairs, suggesting conformational dynamics at the damaged site of these duplexes. Assignment of the NOESY spectra shows that all unmodified nucleotides are in the *anti* conformation around the χ torsion angle, forming regular WC alignments. The damaged dG residues adopt a *syn* conformation and, along with their dC or dA partners, are displaced into the major groove of the helix, facilitating the intercalation of the aromatic 1-AP moiety inside the helix. Refined three-dimensional models of the 1-AP-C^8-dG·dC duplex show the damaged residue tilted toward the 3′-side of the lesion-containing strand and the 1-AP ring system stacking properly with flanking dC·dG base pairs (Figure 39). In the 1-AP-C^8-dG·dA duplex (PDB code 1AXU), the 1-AP moiety stacks with flanking base pairs also, but in this case, the damaged dG residue is tilted toward the 3′-side of the lesion-containing strand. Furthermore, the refined structure shows the lesion partner dA is almost coplanar to the 5′-flanking C·G pair, though direct NMR evidence for this alignment was lacking due to signal broadening at the lesion site.

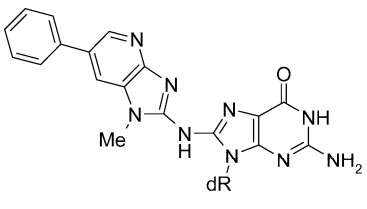
It is informative to compare the structures of 1-AP-C^8-dG-containing duplexes with those of the 2-AF-C^8-dG lesion paired to dC or dA in the same sequence context.[427,429] With dC partners, 2-AF-C^8-dG and 1-AP-C^8-dG duplexes show almost identical intercalative conformations at the lesion site, differing only with the fraction of duplex adopting this structure, which is about 100% for 1-AP[439] and 70% for AF.[427] On the contrary, the mismatched duplexes adopt completely different lesion site structures. The predominant conformation of the 2-AF-C^8-dG·dA duplex shows the AF moiety displaced into the minor groove, with the mismatched bases inside the helix and stacking between flanking base pairs,[429] while 1-AP remains intercalated in the 1-AP-C^8-dG·dA duplex.[440] Thus, whether resulting from favorable stacking interactions inside the duplex or entropically driven by the aqueous media, it seems clear that the larger 1-AP moiety prefers to intercalate in the helix. It must be noted, however, that in a different sequence context the absorption and fluorescence spectra showed the presence of several conformations, with at least one conformer in which the 1-AP is externally bound, indicating that duplex composition plays a role also in lesion site structure.[441]

It is also informative to compare the structures of 2-AF-C^8-dG-containing duplexes with those of 4-aminobiphenyl-C^8-dG lesions (4-ABP-C^8-dG). 4-ABP **164** (Chart 61) is structurally similar to AF but lacks the C9 methylene bridge that links the two phenyl rings, causing an increase of flexibility and a loss of planarity of the aromatic moiety. An early NMR study characterized a 15-mer duplex having a 4-ABP-C^8-dG lesion paired to dC in the d(T G* A) sequence context.[442] The NMR properties established that the duplex adopts a predominant helical conformation with WC alignments throughout, including at the 4-ABP-C^8-dG·dC base pair, and the 4-ABP moiety located in the major groove of the helix, where it is rapidly rotating. In additon, the data showed the presence of a minor conformation, about 5−10% of the sample, which was correlated with an intercalative 4-ABP structure. Interestingly, in the same lesion context, the 2-AF-C^8-dG·dC, containing duplex has a larger population fraction, about 30−40%, intercalated inside the duplex,[425] suggesting an energetic penalty associated with intercalation of the nonplanar 4-ABP ring system. Table 7 lists the structure types of duplexes containing 1-AP and 4-ABF lesions.

8.6.4. N-Methyl-2-amino-phenylimidazopyridine (PhIP)

Cooking of meat, chicken, and fish produces a series of carcinogenic heterocyclic amines, generally called food mutagens, including IQ, MeIQ, MeIQx, and PhIP, which after metabolic activation can generate predominantly C8 guanine lesions.[320] Despite the relevance for human health, preparation of large quantities of adduct-containing oligodeoxynuclotides has been very difficult and little is known about the structure of DNA containing these lesions. To date, only the NMR structure of an 11-mer duplex having an *N*-methyl-2-amino-6-phenylimidazo[4,5]pyridine-C^8-dG **165** (Chart 62) lesion (PhIP-C^8-dG) paired to dC has been

Chart 62. PhIP-C^8-dG Adduct



165

solved.[443] Topologically, PhIP consists of an *N*-methylimidazopyridine moiety substituted with a phenyl ring that breaks its planarity. Assignment of the NMR spectra showed the presence of a right-handed helix with, as observed with other bulky C8 guanine adducts, two slow exchanging

Figure 40. Structure of a DNA duplex containing a PhIP-C^8-dG **165** lesion paired to dC. The picture shows the lesion-containing base pair residues (in red) displaced in the major groove and the intercalated PhIP moiety (in green) protruding in the minor groove. Prepared from PDB entry 1HZ0[443] using MidasPlus.[556]

conformations at the lesion site. The predominant form has all unmodified residues in the *anti* conformation around the χ torsion angle, forming regular WC alignments, including at the lesion-flanking base pairs. The damaged dG is *syn* and, along with its partner dC residue, appears displaced into the major groove of the duplex, facilitating intercalation of PhIP inside the duplex. The refined three-dimensional model of the intercalated structure (PDB code 1HZ0) has the planar imidazopyrimidine ring stacked between flanking base pairs with the phenyl ring and the *N*-methyl group positioned in the minor groove of the duplex, inducing a significant bend in the helix (Figure 40). A refined model of the minor conformer, which depending on the temperature accounted for about 10−30% of the sample, was not obtained, but the analysis of chemical shift perturbations suggested that it has the damaged dG residue in the *anti* conformation with the PhIP moiety located in the major groove of the helix. Thus, as seen with other bulky C8 guanine adducts, the PhIP-C^8-dG lesion exchanges between a base-displaced intercalated conformation and an external major groove conformation, a fact that may help to explain its mutagenic profile.

9. Nucleotide Analogues

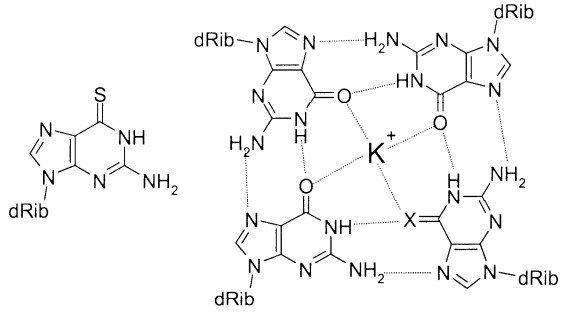
DNA damaging agents introduce chemical modifications on natural nucleotides, resulting in a wide variety of lesions, whose effects on DNA structure have been extensively discussed in previous sections of this review. In this section, we will discuss duplex structures containing artificial inserts, which do not appear as the result of a damaging process but, nevertheless, may appear inside the cell following DNA replication. During the last four decades, a large number of nucleoside analogues have drawn attention as promising anticancer and antiviral compounds. These agents generally inhibit DNA replication either directly, preventing chain elongation during polymerase-catalyzed replication, or indirectly, by depleting mono- and thiphosphate deoxynucleoside pools. The majority of them require activation before they can block polymerase activity. Cellular metabolism converts these compounds into the corresponding 5′-triphosphate nucleoside derivatives that are subsequently incorporated in DNA by replicative or repair polymerases, giving rise to the appearance of non-natural nucleotides. Additionally, we will discuss in this section duplex structures having point modifications that cannot appear during enzymatic DNA replication but are structurally related to lesions discussed in this section.

9.1. Modified Bases with Altered Hydrogen-Bonding Ability

9.1.1. Thioguanosine

Thioguanine is a good example of a small chemical modification in DNA that can trigger important consequencesfor living organisms. An early X-ray characterization showed that 2′-deoxy-6-thioguanosine (S^6-dG) **166** (Chart 63) is almost identical to 2′-deoxyguanosine, differing only

Chart 63. S^6-dG Lesion and a G4 Tetrad



in the length of the C−S double bond, which is about 0.4 Å longer, and its van der Waals radius, which is bigger for the sulfur atom. Surprisingly, these subtle differences affect a long list of biological processes, explaining its clinical use as an antileukemic, immunomodulating, and anti-inflamatory agent (for recent reviews, see refs 444 and 445). However, what causes the initial recognition of S^6-dG as abnormal is a basic question that only recently is finding some answers. An initial NMR study characterized the DNA structure of two different S^6-dG-containing telomeric samples and of the self-complementary Dickerson dodecamer with the lesion at the fourth position, resulting in two S^6-dG·dC pairs present on the same molecule separated by four dA·T base pairs.[446] The duplex T_m of the damaged duplex, as measured by analysis of UV melting profiles, is 15 °C lower than that of the unmodified control, indicating that S^6-dG affects the thermal stability of DNA. However, observation of imino proton signals resonating downfield of 12.5 ppm establishes the formation of a double-stranded duplex with WC alignments on all canonical base pairs, including all four dA·T pairs that separate the lesion sites. On the basis of the absence of the S^6-dG imino proton signal and the decrease in duplex thermal stability, the authors concluded that S^6-dG blocks the formation of WC alignments and must be considered a nonconservative base modification. In the case of samples containing S^6-dG within a telomeric sequence, the NMR spectra obtained at 80 °C, when only a random conformation is present in solution, were essentially identical to those

Table 9. Duplex and Base Pair[a] Stability Effects of S^6-dG[451]

duplex	G·C	S^6-G·C	G·T	S^6-G·T
$\Delta G°_{(37\ °C)}$ (kcal/mol)	-8.1 ± 0.5	-7.4 ± 0.2	-5.2 ± 0.3	-5.8 ± 0.2
τ_0 (ms)	126 ± 8	7.3 ± 1.2	0.1 ± 0.3	2.4 ± 0.9
K_d ($\times 10^{-7}$ mol^{-1})	5.4 ± 0.3	52 ± 4	1760 ± 55	1534 ± 220
τ_{open} (ns)	68 ± 3	38 ± 4	17 ± 12	366 ± 91

[a] Kinetic parameters measured at 5 °C. Reported errors are standard deviations. τ_0, base pair lifetime; K_d, base pair dissociation constant; τ_{open}, lifetime of the opened base pair state.

recorded at room temperature. In contrast, the control sequences lacking S^6-dG exhibited dramatic spectral changes below 50 °C, indicating the formation of G4 tetrads. Thus, it was concluded that S^6-dG blocks the formation of G4 tetrads. It is worth noting that factors affecting the stability of WC base pairs and G4 tetrads are different. It is known that, in the latter case, the efficient assembly of G4 tetrads **167** requires the formation of a complex with a sodium or potassium ion.[447] In this complex, the metal ion is embedded between the planes of two adjacent G4 tetrad layers, forming an almost perfect octahedral complex with the O6 groups of all eight guanine residues. Even a single bulkier and polarizable sulfur moiety per tetrad layer, which cannot bind to alkali ions efficiently, disrupts the symmetry of this highly cooperative structure, thus blocking its formation.[446]

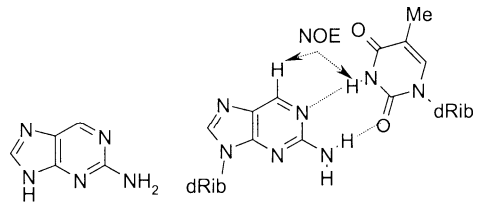
Two independent studies have challenged the conclusion that S^6-dG residues impede formation of WC alignments.[448,449] In both studies, a single S^6-dG·dC base pair was introduced at the center of non-self-complementary duplexes, and in one case,[449] S^6-dG was also paired to T, forming an S^6-dG·T mismatch. As reported previously, the presence of a single S^6-dG·dC base pair decreases the duplex thermal stability but, in both studies, the imino proton signal of S^6-dG appears as a broad resonance around 11.5 ppm. In addition, observation of NOE cross-peaks across the S^6-dG·dC pair and between S^6-dG H1 and the imino protons of lesion-flanking base pairs positively establishes the presence of a WC S^6-dG·dC base pair alignment. Furthermore, the RMD structures showed that both duplexes are minimally perturbed right-handed helices with regular S^6-dG·dC base pair stacking inside the helix. The only structural change identified in both studies is a small opening of the S^6-dG·dC base pair toward the major groove, which results from the presence of the bulkier sulfur atom (PDB codes 1N17 and 1KB1). Similarly, the NMR characterization and RMD structure (PDB code 1KBM) of the S^6-dG·T duplex show a regular B-type conformation with formation of a wobble S^6-dG·T alignment, almost identical to that seen on dG·T mismatches[157,450] (PDB codes 113D and 1BJD) but with increased opening of the mismatch toward the major groove. Therefore, it was concluded that S^6-dG is a highly conservative base modification that does not affect duplex structure and whose biological consequences are most likely related to its stability effects.

Table 9 summarizes the results of a recent publication reporting on the effect of S^6-dG on global duplex and local base pair stability.[451] Van't Hoff analysis of ultraviolet melting curves indicated that a sample containing a single S^6-dG·dC base pair has an about 1 kcal/mol less favorable $\text{Gibbs}_{(37\ °C)}$ free energy of duplex formation than the corresponding dG·dC control. In contrast, the $\text{Gibbs}_{(37\ °C)}$ free energy of the S^6-dG·T and dG·T mismatch-containing duplexes is almost identical and approximately 3 kcal/mol less favorable than that of the control. In addition, NMR measurement of imino proton exchange rates indicated that in the sequence context used for the study the lifetime of an unmodified dG·dC base pair is 126 ms, while that of the dG·T mismatch is about 0.1 ms, right at the detection limit of the method. The lifetimes of the S^6-dG·dC and S^6-dG·T base pairs are about 7 and 2 ms, respectively, indicating that S^6-dG decreases the stability of the pairing with dC but it slightly increases that of a mismatch. Furthermore, the S^6-dG substitution increases the dissociation constant of the dG·dC base pair 1 order of magnitude, making it spend a longer fraction of time in an open state, but has no effect on that of the dG·T mismatch. Taken together, these studies indicate that while S^6-dG causes no significant perturbations of the DNA conformation, it profoundly affects the local base pair stability of DNA duplexes.

9.1.2. 2-Aminopurine

Two relevant features explain the wide use of 2-aminopurine **168** (2AP) (Chart 64) for biophysical and molecular

Chart 64. 2-Aminopurine and d2AP·T Base Pair



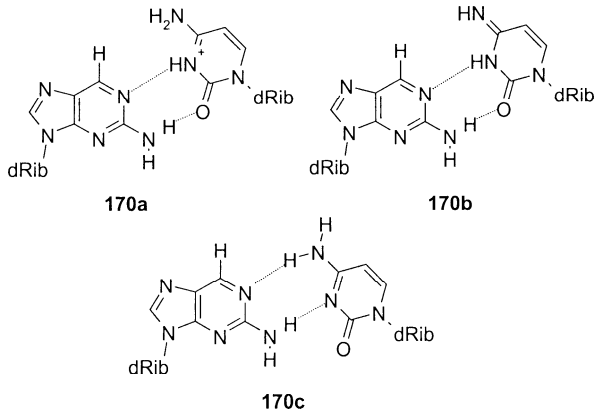
168 169

biology studies of double-stranded DNA. First, 2AP is an environmentally sensitive fluorophore that induces small changes in the DNA conformation and, thus, is a suitable probe for studying nucleic acid structure and dynamics.[452] Second, 2-aminopurine is a moderate mutagen, and duplexes containing 2′-deoxy-2-aminopurine (d2AP) residues have been the target of early *in vitro* and *in vivo* DNA replication studies.[453] When present in the template strand or as the incoming residue, 2AP pairs predominately with thymine. However, in the presence of d(2AP)TP nucleosides, T4 DNA polymerase induces up to 6% C·G → T·A transitions, suggesting that 2AP can also pair to dC.[454]

An early NMR study characterized the alignment of d2AP·T and d2AP·dC lesion-containing base pairs located at the center of 7-mer duplexes. In the d2AP·T duplex, the T H3 proton of the lesion-containing base pair appears around 13.6 ppm and displays strong NOE interactions with the 2AP H6 and the imino protons of adjacent base pairs, indicating that the d2AP·T base pair **169** adopts a WC geometry with formation of two hydrogen bonds.[455] A subsequent saturation-recovery study of a self-complementary decamer duplex showed that d2AP substitution of the innermost dA residue increases the base pair opening rate by 6-fold and diminishes the duplex T_m by 10 °C. Thus, despite the WC geometry across the d2AP·T base pair, d2AP residues diminish local and global stability of the duplex.[456]

Proton spectra of the d2AP·dC duplex show an exchangeable signal around 9.4 ppm that originated from the damage-containing base pair. Since in the 1D NOE difference spectrum this signal exhibits strong interaction with 2AP H6 and the imino protons of adjacent base pairs, it was concluded that the dC residue is mostly protonated at pH 7.2, forming a d2AP·dC^+ alignment **170a** (Chart 65) that retains a regular WC geometry.[455] However a subsequent study using dC

Chart 65. Possible Alignments of the d2AP·dC Base Pair

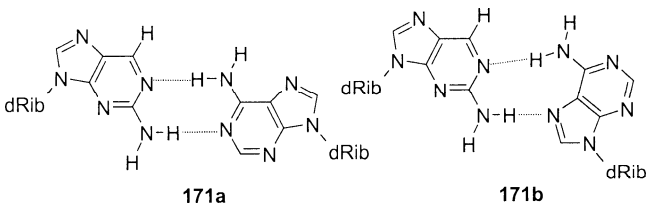


residues with its amino group[15]N enriched showed that the lesion-partner cytosine amino protons resonate as two separated signals, ruling out the presence of both the protonated **170a** and the rare tautomeric form **170b** of cytosine. In addition, chemical shifts of the cytosine amino proton signals indicated that one of them is hydrogen bonded and the intensity of NOE cross-peaks across the d2AP·dC base pair suggested the formation of a wobble alignment **170c**.[457,458]

An additional NMR study of the d2AP·dC base pair revealed that the transition from a wobble **170c** to a dC^+ protonated **170a** alignment does take place at low pH and, based on the pK_a value of cytosine protonation, the ratio between d2AP·dC^+ **170a** and d2AP·dC wobble **170c** alignments was calculated to be 1:25 at physiological pH. Interestingly, this value corresponds to the frequency of A·T → G·C transitions observed when d2AP residues are present in the template strand during DNA replication studies *in vitro*. Of the basis of these observations, the authors proposed that the presence of small amounts of 2AP·C^+ pairs, which have the proper WC-type geometry, might account for the mutagenicity of 2AP.[459] It is interesting to compare the behavior of cytosine amino groups in the wobble d2AP·dC **170c**[458] and O^6-alk-dG·dC **57**[177] base pairs. At the same pH and temperature conditions, cytosine amino protons resonate separately around 9 and 6 ppm in the **170c** pair, whereas they appear as a single signal around 7.70 ppm in **57**. The most likely explanation for this discrepancy is the fast rotation, on an NMR time scale, of the cytosine amino group on the latter wobble pair, a feature that reflects hydrogen bond strength differences between these two pairs, probably due to a lower basicity of the O^6-alk-dG N1 nitrogen.

Besides thymine and cytosine, 2AP can pair with adenine during DNA replication.[460] The NOESY spectrum of a 7-mer duplex containing a d2AP·dA base pair demonstrated that the duplex adopts a characteristic B-DNA conformation with both lesion site residues stacked inside the helix. The authors examined two possibilities for the d2AP·dA alignment (Chart 66): a wobble pair **171a**, with both residues facing each other

Chart 66. Possible Alignments of the d2AP·dA Base Pair



in the *anti* conformation, and a WC geometry **171b**, with the WC edge of d 2AP facing the Hoogsten side of dA. Sequential NOE cross-peaks at the lesion site established the presence of a wobble d2AP·dA pair **171a** in solution.[461]

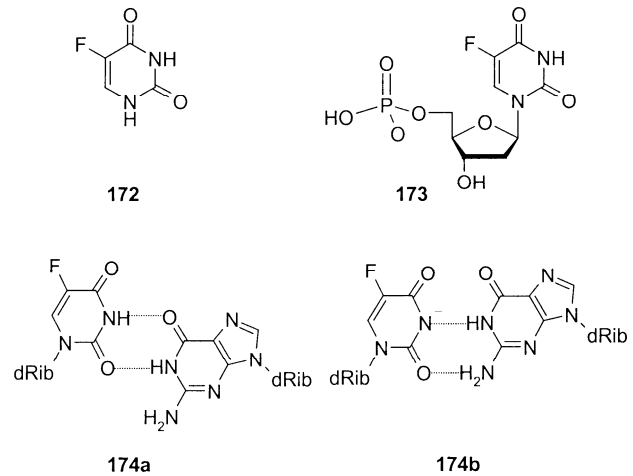
9.1.3. 5-Fluorouracil

5-Fluorouracil **172** (U^F), a thymine analogue, is used in combination therapy for the treatment gastrointestinal tumors and other types of cancer.[462,463] The therapeutic effects of U^F result from its ability to alter nucleotide pools, disrupting nucleic acids synthesis and impeding growth and division of tumor cells. The cellular metabolism converts 5-fluorouracil into 5-fluoro-2′-deoxyuridine 5′-phosphate **173** (5-dU^F-MP), which inhibits thymidilate synthase, causing a reduction of TTP levels.[464] 5-dU^FMP is further transformed to dU^F-TP, its triphosphate derivative, which, after incorporation into DNA by replicative polymerases, inhibits chain elongation and causes strand breaks and mutations.

In spite of their structural similarity, the different dipole moments of thymidine and dU^F may cause distinct packing in DNA. In addition, the imino nitrogen acidity of dU^F (pK_a = 8.15) is greater that of T (pK_a = 9.9), affecting the hydrogen-bonding properties of dU^F. NMR spectroscopy was used to characterize the effect that dU^F has on duplex structure and dynamics.[465−469] A self-complementary octamer duplex having the *Eco*RI recognition sequence with two symmetrical dU^F·dA base pairs mostly adopts a regular right-handed conformation with WC base pair alignments throughout. The imino proton signal of dU^F is broadened by solvent exchange with a water exchange rate similar to that of terminal base pairs. The temperature dependence of imino proton signals showed that dU^F H1, while broad, is visible up to the duplex melting temperature, indicating that the local stability of the dA·dU^F pair was not affected.[468] Subsequently, an NMR study compared the structure of self-complementary 8-mer duplexes having zero, one, or two internal dA·dU^F base pairs. It was found that the dU^F presence increases the roll angle value by about 10°, causing a slight bend in the double helix at the lesion site but without affecting base inclination or tip angles.[469]

It has been proposed that the increased acidity of dU^F is responsible for its ability to form almost equally stable pairs with dA and dG[470] and explains its mutagenic activity.[471] Several schemes (Chart 67) have been suggested for the

Chart 67. Fluorouracil 172 and Its Base Pair Alignments 174a,b

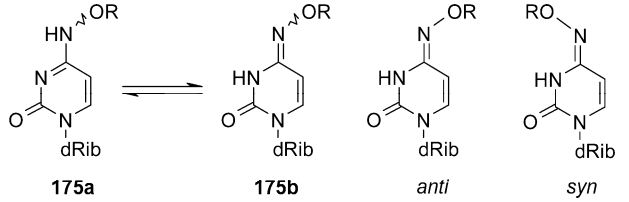


alignment of dU^F·dG base pairs, including a wobble type **174a**, with the lesion displaced toward the major groove, and the WC-geometry **174b**, with U^F present in the rare tautomeric form. The NMR study of a heptamer duplex having a central dU^F·dG mismatched revealed that three exchangeable proton signals originate from the lesion-containing pair, in neutral and slightly acidic media. Two of them are bound to dG N1 and dU^F N3 imino nitrogens, and the other belongs to the freely rotating guanine amino protons. These findings are consistent only with the presence of a wobble dU^F·dG mismatch alignment **174a**. As pH increases, the ^{19}F chemical shift of dU^F and the chemical shifts of protons at and adjacent to the lesion site show changes that are typical sigmoid titration curves, with a midpoint around pH 8.3. In addition, the line width of the ^{19}F signal reaches a maximum at the same pH. These observations clearly suggest that ionization of the U^F moiety is taking place at basic pH values with a pK_a of 8.3, close to that found for free dU^F (pK_a = 8.15). A NOESY spectrum recorded at pH 9.5 displays regular NOE connectivities for both strands of the duplex, demonstrating the presence of B-form conformation. On the basis of these findings, the authors proposed that ionization of U^F at basic pH triggers the transition from a wobble dU^F·dG pair **174a** to a WC alignment **174b**, where U^F is in an anionic state. Since the pK_a for ionization of U^F in DNA is 8.3, about 10% of dU^F residues are dissociated at physiological pH values, suggesting that the anionic form of U^F may account for the observed levels of dG incorporation.[472] A similar transition from a wobble alignment to a WC ionized base pair has been observed for duplexes containing a single 5-bromo-dU·dG pair, upon pH increase. However, due to the lower acidity of bromouracil, the transition midpoint is around pH 8.6.[473]

9.1.4. N-Alkoxycytosines: Degenerated Bases Forming WC Hydrogen Bonds

Methoxyamine, *O*-methoxyhydroxylamine, can substitute the amino group of cytidine in DNA, forming *N*^{4}-methoxycytidine (*N*^{4}-MeOdC), a highly mutagenic residue that induces mainly G·C → A·T transition mutations.[474,475] A striking feature of this modified nucleotide is its ability to form stable base pairs with both adenine and guanine bases. The thermal and thermodynamics stabilities of 17-mer duplexes having a single *N*^{4}-MeOdC·dG or *N*^{4}-MeOdC·dA base pair are very similar to each other and close to those of corresponding duplexes with normal dC·dG and dA·dT pairs, but they are considerably higher than the stabilities of the mismatch-containing duplexes.[476] Since the equilibrium constant of the amino and imino forms of *N*^{4}-methoxycytosine (Scheme 16) is close to unity,[477] it was proposed that

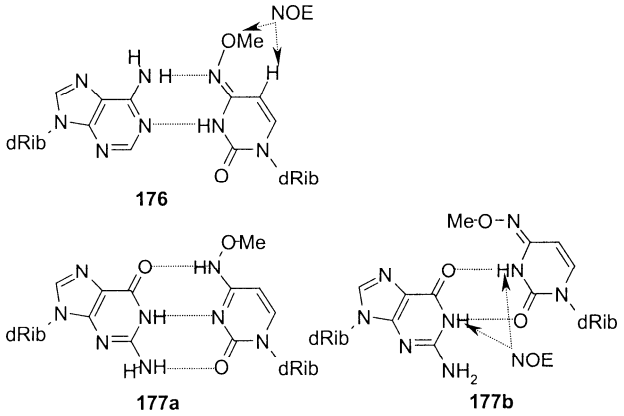
Scheme 16. *N*^{4}-Alkoxy-dC Tautomers and the Isomerism of Their Imino Forms



the identity of the lesion-partner base could tip this balance and, as a result, different tautomeric forms **175a** and **175b** would be pairing to dG and dA. In addition, several alignments are hypothetically possible for *N*^{4}-MeOdC·A and *N*^{4}-MeOdC·G, including wobble alignments, base pairs with normal WC geometry, and pairs where the *N*^{4}-methoxy group is *syn* or *anti* with respect to the lesion imino nitrogen.

NMR spectroscopy was used to characterize the structures of duplexes containing a single *N*^{4}-MeOdC lesion opposite dA or dG.[478−481] The NMR spectra showed the presence of regular B-form duplexes with *N*^{4}-MeOdC and partner residues properly stacked inside the double helix but with a different *N*^{4}-methoxycytosine tautomeric state in each case. Analysis of the exchangeable proton spectra of the *N*^{4}-MeOdC·dA duplex established the predominant presence of the lesion imino tautomer **175b** with its *N*^{4}-MeOdC H3 hydrogen bonded to the dA imino nitrogen, forming a WC-type *N*^{4}-MeOdC·dA alignment **176** (Chart 68). In addition,

Chart 68. *N*^{4}-MeOdC·dA 176 and *N*^{4}-MeOdC·dG 177a,b Pair Alignments



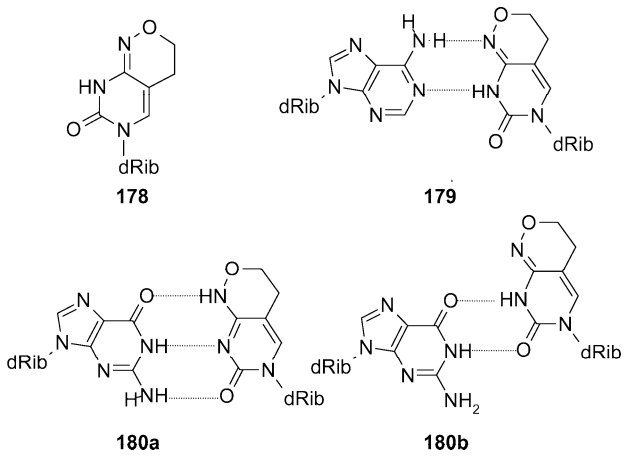
strong NOE connectivities between O-methyl and H5 lesion protons, and between O-methyl and aromatic protons of adjacent bases, indicated that the *N*^{4}-methoxy group is *anti*, avoiding steric clashes during *N*^{4}-MeOdC·dA base pair formation.[478,480,481] In one of these studies, the experimental data suggested the presence of a minor conformation having a wobble *N*^{4}-MeOdC·dA pair with the adenine residue displaced to the minor groove in fast exchange with the WC alignment.[478] However, the NMR characterization of another duplex, having a single *N*^{4}-MeOdC·dA pair in a different sequence context, established the presence of only the WC-type **176** alignment, suggesting that flanking residues may influence the alignment of the lesion-containing base pair.[481]

Pairing of *N*^{4}-methoxycytosine with dG is somewhat more complicated. The exchangeable proton spectra of duplexes containing *N*^{4}-MeOdC·dG pairs show two sets of resonances, with approximately a 1:1 ratio, indicating that the *N*^{4}-MeOdC·dG base pair exists as two slowly interchanging conformations. The existence of two downfield imino proton signals with a strong NOE interaction between them and a non-hydrogen-bonded guanine amino group establishes the presence of a wobble *N*^{4}-MeOdC·dG pair **177b** for one of these conformations. The second form corresponds to a duplex having the amino tautomer of *N*^{4}-methoxycytosine forming an *N*^{4}-MeOdC·dG pair with WC geometry **177a**.[480]

As in the case of *O*^{6}-alkylguanine, the lesion imino tautomer prefers a *syn* methoxy group configuration in free nucleotides and single-stranded DNA. Since this configuration causes no steric clashes in the wobble *N*^{4}-MeOdC·dG pair, the imino tautomer readily forms this alignment, keeping the methoxy group in the *syn* orientation. However, a *syn* methoxy group configuration hinders the WC-type alignment,

and consequently, change to an *anti* orientation must occur before the lesion forms a WC N^4-MeOdC·dG pair. This transition proceeds slowly and is the rate-limiting step of duplex formation.[478] A close analogue of N^4-MeOdC with the aminoalkoxy function incorporated in a six-member ring system that freezes the conformation of the methoxy group to *anti*, named P-nucleotide **178**, was synthesized in order to eliminate this effect.[482] The NMR spectra of duplexes containing a T·dA or P·dA base pair overlap almost completely, indicating that these duplexes are isostructural. In addition, the thermal stabilities of T·dA- or P·dA-containing duplexes are virtually indistinguishable, suggesting that the P nucleotides form a normal P·dA WC alignment **179** and are a better T substitute than N^4-MeOdC.[478] However, even in the absence of *syn*−*anti* isomerism, pairing of the P-nucleotide to guanine is more complicated thanpairing to adenine (Chart 69). The NMR spectra of 8-mer

Chart 69. P-Nucleoside 178 and Its Alignment with dA 179 and dG 180a,b



duplexes containing two P·dG base pairs showed that both the WC **180a** and the wobble **180b** alignments are present in solution in slow exchange on the NMR time scale. However, the duplex population with a P·dG wobble pair was considerably smaller than that observed for the N^4-MeOdC·dG duplex. Despite the conformational and tautomeric inhomogenity of P·dG pairs, the P·dG duplex is slightly less stable than a control having unmodified residues and more stable than duplexes having T·dG or N^4-MeOdC·dG pairs. Thus, since the P nucleoside can pair with all natural nucleotides with minor consequences for duplex stability and conformation, it has become an ideal degenerated base for promising applications for molecular biology and genetic engineering.[479] It is worth noting that the existence of both WC and wobble P·dG base pair alignments was confirmed later by X-ray crystallography.[483]

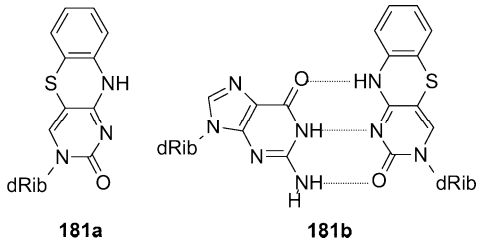
The existence of low levels (10^{-4}−10^{-5}) of rare base tautomeric forms in DNA, which can form mutagenic intermediates with correct WC geometry, is one of the main causes of genome instability.[484] However, in the case of DNA lesions, base tautomerism is seldom responsible for the appearance of mutations. For all base modifications described in the previous chapters, base pairing of mutagenic intermediates results from functionality changes of hydrogen bond donors and acceptors or from strong hydrophobic interactions in double-stranded DNA. To the best of our knowledge, alkoxycytidine derivatives are the first example of a base modification for which tautomerism plays a decisive role in the alignment of a mutagenic base pair.

9.2. Nucleotides Lacking Hydrogen-Bonding Ability

9.2.1. 3-Nitropyrrol Nucleotide

Since the discovery of the double-helical nature of DNA, hydrogen-bonding interactions between complementary bases are believed to play a pivotal role in stabilizing the DNA structure. Despite the high water capacity to form hydrogen bonds that might compensate a decrease of interstrand hydrogen-bonding energy, the effect of base damages on DNA stability and conformation is frequently explained in terms of hydrogen bond disruption at the lesion site. The contribution that "stacking forces", a group of effects that combine hydrophobic, van der Waals, and $\pi-\pi$ interactions between aromatic moieties, make to duplex conformation and stability has received much less attention. Since it is very difficult to experimentally measure stacking interactions, the relative contribution of each of its components is not completely understood.[485,486] However, they are important for understanding DNA stability and sequence dependent structural variations.[487] It has been established that an artificial nucleotide having a modified cytosine moiety **181a** with enhanced stacking potential and normal hydrogen bond potential forms a canonical WC base pair alignment with dG **181b**, increasing duplex thermal stability and without distorting the double-stranded helical structure[488] (PDB code 1TUQ) (Chart 70).

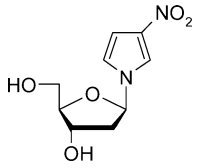
Chart 70. Tricyclic dC 181a and Its Alignment with dG 181b



However, whether stacking interactions alone are sufficient for the assembly of a double-helical DNA structure is still an open question.

Studies of duplexes having a 1-(2′-deoxy-β-ribofuranosyl)-3-nitropyrrol (**182**) residue (M) (Chart 71) initially tested

Chart 71. M Nucleoside



the idea that a nucleotide analogue, able to stack within the DNA helix without forming hydrogen bonds, may serve as a "universal" nucleotide.[489] The electronic density and π-electron distribution of 3-nitropyrrol, necessary for effective stacking, closely resemble those of natural nucleobases, but it lacks hydrogen bond donors and only poor acceptors are present. When incorporated into DNA, M nucleotides cause a decrease of duplex T_m that is mostly independent of the partner base, and they behave as an almost perfect

"wildcard" during sequencing reactions.[490] The NMR spectra of a 9-mer duplex containing a centrally located M·dA pair indicate the presence of a regular B-form conformation with unmodified residues adopting normal WC base pair alignments. The RMD structure (PDB code 1DK6) showed that the M residue adopts a *syn* conformation with the 3-nitropyrrol moiety inside the helix directing its nitro group toward the major groove. Stacking interactions of 3-nitropyrrol with flanking bases are not as pronounced as those of a corresponding thymine base in the parent duplex. The M·dA pair shows a small buckle, and phosphate groups around the modified nucleotide are slightly bulged (Figure 41). In

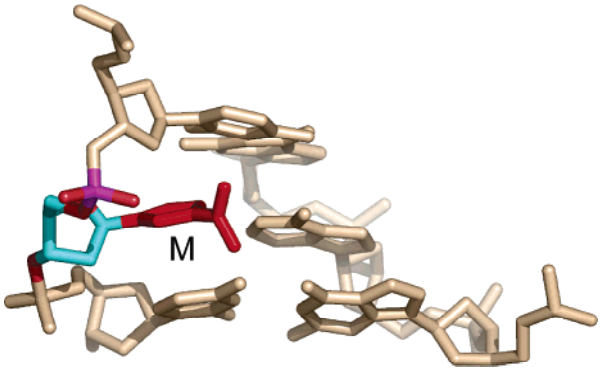


Figure 41. Structure of the central three-base-pair segment of a duplex containing a nitropyrrol nucleotide **182** viewed with the major groove prominent. The nitropyrrol moiety is colored red. Hydrogen atoms are not shown. Prepared from PDB entry 1DK6[491] using PyMol.[555]

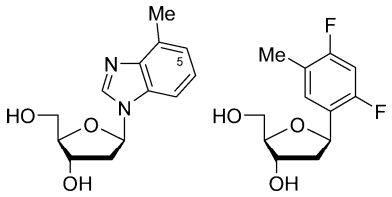
addition, the RMD structure suggests that the M nitro group is not coplanar with the pyrrol ring. On the basis of these observations, the authors concluded that the presence of the bulky, non-coplanar nitro group hinders optimal stacking of the pyrrol moiety with flanking bases.[491] Furthermore, the bulkiness of the 3-nitropyrrol moiety causes the disruption of DNA triple helices when an M residue is present in the Hoogsten third strand.[492]

9.2.2. Non-Hydrogen Bonding A and T Isosters

Two aromatic moieties, 4-methylbenzimidazole (Z) **183** and 2,4-difluorotoluene (F) **184**, have been examined as isosteric analogues of adenine and thymine bases, respectively. Thermodynamics studies demonstrated that dA·F, Z·T, and Z·F pairs destabilize duplex DNA by approximately 4−5 kcal/mol with respect to the dA·T-containing control sample, which is not surprising taking into account the almost complete inability of Z and F to form hydrogen bonds.[493] However, the behavior of these nucleotide analogues during DNA replication is much more unexpected. During template-directed synthesis, the Klenow fragment of DNA Pol I incorporates Z and F nucleotides opposite to T and A, respectively, with similar selectivity to incorporation of natural nucleotides. Furthermore, the Pol I−Klenow fragment catalyzes the selective formation of Z·F base pairs, indicating that, at the active center of DNA polymerases, specific nucleotide interactions are possible without the formation of hydrogen bonds, provided that the geometry match of the alignment is correct.[494−496]

To verify the hypothesis, NMR studies have characterized the structure of dodecamer duplexes containing an A·F (PDB code 1BW7)[497] or Z·F (PDB code 1EEK)[498] pair (Chart 72). The presence of Z and F residues does not perturb the global

Chart 72. Z 183 and F 184 Nucleosides



183 **184**

C G C A T F G T T A C C
G C G T A X C A A T G G

F = 184
X = dA, 183

duplex conformation, which retains all the essential features of B-form DNA. In both duplexes, bases of the modified pairs adopt an *anti* conformation around the χ angle and are intrahelical, displaying a small propeller twist that causes no major distortions on flanking base pairs. Steric clashes between the Z H5 proton, which is absent in the parent adenine base, and F H3 cause a slight displacement of Z toward the minor groove of the duplex and an increase of about 0.8 Å in the C1′−C1′ distance. Broadening of imino proton signals adjacent to the artificial pair indicates increased dynamics at the 5′-side of the modification site in both the Z·F- and dA·F-containing duplexes. However, the absence of imino proton signals at the modified base pairs precluded the NMR determination of their opening rates. The RMD models show that the A·T, Z·F, and A·F pairs have very similar geometries (Figure 42), providing structural

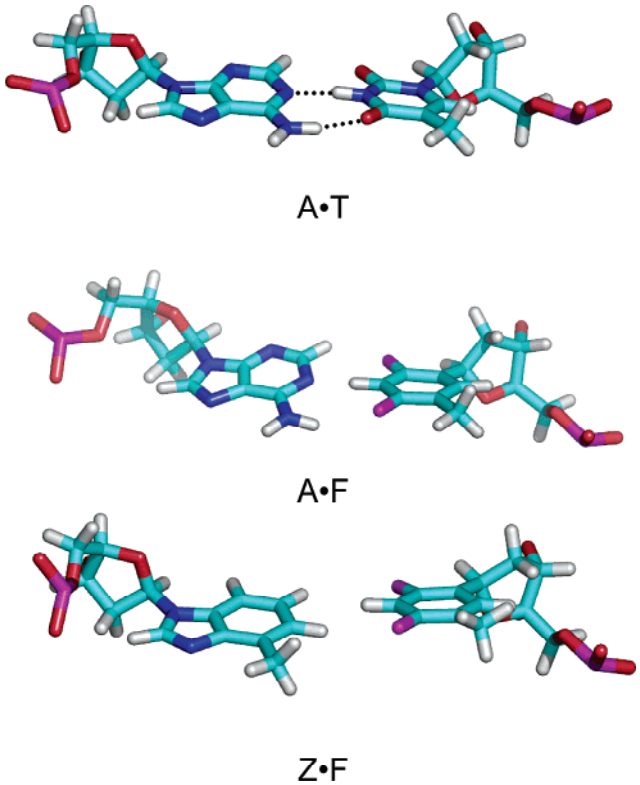


Figure 42. Alignment of the dA·T, dA·F, and Z·F, pairs in DNA duplexes. Prepared from PDB entries 1BW7[497] and 1EEK[498] using PyMol.[555]

support for the hypothesis that a steric match between the incoming and template nucleotides at the polymerase active center is necessary and sufficient for DNA synthesis.

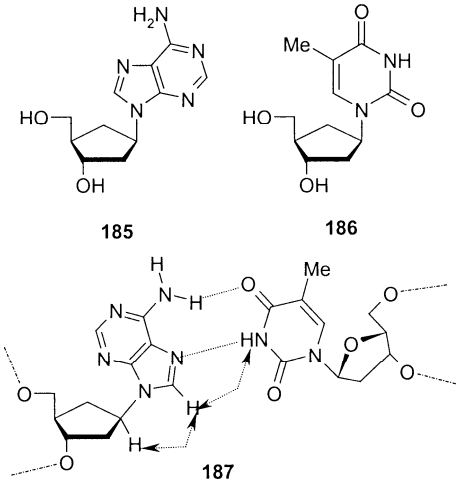
9.3. Sugar- and Phosphate-Modified Nucleotides

9.3.1. Aristeromycin and Other Carbocyclic Nucleotides

Carbocyclic nucleosides lack the endocyclic oxygen of the sugar ring. This absence leads to the disappearance of three important stereoelectronic features, namely the anomeric effect[499] and 3′,4′-*gauche* and 4′,5′-*gauche* effects.[500,501] In normal nucleotides, the anomeric effect governs primarily the sugar pseudorotation angle and the two *gauche* effects drive the 3′- and 5′-exocyclic sugar oxygens into a conformation where they have *gauche* orientations relative to O4′. As a result, the furanose ring of 2′-deoxynucleosides adopts a 2′-endo/3′-exo conformation and the 5′-hydroxymethyl group assumes a *gauche*$^+$ or *gauche*$^-$ conformation. Since carbocyclic nucleotides lack these properties, incorporation of these analogues in double-stranded duplexes illustrates the contribution that stereoelectronic effects have in stabilization of the DNA structure. Carbocyclic analogues are absolutely resistant to DNA glycosylase activity and have been used for biochemical and molecular biology studies.[502,503]

NMR spectroscopy has evaluated the impact on duplex structure following the incorporation of a single 2′-deoxyaristeromycin **185** (dAr) residue paired to thymidine[504] (PDB code 1EXL) (Chart 73). The NMR spectra indicated

Chart 73. dAr 185, Carbocyclic-T 186, and Hoogsten dAr·T Pair 187, with Relevant NOE Contacts Indicated by Dashed Arrows



that the presence of a dAr·T base pair produces minimal perturbations of the right-handed helix that remains fully WC hydrogen bonded. The RMD structure displays the dAr residue with a 1′-exo conformation on its cyclopentane ring, which is entirely consistent with the lack of anomeric and endocyclic *gauche* effects. The γ backbone torsion angle of the modified nucleotide, as well as those of adjacent residues, has typical B-form DNA values, and the 3′ and 5′ dAr-flanking phosphate groups show a B_I conformation. The dAr·T base pair alignment displays a moderate (26°) propeller twist that has no impact on the formation of WC hydrogen bonds.

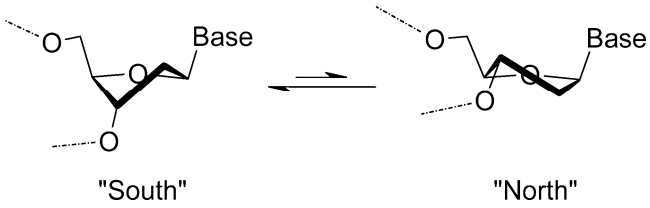
The question of whether the small structural perturbations caused by a single carbocyclic analogue can be amplified by the presence of another nearby carbocyclic nucleotide was investigated by characterizing the structure of a Dickerson dodecamer containing two carbocyclic thymidine nucleotides **186**.[505] Despite the presence of four consecutive carbocyclic-T·dA base pairs, the NMR spectra suggest no significant violations of the global DNA geometry. Analysis of RMD models shows that duplexes having **186** have A- and B-form conformational features, with only a moderate flexibility increase in the carbocyclic moiety. Parent control duplexes also showed mixed A- and B-form DNA properties. As in the single dAr residue case, it was concluded that the carbocyclic T residue is a good structural mimic for sequences containing stretches of T.

However, this conclusion seems to be not universal. The NMR NOESY spectra of the Dickerson dodecamer having two adjacent dAr·T base pairs at its center show a double set of cross-peaks for interactions of the eight centrally located nucleotides.[506] After raising the temperature above 50 °C, these two sets of signals coalesce into one set, indicating that at room temperature they are in slow exchange on the NMR time scale. Assignment of the spectra shows that one set corresponds to a regular B-form duplex with canonical WC alignments at the adjacent dAr·T base pairs. Analysis of the second set of NOE peaks suggests that the alternative conformation of the duplex has the dAr·T pair forming Hoogsten alignment **187**. Interestingly, both dAr·T base pairs of the duplex simultaneously adopt the same alignment type, demonstrating a high cooperativity of the WC ↔ Hoogsten transition. The energetic difference between the two conformations favors WC alignments by about 0.24 kcal/mol, resulting in a 2:1 duplex population ratio. RMD structures of each form showed that, while WC dAr·T alignments do not perturb the duplex, formation of two adjacent Hoogsten pairs causes significant distortion of the double helix. In this case, the buckle, propeller twist, open, tilt, roll, helical twist, inclination, and tip change considerably at the center of the duplex, resulting in a narrowing of the minor groove by 2−3 Å.

9.3.2. Nucleotides with the Frozen Sugar Conformation

Stereoelectronic effects drive the conformational equilibrium of 2′-deoxyribose to the 2′-endo/3′-exo ("south") range. Nevertheless, the sugar ring remains flexible in duplex DNA, exchanging mainly with conformations in the 2′-exo/3′-endo ("north") range[507] (Scheme 17). This interconversion between

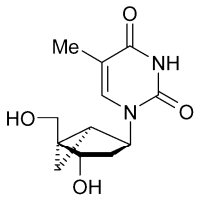
Scheme 17. Preferred 2′-Deoxyribose Conformations in DNA



sugar conformations partially accounts for the global flexibility of DNA, and it is associated with kinks of the DNA helix found in many DNA−protein complexes.[508] However, the question of whether a single south → north transition *per se* is sufficient for the appearance of a kink in DNA can hardly be answered, because in most DNA−protein complexes, or DNA duplexes with point modifications, it is almost impossible to separate purely conformational effects from other interactions.

To separate these effects, a thymidine nucleotide analogue (bh-T) **188** (Chart 74) having a bicyclo[3.1.0]hexane ring as a sugar mimic, has been synthesized. The cyclopropane ring of **188** locks the nucleotide conformation to the "north"

Chart 74. Bicyclo[3.1.0]hexane Thymidine Nucleoside 188



188

C G C G A A X Y C G C G
G C G C Y X A A G C G C

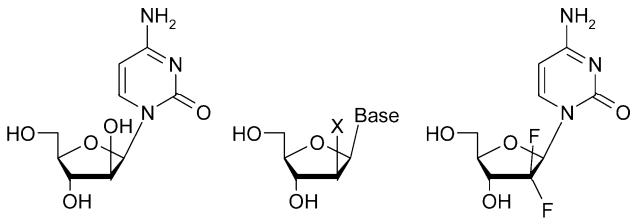
189a: X=T, Y=T
189b: X=T, Y=**188**
189c: X=**188**, Y=T
189d: X=Y=**188**

range without causing additional effects[509] and, therefore, facilitates the identification of conformational features in DNA that correlate solely with the sugar pseudorotation angle value. As we discussed previously, the accuracy in the determination of the global DNA shape improves considerably by running RMD with NOE distances and RDC restrains.[510] A combined NOESY−RDC NMR study of duplexes containing bh-T residues **189a**−**d** used one-bond natural-abundance $^{13}C-^{1}H$ couplings in anisotropic Pf1 phage solutions to obtain helical bend values with absolute errors smaller than 3°.[511] Dodecamers **189c** and **189b** having a single **188** residue at position 7 or 8 exhibit a helix bend toward the major groove of 10° or 5°, respectively. Furthermore, the double T substitution in **189d** is additive, displaying a bend value of 14.4° at the center of the duplex. In addition, since in the duplex **189c** the modified residues are adjacent and bend the helix in the same direction, the authors estimate that the transition of a single deoxyribose to a north conformation bends DNA by 5°. These data demonstrate that sugar transitions to a north conformation are sufficient to induce a local kink in DNA.

9.3.3. Arabinonucleosides and Their Derivatives inside a Duplex

It has been known for more than 40 years that β-arabinofuranosyl nucleotides and their derivatives have important anticancer and antiviral activity. Cytarabine **190** (Chart 75),

Chart 75. Arabinocytosine 190 and Its Analogues



190 **191a** : X = F **192**
191b : X = OAlk
191c : X = OMe, Base = T

β-arabinofuranosyl cytosine (araC), the most prominent representative of this family, is a potent antileukemic drug widely in use today in chemotherapy.[512] AraC undergoes metabolic activation to the corresponding 2′-deoxy-5′-triphosphate derivative that inhibits viral, bacterial, and mammalian DNA replication. DNA polymerases introduce araC residues in DNA, and the biological effects of the drug correlate with the amount of araC present in cellular DNA.[513] In contrast to terminators of DNA chain elongation, such as AZT or 2′,3′-dideoxyinosine, araC possesses a 3′-hydroxyl group capable of coupling with an incoming dNTP and, therefore, is only a partial block for DNA synthesis.[514]

As opposed to the more flexible 2′-deoxyribose moiety, the ribofuranose ring is more rigid and exists exclusively in a 2′-exo conformation, due to the additional *gauche* effect between the 2′-OH group and the 4′-oxygen. For the same reasons and, also, due to the formation of intramolecular hydrogen bonds,[515] araC,[516] 2′-deoxy-2′-fluoroarabinonucleotides **191a**,[517] and 2′-*O*-alkylarabinonucleotides **191b** mostly adopt a sugar conformation in the south range. As a result, incorporation of such analogues in duplex DNA should induce only small structural changes.

To elucidate the molecular basis of DNA synthesis inhibition, an NMR study[518] has compared the structure of the Dickerson dodecamer having araC at the ninth position (PDB code 170D) with that of the unmodified parent duplex (PDB code 171D). The NMR spectra demonstrated that the global conformation of the araC-containing duplex is B-type with WC base pair alignments throughout. However, despite the conformational similarity between free 2′-deoxy and arabino nucleotides, the presence of the 2′-OH group of araC causes intra- and internucleotide interactions that induce local structural perturbations at the lesion site. On the basis of the deshielding experienced by the ^{31}P signal of the 3′-phosphate and the proximity of the phosphate oxygen and the araC hydroxyl group, the authors proposed the existence of a hydrogen bond between these two groups. As the refined structure shows, formation of this hydrogen bond occurs because araC adopts a C4′-exo conformation, with a 55° pseudorotation angle, which makes the 2′-OH and 3′-phosphate groups pseudoequatorial and spatially close. In addition, the base pair rise at the araC9-G10 step increases to 4 Å, relieving steric clashes between araC 2′-OH and 3′-flanking gaunine H8 (Figure 43A and B). The same interaction is probably responsible for moderate variations of the propeller twist and buckle of lesion-flanking base pairs.

The above-described propeller twist and buckle changes are certainly small, and it is difficult to discriminate whether the presence of the lesion or duplex sequence effects originate them. Thus, it is not surprising that an earlier X-ray structure displayed a 2′-endo sugar conformation for the araC residue and no hydrogen bonding between its 2′-hydroxyl and the 3′-phosphate groups of araC.[519] It is unclear, however, whether crystal-packing forces can explain these differences. Similarly, a "south" sugar conformation is present in the structure of a 2′-*O*-methylarabinofuranosylthymine residue **191c** that replaces the second thymidine on a TATA box-containing duplex.[520] RMD data (PDB code 214D) show only negligible perturbations of the B-form DNA structure, with the south conformation of the modified sugar resulting in pseudoaxial orientations for both the 2′-methoxy and the 3′-flanking phosphate groups. It is worth noting that the 2′-methoxy group in **191c** is not a hydrogen donor and thus cannot form a hydrogen bond with the 3′-phosphate (Figure 43).

It has been observed that, during DNA replication, araC and its derivatives preferentially inhibit lagging strand synthesis[521,522] that progresses by ligation of relatively short DNA segments called Okazaki fragments.[523] Since the synthesis of Okazaki fragments starts from an RNA primer, its 5′-terminal side contains a junction between a chimeric RNA/

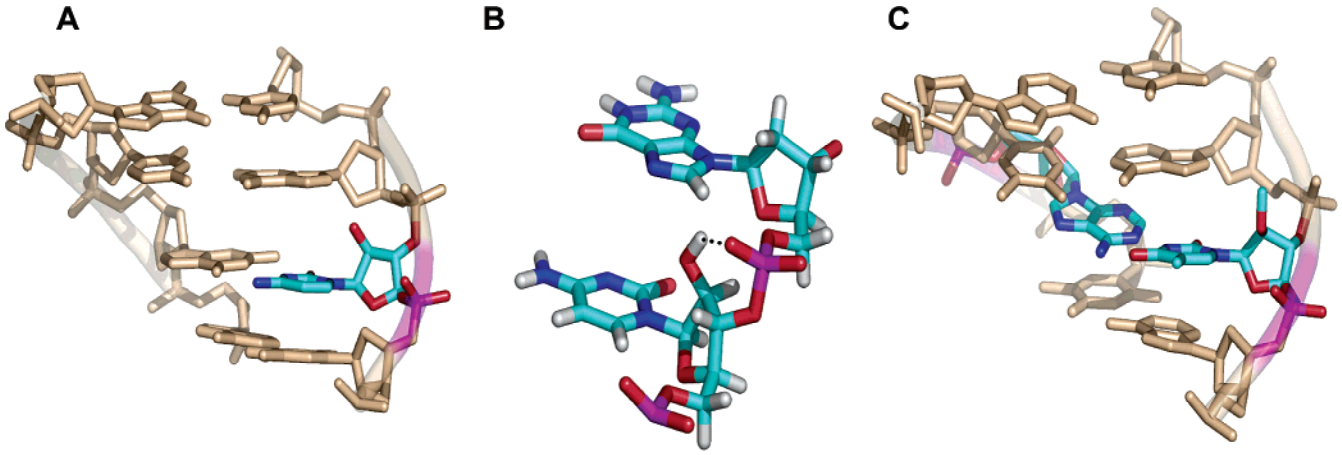


Figure 43. Structure of DNA duplexes containing modified sugar residues. (A) The central four-base-pair segment of a duplex having the cytarabine (araC) moiety **190** viewed from the major groove. Hydrogen atoms are not shown. (B) Close-up view of the araCpdG step showing the proximity of araC 2′-OH and guanine H^8 atoms and the putative hydrogen bond between araC 2′-OH and a 3′-phosphate oxygen. Prepared from PDB entry 170D[518] using PyMol.[555] (C) Four-base-pair segment containing the 2′-*O*-arabinofuranosyl thymine moiety **191c** viewed from the major groove. Hydrogen atoms are not shown. Prepared from PDB entry 214D[520] using PyMol.[555]

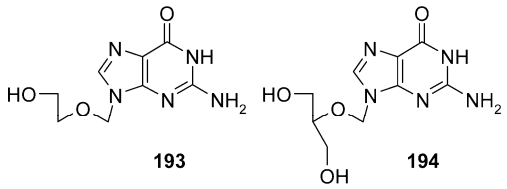
DNA duplex and a normal double-stranded DNA segment,[524] which possesses unique conformational features. Thus, it is conceivable that the biological effects of araC result from structural perturbations of the Okazaki fragment RNA/DNA junctions during lagging strand DNA synthesis. A thermodynamic study of the chimeric 5′-d(gaggaATXTTTG)·5′-d(CAAAGATTCCTC) duplex, where lower case letters denote ribonucleotides and X = araC, showed that the araC substitution increases the free energy of duplex formation by about 2 kcal/mol with respect the case of the parent dC-containing sample. In addition, the imino proton spectra of these duplexes suggested that araC destabilizes two base pairs adjacent to its 3′-side.[525] The subsequent NMR characterization of these samples establishes the presence regular helical conformations with WC alignments throughout and shows small local conformational differences between the parent and araC-containing Okazaki fragments.[526] The RMD structures demonstrate that araC increases an intrinsic 22° helix bend present on the unmodified RNA/DNA−DNA junction to 41°. Interestingly, the sugar of araC remains within the south conformational range and does not cause significant distortions of the structure of the DNA segment. The authors proposed that the change induced on the global shape of Okazaki fragments after araC incorporation might contribute to the inhibition of lagging strand synthesis and, thus, explain the cytotoxic effect of **190**.

The same group also studied the effect of 2′,2′-difluorocytidine **192** (gemcitabine) incorporation on the conformation of the same Okazaki fragment model.[527] Although gemcitabine, another nucleotide anticancer agent of current use, is believed to act by similar mechanisms as araC, the NMR characterization of the Okazaki fragment model demonstrated that, in contrast to araC, the gemcitabine sugar ring adopts a north conformation. According to NMR data, incorporation of **192** perturbs the local conformation of DNA and affects base stacking at the DNA·DNA/DNA·RNA junction but induces a smaller helix bend than araC.[527]

9.3.4. Ganciclovir—A Flexible Sugar Analogue

Acyclic guanosine analogues, such as acyclovir **193** and ganciclovir **194** (Chart 76), are nucleoside antiviral agents widely used for treatment of herpes and cytomegalovirus virus infections. This type of antiviral drugs requires

Chart 76. Acyclovir 193 and Ganciclovir 194



transformation to triphosphate derivatives, a process initiated by thymidine kinases, before they can inhibit DNA synthesis by terminating DNA chain elongation.[528] Since human thymidine kinase has reduced affinity for theses drugs, their metabolic activation reaches meaningful levels only in the presence of viral enzymes, explaining the selective inhibition of DNA synthesis in the virus-infected cells.[529,530] In contrast to acyclovir, which lacks a 3′-OH group and blocks chain elongation immediately after its DNA incorporation, ganciclovir **194** frequently allows incorporation of one or more natural nucleotides, resulting in the appearance of ganciclovir-containing duplex fragments before inhibition of DNA synthesis takes place.[531]

To characterize the thermodynamics and structural impact of ganciclovir **194** (GAN) incorporation in DNA, a self-complementary DNA duplex containing two symmetrically positioned GAN residues opposite dC has been prepared.[532] Van't Hoff analysis of UV melting curves showed that the presence of two GAN·dC pairs reduces duplex T_m by 13 °C and increases the free energy of duplex formation by 3.3 kcal/mol. The NMR structure of this duplex (PDB code 1AC9) has the guanine base of the lesion properly stacked inside the helix and forming normal WC hydrogen bonds with the partner cytosine on the complementary strand (Figure 44). Although the duplex did not show major perturbations of the global conformation, some structural parameters around the lesion site change significantly. The 3′-lesion-flanking base pair displayed increased flexibility and decreased twist angle values, and one of the sugar residues adopts a north conformation. On the basis of these observations, the authors propose that these transitions at the 3′-side of the lesion may account for the drug inhibitory effects on DNA synthesis.[533]

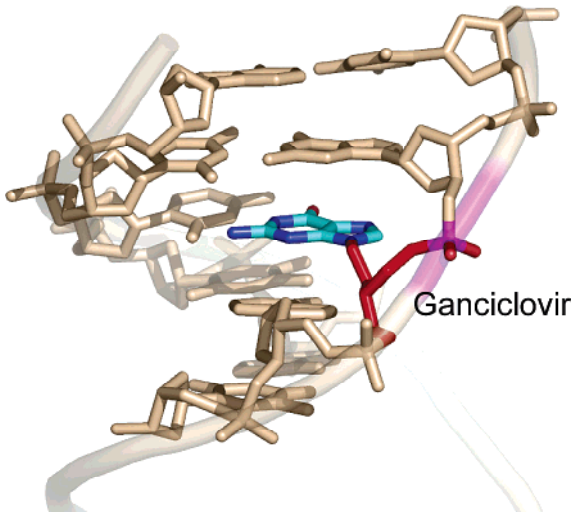


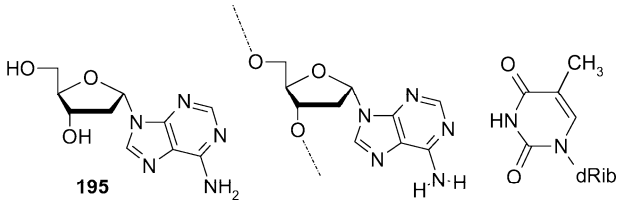
Figure 44. Five-base-pair segment of a DNA duplex containing a ganciclovir residue. The figure shows the modified residue in red with its guanine base in blue. Hydrogen atoms are not shown. Prepared from PDB entry 1AC9[533] using PyMol.[555]

9.3.5. Stereoisomers of Natural Nucleotides—α-Nucleosides

One consequence of free radical attack on DNA is the abstraction of the sugar 1′-hydrogen, which under anoxic conditions may lead to an inversion of the configuration of the anomeric carbon.[534] In the absence of molecular oxygen, α-2′-deoxyadenosine **195** (α-dA) is the main dA lesion formed after γ-irradiation of salmon testis DNA.[535] Notwithstanding a considerable replication delay, *E. coli* can bypass α-dA lesions by causing one-base deletion mutations in a sequence dependent manner.[536] In addition, α-dA·T pairs moderately affect thermodynamic properties, destabilizing the duplex by less than 2 kcal/mol.[537,538]

To understand the effects of α-dA in DNA structure and dynamics, NMR spectroscopy was used to characterize the solution structure of a decamer duplex having a single α-dA·T base pair at the fifth position.[538] The RMD structures (PDB codes 1S0T, 1S74, and 1S75) showed that duplex segments flanking the α-dA·T pair adopt a regular B-form conformation with sugar puckers in a south range and canonical WC base pair alignments. While the relative orientation of α-dA and T favors hydrogen bonding (Chart 77), the distance

Chart 77. α-Deoxyadenosine 195 and Its Alignment with T



between T H1 and α-dA N3 is about 0.3 Å longer than that in canonical base pairs, a fact that is consistent with the fast imino proton exchange rate observed for the α-dA·T base pair. The α configuration of the lesion anomeric carbon produces considerable displacement of α-dA toward the minor groove, which reduces stacking with the 5′-flanking dC residue but, simultaneously, improves the interaction with a cytosine at the 3′-side (Figure 45). The RMD structures also show a change in the ζ backbone torsion angle, that changes from a *gauche*$^-$ to an *anti* range, allowing the minor

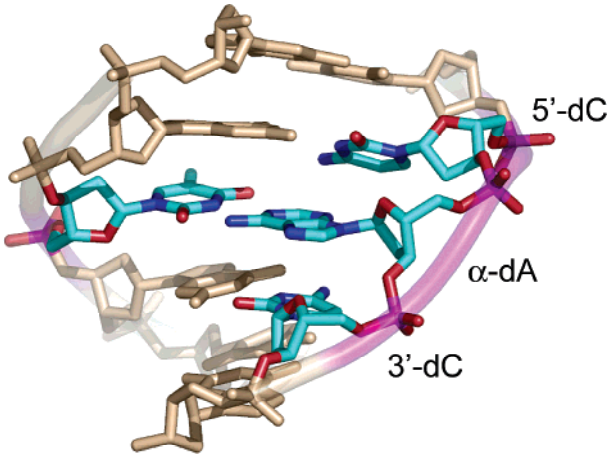
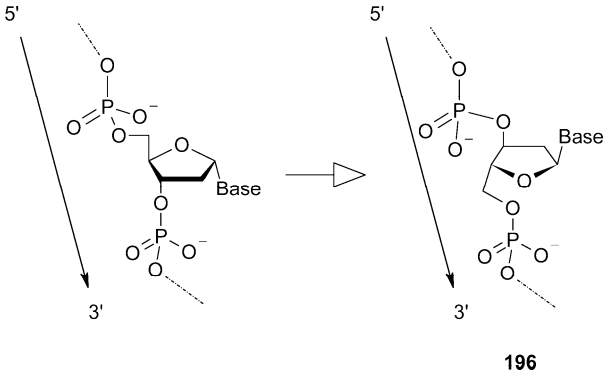


Figure 45. Structure of a five-base-pair segment of DNA containing α-deoxyadenosine opposite T. The picture displays a five-base-pair segment showing stacking of α-deoxyadenosine with the 3′-flanking dC and not with the 5′-flanking residue. Hydrogen atoms are not shown. Prepared from PDB entry 1S0T[538] using PyMol.[555]

groove displacement of α-dA. Since α-dA is poorly stacked with its 5′-neighbor, there is an increase of the base roll at this step and a decrease of the twist angle value that translate into a considerable global shape change. Molecular dynamics simulation using conventional NOE interproton distance bounds and 25 additional RDC restraints revealed the presence at the damage site of an 18° helix bend in the direction of the major groove. The authors concluded that these changes in global duplex shape might contribute to the recognition of α-dA residues by *E. coli* endonuclease IV, which is able to excise this lesion from DNA duplexes.[539]

The consequences of changing the anomeric carbon configuration may be compensated via polarity reversal of the α-nucleotide, by which it is inserted in the DNA backbone through 3′−3′ and 5′−5′ phosphodiester links **196** (Chart 78).[540] While such artificial constructs have never been

Chart 78. Concept of Polarity Reversal



observed inside the cells, they have drawn some attention in the area of antisense technology since oligonucleotides resistance to nuclease activity appears to increase by their presence.[541,542] The thermodynamics and structural properties of four self-complementary dodecamers containing iα-T·dA,[543] iα-dA·T, iα-dG·dC, and iα-dC·dG[544] pairs with polarity reversal on the α-nucleotides have been reported. In general, these duplexes exhibit moderate reduction in thermal stability with T_m values 5−10 °C smaller than those of unmodified control samples. Imino proton resonances appear as sharp signals in the low field region of the spectra, indicating that internal base pairs of the duplex, including

those containing α-nucleotides, are hydrogen bonded. Analysis of scalar coupling constants, as well as NOE interactions, indicates the presence of regular B-type helices with sugar rings of unmodified residues in a south conformation. The only exception is the iα-dG·dC duplex, for which the experimental data predict a significant perturbation of the conformation of the sugar−phosphate backbone that is probably correlated with the low thermal stability of this sample.

Due to the reversed backbone linkage and anomeric configuration of the α-nucleotides, the direct comparison of standard spectral features and structural parameters between normal and modified duplexes, such as glycosidic torsion angle values or NOESY patterns, is of little relevance. Thus, while sequential NOE interactions between the aromatic proton of iα-nucleotides and H2′/H2″ protons of the 3′-neighbor are broken, they are normal with the H1′ proton, indicating no overall violations of helix geometry. Similarly, the conformation of the inverted nucleotide sugar is not a proper criterion to evaluate helix type. For example, RMD structures of the iα-T·dA duplex **197** (PDB code 1BX5) (Figure 46)

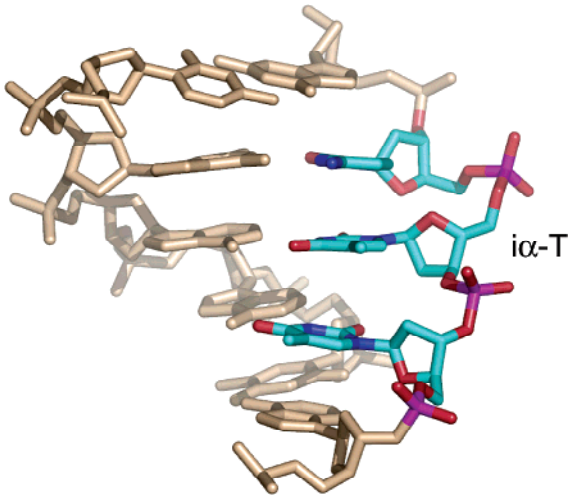


Figure 46. Five-base-pair segment of the DNA duplex containing an incorporated iα-T·dA pair viewed from the major groove. Hydrogen atoms are not shown. Prepared from PDB entry 1BX5[545] using PyMol.[555]

show minimal perturbations at the modification site, despite the unusual $\chi = 140°$ and pseudorotation angle in the 200−225° range.[545] In addition, the comparison of experimental coupling constant values with those derived from RMD structures shows that the dC residue attached to the 5′-side of iα-T is present with almost equal fractions of south and north conformers. Therefore, to better describe the conformation and dynamics of the iα-T·dA duplex, the authors used time-averaged RMD (TAR), which enforces experimental restraints over a period of time, rather than at each simulation step, increasing the agreement with experimental data.[546] The TAR simulations confirmed the large conformational flexibility of the d*C*[8] residue forecasted by RMD and demonstrated that the 3′−3′ phosphodiester linkage is comparatively rigid in contrast to the highly flexible 5′−5′ linkage.[547]

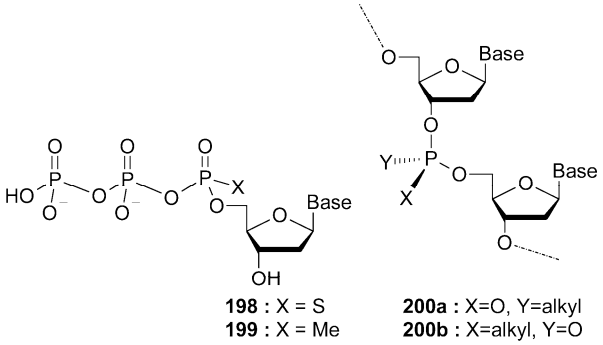
9.3.6. Modification of Phosphodiester Bonds

During the period of increased popularity of antisense DNA approaches, extensive biophysical and biochemical studies have characterized the properties of oligodeoxynucleotides having modified phosphodiester bonds. Several phosphate modifications have been evaluated for increasing resistance to nuclease degradation, better hybridization abilit,y or cell penetration properties.[548] Far less attention has been paid, however, to the effect of point phosphate modifications in DNA structure and dynamics.

Phosphodiesters are the most chemically stable phosphate esters, and pentavalent phosphorus is extremely resistant to oxidation or reduction. As a result, changes of the DNA phosphodiester bond by the usual damaging agents are not frequent. In addition, DNA polymerases are especially sensitive to changes in the dNTP α-phosphate, and hence, the appearance of phosphate modifications in DNA following the enzymatic incorporation of non-natural nucleotide triphosphates is possible only for a small set of analogues. The effective enzymatic incorporations of phosphorothioate- and methylphosphonate-containing[549,550] dNTP analogues **198** and **199** are rare examples of such events.

The substitution of negatively charged phosphate oxygen with a neutral hydrophobic alkyl group significantly affects the hydration shell of DNA and electrostatic strand interactions. NMR spectroscopy has been used in order to evaluate the structural impact of having all *S* or *R* methyl phosphonate moieties in the 7-mer strand of 7-mer−8-mer duplexes. The presence of six methyl phosphonate substitutions affects significantly the hybridization and conformational properties of the DNA, which differ from those of the unmodified parent duplex and between the *S* and *R* methyl phosphonate isomers[551] (PDB codes 1K1R and 1K1H). The structural effects of a single phosphonate modification on DNA have been studied on decamer duplexes having a centrally located *R* or *S* octyl phosphonate group in one strand (**200a**,**b**, Chart 79).[552] The RMD structures (PDB codes 1IEY and 1IEK,

Chart 79. Modified Phosphodiester Bond in Nucleotide Analogues 198 and 199 and DNA 200a,b



correspondingly) indicate that both duplexes have the essential features of B-form DNA but with an increase of the north population of the phosphonate-linked sugar residues. In both duplexes, the octyl chain is highly flexible and adopts multiple orientations but without affecting duplex conformation, suggesting that the configuration of the phophonate group plays a secondary role in duplex structure. In addition, the RMD models show the presence of a helix bend at the center of these duplexes. On the basis of these data the authors proposed that the phosphonate link increases both local nucleotide and global duplex flexibility at the modification site, supporting the earlier hypothesis that nonuniform neutral substitution of negatively charged phosphates causes asymmetrical cancellation of electrostatic repulsions and induces DNA bends.[553]

10. Concluding Remarks

Research in the last two decades has greatly increased our knowledge of damaged DNA structure. This bulk of information allowed the identification of common conformational features that describe basic principles of damaged DNA conformation, intermediates of lesion-induced mutagenesis, and lesion recognition by repair proteins. We briefly describe some of these principles below.

Most Damages Cause Local Perturbations of the DNA Structure. With only a few exceptions, the structures reported to date show that damages cause perturbations at the lesion site and, frequently, flanking base pairs. In some instances, structural perturbations going beyond this point have been reported, but they are always minor and generally within the experimental error of the NMR methods used for structure determination. Thus, whether these long-range structural perturbations result from the presence of the lesion or express sequence dependent effects is, in many cases, still an open question. The plasticity of DNA, which limits the extent of lesion-induced structural perturbations, has important biological consequences. On one hand, DNA segments near the lesion site conserve to some extent the ability to participate of many metabolic transactions, including transcription and replication, a fact that reduces the acute toxicity that DNA damages have on the cell. Even when the lesion is present in a coding region, the existence of a transcription-coupled repair mechanism, which specifically repairs lesions present on the transcribed strand, further ensures cellular activity in the presence of low DNA damage levels. On the other hand, the moderate changes that most damages induce on the structure of DNA require a sophisticated mechanism of lesion recognition by the DNA repair machinery. Since repair mechanims are not 100% efficient, a background level of lesion-induced mutations are always present, a fact that correlates with disease etiology and the aging process.

Not All Lesions Destabilize the Double-Helical Structure. While the large majority of damages cause a reduction in the stability of DNA duplexes, some lesions produce the opposite effect. Reflecting the NMR ability to observe hydrogen-bonded proton signals, lesion-induced stability effects are frequently explained in terms of gain or loss in hydrogen-bonding interactions at the lesion site. However, this bias generates a simplified view of DNA stability, and in most instances, the free energy of the duplex is determined by its entropy content. When damages stabilize the double helix, biological consequences may be of paramount importance because lesion recognition and/or processing by the repair pathways is affected by duplex stability, and thus, persistence of these lesions in the cellular genome would increase the onset of lesion-induced mutations.

Lesion-Induced Mutagenesis during DNA Replication Reflects WC-like Geometry. The structure of many lesion-containing mismatches revealed that those retaining a regular structure on the sugar−phosphate backbone have a tendency to appear *in vivo* or *in vitro* during DNA replication. Conversely, damaged mismatches with distorted geometry occur very seldom regardless of their stability. These observations suggest that the structures of these damaged duplexes may serve as satisfactory models for mutagenic replication intermediates and, importantly, are consistent with the concept that replicative DNA polymerases favor the insertion of dNTPs that can adopt a WC-like geometry at their active sites. However, many lesions, including bulky adducts, pyrimidine photoproducts, or exocyclic lesions, mostly block DNA synthesis catalyzed by replicative polymerases, so the involvement of translesion synthesis (TLS) polymerases is needed. Since TLS polymerases have a wider active site, the need for a WC-like geometry for mutagenic intermediates is less imperative and NMR studies of lesion-containing mismatches provide, in this case, limited information about the structural basis of mutations. Nevertheless, the structure of these mutagenic intermediates frequently shows specific hydrogen-bonding interactions, suggesting that stabilization of lesion-containing mismatches plays a role during inception or for propagation of chemical mutagenesis.

Lesion Site Conformation Is Insufficient for a Complete Understanding of Damage Recognition. Cell repair systems have to recognize and eliminate any deviation from the canonical DNA structure. Several X-ray structures have provided many insights into mechanisms of damaged recognition and excision. Before removal of the modified base or incision of the DNA backbone, the BER system kinks the helix extruding the damaged residue from the duplex and into the enzyme active site. Independently, the NER apparatus creates a long internal loop and treats the damaged and intact strands of DNA separately. These protein−DNA interactions in many aspects are similar to those seen in other catalytic processes. For example, steric constraints present in the catalytic pocket of UDG explain the exclusion of thymine as its substrate. However, the whole spectrum of lesions that are recognized by some glycosylases cannot be explained by specific active site interactions. In many instances, some lesions that are not substrates for a specific DNA glycosylase can be readily modeled in the protein active site, suggesting that they are screened out during the initial recognition process. In contrast to latter steps of lesion processing, its initial recognition can hardly be described by the classical "key and lock" concept and, therefore, is much less understood. Obscuring this issue is the fact that DNA damages recognized by different enzymes generally show only minor structural perturbations at the lesion site and exhibit small variations in their shape. Furthermore, what makes a lesion detected as damage against the predominant background of unmodified DNA residues is still an open question, and it is almost certain that lesion-induced dynamics is an important component in this process. Similarly, while the NER system recognizes a broad range of chemical modifications, the large variations on processing efficiency suggest that the duplex stability, its global shape, and the local conformation affect the initial recognition of lesions. Thus, the structure and dynamics of lesion-containing duplexes provide elements required for their recognition by the repair systems. Future research on lesion-induced changes in global shape and local stability of DNA and variations of its dynamic behavior are needed in order to fully characterize lesion recognition processes.

In addition to the vast accomplishments of the last 20 years, improved spectroscopic methods available today and implementation of complementary experimental approaches promise to generate a new wealth of structural information in the field of damaged DNA structure. Development of novel methods for ^{13}C-, ^{15}N-, and ^{2}H-labeling of undamaged and damaged nucleotides will improve the accuracy of global shape features determined by NMR spectroscopy and will facilitate studies of damaged DNA−protein complexes. Experimental and computational studies of duplex dynamics will identify motional modes that can correlate with conformational rearrangements accompanying the formation of

protein−DNA complexes. Structural and thermodynamics characterization of lesion-containing duplexes and protein-damaged DNA complexes is certain to uncover novel structure−function relationships. X-ray structures of damaged DNA−protein complexes will further explain mechanisms of mutagenesis and repair. All these studies promise to answer in the nondistant future most of the outstanding questions still remaining on the structural biology of damaged DNA.

11. Acknowledgments

We thank Tanya Zaliznyak for helping with the preparation of many figures and Dr. Holly Miller for critical reading of this review. NIH Grants CA47995 and CA77094 supported his work.

12. References

(1) http://www.rcsb.org/pdb/holdings.html (March 8, 2005).
(2) Ferre-D'Amare, A. R.; Zhou, K.; Doudna, J. A. *J. Mol. Biol.* **1998**, *279*, 621.
(3) Batey, R. T.; Battiste, J. L.; Williamson, J. R. *Methods Enzymol.* **1995**, *261*, 300.
(4) Zimmer, D. P.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3091.
(5) Werner, M. H.; Gupta, V.; Lambert, L. J.; Nagata, T. *Methods Enzymol.* **2001**, *338*, 283.
(6) Varani, G.; Aboul-ela, F.; Allain, F. H.-T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 51.
(7) Majumdar, A.; Kettani, A.; Skripkin, E. *J. Biomol. NMR* **1999**, *14*, 67.
(8) Majumdar, A.; Kettani, A.; Skripkin, E.; Patel, D. J. *J. Biomol. NMR* **1999**, *15*, 211.
(9) Zidek, L.; Stefl, R.; Sklenar, V. *Curr. Opin. Struct. Biol.* **2001**, *11*, 275.
(10) Trantírek, L.; Štefl, R.; Masse, J. E.; Feigon, J.; Sklenar, V. *J. Biomol. NMR* **2002**, *23*, 1.
(11) Orrell, K. G.; Sik, V.; Stephenson, D. *Prog. Nucl. Magn. Reson. Spectrosc.* **1990**, *22*, 141.
(12) Lane, A. N. *Prog. Nucl. Magn. Reson. Spectrosc.* **1993**, *25*, 481.
(13) Ravindranathan, S.; Kim, C. H.; Bodenhausen, G. *J. Biomol. NMR* **2003**, *27*, 365.
(14) Zhang, Q.; Throolin, R.; Pitt, S. W.; Serganov, A.; Al-Hashimi, H. M. *J. Am. Chem. Soc.* **2003**, *125*, 10530.
(15) O'Neil-Cabello, E.; Bryce, D. L.; Nikonowicz, E. P.; Bax, A. *J. Am. Chem. Soc.* **2004**, *126*, 66.
(16) Boisbouvier, J.; Wu, Z. R.; Ono, A.; Kainosho, M.; Bax, A. *J. Biomol. NMR* **2003**, *27*, 133.
(17) Lipsitz, R. S.; Tjandra, N. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 387.
(18) Kojima, C.; Ono, A.; Ono, A.; Kainosho, M. *Methods Enzymol.* **2001**, *338*, 261.
(19) Patel, D. J.; Shapiro, L.; Hare, D. *Q. Rev. Biophys.* **1987**, *20*, 35.
(20) Van de Ven, F. J. M.; Hilbers, C. W. *Eur. J. Biochem.* **1988**, *178*, 1.
(21) Majumdar, A.; Hosur, R. V. *Prog. Nucl. Magn. Reson. Spectrosc.* **1992**, *24*, 109.
(22) de los Santos, C. In *Comprehensive Natural Products Chemistry*; Kool, E., volume Ed.; Elsevier Science Ltd.: Oxford, 1999; Vol. 7, p 55.
(23) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. *Chem. Rev.* **2004**, *104*, 3519.
(24) Moe, J. G.; Russu, I. M. *Biochemistry* **1992**, *31*, 8421.
(25) Goljer, I.; Withka, J. M.; Kao, J. Y.; Bolton, P. H. *Biochemistry* **1992**, *31*, 11614.
(26) Gonzalez, C.; Stec, W.; Kobylanska, A.; Hogrefe, R. I.; Reynolds, M.; James, T. L. *Biochemistry* **1994**, *33*, 11062.
(27) Leijon, M.; Leroy, J. L. *Biochimie* **1997**, *79*, 775.
(28) Palmer, A. G., III. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 129.
(29) Isaacs, R. J.; Spielmann, H. P. *DNA Repair* **2004**, *3*, 455.
(30) Fazakerley, G. V.; Boulard, Y. *Methods Enzymol.* **1995**, *261*, 145.
(31) Chou, S.-H.; Chin, K.-H.; Wang, A. H.-J. *Nucleic Acids Res.* **2003**, *31*, 2461.
(32) IUPAC−IUB Joint Commission on Biochemical Nomenclature. *Eur. J. Biochem.* **1983**, *131*, 9.
(33) Demple, B.; Harrison, L. *Annu. Rev. Biochem.* **1994**, *63*, 9154.
(34) Nakamura, J.; Walker, V. E.; Upton, P. B.; Chiang, S. Y.; Kow, Y. W.; Swenberg, J. A. *Cancer Res.* **1998**, *58*, 222.
(35) Fromme, J. C.; Banerjee, A.; Verdine, G. L. *Curr. Opin. Struct. Biol.* **2004**, *14*, 43.
(36) McCullough, A. K.; Dodson, M. L.; Lloyd, R. S. *Annu. Rev. Biochem.* **1999**, *68*, 255.
(37) Krokan, H. E.; Standal, R.; Slupphaug, G. *Biochem. J.* **1997**, *325*, 1.
(38) Lindahl, T.; Nyberg, B. *Biochemistry* **1972**, *11*, 3610.
(39) Kappen, L. S.; Goldberg, I. H. *Biochemistry* **1989**, *28*, 1027.
(40) Pieters, J. M. L.; Manst, R. M. W.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H.; Altona, C. *Nucleic Acids Res.* **1989**, *17*, 4551.
(41) Snowden-Ifft, E. A.; Wemmer, D. E. *Biochemistry* **1990**, *29*, 6017.
(42) Kozerski, L.; Mazurek, A. P.; Kawecki, R.; Bocian, W.; Krajewski, P.; Bednarek, E.; Sitkowski, J.; Williamson, M. P.; Moir, A. J. G.; Hansen, P. E. *Nucleic Acids Res.* **2001**, *29*, 1132.
(43) Mills, J. B.; Cooper, J. P.; Hagerman, P. J. *Biochemistry* **1994**, *33*, 1797.
(44) Singh, S.; Patel, P. K.; Hosur, R. V. *Biochemistry* **1997**, *36*, 13214.
(45) Roll, C.; Ketterlé, C.; Faibis, V.; Fazakerley, G. V.; Boulard, Y. *Biochemistry* **1998**, *37*, 4059.
(46) Roll, C.; Ketterlé, C.; Boulard, Y.; Fazakerley, G. V. *Biochimie* **1997**, *79*, 593.
(47) Roll, C.; Ketterlé, C.; Faibis, V.; Fazakerley, G. V.; Boulard, Y. *Eur. J. Biochem.* **1999**, *264*, 120.
(48) Claussen, C. A.; Long, E. C. *Chem. Rev.* **1999**, *99*, 2797.
(49) Junker, H.-D.; Hoehn, S. T.; Bunt, R. C.; Marathius, V.; Chen, J.; Turner, C. J.; Stubbe, J. *Nucleic Acids Res.* **2002**, *30*, 5497.
(50) Guo, H.; Tullius, T. D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3743.
(51) Singh, M. P.; Hill, G. C.; Péoc'h, D.; Rayner, B.; Imbach, J. L.; Lown, J. W. *Biochemistry* **1994**, *33*, 10271.
(52) Manoharan, M.; Mazumder, A.; Ransom, S. C.; Gerlt, J. A.; Bolton, P. H. *J. Am. Chem. Soc.* **1988**, *110*, 2690.
(53) Kalnik, M. W.; Chang, C.-N.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1989**, *28*, 3373.
(54) Vesnaver, G.; Chang, C.-N.; Eisenberg, M.; Grollman, A. R.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3614.
(55) Kalnik, M. W.; Chang, C. N.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1988**, *27*, 924.
(56) de los Santos, C.; El-khateeb, M.; Rege, P.; Tian, K.; Johnson, F. *Biochemistry* **2004**, *43*, 15349.
(57) Manoharan, M.; Ransom, S. C.; Mazumder, A.; Gerlt, J. A.; Wilde, J. A.; Withka, J. A.; Bolton, P. H. *J. Am. Chem. Soc.* **1988**, *110*, 1620.
(58) Wilde, J. A.; Bolton, P. H.; Mazumder, A.; Manoharan, M.; Gerlt, J. A. *J. Am. Chem. Soc.* **1989**, *111*, 1894.
(59) Beger, R. D.; Bolton, P. H. *J. Biol. Chem.* **1998**, *273*, 15565.
(60) Goljer, I.; Kumar, S.; Bolton, P. H. *J. Biol. Chem.* **1995**, *270*, 22980.
(61) Wang, K. Y.; Parker, S. A.; Goljer, I.; Bolton, P. H. *Biochemistry* **1997**, *36*, 11629.
(62) Hoehn, S. T.; Turner, C. J.; Stubbe, J. *Nucleic Acids Res.* **2001**, *29*, 3413.
(63) Cuinasse, P.; Fazakerley, G. V.; Guschlbauer, W.; Kaplan, B. E.; Sowers, L. S. *J. Mol. Biol.* **1990**, *213*, 303.
(64) Cuinasse, P.; Sowers, L. S.; Eritja, R.; Kaplan, B. E.; Goodman, M. F.; Cognet, J. A. H.; LeBret, M.; Guschlbauer, W.; Fazakerley, G. V. *Nucleic Acids Res.* **1987**, *15*, 8003.
(65) Coppel, Y.; Berthet, N.; Coulombeau, C.; Coulombeau, C.; Garcia, J.; Lhomme, J. *Biochemistry* **1997**, *36*, 4817.
(66) Wilson, D. M., III; Takeshita, M.; Grollman, A. P.; Demple, B. *J. Biol. Chem.* **1995**, *270*, 16002.
(67) Mol, C. D.; Izumi, T.; Mitra, S.; Tainer, J. *Nature* **2000**, *403*, 451.
(68) Nakano, S.-I.; Chadalavada, D. M.; Bevilaqua, P. C. *Science* **2000**, *287*, 1493.
(69) Mao, H.; Schnetz-Boutaud, N. C.; Weisenseel, J. P.; Marnett, L. J.; Stone, M. P. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6615.
(70) Jourdan, M.; Garcia, J.; Defrancq, E.; Kotera, M.; Lhomme, J. *Biochemistry* **1999**, *38*, 3985.
(71) Iakubovskaia, E. A.; Gabibov, A. G. *Mol. Biol. (Mosk.)* **1999**, *33*, 368.
(72) Kingma, P. S.; Osheroff, N. *J. Biol. Chem.* **1997**, *272*, 1148.
(73) Cline, S. D.; Jones, W. R.; Stone, M. P.; Osheroff, N. *Biochemistry* **1999**, *38*, 15500.
(74) Zahn, K.; Blattner, F. R. *Science* **1987**, *236*, 416.
(75) Schroth, G. P.; Siino, J. S.; Cooney, C. A.; Th'ng, J. P. H.; Ho, P. S.; Bradbury, E. M. *J. Biol. Chem.* **1992**, *267*, 9958.
(76) Ohyama, T. *BioEssays* **2001**, *23*, 708.
(77) MacDonald, D.; Herbert, K.; Zhang, X.; Pologruto, T.; Lu, P. *J. Mol. Biol.* **2001**, *306*, 1081.
(78) Wilson, D. M., III; Barsky, D. *Mutat. Res.* **2001**, *485*, 283.
(79) Klinedinst, D. K.; Drinkwater, N. R. *Mol. Carcinog.* **1992**, *6*, 32. Lawrence, C. W.; Borden, A.; Banerjee, S. K.; LeClerc, J. E. *Nucleic Acids Res.* **1990**, *18*, 2153.

(80) Cuniasse, Ph.; Sowers, L. C.; Eritja, R.; Kaplan, B.; Goodman, N. F.; Cognet, J. A. H.; LeBret, M.; Guschlbauer, W.; Fazakerly, G. V. *Biochemistry* **1989**, *28*, 2018.
(81) Lin, Z.; Hung, K.-N.; Grollman, A. P.; de los Santos, C. *Nucleic Acids Res.* **1998**, *10*, 2385.
(82) Ward, J. F. *Prog. Nucleic Acids Res. Mol. Biol.* **1988**, *35*, 95.
(83) Sutherland, B. M.; Bennett, P. V.; Sidorkina, O.; Laval, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 103.
(84) Chaudhry, M. A.; Weinfeld, M. *J. Biol. Chem.* **1997**, *272*, 15650.
(85) David-Cordonnier, M. H.; Cunniffe, S. M.; Hickson, I. D.; O'Neill, P. *Biochemistry* **2002**, *41*, 634.
(86) Lomax, M. E.; Cunniffe, S.; O'Neill, P. *Biochemistry* **2004**, *43*, 11017.
(87) Lomax, M. E.; Salje, H.; Cunniffe, S.; O'Neill, P. *Radiat. Res.* **2005**, *163*, 79.
(88) Lin, Z.; de los Santos, C. *J. Mol. Biol.* **2001**, *308*, 341.
(89) Hazel, R. D.; Tian, K.; de los Santos, C. Submitted.
(90) Woodson, S. A.; Crothers, D. M. *Biochemistry* **1988**, *27*, 8904.
(91) Kappen, L. S.; Goldberg, I. H. *Biochemistry* **1997**, *36*, 14861.
(92) Williams, H. E. L.; Colgrave, M. L.; Searle, M. S. *Eur. J. Biochem.* **2002**, *269*, 1726.
(93) Williams, H. E.; Searle, M. S. *J. Mol. Biol.* **1999**, *290*, 699.
(94) Zhang, X. L.; Patel, D. J. *Biochemistry* **1990**, *29*, 9451.
(95) Hoehn, S. T.; Junker, H. D.; Bunt, R. C.; Turner, C. J.; Stubbe, J. *Biochemistry* **2001**, *40*, 5894.
(96) Bertrand, J.-R.; Vasseur, J.-J.; Gouyette, A.; Rayner, B.; Imbach, J.-L.; Paoletti, C.; Malvy, C. *J. Biol. Chem.* **1989**, *264*, 14172.
(97) Matray, T. J.; Kool, E. T. *J. Am. Chem. Soc.* **1998**, *120*, 6191.
(98) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704.
(99) Smirnov, S.; Matray, T. J.; Kool, E. T.; de los Santos, C. *Nucleic Acids Res.* **2002**, *30*, 5561.
(100) Demple, B.; Harrison, L. *Annu. Rev. Biochem.* **1994**, *63*, 915.
(101) Collins, A. R.; Cadet, J.; Moller, L.; Poulsen, H. E.; Vina, J. *Arch. Biochem. Biophys.* **2004**, *423*, 57.
(102) Dizdaroglu, M. *Mutat. Res.* **2003**, *531*, 109.
(103) Téoule, R. *Int. J. Radiat. Res.* **1987**, *51*, 573.
(104) Rouet, P.; Essigmann, J. M. *Cancer Res.* **1985**, *45*, 6113.
(105) Ide, H.; Kow, Y. W.; Wallace, S. S. *Nucleic Acids Res.* **1985**, *13*, 8035.
(106) Clark, J. M.; Beardsley, G. P. *Nucleic Acids Res.* **1986**, *14*, 737.
(107) Basu, A. K.; Loechler, E. L.; Leadon, S. A.; Essigmann, J. M. *Proc. Natl. Acad. Scl. U.S.A.* **1989**, *86*, 7677.
(108) Kao, J. E.; Goljer, I.; Phan, T. A.; Bolton, P. H. *J. Biol. Chem.* **1993**, *268*, 17787.
(109) Kung, H. C.; Bolton, P. H. *J. Biol. Chem.* **1997**, *272*, 9227.
(110) Lustig, M. J.; Cadet, J.; Boorstein, R. J.; Teebor, G. W. *Nucleic Acids Res.* **1992**, *25*, 4839.
(111) Henderson, P. T.; Neeley, W. L.; Delaney, J. C.; Gu, F.; Niles, J. C.; Hah, S. S.; Tannenbaum, S. R.; Essigmann, J. M. *Chem. Res. Toxicol.* **2005**, *18*, 12.
(112) Cadet, J.; Nardin, R.; Voituriez, L.; Remin, M.; Hruska, F. E. *Can. J. Chem.* **1981**, *59*, 3313.
(113) Gervais, V.; Guy, A.; Teoule, R.; Fazakerley, G. V. *Nucleic Acids Res.* **1992**, *20*, 6455.
(114) Guy, A.; Duplaa, A. M.; Ulrich, J.; Téoule, R. *Nucleic Acids Res.* **1991**, *19*, 5815.
(115) Maufrais, C.; Boulard, Y. *Can. J. Physiol. Pharmacol.* **2002**, *80*, 609.
(116) Maufrais, C.; Fazakerley, G. V.; Cadet, J.; Boulard, Y. *Biochemistry* **2000**, *39*, 5614.
(117) Maufrais, C.; Fazakerley, G. V.; Guy, A.; Cadet, J.; Boulard, Y. *Biochimie* **2000**, *82*, 65.
(118) Maufrais, C.; Fazakerley, G. V.; Cadet, J.; Boulard, Y. *Nucleic Acids Res.* **2003**, *31*, 5930.
(119) Lewis, H. L.; Muhelman, D. R.; Ward, J. F. *Radiat. Res.* **1978**, *75*, 305.
(120) Teebor, G. W.; Frenkel, K.; Goldstein, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 318.
(121) Wagner, R. W.; Matteucci, M. D.; Lewis, J. C.; Gutierrez, A. J.; Moulds, C.; Froehler, B. C. *Science* **1993**, *260*, 1510.
(122) Langer, P. R.; Waldrop, A. A.; Ward, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 6633.
(123) Alexandrova, L. A.; Lukin, M. A.; Rozovskaia, T. A.; Atrazhev, A. M.; Kukhanova, M. K.; Kraevskii, A. A. *Mol. Biol. (Mosk.)* **1990**, *24*, 1100.
(124) Kallen, R.; Simon, M.; Marmur, J. *J. Mol. Biol.* **1962**, *5*, 248.
(125) Boorstein, R. J.; Teebor, G. W. *Cancer Res.* **1988**, *48*, 5466.
(126) Cannon-Carlson, S. V.; Gokhale, H.; Teebor, G. W. *J. Biol. Chem.* **1989**, *264*, 13306.
(127) Mellac, S.; Fazakerley, G. V.; Sowers, L. S. *Biochemistry* **1993**, *32*, 7779.
(128) Pasternack, L. B.; Bramham, J.; Mayol, L.; Galeone, A.; Xin Jia, X.; Kearns, D. R. *Nucleic Acids Res.* **1996**, *24*, 2740.
(129) Vu, H. M.; Pepe, A.; Mayol, L.; Kearns, D. R. *Nucleic Acids Res.* **1999**, *27*, 4143.
(130) Greene, J.; Geiduschek, E. P. *EMBO. J.* **1985**, *4*, 1345.
(131) Geiduschek, E. P.; Schneider, G. J.; Sayre, M. H. *J. Struct. Biol.* **1990**, *104*, 84.
(132) Wagner, J. R.; Hu, C. C.; Ames, B. N. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3380.
(133) Kreutzer, D. A.; Essigmann, J. M. *Proc Natl. Acad. Sci. U.S.A.* **1998**, *95*, 3578.
(134) Thiviyanathan, V.; Somasunderam, A.; Volk, D. E.; Gorenstein, D. G. *Chem. Commun.* **2005**, 400.
(135) Martinez, G. R.; Loureiro, A. P. M.; Marques, S. A.; Miyamoto, S.; Yamaguchi, L. F.; Onuki, J.; Almeida, E. A.; Garcia, C. C. M.; Barbosa, L. F.; Medeiros, M. H. G.; Di Mascio, P. *Mutat. Res.* **2003**, *544*, 115.
(136) Shibutani, S.; Takeshita, M.; Grollman, A. P. *Nature* **1991**, *349*, 431.
(137) Avkin, S.; Livneh, Z. *Mutat. Res.* **2002**, *510*, 81.
(138) Guschlbauer, W.; Duplaa, A.-M.; Guy, A.; Téoule, R.; Fazakerley, G. V. *Nucleic Acids Res.* **1991**, *19*, 1753.
(139) Wood, M. L.; Esteve, A.; Morningstar, M. L.; Kuziemko, G. M.; Essigmann, J. M. *Nucleic Acids Res.* **1992**, *20*, 6023.
(140) Kamiya, H.; Miura, H.; Murata-Kamiya, N.; Ishikawa, H.; Sakaguchi, T.; Inoue, H.; Sasaki, T.; Masutanil, C.; Hanaoka, F.; Nishimura, S.; Ohtsuka, E. *Nucleic Acids Res.* **1995**, *23*, 2893.
(141) Gannett, P. M.; Sura, T. P. *Chem. Res. Toxicol.* **1993**, *6*, 690.
(142) Cho, B. P.; Kadlubar, F. F.; Culp, S. J.; Evans, F. E. *Chem. Res. Toxicol.* **1990**, *3*, 445.
(143) Oda, Y.; Uesugi, S.; Ikehara, M.; Nishimura, S.; Kawase, Y.; Ishikawa, H.; Inoue, H.; Ohtsuka, E. *Nucleic Acids Res.* **1991**, *19*, 1407.
(144) Kouchakdjian, M.; Bodepudi, V.; Shibutani, S.; Eisenberg, M.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1991**, *30*, 1403.
(145) Thiviyanathan, V.; Somasunderam, A.; Hazra, T. K.; Mitra, S.; Gorenstein, D. G. *J. Mol. Biol.* **2003**, *325*, 433.
(146) Chen, H.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Magn. Reson. Chem.* **1996**, *34*, 23.
(147) Moran, S.; Ren, R. X.-F.; Rumney, S., IV; Kool, E. T. *J. Am. Chem. Soc.* **1997**, *119*, 2056.
(148) Delaney, J. C.; Henderson, P. T.; Helquist, S. A.; Morales, J. C.; Essigmann J. M.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4469.
(149) Yang, X. L.; Sugiyama, H.; Ikeda, S.; Saito, I.; Wang, A. H. J. *Biophys. J.* **1998**, *75*, 1163.
(150) Sepiol, J.; Kazimierczuk, Z.; Shugar, D. *Z. Naturforsch.* **1976**, *31C*, 361.
(151) Robinson, H.; Gao, Y. G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H.-J. *Biochemistry* **1998**, *37*, 10897.
(152) Bukowska, A. M.; Kusmierek, J. T. *Acta Biochim. Pol.* **1996**, *43*, 247.
(153) Shapiro, R.; Klein, R. S. *Biochemistry* **1966**, *5*, 2358.
(154) Carbonnaux, C.; Fazakerley, G. V.; Sowers, L. C. *Nucleic Acids Res.* **1990**, *18*, 4075.
(155) Patel, D. J.; Kozlowski, S. A.; Marky, L. A.; Rice, J. A.; Broka, C.; Dallas, J.; Itakura, K.; Breslauer, K. J. *Biochemistry* **1982**, *21*, 437.
(156) Quignard, E.; Fazakerley, G. V.; van der Marel, G.; van Boom, J. H.; Guschlbauer, W. *Nucleic Acids Res.* **1987**, *15*, 3397.
(157) Hunter, W. N.; Brown, T.; Kneale, G.; Anand, N. N.; Rabinovich, D.; Kennard, O. *J. Biol. Chem.* **1987**, *262*, 9962.
(158) Pearl, L. H. *Mutat. Res.* **2000**, *460*, 165.
(159) Kumar, N. V.; Varshney, U. *Nucleic Acids Res.* **1997**, *25*, 2336.
(160) Ghosh, M.; Kumar, N. V.; Varshney, U.; Chary, K. V. R. *Nucleic Acids Res.* **1999**, *27*, 3938.
(161) Ghosh, M.; Kumar, N. V.; Varshney, U.; Chary, K. V. R. *Nucleic Acids Res.* **2000**, *28*, 1906.
(162) Ghosh, M.; Rumpal, N.; Varshney, U.; Chary, K. V. R. *Eur. J. Biochem.* **2002**, *269*, 1886.
(163) Pearl, L. H. *Mutat. Res.* **2000**, *460*, 165.
(164) Jeltsch, A. *ChemBioChem* **2002**, *3*, 274.
(165) Bae, S.-H.; Cheong, H.-K.; Cheong, C.; Kang, S.; Hwang, D. S.; Choi, B.-S. *J. Biol. Chem.* **2003**, *278*, 45987.
(166) Lingbeck, J.; Kubinec, M. G.; Miller, J.; Reid, B. R.; Drobny, G. P.; Kennedy, M. A. *Biochemistry* **1996**, *35*, 719.
(167) Margison, G. P.; Santibanez-Koref, M. F. *BioEssays* **2002**, *24*, 255.
(168) Ezaz-Nikpay, K.; Verdine, G. L. *Chem. Biol.* **1994**, *1*, 235.
(169) Loveless, A. *Nature* **1969**, *223*, 206.
(170) Bignami, M.; O'Driscoll, M.; Aquilina, G.; Karran, P. *Mutat. Res.* **2000**, *462*, 71.
(171) Gaffney, B. L.; Marky, L. A.; Jones, R. A. *Biochemistry* **1984**, *23*, 5686.
(172) Spratt, T. E.; Levy, D. E. *Nucleic Acids Res.* **1997**, *25*, 3354.
(173) Basu, A. K.; Essigmann, J. M. *Chem. Res. Toxicol.* **1988**, *1*, 1.
(174) Patel, D. J.; Shapiro, L.; Kozlowski, S. A.; Gaffney, B. L.; Jones, R. A. *Biochemistry* **1986**, *25*, 1036.

(175) Kalnik, M. W.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *Biochemistry* **1989**, *28*, 6170.
(176) Patel, D. J.; Shapiro, L.; Kozlowski, S. A.; Gaffney, B. L.; Jones, R. A. *Biochemistry* **1986**, *25*, 1027.
(177) Kalnik, M. W.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *Biochemistry* **1989**, *28*, 6182.
(178) Leonard, G. A.; Thomson, J.; Watson, W. P.; Brown, T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9573.
(179) Gaffney, B. L.; Goswami, B.; Jones, R. A. *J. Am. Chem. Soc.* **1993**, *115*, 12607.
(180) Ginell, S. L.; Kuzmich, S.; Jones, R. A.; Berman, H. M. *Biochemistry* **1990**, *29*, 10461.
(181) Patel, D. J.; Shapiro, L.; Kozlowski, S. A.; Gaffney, B. L.; Jones, R. A. *J. Mol. Biol.* **1986**, *188*, 677.
(182) Peterson, L. A.; Vu, Ch.; Hingerty, B. E.; Broyde, S.; Cosman, M. *Biochemistry* **2003**, *42*, 13134.
(183) Aboul-ela, F.; Koh, D.; Tinoco, I. J.; Martin, F. H. *Nucleic Acids Res.* **1985**, *13*, 4811.
(184) Swann, P. F. *Mutat. Res*. **1990**, *233*, 81.
(185) Goodman, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10493.
(186) Kalnik, M. W.; Kouchakdjian, M.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *Biochemistry* **1988**, *27*, 108.
(187) Kalnik, M. W.; Kouchakdjian, M.; Li, B. F. L.; Swann, P. F.; Patel, D. J. *Biochemistry* **1988**, *27*, 100.
(188) Li, F.; Segal, A.; Solomon, J. J. *Chem. Biol. Interact.* **1992**, *83*, 35.
(189) Bhanotl, O. S.; Singhs, U. S.; Solomon, J. J. *J. Biol. Chem.* **1994**, *269*, 30056.
(190) Boulard, Y.; Fazakerley, G. V.; Sowers, L. C. *Nucleic Acids Res.* **2002**, *30*, 1371.
(191) Csanady, G. A.; Guengerich, F. P.; Bond, J. A. *Carcinogenesis* **1992**, *13*, 1143.
(192) Selzer, R. R.; Elfarra, A. A. *Chem. Res. Toxicol.* **1996**, *9*, 875.
(193) Qian, C.; Dipple, A. *Chem. Res. Toxicol.* **1995**, *8*, 389.
(194) Kanuri, M.; Nechev, L. V.; Tamura, P. J.; Harris, C. M.; Harris, T. M.; Lloyd, S. R. *Chem. Res. Toxicol.* **2002**, *15*, 1572.
(195) Scholdberg, T. A.; Merritt, W. K.; Dean, S. M.; Kowalcyzk, A.; Harris, C. M.; Harris, T. M.; Rizzo, C. J.; Lloyd, R. S.; Stone, M. P. *Biochemistry* **2005**, *44*, 3327.
(196) Weisenseel, J. P.; Reddy, G. R.; Marnett, L. J.; Stone, M. P. *Chem. Res. Toxicol.* **2002**, *15*, 127.
(197) Carmical, J. R.; Nechev, L. V.; Harris, C. M.; Harris, T. M.; Lloyd, R. S. *Environ. Mol. Mutagen.* **2000**, *35*, 48.
(198) Merritt, W. K.; Scholdberg, T. A.; Nechev, L. V.; Harris, T. M.; Harris, C. M.; Lloyd, R. S.; Stone, M. P. *Chem. Res. Toxicol.* **2004**, *17*, 1007.
(199) Scholdberg, T. A.; Nechev, L. V.; Merritt, W. K.; Harris, T. M.; Harris, C. M.; Lloyd, R. S.; Stone, M. P. *Chem. Res. Toxicol.* **2004**, *17*, 717.
(200) Carmical, J. R.; Nechev, L. V.; Harris, C. M.; Harris, T. M.; Lloyd, R. S. *Environ. Mol. Mutagen.* **2000**, *35*, 48.
(201) Scholdberg, T. A.; Nechev, L. V.; Merritt, W. K.; Harris, T. M.; Harris, C. M.; Lloyd, R. S.; Stone, M. P. *Chem. Res. Toxicol.* **2005**, *18*, 145.
(202) Greene, K. L.; Jones, R. L.; Li, Y.; Robinson, J. H.; Wang, A. H.-J.; Zen, G.; Wilson, W. D. *Biochemistry* **1994**, *33*, 1053.
(203) Barbin, A.; Bartsch, H. *The Role of Cyclic Nucleic Acid Adducts in Carcinogenesis and Mutagenesis*; I.A.R.C.: Lyon, France, 1986; Vol. 70, p 345.
(204) Marnett, L. J. *Carcinogenesis* **2000**, *21*, 361.
(205) Kouchakdjian, M.; Eisenberg, M.; Yarema, K.; Basu, A.; Essigmann, J.; Patel, D. J. *Biochemistry* **1991**, *30*, 1820.
(206) de los Santos C.; Kouchakdjian, M.; Yarema, K.; Basu, A.; Essigmann, J.; Patel, D. J. *Biochemistry* **1991**, *30*, 1828.
(207) Leonard, G. A.; McAuley-Hecht, K. E.; Gibson, N. J.; Brown, T.; Watson, W. P.; Hunter, W. N. *Biochemistry* **1994**, *33*, 4755.
(208) Korobka, A.; Cullinan, D.; Cosman, M.; Grollman, A. P.; Patel, D. J.; Eisenberg, M.; de los Santos, C. *Biochemistry* **1996**, *35*, 13310.
(209) Cullinan, D.; Korobka, A.; Grollman, A. P.; Patel, D. J.; Eisenberg, M.; de los Santos, C. *Biochemistry* **1996**, *35*, 13319.
(210) Cullinan, D.; Grollman, A. P.; Eisenberg, M.; de los Santos, C. *Biochemistry* **1997**, *36*, 11933.
(211) Cullinan, D.; Johnson, F.; de los Santos, C. *J. Mol. Biol.* **2000**, *296*, 851.
(212) Saparbaev, M.; Laval, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8508.
(213) Barrett, T. E.; Schaerer, O. D.; Savva, R.; Brown, T.; Jiricny, J.; Verdine, G. L.; Pearl, L. H. *EMBO J.* **1999**, *18*, 6599.
(214) Kouchakdjian, M.; Marinelli, E.; Gao, X.; Johnson, F.; Grollman, A.; Patel, D. *Biochemistry* **1989**, *28*, 5647.
(215) Kouchakdjian, M.; Eisenberg, M.; Live, D.; Marinelli, E.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1990**, *29*, 4456.
(216) Huang, P.; Eisenberg, M. *Biochemistry* **1992**, *31*, 6518.
(217) Huang, P.; Patel, D. J.; Eisenberg, M. *Biochemistry* **1993**, *32*, 3852.
(218) Singh, U. S.; Moe, J. G.; Reddy, G. R.; Weisenseel, J. P.; Marnett, L. J.; Stone, M. P. *Chem. Res. Toxicol.* **1993**, *6*, 825.
(219) Kouchakdjian, M.; Eisenberg, M.; Johnson, F.; Grollman, A. P.; Patel, D. J. *Biochemistry* **1991**, *30*, 3262.
(220) Moe, J. G.; Reddy, G. R.; Marnett, L. J.; Stone, M. P. *Chem. Res. Toxicol*. **1994**, *7*, 319.
(221) Weisenseel, J. P.; Moe, J. G.; Reddy, G. R.; Marnett, L. J.; Stone, M. P. *Biochemistry* **1995**, *34*, 50.
(222) Weisenseel, J. P.; Reddy, G. R.; Marnett, L. J.; Stone, M. P. *Chem. Res. Toxicol.* **2002**, *15*, 140.
(223) (a) Mao, H.; Reddy, G. R.; Marnett, L. J.; Stone, M. P. *Biochemistry* **1999**, *38*, 13491. (b) Hashim, M. F.; Riggins, J. N.; Schnetz-Boutaud, N.; Voehler, M.; Stone, M. P.; Marnett, L. J. *Biochemistry* **2004**, *43*, 11828.
(224) Schnetz-Boutaud, N. C.; Saleh, S.; Marnett, L. J.; Stone, M. P. *Biochemistry* **2001**, *40*, 15638.
(225) de los Santos, C.; Zaliznyak, T.; Johnson, F. *J. Biol. Chem.* **2000**, *276*, 9077.
(226) Zaliznyak, T.; Johnson, F.; de los Santos, C. In preparation.
(227) Mies, L. G.; Dronkert, M. L. G.; Kanaar, R. *Mutat. Res.* **2001**, *486*, 217.
(228) Noronha, A. M.; Noll, D. M.; Wilds, C. J.; Miller, P. S. *Biochemistry* **2002**, *41*, 760.
(229) Wilds, C. J.; Noronha, A. M.; Robidoux, S.; Miller, P. S. *J. Am. Chem. Soc.* **2004**, *126*, 9257.
(230) da Silva, M. W.; Wilds, C. J.; Noronha, A. M.; Colvin, O. M.; Miller, P. S.; Gamcsik, M. P. *Biochemistry* **2004**, *43*, 12549.
(231) da Silva, M. W.; Noronha, A. M.; Noll, D. M.; Miller, P. S.; Colvin, O. M.; Gamcsik, M. P. *Biochemistry* **2002**, *41*, 15181.
(232) Niedernhofer, L. J.; Daniels, J. S.; Rouzer, C. A.; Greene, R. E.; Marnett, L. J. *J. Biol. Chem.* **2003**, *278*, 31426.
(233) Dooley, P. A.; Tsarouhtsis, D.; Korbel, G. A.; Nechev, L. V.; Shearer, J.; Zegar, I. S.; Harris, C. M.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **2001**, *123*, 1730.
(234) Dooley, P. A.; Zhang, M.; Korbel, G. A.; Nechev, L. V.; Harris, C. M.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **2003**, *125*, 62.
(235) Norman, D.; Live, D.; Sastry, M.; Lipman, R.; Hingerty, B. E.; Tomasz, M.; Broyde, S.; Patel, D. J. *Biochemistry* **1990**, *29*, 2861.
(236) Fagan, P. A.; Spielmann, H. P.; Sigurdsson, S. T.; Rink, S. M.; Hopkins, P. B.; Wemmer, D. E. *Nucleic Acids Res.* **1996**, *24*, 1566.
(237) Perdiz, D.; Gróf, P.; Mezzina, M.; Nikaido, O.; Moustacchi, E.; Sage, E. *J. Biol. Chem.* **2000**, *275*, 26732.
(238) Kao, J. L.-F.; Nadji, S.; Taylor, J. S. *Chem. Res. Toxicol.* **1993**, *6*, 561.
(239) Lingbeck, J. M.; Taylor, J.-S. *Biochemistry* **1999**, *38*, 13717.
(240) Lemaire, D. G. E.; Ruzsicska, B. P. *Biochemistry* **1993**, *32*, 2525.
(241) Glickman, B. W.; Schaaper, R. M.; Haseltine, W. A.; Dunn, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6945.
(242) Kim, J.-K.; Soni, S.-D.; Wallace, J. C.; Alderfer, J. C. *Nucleic Acids Res.* **1993**, *21*, 2755.
(243) Rycyna, R. E.; Alderfer, J. L. *Nucleic Acids Res*. **1985**, *13*, 5949.
(244) Taylor, J.-S.; Cohrs, M. P. *J. Am. Chem. Soc.* **1987**, *109*, 2834.
(245) Taylor, J.-S.; Brockie, I. R. *Nucleic Acids Res.* **1988**, *16*, 5123.
(246) Iwai, S.; Mizukishi, T.; Fujiwara, Y.; Masutani, C.; Hanaoka, F.; Hayakawa, Y. *Nucleic Acids Res.* **1999**, *27*, 2299.
(247) Hruska, F. E.; Wood, D. J.; Ogilvie, K. K.; Charlton, J. L. *Can. J. Chem.* **1975**, *53*, 1193.
(248) Cadet, J.; Voituriez, L.; Hruska, F. E.; Grand, A. *Biopolymers* **1985**, *24*, 897.
(249) Kan, L. S.; Voituriez, L.; Cadet, J. *Biochemistry* **1988**, *27*, 5796.
(250) Kemmink, J.; Boeleus, R.; Koning, T. M. G.; Kaptein, R.; van der Marel, G.; van Boom, J. H. *Eur. J. Biochem.* **1987**, *162*, 37.
(251) Taylor, J.-S.; Garrett, D. S.; Brockie, I. R.; Svoboda, D. L.; Telser, J. *Biochemistry* **1990**, *29*, 8858.
(252) Kim, J.-K.; Patel, D.; Choi, B.-S. *Photochem. Photobiol.* **1995**, *62*, 44.
(253) Rao, S. N.; Keepers, J. W.; Kollman, P. *Nucleic Acids Res.* **1984**, *12*, 4789.
(254) Lee, H.-M.; Kim, J.-K. *Photochem. Photobiol.* **2002**, *76*, 417.
(255) Kemmink, J.; Boelens, R.; Koning, T.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *Nucleic Acids Res.* **1987**, *15*, 4645.
(256) Kim, J. K.; Choi, B. S. *Eur. J. Biochem.* **1995**, *228*, 849.
(257) Hwang, G.-S.; Kim, J.-K.; Choi, B.-S. *Eur. J. Biochem.* **1996**, *235*, 359.
(258) Lee, J.-H.; Hwang, G.-S.; Kim, J.-K.; Choi, B.-S. *FEBS Lett.* **1998**, *428*, 269.
(259) Jing, Y.; Kao, J. F.-L.; Taylor, J.-S. *Nucleic Acids Res.* **1998**, *26*, 3845.
(260) Lee, J. H.; Park, C.-J.; Shin, J.-S.; Ikegami, T.; Akutsu, H.; Choi, B. S. *Nucleic Acids Res.* **2004**, *32*, 2474.
(261) Lee, J.-H.; Choi, Y.-J.; Choi, B.-S. *Nucleic Acids Res.* **2000**, *28*, 1794.
(262) Smith, C. A.; Wang, M.; Jiang, N.; Che, L.; Zhao, X.; Taylor, J.-S. *Biochemistry* **1996**, *35*, 4146.

(263) Lee, J.-H.; Hwang, G.-S.; Choi, B.-S. *Proc. Natl. Acad. Sci. U S A.* **1999**, *96*, 6632.
(264) Lee, J.-H.; Bae, S.-H.; Choi, B.-S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 4591.
(265) McAteer, K.; Jing, Y.; Kao, J.; Taylor, J.-S.; Kennedy, M. A. *J. Mol. Biol.* **1998**, *282*, 1013.
(266) Wang, C. I.; Taylor, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9072.
(267) Rao, S. N.; Keepers, J. W.; Kollman, P. *Nucleic Acids Res.* **1984**, *12*, 4789.
(268) Miller, J.; Cooney, M. *Nucleic Acids Res.* **1997**, *25*, 1432.
(269) Pearlman, D. A.; Holbrook, S. R.; Pirkle, D. H.; Kim, S. H. *Science* **1985**, *227*, 1304.
(270) Husain, I.; Griffith, J.; Sankar, A. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2558.
(271) Park, H. J.; Zhang, K.; Ren, Y.; Nadji, S.; Sinha, N.; Taylor, J.-S.; Kang, C. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15965.
(272) DiGabriele, A. D.; Sanderson, M. R.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1816.
(273) DiGabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *131*, 1024.
(274) Koo, H.-S.; Drak, J.; Rice, J. A.; Crothers, D. M. *Biochemistry* **1990**, *29*, 4227.
(275) Dickerson, R. E.; Goodsell, D. S.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3579.
(276) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins* **1993**, *17*, 412.
(277) Yamaguchi, H.; van Aalten, D. M.; Pinak, M.; Furukawa, A.; Osman, R. *Nucleic Acids Res.* **1998**, *26*, 1939.
(278) Straub, K.; Kanne, D.; Hearst, J.; Rapoport, H. *J. Am. Chem. Soc.* **1981**, *103*, 2347.
(279) Spielman, H. P.; Sastry, S. S.; Hearst, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4514.
(280) Tomic, M. T.; Wemmer, D. E.; Kim, S.-H. *Science* **1987**, *238*, 1722.
(281) Hwang, G.-S.; Kim, J.-K.; Choi, B. S. *Biochim. Biophys. Res. Commun.* **1996**, *219*, 191.
(282) Spielmann, H. P.; Dwyer, T. J.; Sastry, S. S.; Hearst, J. E.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2345.
(283) Spielmann, H. P.; Dwyer, T. D.; Hearst, J. E.; Wemmer, D. E. *Biochemistry* **1995**, *34*, 12937.
(284) Shi, Y.-B.; Hearst, J. E. *Biochemistry* **1986**, *25*, 5895.
(285) Wiesehahn, G.; Hearst, J. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2703.
(286) Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, *99*, 2467.
(287) Fichtinger-Schepman, A. M. J.; van der Veer, J. L.; den Hartog, J. H. J.; Lohman, P. H. M.; Reedijk, J. *Biochemistry* **1985**, *24*, 707.
(288) Eastman, A. *Biochemistry* **1986**, *25*, 3912.
(289) Brabec, V.; Leng, M, *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5345.
(290) Comess, K. M.; Costello, C. E.; Lippard, S. J. *Biochemistry* **1990**, *29*, 2102.
(291) Sherman, S. E.; Gibson, D.; Wang, A. H.-J.; Lippard, S. J. *Science* **1985**, *230*, 412.
(292) Kline, T. P.; Marzilli, L. G.; Live, D.; Zon, G. *J. Am. Chem. Soc.* **1989**, *111*, 7057.
(293) Bancroft, D. P.; Lepre, C. A.; Lippard, S. J. *J. Am. Chem. Soc.* **1990**, *112*, 6860.
(294) Wu, Y.; Pradhan, P.; Havener, J.; Boysen, G.; Swenberg, J. A.; Campbell, Sh. L.; Chaney, S. G. *J. Mol. Biol.* **2004**, *341*, 1251.
(295) Spingler, B.; Whittington, D. A.; Lippard, S. J. *Inorg. Chem.* **2001**, *40*, 5596.
(296) Marzilli, L. G.; Saad, J. S.; Kuklenyik, Z.; Keating, K. A.; Xu, Y. *J. Am. Chem. Soc.* **2001**, *123*, 2764.
(297) Yang, D.; van Boom, S. S. G. E.; Reedijk, J.; van Boom, J. H.; Wang, A. H.-J. *Biochemistry* **1995**, *34*, 12912.
(298) Gelasco, A.; Lippard, S. J. *Biochemistry* **1998**, *37*, 9230.
(299) Takahara, P. M.; Frederick, C. A.; Lippard, S. J. *J. Am. Chem. Soc.* **1996**, *118*, 12309.
(300) Parkinson, J. A.; Chen, Y.; Murdoch, P. del S.; Guo, Z.; Berners-Price, S. J.; Brown, T.; Sadler, P. J. *Chem.—Eur. J.* **2000**, *6*, 3636.
(301) Dunham, Sh. U.; Dunham, S. U.; Turner, C. J.; Lippard, S. J. *J. Am. Chem. Soc.* **1998**, *120*, 5395.
(302) Fouchet, M.-H.; Guittet, E.; Cognet, J. A. H.; Kozelka, J.; Gauthier, C.; Le Bret, M.; Zimmermann, K.; Chottard, J.-C. *J. Bioinorg. Chem.* **1997**, *2*, 83.
(303) Rice, J. A.; Crothers, D. M.; Pinto, A. L.; Lippard, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4158.
(304) Dunham, S. U.; Lippard, S. J. *J. Am. Chem. Soc.* **1995**, *117*, 10702.
(305) Huang, H.; Zhu, L.; Reid, B. R.; Drobny, G. P.; Hopkins, P. B. *Science* **1995**, *270*, 1842.
(306) Paquet, F.; Perez, C.; Leng, M.; Lancelot, G.; Malonge, J. M. *J. Biomol. Struct. Dyn.* **1996**, *14*, 67.
(307) Coste, F.; Malinge, J. M.; Serre, L.; Shepard, W.; Roth, M.; Leng, M.; Zelwer, C. *Nucleic Acids Res.* **1999**, *27*, 1837.
(308) Paquet, F.; Boudvillain, M.; Lancelot, G.; Leng, M. *Nucleic Acids Res.* **1999**, *27*, 4261.
(309) Andersen, B.; Sletten, E. *J. Inorg. Biochem.* **2000**, *79*, 353.
(310) Aich, P.; Labiuk, S. L.; Tari, L. W.; Delbaere, L. J. T.; Roesler, W. J.; Falk, K. J.; Steer, R. P.; Lee, J. S. *J. Mol. Biol.* **1999**, *294*, 477.
(311) Baruah, H.; Barry, C. G.; Bierbach, U. *Curr. Top. Med. Chem.* **2004**, 1537.
(312) Baruah, H.; Rector, C. L.; Monnier, S. M.; Bierbach, U. *Biochem. Pharmacol.* **2002**, *64*, 191.
(313) Baruah, H.; Wright, M. W.; Bierbach, U. *Biochemistry* **2005**, *44*, 6059.
(314) Wood, R. D. *J. Biol. Chem.* **1997**, *272*, 23465.
(315) Geacintov, N. E.; Cosman, M.; Hingerty, B. E.; Amin, S.; Broyde, S.; Patel, D. J. *Chem. Res. Toxicol.* **1997**, *10*, 111.
(316) Conney, A. H. *Cancer Res.* **1982**, *42*, 4875.
(317) Beland, F. A.; Kadlubar, F. F. *Handbook on Experimental Pharmacology*; 1990; Vol. 94/1, p 267.
(318) Wolfe, A. R.; Smith, T. J.; Meehan, T. *Chem. Res. Toxicol.* **2004**, *17*, 476.
(319) Purohit, V.; Basu, A. K. *Chem. Res. Toxicol.* **2000**, *13*, 673.
(320) Schut, H. A. J.; Snyderwine, E. G. *Carcinogenesis* **1999**, *20*, 353.
(321) Wood, A. W.; Levin, W.; Chang, G. L.; Lehr, R. E.; Schaefer-Ridder, M.; Kerle, J.; Jerina, D. M.; Conney, A. H. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 3176.
(322) Geacintov, N. E. *Carcinogenesis* **1986**, *7*, 759.
(323) Wood, A. W.; Wislocki, P. G.; Chang, R. L.; Levin, W.; Lu, A. Y. H.; Yagi, H.; Hernandez, O.; Jerina, D. M.; Conney, A. H. *Cancer Res.* **1976**, *36*, 3358.
(324) Xue, W.; Warshawsky, D. *Toxicol. Appl. Pharmacol.* **2005**, *206*, 73.
(325) Cosman, M.; de los Santos, C.; Fiala, R.; Hingerty, B. E.; Ibanez, V.; Margulis, L. A.; Live, D.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1914.
(326) de los Santos, C.; Cosman, M.; Hingerty, B. E.; Ibanez, V.; Margulis, L. A.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1992**, *31*, 5245.
(327) Graves, D. E.; Stone, M. P.; Krugh, T. R. *Biochemistry* **1985**, *24*, 7573.
(328) Cheatham, S.; Kook, A.; Hurley, L. H.; Barkley, M. D.; Remers, W. *J. Med. Chem.* **1988**, *31*, 583.
(329) Krugh, T. R.; Graves, D. E.; Stone, M. P. *Biochemistry* **1989**, *28*, 9988.
(330) Boyd, F. L.; Stewart, D.; Remers, W. A.; Barkley, M. D.; Hurley, L. H. *Biochemistry* **1990**, *29*, 2387.
(331) Cosman, M.; de los Santos, C.; Fiala, R.; Hingerty, B. E.; Ibanez, V.; Luna, E.; Harvey, R. G.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1993**, *32*, 4145.
(332) Cosman, M.; Hingerty, B. E.; Luneva, N.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1996**, *35*, 9850.
(333) Fountain, M. A.; Krugh, T. R. *Biochemistry* **1995**, *34*, 3152.
(334) Zhang, N.; Lin, C.; Huang, X.; Kolbanovskiy, A.; Hingerty, B. E.; Amin, S.; Broyde, S.; Geacintov, N. E.; Patel, D. J. *J. Mol. Biol.* **2005**, *346*, 951.
(335) Cosman, M.; Fiala, R.; Hingerty, B. E.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1994**, *33*, 11507.
(336) Feng, B.; Gorin, A.; Kolbanovskiy, A.; Hingerty, B. E.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1997**, *36*, 13780.
(337) Cosman, M.; Fiala, R.; Hingerty, B. E.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1994**, *33*, 11518.
(338) Ya, N.-Q.; Smirnov, S.; Cosman, M.; Bhanot, S.; Ibanez, V.; Geacintov, N. E. In *Structural Biology: The State of the Art. Proceedings of the 8th Conversation*; Sarma, R. H., Sarma, M. H., Eds.; Adenine Press: Schenectady, NY, 1994; Vol. 2, p 349.
(339) Suzuki, N.; Ohashi, E.; Kolbanovskiy, A.; Geacintov, N. E.; Grollman, A. P.; Ohmori, H.; Shibutani, S. *Biochemistry* **2002**, *41*, 6100.
(340) Cosman, M.; Hingerty, B.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1995**, *34*, 15334.
(341) Feng, B.; Gorin, A.; Hingerty, B. E.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1997**, *36*, 13769.
(342) Hsu, G. W.; Huang, X.; Luneva, N. P.; Geacintov, N. E.; Beese, L. S. *J. Biol. Chem.* **1995**, *280*, 3764.
(343) Kim, H.-Y. H.; Wilkinson, A. S.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **2003**, *42*, 2328.
(344) Cosman, M.; Xu, R.; Hingerty, B. E.; Amin, S.; Harvey, R. G.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1995**, *34*, 6247.
(345) Lin, C. H.; Huang, X.; Kolbanovskii, A.; Hingerty, B. E.; Amin, S.; Broyde, S.; Geacintov, N. E.; Patel, D. J. *J. Mol. Biol.* **2001**, *306*, 1059.
(346) Zegar, I. S.; Setayesh, F. R.; DeCorte, B. L.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1996**, *35*, 4334.
(347) Setayesh, F. R.; DeCorte, B. L.; Horton, P.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Chem. Res. Toxicol.* **1998**, *11*, 766.

(348) Shimotakahara, S.; Gorin, A.; Kolbanovskiy, A.; Kettani, A.; Hingerty, B. E.; Amin, S.; Broyde, S.; Geacintov, N.; Patel, D. J. *J. Mol. Biol.* **2000**, *302*, 377.
(349) Zaliznyak, T.; Johnson, F.; de los Santos, C. Submitted.
(350) Cui, X. S.; Eriksson, L. C.; Moller, L. *Mutat. Res.* **1999**, *442*, 9.
(351) Geacintov, N. E.; Broyde, S.; Buterin, T.; Naegeli, H.; Wu, M.; Yan, S.; Patel, D. J. *Biopolymers* **2002**, *65*, 202.
(352) Sastry, M.; Fiala, R.; Lipman, R.; Tomasz, M.; Patel, D. J. *J. Mol. Biol.* **1995**, *247*, 338.
(353) Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Munk, S. A.; Kitos, P. A.; Suntornwatt, O. *J. Am. Chem. Soc.* **1990**, *112*, 8961.
(354) Hurley, L. H.; Lee, C.-S.; McGovren, J. P.; Warpehoski, M. A.; Mitchel1, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886.
(355) Lin, C. H.; Hurley, L. H. *Biochemistry* **1990**, *29*, 9503.
(356) Scahill, T. A.; Jensen, R. M.; Swenson, D. H.; Hatzenbuhler, N. T.; Petzold, G.; Wierenga, W.; Brahmel, N. D. *Biochemistry* **1990**, *29*, 2852.
(357) Lin, C. H.; Beale, J. M.; Hurley, L. H. *Biochemistry* **1991**, *30*, 3597.
(358) Powers, R.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 9994.
(359) Lin, C. H.; Patel, D. J. *J. Mol. Biol.* **1995**, *248*, 162.
(360) Zegar, I. S.; Kim, S. J.; Johansen, T. N.; Horton, P. J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1996**, *35*, 6212.
(361) Zegar, I. S.; Chary, P.; Jabil, R. J.; Tamura, P. J.; Johansen, T. N.; Lloyd, R. S.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1998**, *37*, 16516.
(362) Volk, D. E.; Rice, J. S.; Luxon, B. A.; Yeh, H. J. C.; Liang, C.; Xie, G.; Sayer, J. M.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **2000**, *39*, 14040.
(363) Mao, B.; Gu, Z.; Gorin, A.; Chen, J.; Hingerty, B. E.; Amin, S.; Broyde, S.; Geacintov, N. E.; Patel, D. J. *Biochemistry* **1999**, *38*, 10831.
(364) Volk, D. E.; Thiviyanathan, V.; Rice, J. S.; Luxon, B. A.; Shah, J. H.; Yagi, H.; Sayer, J. M.; Yeh, H. J. C.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **2003**, *42*, 1410.
(365) Schurter, E. J.; Sayer, J. M.; Oh-hara, T.; Yeh, H. J. C.; Yagi, H.; Luxon, B. A.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **1995**, *34*, 9009.
(366) Pradhan, P.; Tirumala, S.; Liu, X.; Sayer, J. M.; Jerina, D. M.; Yeh, H. J. C. *Biochemistry* **2001**, *40*, 5870.
(367) Schurter, E. J.; Yeh, H. J. C.; Sayer, J. M.; Lakshman, M. K.; Yagi, H.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **1995**, *34*, 1364.
(368) Yeh, H. J. C.; Sayer, J. M.; Liu, X.; Altieri, A. A.; Byrd, R. A.; Lakshman, M. K.; Yagi, H.; Schurter, E. J.; Gorenstein, D. G.; Jerina, D. M. *Biochemistry* **1995**, *34*, 13570.
(369) Schwartz, J. L.; Rice, J. S.; Luxon, B. A.; Sayer, J. M.; Xie, G.; Yeh, H. J. C.; Liu, X.; Jerina, D. M.; Gorenstein, D. G. *Biochemistry* **1997**, *36*, 11069.
(370) Li, Z.; Mao, H.; Kim, H.-Y.; Tamura, P. J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1999**, *38*, 2969.
(371) Li, Z.; Kim, H.-Y.; Tamura, P. J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1999**, *38*, 16045.
(372) Li, Z.; Tamura, P. J.; Wilkinson, A. S.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **2001**, *40*, 6743.
(373) Li, Z.; Kim, H.-Y.; Tamura, P. J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1999**, *38*, 14820.
(374) Feng, B.; Zhou, L.; Passarelli, M.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1995**, *34*, 14021.
(375) Feng, B.; Voehler, M.; Zhou, L.; Passarelli, M.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1996**, *35*, 7316.
(376) Painter, S. L.; Zegar, I. S.; Tamura, P. J.; Bluhm, S.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **1999**, *38*, 8635.
(377) Simeonov, M. F.; Tamura, P. J.; Wilkinson, A. S.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **2000**, *39*, 924.
(378) Hennard, C.; Finneman, J.; Harris, C. M.; Harris, T. M.; Stone, M. P. *Biochemistry* **2001**, *40*, 9780.
(379) Cosman, M.; Fiala, R.; Hingerty, B.; Laryea, A.; Lee, H.; Harvey, R. G.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. *Biochemistry* **1993**, *32*, 12488.
(380) Cosman, M.; Laryea, A.; Fiala, R.; Hingerty, B. E.; Amin, S.; Geacintov, N. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1995**, *34*, 1295.
(381) Suri, A. K.; Mao, B.; Amin, S.; Geacintov, N. E.; Patel, D. J. *J. Mol. Biol.* **1999**, *292*, 289.
(382) Stoychev, G.; Kierdaszuk, B.; Shugar, D. *Biochim. Biophys. Acta* **2001**, *1544*, 74.
(383) Tomasz, M.; Lipman, R.; Lee, M. S.; Verdine, G. L.; Nakanishi, K. *Biochemistry* **1987**, *26*, 2010.
(384) Smela, M. E.; Currier, S. S.; Bailey, E. A.; Essigmann, J. M. *Carcinogenesis* **2001**, *22*, 535.
(385) Gopalakrishnan, S.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **1989**, *111*, 7232.
(386) Gopalakrishnan, S.; Harris, T. M.; Stone, M. P. *Biochemistry* **1990**, *29*, 10438.
(387) Jones, W. R.; Johnston, D. S.; Stone, M. P. *Chem. Res. Toxicol.* **1998**, *11*, 873.
(388) Giri, I.; Jenkins, M. D.; Schnetz-Boutaud, N. C.; Stone, M. P. *Chem. Res. Toxicol.* **2002**, *15*, 638.
(389) Giri, I.; Stone, M. P. *Biochemistry* **2003**, *42*, 7023.
(390) Giri, I.; Johnston, D. S.; Stone, M. P. *Biochemistry* **2002**, *41*, 5462.
(391) Johnston, D. S.; Stone, M. P. *Biochemistry* **1995**, *34*, 14037.
(392) Gopalakrishnan, S.; Patel, D. J. *J. Am. Chem. Soc.* **1993**, *115*, 9321.
(393) Mao, H.; Deng, Z.; Wang, F.; Harris, T. M.; Stone, M. P. *Biochemistry* **1998**, *37*, 4374.
(394) Alekseyev, Y. O.; Hamm, M. L.; Essigmann, J. M. *Carcinogenesis* **2004**, *25*, 1045.
(395) Mori, H.; Sugie, S.; Yoshimi, N.; Kuniyasu, T.; Iwata, H.; Kawai, K.; Hamasaki, T. *Carcinogenesis* **1988**, *9*, 1039.
(396) Krivobok, S.; Oliver, P.; Marzin, D. R.; Seigle-Murandi, F.; Steiman, R. *Mutagenesis* **1987**, *2*, 433.
(397) Gopalakrishnan, S.; Xiucai, L.; Patel, D. J. *Biochemistry* **1992**, *31*, 10790.
(398) Tomasz, M.; Olson, J.; Mercado, C. M. *Biochemistry* **1972**, *11*, 1235.
(399) Yasuzawa, T.; Saitoh, Y.; Sano, H. *J. Antibiot. (Tokyo)* **1990**, *43*, 485.
(400) Bennet, G. N. *Nucleic Acids Res.* **1982**, *10*, 4581.
(401) Sun, D.; Hansen, M.; Hurley, L. H. *J. Am. Chem. Soc.* **1995**, *117*, 2430.
(402) Sun, D.; Hansen, M.; Clement, J. J.; Hurley, L. H. *Biochemistry* **1993**, *32*, 8068.
(403) Hansen, M.; Hurley, L. *J. Am. Chem. Soc.* **1995**, *117*, 2421.
(404) Hansen, M.; Yun, S.; Hurley, L. *Chem. Biol.* **1995**, Hansen, M.; Yun, S.; Hurley, L. *Chem. Biol.* **1995**, *2*, 229.
(405) Pavlopoulos, S.; Bicknell, W.; Craik, D. J.; Wickham, G. *Biochemistry* **1996**, *35*, 9314.
(406) Owen, E. A.; Burley, G. A.; Carver, J. A.; Wickham, G.; Keniry, M. A. *Biochim. Biophys. Res. Commun.* **2002**, *290*, 1602.
(407) Palom, Y.; Belcourt, M. F.; Kumar, G. S.; Arai, H.; Kasai, M.; Sartorelli, A. C.; Rockwell, S.; Tomasz, M. *Oncol. Res.* **1998**, *10*, 509.
(408) Subramaniam, G.; Paz, M. M.; Kumar, G. S.; Das, A.; Palom, Y.; Clement, C. C.; Patel, D. J.; Tomasz, M. *Biochemistry* **2001**, *40*, 10473.
(409) Palom, Y.; Belcourt, M. F.; Tang, L. Q.; Mehta, S. S.; Sartorelli, A. C.; Pritsos, C. A.; Pritsos, K. L.; Rockwell, S.; Tomasz, M. *Biochem. Pharmacol.* **2001**, *61*, 1517.
(410) Beland, R. A.; Marques, M. M. *IARC Sci. Publ.* **1994**, *125*, 229.
(411) Heflich, R. H.; Neft, R. E. *Mutat. Res.* **1994**, *318*, 73.
(412) Purohit, V.; Basu, A. K. *Chem. Res. Toxicol.* **2000**, *13*, 675.
(413) Hoffman, G. R.; Fuchs, R. P. P. *Chem. Res. Toxicol.* **1997**, *10*, 347.
(414) Felton, J. S.; Malfatti, M. A.; Knize, M. G.; Salmon, C. P.; Hopmans, E. C.; Wu, R. W. *Mutat. Res.* **1997**, *376*, 37.
(415) Patel, D. J.; Mao, B.; Gu, Z.; Hingerty, B. E.; Gorin, A.; Basu, A. K.; Broyde, S. *Chem. Res. Toxicol.* **1998**, *11*, 391.
(416) O'Handley, S. F.; Sanford, D. G.; Xu, R.; Lester, C. C.; Hingerty, B. E.; Broyde, S.; Krugh, T. R. *Biochemistry* **1993**, *32*, 2481.
(417) Grunberger, D.; Nelson, J. H.; Cantor, C. R.; Weinstein, I. B. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *66*, 488.
(418) Fuchs, R. P. P.; Daune, M. *FEBS Lett.* **1971**, *14*, 206.
(419) Fuchs, R. P. P.; Daune, M. *Biochemistry* **1972**, *11*, 2659.
(420) Cho, B. P.; Zhou, L. *Biochemistry* **1999**, *38*, 7572.
(421) Milhé, C.; Dhalluin, C.; Fuchs, R. P. P.; Lefevre, J. F. *Nucleic Acids Res.* **1994**, *22*, 4646.
(422) Milhé, C.; Fuchs, R. P. P.; Lefevre, J. F. *Eur. J. Biochem.* **1996**, *235*, 120.
(423) Eckel, L. M.; Krugh, T. R. *Nature Struct. Biol.* **1994**, *1*, 89.
(424) Eckel, L. M.; Krugh, T. R. *Biochemistry* **1994**, *33*, 13611.
(425) Cho, B. P.; Beland, F. A.; Marques, M. M. *Biochemistry* **1994**, *33*, 1373.
(426) Zhou, L.; Rajabzadeh, G.; Traficante, D. D.; Cho, B. P. *J. Am. Chem. Soc.* **1997**, *119*, 5384.
(427) Mao, B.; Gu, Z.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1998**, *37*, 81.
(428) Mao, B.; Gu, Z.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1998**, *37*, 95.
(429) Norman, D.; Abuaf, P.; Hingerty, B. E.; Live, D.; Grunberger, D.; Broyde, S.; Patel, D. J. *Biochemistry* **1989**, *28*, 7462.
(430) Abuaf, P.; Hingerty, B. E.; Broyde, S.; Grunberger, D. *Chem. Res. Toxicol.* **1995**, *8*, 369.
(431) Mao, B.; Cosman, M.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1995**, *34*, 6226.
(432) Mao, B.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1995**, *34*, 16641.
(433) Mao, B.; Gorin, A.; Gu, Z.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1997**, *36*, 14479.

(434) Bichara, M.; Fuchs, R. P. P. *J. Mol. Biol.* **1985**, *183*, 341.
(435) Carothers, A. M.; Urlaub, G.; Mucha, J.; Yuan, W.; Chasin, L. A.; Grunberger, D. *Carcinogenesis* **1993**, *14*, 2181.
(436) Mao, B.; Gu, Z.; Gorin, A.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1997**, *36*, 14491.
(437) Mao, B.; Gu, Z.; Gorin, A.; Hingerty, B. E.; Broyde, S.; Patel, D. J. *Biochemistry* **1999**, *38*, 10855.
(438) Hsu, G. W.; Kiefer, J. R.; Burnouf, D.; Becherel, O. J.; Fuchs, R. P. P.; Beese, L. S. *J. Biol. Chem.* **2004**, *279*, 50280.
(439) Mao, B.; Vyas, R. R.; Hingerty, B. E.; Broyde, S.; Basu, A. K.; Patel, D. J. *Biochemistry* **1996**, *35*, 12659.
(440) Gu, Z.; Gorin, A.; Krishnasamy, R.; Hingerty, B. E.; Basu, A. K.; Broyde, S.; Patel, D. J. *Biochemistry* **1999**, *38*, 10843.
(441) Nolan, S. J.; Vyas, R. R.; Hingerty, B. E.; Ellis, S.; Broyde, S.; Shapiro, R.; Basu, A. K. *Carcinogenesis* **1996**, *17*, 133.
(442) Cho, B. P.; Beland, F. A.; Marques, M. M. *Biochemistry* **1992**, *31*, 9587.
(443) Brown, K.; Hingerty, B. E.; Guenther, E. A.; Krishnan, V. V.; Broyde, S.; Turteltaub, K. W.; Cosman, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8507.
(444) Cara, C. J.; Pena, A. S.; Sans, M.; Rodrigo, L.; Guerrero-Esteo, M.; Hinojosa, J.; Garcia-Paredes, J.; Guijarro, L. G. *Med. Sci. Monit.* **2004**, *10*, 247.
(445) Karran, P.; Offman, J.; Bignami, M. *Biochimie* **2003**, *85*, 1149.
(446) Marathias, V. M.; Sawicki, M. J.; Bolton, P. H. *Nucleic Acids Res.* **1999**, *27*, 2860.
(447) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871.
(448) Sommervile, L.; Krynetski, E. Y.; Krynetskaia, N. F.; Beger, R. D.; Zhang, W.; Marhefka, C. A.; Evans, W. E.; Kriwacki, R. W. *J. Biol. Chem.* **2003**, *278*, 1005.
(449) Bohon, J.; de los Santos, C. *Nucleic Acids Res.* **2003**, *31*, 1331.
(450) Allawi, H. T.; Santa Lucia, J. *Nucleic Acids Res.* **1998**, *26*, 4925.
(451) Bohon, J.; de los Santos, C. *Nucleic Acids Res.* **2005**, *33*, 2880.
(452) Walter, N. G.; Harris, D. A.; Pereira, M. J.; Rueda, D. *Biopolymers* **2001**−**2002**, *61*, 224.
(453) Ronen, A. *Mutat. Res.* **1980**, *75*, 1.
(454) Mhaskar, D. N.; Goodman, M. F. *J. Biol. Chem*. **1984**, *259*, 11713.
(455) Sowers, L. C.; Fazakarley, G. V.; Eritja, R.; Kaplan, B. E.; Goodman, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 5434.
(456) Lycksell, P.-O.; Graslund, A.; Claesens, F.; McLaughlin, L. W.; Larsson, U.; Rigler, R. *Nucleic Acids Res*. **1987**, *15*, 9011.
(457) Sowers, L. C.; Eritja, R.; Chen, F. M.; Khwaja, T.; Kaplan, B. E.; Goodman, M. F.; Fazakerley, G. V. *Biochem. Biophys. Res. Commun.* **1989**, *165*, 89.
(458) Fagan, P. A.; Fabrega, C.; Eritja, R.; Goodman, M. F.; Wemmer, D. E. *Biochemistry* **1996**, *35*, 4026.
(459) Sowers, L. C.; Boulard, Y.; Fazakerley, G. V. *Biochemistry* **2000**, *39*, 7613.
(460) Eritja, R.; Kaplan, B. E.; Mhaskar, D.; Sowers, L. C.; Petruska, J.; Goodman, M. F. *Nucleic Acid Res.* **1986**, *14*, 5869.
(461) Fazakerley, G. V.; Sowers, L. C.; Eritja, R.; Kaplan, B. E.; Goodman, M. F. *Biochemistry* **1987**, *26*, 5641.
(462) Pinedo, H. M.; Peters, G. F. *J. Clin. Oncol.* **1988**, *6*, 1653.
(463) Nguyen, H. N.; Nordqvist, S. R. *Semin. Surg. Oncol*. **1999**, *16*, 247.
(464) Weckbecker, G. *Pharmacol. Ther.* **1991**, *50*, 367.
(465) Kremer, A. B.; Mikita, T.; Beardsley, G. P. *Biochemistry* **1987**, *26*, 391.
(466) Sowers, L. C.; Eritja, R.; Kaplan, B. E.; Goodman, M. F.; Fazakerley, G. V. *J. Biol. Chem.* **1987**, *262*, 15436.
(467) Kerwood, D. J.; Zon, G.; James, T. L. *Eur. J. Biochem.* **1991**, *197*, 583.
(468) Stolarski, R.; Egan, W.; James, T. L. *Biochemistry* **1992**, *31*, 7027.
(469) Sahasrabudhe, P. V.; Pon, R. T.; Gmeiner, W. H. *Biochemistry* **1996**, *35*, 13597.
(470) Hebener, J. F.; Vo, C. D.; Dinh, B. L.; Gryan, G. P.; Ercolani, L.; Wang, A. H.-J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1735.
(471) Pringle, C. R. *J. Virol.* **1970**, *5*, 559.
(472) Sowers, L. C.; Eritja, R.; Kaplan, B.; Goodman, M. F.; Fazakerley, G. V. *J. Biol. Chem.* **1988**, *263*, 14794.
(473) Sowers, L. C.; Goodman, M. F.; Eritja, R.; Kaplan, B.; Fazakerley, G. V. *J. Mol. Biol.* **1989**, *205*, 437.
(474) Budowsky, E. I. *Prog. Nucleic Acid Res. Mol. Biol.* **1976**, *16*, 125.
(475) Budowsky, E. I.; Sverdlov, E. D.; Spasokukotskaya, T. N. *FEBS Lett.* **1971**, *17*, 336.
(476) Anand, N. N.; Brown, D. M.; Salisbury, S. A. *Nucleic Acids Res.* **1987**, *15*, 8167.
(477) Morozov, Y. V.; Savin, F. A.; Chekhov, V. O.; Budowsky, E. I.; Yakovlev, D. Y. *J. Photochem.* **1982**, *20*, 229.
(478) Stone, M. J.; Nedderman, A. N.; Williams, D. H.; Lin, P. K. T.; Brown, D. M. *J. Mol. Biol.* **1991**, *222*, 711.
(479) Nedderman, A. N.; Stone, M. J.; Williams, D. H.; Lin, P. K. T.; Brown, D. M. *J. Mol. Biol.* **1993**, *230*, 1068.
(480) Fazakerley, G. V.; Gdaniec, Z.; Sowers, L. C. *J. Mol. Biol.* **1993**, *230*, 6.
(481) Gdaniec, Z.; Ban, B.; Sowers, L. C.; Fazakerley, G. V. *Eur. J. Biochem.* **1996**, *242*, 271.
(482) Lin, P. K. T.; Brown, D. M. *Nucleic Acids Res.* **1992**, *20*, 5149.
(483) Moore, M. H.; Van Meervel, L.; Salisbury, S. A.; Lin, P. K. T.; Brown, D. M. *J. Mol. Biol.* **1995**, *251*, 665.
(484) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285.
(485) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4993.
(486) Pang, Y.-P.; Miller, J. L.; Kollman, P. J. *J. Am. Chem. Soc.* **1999**, *121*, 1717.
(487) Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025.
(488) Engman, K. C.; Sandin, P.; Osborne, S.; Brown, T.; Billeter, M.; Lincoln, P.; Norde, B.; Albinsson, B.; Wilhelmsson, M. *Nucleic Acids Res.* **2004**, *32*, 5087.
(489) Nichols, R.; Andrews, P. C.; Zhang, P.; Bergstrom, D. E. *Nature* **1994**, *396*, 492.
(490) Bergstrom, D. E.; Zhang, P.; Toma, P.; Andrews, P. C.; Nichols, R. *J. Am. Chem. Soc.* **1995**, *117*, 1201.
(491) Klewer, D. A.; Hoskins, A.; Zhang, P.; Davisson, V. J.; Bergstrom, D. E.; LiWang, A. C. *Nucleic Acids Res.* **2000**, *28*, 4514.
(492) Amosova, O.; George, J.; Fresco, J. R. *Nucleic Acids Res*. **1997**, *25*, 1930.
(493) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990.
(494) Moran, S.; Ren, R. X.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10506.
(495) Morales, J. C.; Kool, E. T. *Nat. Struct. Biol*. **1998**, *5*, 950.
(496) Kool, E. T. *Biopolymers* **1998**, *48*, 3.
(497) Guckian, K. M.; Krugh, T. R.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *5*, 954.
(498) Guckian, K. M.; Krugh, T. R.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 6841.
(499) Plavec, J.; Tong, W.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **1993**, *115*, 9734.
(500) Wiberg, K. B.; Murcko, M. A.; Laidig, K. E.; MacDougall, P. J. *J. Phys. Chem.* **1990**, *94*, 6956.
(501) Guschlbauer, W.; Jankowski, K. *Nucleic Acids Res.* **1980**, *8*, 1421.
(502) Bulychev, N. V.; Chamakura, V. V.; Dorman, G.; Miller, J. H.; Eisenberg, M.; Grollman, A. P.; Johnson, F. *Biochemistry* **1996**, *35*, 13147.
(503) De Mesmaeker, A.; Häner, R.; Martin, P.; Moser, H. E. *Adv. Chem. Res.* **1995**, *28*, 366.
(504) Smirnov, S.; Johnson, F.; Marumoto, R.; de los Santos, C. *J. Biomol. Struct. Dyn.* **2000**, *6*, 981.
(505) Denisov, A. Y.; Zamaratski, E. V.; Maltseva, T. V.; Sandstrom, A.; Bekiroglu, S.; Altmann, K. H.; Egli, M.; Chattopadhyaya, J. *J. Biomol. Struct. Dyn.* **1998**, *3*, 547.
(506) Isaksson, J.; Zamaratski, E.; Maltseva, T. V.; Agback, P.; Kumar, A.; Chattopadhyaya, J. *J. Biomol. Struct. Dyn.* **2001**, *6*, 783.
(507) Tonelli, M.; James, T. L. *Biochemistry* **1998**, *37*, 11478.
(508) Tolstorukov, M. Y.; Jernigan, R. L.; Zhurkin, V. B. *J. Mol. Biol.* **2004**, *337*, 65.
(509) Marquez, V. E.; Ezzitouni, A.; Russ, P.; Siddiqui, M. A.; Ford, H.; Feldman, R. J.; Mitsuya, H.; George, C.; Barchi, J. J. *J. Am. Chem. Soc.* **1998**, *120*, 2780.
(510) Tjandra, N.; Tate, S.-I.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190.
(511) Wu, Z.; Maderia, M.; Barchi, J. J., Jr.; Marquez, V. E.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 24.
(512) Frei, E., III; Bickers, J.; Hewlett, J.; Lane, M.; Leary, W.; Talley, R. *Cancer Res.* **1969**, *29*, 1325.
(513) Kufe, D. W.; Major, P. P.; Egan, E. M.; Beardsley, G. P. *J. Biol. Chem.* **1980**, *255*, 8997.
(514) Townsend, A. J.; Cheng, Y.-C. *Mol. Pharmacol.* **1987**, *32*, 330.
(515) Tougard, P.; Lefebvre-Soubeyran, O. *Acta Crystallogr*. **1974**, *B30*, 86.
(516) Chwang, A. K.; Sundaralingam, M. *Nature (London)* **1973**, *243*, 78.
(517) Wilds, C. J.; Damha, M. J. *Nucleic Acids Res.* **2000**, *28*, 3625.
(518) Schweitzer, B. I.; Mikita, T.; Kellogg, G. W.; Gardner, K. H.; Beardsley, G. P. *Biochemistry* **1994**, *33*, 11460.
(519) Gao, Y. G.; van der Marel, G. A.; van Boom, J. H.; Wang, A. H.-J. *Biochemistry* **1991**, *30*, 9922.
(520) Gotfredsen, C. H.; Spielmann, H. P.; Wengel, J.; Jacobsen, J. P. *Bioconjugate Chem*. **1996**, *7*, 680.
(521) Higashigawa, M.; Ido, M.; Kuwabara, H.; Hori, H.; Ohkubo, T.; Kawasaki, H.; Sakurai, M.; Taniguchi, K.; Hamazaki, M. *Leukemia Res*. **1991**, *15*, 255.
(522) Ross, D. D.; Cuddy, D. P.; Cohen, N.; Hensley, D. R. *Cancer Chemother. Pharmacol*. **1992**, *31*, 61.
(523) Ogawa, T.; Okazaki, T. *Annu. Rev. Biochem.* **1980**, *49*, 421.
(524) Tomkinson, A. E.; Mackey, Z. B. *Mutat. Res.* **1998**, *407*, 1.

(525) Gmeiner, W. H.; Skradis, A.; Pon, R. T.; Liu, J. *Nucleic Acids Res.* **1998**, *26*, 2359.
(526) Gmeiner, W. H.; Konerding, D.; James, T. L. *Biochemistry* **1999**, *38*, 1166.
(527) Konerding, D.; James, T. L.; Trump, E.; Soto, A. M.; Marky, L. A.; Gmeiner, W. H. *Biochemistry* **2002**, *41*, 839.
(528) Frank, K. B.; Chiou, J. F.; Cheng, Y. C. *J. Biol. Chem.* **1984**, *259*, 1566.
(529) Littler, E.; Stuart, A. D.; Chee, M. S. *Nature* **1992**, *358*, 160.
(530) Sullivan, V.; Talarico, C. L.; Stanat, S. C.; Davis, M.; Coen, D. M.; Biron, K. K. *Nature* **1992**, *358*, 162.
(531) Ilsley, D. D.; Lee, S. H.; Miller, W. H.; Kuchta, R. D. *Biochemistry* **1995**, *34*, 2504.
(532) Marshalko, S. J.; Schweitzer, B. I.; Beardsley, G. P. *Biochemistry* **1995**, *34*, 9235.
(533) Foti, M.; Marshalko, S.; Schurter, E.; Kumar, S.; Beardsley, G. P.; Schweitzer, B. I. *Biochemistry* **1997**, *36*, 5336.
(534) Mariaggi, N.; Teoule, R.; Cadet, J.; Dickie, H.; Hugues, E. *Radiat. Res.* **1979**, *79*, 431.
(535) Lesiak, K. B.; Wheeler, K. T. *Radiat. Res.* **1990**, *121*, 328.
(536) Shimizu, H.; Yagi, R.; Kimura, Y.; Makino, K.; Terato, H.; Ohyama, Y.; Ide, H. *Nucleic Acids Res.* **1997**, *25*, 597.
(537) Ide, H.; Shimizu, H.; Kimura, Y.; Sakamoto, S.; Makino, K.; Glackin, M.; Wallace, S. S.; Nakamuta, H.; Sasaki, M.; Sugimoto, N. *Biochemistry* **1995**, *34*, 6947.
(538) Aramini, J. M.; Cleaver, S. H.; Pon, R. T.; Cunningham, R. P.; Germann, M. W. *J. Mol. Biol.* **2004**, *338*, 77.
(539) Ide, H.; Tedzuka, K.; Shimizu, H.; Kimura, Y.; Purmal, A. A.; Wallace, S. S.; Kow, Y. W. *Biochemistry* **1994**, *33*, 7842.
(540) Koga, M.; Moore, M. F.; Beaucage, S. L. *J. Org. Chem.* **1991**, *56*, 3757.
(541) Vichier-Guerre, S.; Pompon, A.; Lefebvre, I.; Imbach, J. L. *Antisense Res.* **1994**, *4*, 9.
(542) Koga, M.; Wilk, A.; Moore, M. F.; Scremin, C. L.; Zhou, L.; Beaucage, S. L. *J. Org. Chem.* **1995**, *60*, 1520.
(543) Aramini, J. M.; Kalisch, B. W.; Pon, R. T.; van de Sande, J. H.; Germann, M. W. *Biochemistry* **1996**, *35*, 9355.
(544) Aramini, J. M.; van de Sande, J. H.; Germann, M. W. *Biochemistry* **1997**, *36*, 9715.
(545) Aramini, J. M.; Mujeeb, A.; Germann, M. W. *Nucleic Acids Res.* **1998**, *26*, 5644.
(546) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *J. Biol. Mol.* **1990**, *214*, 223.
(547) Aramini, J. M.; Mujeebb, A.; Ulyanov, N. B.; Germann, M. W. *J. Biomol. NMR* **2000**, *18*, 287.
(548) Nielsen, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167.
(549) Vosberg, H. P.; Eckstein, F. *Biochemistry* **1977**, *16*, 3633.
(550) Viktorova, L. S.; Dyatkina, N. B.; Mozzherin, D. Ju.; Atrazhev, A. M.; Krayevskii, A. A.; Kukhanova, M. K. *Nucleic Acids Res.* **1992**, *20*, 783.
(551) Thiviyanathan, V.; Vyazovkina, K. V.; Gozansky, E. K.; Bichenchova, B.; Abramova, T. V.; Luxon, B. A.; Lebedev, A. V.; Gorenstein, D. G. *Biochemistry* **2002**, *41*, 827.
(552) Soliva, R.; Monaco, V.; Gomez-Pinto, I.; Meeuwenoord, N. J.; Marel, G. A.; van Boom, J. H.; Gonzalez, C.; Orozco, M. *Nucleic Acids Res.* **2001**, *29*, 2973.
(553) Strauss, J. K.; Maher, L. J., III. *Science* **1994**, *266*, 1829.
(554) Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J. *Nature* **1995**, *377*, 649.
(555) DeLano Scientific LLC, 400 Oyster Point Blvd., Suite 213, South San Francisco, CA 94080-1918, USA. http://www.pymol.org.
(556) MidasPlus: a molecular image representation program from the Computer Graphics Laboratory, University of California, San Francisco. Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langride, R. *J. Mol. Graphics* **1988**, *6*, 13.

CR0404646