J|A|C|S ARTICLES

Published on Web 07/14/2009

Scope and Mechanism of Interstrand Cross-Link Formation by the C4′-Oxidized Abasic Site

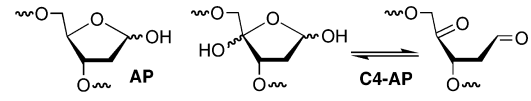
Jonathan T. Sczepanski,[†] Aaron C. Jacobs,[†] Ananya Majumdar,[‡] and Marc M. Greenberg*,[†]

Department of Chemistry and Biomolecular NMR Center, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218

Received April 27, 2009; E-mail: mgreenberg@jhu.edu

Abstract: The C4′-oxidized abasic site (C4-AP) is a commonly formed DNA lesion, which generates two types of interstrand cross-links (ICLs). The kinetically favored cross-link consists of two full length strands and forms reversibly and exclusively with dA. Cross-link formation is attributed to condensation of C4-AP with the N6-amino group of dA. Formation of the thermodynamic ICL involves cleavage of the strand containing C4-AP on the 3′-side of the lesion. The ratios and yields of the ICLs are highly dependent upon the local sequence. Product analysis of enzyme-digested material reveals that the ICL with dA is a cyclic adduct. Formation of the thermodynamically favored cross-link is catalyzed by the surrounding DNA sequence and occurs favorably with dC and dA but not with dG or dT. Mechanistic studies indicate that β-elimination from C4-AP is the rate-limiting step in the formation of the thermodynamic ICL and that the local DNA environment determines the rate constant for this reaction. The efficiency of ICL formation, the stability of the thermodynamic products, and their possible formation in cells (Regelus, P.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14032) suggest that these lesions will be deleterious to the biological system in which they are produced.

DNA interstrand cross-links (ICLs) block replication and transcription. Consequently, they are associated with the cytotoxic effects of chemotherapeutic agents that are bis-alkylators, such as the mitomycins and nitrogen mustards.[1−5] Purposeful interstrand cross-linking is also useful as a tool for probing nucleic acid structure and in biotechnology.[6−13] More recent investigations implicate endogenously produced bis-alkylating agents resulting from lipid and DNA oxidation in ICL formation.[14] In other instances, environmental pollutants and/or their metabolites have been found to produce ICLs.[15,16] Each of these examples involves covalent linkage of the opposing DNA strands by a bridging molecule. The formation of ICLs by direct reaction of one nucleic acid strand with its hybridization partner was firmly established in 2005. In one instance, a DNA radical, generated by reaction with hydroxyl radical, produced ICLs with the opposing dA.[17−19] The Gates group was the first to unequivocally describe the reaction of an abasic site with one or more nucleotides on the opposing strand.[20] They showed that the ubiquitous abasic site (AP) reacts with the deoxyguanosine opposite the 5′-flanking dC. We recently reported that the C4′-oxidized abasic site (C4-AP) also produces ICLs.[21] In this article, we elaborate on the scope of this reaction and explore its mechanism.



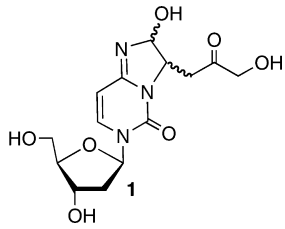
C4-AP is a major product of bleomycin-mediated DNA damage and is produced by other antitumor agents that oxi-

[†] Department of Chemistry.
[‡] Biomolecular NMR Center.

(1) Noll, D. M.; Mason, T. M.; Miller, P. S. *Chem. Rev.* **2006**, *106*, 277–301.
(2) Schärer, O. D. *ChemBioChem.* **2005**, *6*, 27–32.
(3) Martin, C.; Ellis, T.; McGurk, C. J.; Jenkins, T. C.; Hartley, J. A.; Waring, M. J.; Thurston, D. E. *Biochemistry* **2005**, *44*, 4135–4147.
(4) Tomasz, M.; Palom, Y. *Pharmacology & Therapeutics* **1997**, *76*, 73–87.
(5) Paz, M. M.; Hopkins, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 5999–6005.
(6) Puri, N.; Majumdar, A.; Cuenoud, B.; Miller, P. S.; Seidman, M. M. *Biochemistry* **2004**, *43*, 1343–1351.
(7) Thazhathveetil, A. K.; Liu, S.-T.; Indig, F. E.; Seidman, M. M. *Bioconjugate Chem.* **2007**, *18*, 431–437.
(8) Halila, S.; Velasco, T.; De Clercq, P.; Madder, A. *Chem. Commun.* **2005**, 936–938.
(9) Stevens, K.; Madder, A. *Nucleic Acids Res.* **2009**, *37*, 1555–1565.
(10) Osborne, S. E.; Cain, R. J.; Glick, G. D. *J. Am. Chem. Soc.* **1997**, *119*, 1171–1182.
(11) Ferentz, A. E.; Keating, T. A.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 9006–14.
(12) Peng, X.; Greenberg, M. M. *Nucleic Acids Res.* **2008**, *36*, e31.
(13) Peng, X.; Hong, I. S.; Li, H.; Seidman, M. M.; Greenberg, M. M. *J. Am. Chem. Soc.* **2008**, *130*, 10299–10306.
(14) Stone, M. P.; Cho, Y.-J.; Huang, H.; Kim, H.-Y.; Kozekov, I. D.; Kozekova, A.; Wang, H.; Minko, I. G.; Lloyd, R. S.; Harris, T. M.; Rizzo, C. J. *Acc. Chem. Res.* **2008**, *41*, 793–804.
(15) Romano, K. P.; Newman, A. G.; Zahran, R. W.; Millard, J. T. *Chem. Res. Toxicol.* **2007**, *20*, 832–838.
(16) Kozekov, I. D.; Nechev, L. V.; Moseley, M. S.; Harris, C. M.; Rizzo, C. J.; Stone, M. P.; Harris, T. M. *J. Am. Chem. Soc.* **2003**, *125*, 50–61.
(17) Ding, H.; Majumdar, A.; Tolman, J. R.; Greenberg, M. M. *J. Am. Chem. Soc.* **2008**, *130*, 17981–17987.
(18) Hong, I. S.; Ding, H.; Greenberg, M. M. *J. Am. Chem. Soc.* **2006**, *128*, 485–491.

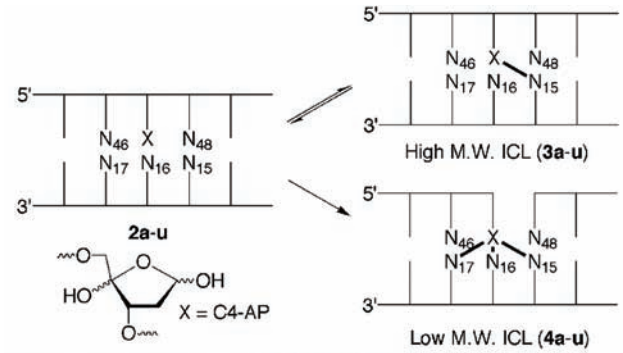
10.1021/ja903404v CCC: $40.75

datively damage DNA.[22−25] The lesion is also produced by ionizing radiation, which generates the reactive oxygen species, hydroxyl radical.[26] Ravanat, Cadet, and co-workers brought to light the possibility that C4-AP forms cross-links.[27] They isolated the dC adduct of the formal C4-AP β-elimination product (**1**) following enzyme digestion of DNA that was treated with bleomycin or ionizing radiation. Because the product was isolated from enzyme-digested DNA, it was not possible to determine whether it resulted from an intra- and/or interstrand reaction between the lesion and native nucleotide.



We investigated the reactivity of C4-AP in a synthetic duplex in which the lesion was embedded in the local sequence 5′-GXT (X = C4-AP).[21] This sequence was identified using pepleomycin as a means for generating C4-AP and determining which sequences yielded cross-links most efficiently. Two types of interstrand cross-links (ICLs), both of which involved exclusive reaction with the dA opposite the 3′-flanking thymidine in **2a** were formed in competition with one another from C4-AP (Scheme 1 and Figure 1). In the kinetically favored ICL, neither DNA strand was cleaved (**3**) but the product was unstable and reverted to C4-AP. The more slowly formed ICL (**4**) was stable and arose formally from reaction of the α,β-unsaturated keto-aldehyde derived from elimination of C4-AP with the same dA in the duplex. The kinetic scheme outline (Scheme 1) involving reversible formation of **3** is described in more detail below. The electrophoretic migrations of the cross-links in denaturing polyacrylamide gels were consistent with their proposed sizes (Figure 1). However, the mechanistic details of ICL formation were uncertain. For instance, formation of the lower molecular weight, more stable cross-link (**4**) was dependent on the identity of the nucleotide opposite C4-AP. ICL formation was severely compromised when C4-AP was opposed by thymidine but was rescued by the addition of exogenous

Scheme 1



Duplex	N46 C4-AP N48 / N17 N16 N15	Duplex	N46 C4-AP N48 / N17 N16 N15
2a	G C4-AP T / C A A	2l	G C4-AP C / C T G
2b	G C4-AP T / C C A	2m	G C4-AP A / C A T
2c	G C4-AP T / C G A	2n	G C4-AP A / C T T
2d	G C4-AP T / C T A	2o	C C4-AP A / G A T
2e	G C4-AP T / C Ne A	2p	C C4-AP A / G T T
2f	G C4-AP T / C A Ne	2q	T C4-AP A / A A T
2g	G C4-AP T / C T Ne	2r	T C4-AP A / A T T
2h	G C4-AP G / C A C	2s	A C4-AP A / T A T
2i	G C4-AP G / C T C	2t	G C4-AP G / C G C
2j	G C4-AP G / C C C	2u	A C4-AP A / T T T
2k	G C4-AP C / C A G		

adenine. The scope of the cross-linking reactions with respect to flanking sequence effects was also largely unexplored.

Results and Discussion

Oligonucleotides containing the photolabile C4-AP precursor were prepared as previously described from **5** (Scheme 2).[21,28] Syntheses of oligonucleotides containing the bis-*o*-nitrobenzyl protected abasic site (**6**) were carried out using shortened acidic detritylation cycles, as reported by Aso to prevent hydrolysis of the acetal and/or ketal components.[29,30] The chemically labile C4-AP lesion was generated as needed via a brief photolysis in a Rayonet photoreactor at λ_{max} = 350 nm from the bis-*o*-nitrobenzyl precursor. Complementary oligonucleotides were hybridized prior to photochemical generation of the oxidized abasic site to minimize adventitious cleavage of the C4-AP lesion. The flanking sequence was varied (**2a**-**u**) to explore the possibility that any native nucleotide containing one or more

(19) Hong, I. S.; Greenberg, M. M. *J. Am. Chem. Soc.* **2005**, *127*, 3692–3693.
(20) Dutta, S.; Chowdhury, G.; Gates, K. S. *J. Am. Chem. Soc.* **2007**, *129*, 1852–1853.
(21) Sczepanski, J. T.; Jacobs, A. C.; Greenberg, M. M. *J. Am. Chem. Soc.* **2008**, *130*, 9646–9647.
(22) Rabow, L. E.; Stubbe, J.; Kozarich, J. W. *J. Am. Chem. Soc.* **1990**, *112*, 3196–3203.
(23) Xi, Z.; Goldberg, I. H. In *Comprehensive Natural Products Chemistry*; Kool, E. T., Ed.; Elsevier: Amsterdam, 1999; Vol. 7, pp 553−592.
(24) Sugiyama, H.; Fujiwara, T.; Kawabata, H.; Yoda, N.; Hirayama, N.; Saito, I. *J. Am. Chem. Soc.* **1992**, *114*, 5573–5578.
(25) Sugiyama, H.; Kawabata, H.; Fujiwara, T.; Dannoue, Y.; Saito, I. *J. Am. Chem. Soc.* **1990**, *112*, 5252–5257.
(26) von Sonntag, C. *The Chemical Basis of Radiation Biology*; Taylor & Francis: London, 1987.
(27) Regulus, P.; Duroux, B.; Bayle, P.-A.; Favier, A.; Cadet, J.; Ravanat, J.-L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14032–14037.
(28) Kim, J.; Gil, J. M.; Greenberg, M. M. *Angew. Chem., Int. Ed.* **2003**, *42*, 5882–5885.
(29) Aso, M.; Usui, K.; Fukuda, M.; Kakihara, Y.; Goromaru, T.; Suemune, H. *Org. Lett.* **2006**, *15*, 3183–3186.
(30) Usiv, K.; Aso, M.; Fukuda, M.; Suemune, H. *J. Org. Chem.* **2008**, *73*, 241–248.

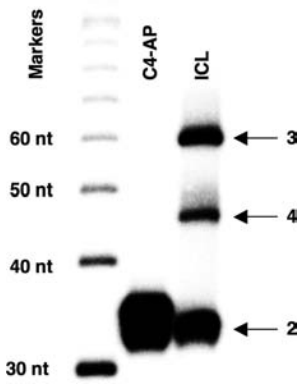


Figure 1. ICL formation from C4-AP containing DNA (**2a**).

Scheme 2

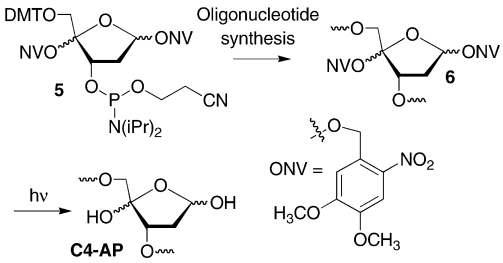


Table 1. Opposing Nucleotide Effect on the Yield of the High Molecular Weight Interstrand Cross-Link at dA_{15} (**3**)

5'-d(CGA CCG GCT CGT AT N_{46} C4-AP N_{48} G TGT GGA GCT GTG G)
3'-d(GCT GGC CGA GCA TA N_{17} N_{16} N_{15} C ACA CCT CGA CAC C)

Substrate	N_{46} C4-AP N_{48} / N_{17} N_{16} N_{15}	Yield **3** (%)[a]	ICL Location
2a[b]	G C4-AP T / C A A	21.3 ± 0.9	A_{15}
2b[b]	G C4-AP T / C C A	16.5 ± 0.9	A_{15}
2c	G C4-AP T / C G A	37.0 ± 0.3	A_{15}
2d[b]	G C4-AP T / C T A	30.9 ± 0.3	A_{15}
2e[b]	G C4-AP T / C Ne A	29.9 ± 0.9	A_{15}

[a] Yields are the average of at least 3 replicates. [b] See ref 21.

nucleophilic nitrogen atoms cross-links with the oxidized abasic site as well as to examine the role of the nucleotide opposing C4-AP on cross-link formation.

Duplex Position and Nucleotide Selective Formation of the Kinetically Favored ICL Product (3). Using hydroxyl radical (HO•) cleavage, we previously showed that ICLs containing an uncleaved C4-AP strand (referred to as the high molecular weight product) involved reaction with a dA_{15} (Table 1). Although the ICL yield varied depending upon the identity of the nucleotide opposite C4-AP in **2a**−**e**, the nucleotide cross-linked was always dA_{15}. Modifying the nucleotide sequence surrounding C4-AP revealed that the reactivity of dA_{15} was distinctive. Incorporation of dC (**2h**−**j**) or nebularine (**Ne**, **2f**, **2g**) at dN_{15} produced low yields (≤1.5%) of the analogous products. The yields of **3f**−**j** were too low to determine the cross-linking site using HO• cleavage. Examination of a variety of other duplexes (**2k**−**r**) revealed that **3** was not formed when either dG (**2k**, **l**) or dT (**2m**−**r**) were present at dN_{15}.

Table 2. Effect of Adenine on the Yield of the High Molecular Weight Interstrand Cross-Link (**3**)

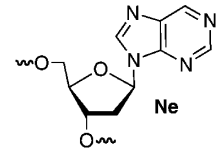
5'-d(CGA CCG GCT CGT AT N_{46} C4-AP N_{48} G TGT GGA GCT GTG G)
3'-d(GCT GGC CGA GCA TA N_{17} N_{16} N_{15} C ACA CCT CGA CAC C)

Substrate	N_{46} C4-AP N_{48} / N_{17} N_{16} N_{15}	Yield **3** (%)[a] No adenine	Yield **3** (%)[a] 0.1 mM Adenine
2b	G C4-AP T / C C A	16.5 ± 0.9	15.3 ± 0.5
2c	G C4-AP T / C G A	37.0 ± 0.3	41.8 ± 0.7
2d	G C4-AP T / C T A	30.9 ± 0.3	16.3 ± 0.6
2i	G C4-AP G / C T C	0.9 ± 0.1	0.6 ± 1.2
2j	G C4-AP G / C C C	1.5 ± 0.1	2.3 ± 0.2

[a] Yields are the average of at least 3 replicates.

Incorporating dA (**2q**, **r**) or dG (**2o**, **p**) at dN_{17} also failed to produce ICL containing full length strands. In contrast to the thermodynamic, low molecular weight ICL (**4**) (below), the yield of the high molecular weight ICL (**3**) experienced little if any increase when adenine was added to the reaction (Table 2), and in one instance (**2d**, Table 2) adenine significantly decreased the ICL yield, which was compensated by an increase in the yield of the lower molecular weight ICL (**4d**).

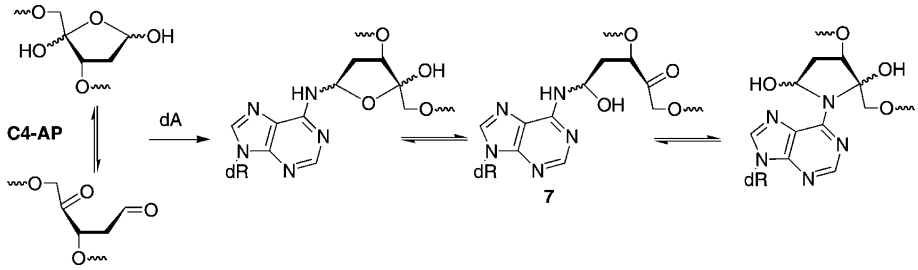
The instability of the high molecular weight ICL precluded unambiguously determining its structure by NMR following its isolation from material digested by enzymes. However, a number of observations lead us to propose that the cross-linked product within **3** is one or more of the cyclic condensation



products (**7**, Scheme 3) involving the exocyclic amine of dA. First, we previously reported that high molecular weight ICL (**3**) is not observed when nebularine (**Ne**) is substituted for dA_{15}.[21] ESI-MS analysis indicates that water is not eliminated when ICL **7** is formed. The lack of reaction with $NaBH_3CN$ is also consistent with the absence of water loss. Preferential formation of **7** over the comparable cross-links with dG at dN_{15} (**2k**, **l**) can be rationalized by the greater proximity of the exocyclic amine in dA to the aldehyde carbon (~7.0 Å for dA_{15} vs ~7.8 Å for dG_{15}) of the acyclic form of C4-AP when the purines are present at N_{15} (Figure 2). However, the lack of reaction with dC at dN_{15} (**2h**−**j**) or dG at dN_{17} (**2o**,**p**) cannot be attributed to proximity effects. For instance, the N2-amino group of dG_{17} is less than 5 Å from the aldehyde carbon of C4-AP (part A of Figure 2). Hence, the differences must be due to inherent reactivity and/or stability of the respective adducts.[20] This reactivity pattern is different than that for most bis-electrophiles, which typically involve a 1,4-conjugate addition.[14,31,32] The formation of **3** (**7**) from C4-AP as proposed in Scheme 3 involves the condensation of two carbonyl groups with an exocyclic amine.

Interstrand Cross-Link Formation and Concomitant Strand Cleavage at C4-AP (4). The C4-AP containing strand is cleaved between the lesion and its 3′-flanking nucleotide in the thermodynamic (low molecular weight) ICL product (**4**). The

Scheme 3



nucleotide opposing C4-AP significantly affects the yield of this ICL (Table 3). Higher yields were obtained when dA (**2a**) or nebularine (**2e**) opposed the lesion compared to dG or a pyrimidine. Unlike the less stable, high molecular weight cross-link (**3**), high yields of the low molecular weight ICL (**4**) were observed with dC as well as dA (but not dG or dT) and at positions other than dN_{15} (Tables 3 and 4). When dA or dC (**2h**−**j**, Table 4) was present at dN_{15}, this was still the preferred reaction site. However, when dA or dC was not present at this position, preferential ICL formation occurred at the opposing nucleotide (dN_{16}) and/or at dN_{17}. In one instance, reaction at dN_{16} occurred when dC was opposite the lesion, albeit in lower yield, even though dA was present at dN_{15} (**2b**, Table 3). Given the facility with which many bis-electrophiles react with dC and dG, the lower yield of **4b** suggests that there is another factor controlling the cross-linking reaction (below).[14,16,33−35] Reaction at dA_{16} was especially prominent when dA or dG was incorporated at dN_{17} and dT was at dN_{15} (**2o**, **2q**, Table 4). However, no reaction was detected at dG even when it was incorporated at dN_{17} (**2o**, **2p**, Table 4) where the exocyclic amine was best positioned to react with C4-AP (Figure 2A). In some instances, small amounts of cross-linking were even detected at dC_{14} (**4b**, **4c**, **4i**, **4o**, and **4q**), which illustrates the inherent high reactivity of dC.[36]

Effect of the Nucleotide Opposite C4-AP on the Yield of the Low Molecular Weight Interstrand Cross-Link (4) and Its Rescue by Additives. The preliminary report concerning ICL formation from C4-AP described the effect of adenine and other additives on the cross-link yield when the abasic site was opposed by thymidine.[21] Adenine rescued ICL formation from **2d** in a dose-dependent manner that showed a saturation effect consistent with a process involving binding. The enhancement in ICL yield by adenine was observed for several other duplexes (Table 5) and was not limited to cross-linking to dA or even to the dN_{15} position. When C4-AP was opposed by a pyrimidine 0.1 mM adenine-enhanced ICL formation between 3.6- and 18-fold. However, the ICL yield increased less than 2-fold when the larger dG was opposite C4-AP (**2c**, **2t**, Table 5). These observations are consistent with the hypothesis that binding by adenine plays a role in the ICL enhancement. In instances where cross-linking was measurable in the absence of adenine, hydroxyl radical cleavage mapping showed that purine addition did not alter the nucleotide position where cross-linking occurred.[36] This suggested that the species responsible for directly forming the ICL was the same whether or not adenine was added (below).

Adenine is not unique in its ability to rescue ICL formation. In general, purines (purine, 2-aminopurine) were more effective than other potential bases or nucleophiles, including cytosine and imidazole.[21] The ability of a more comprehensive collection of purines to induce ICL formation from **2d** was examined (Figure 3). Chloro and methoxy substituents did not improve the ICL yield compared to the unsubstituted purine. In contrast, 2-aminopurine and adenine were more effective than purine at rescuing the ICL product. Most striking was the effect of 0.1 mM 2,6-diaminopurine, which increased the ICL yield almost 6-fold relative to when an equal concentration of purine was added. There are two factors that can account for the large enhancement by 2,6-diaminopurine. The purine was previously proposed to enhance the ability of artificial nucleases to excise abasic sites due to its stronger binding.[37] However, one cannot rule out a more direct involvement of the exocyclic amines in rescuing ICL formation by reacting with the damaged DNA.

Structural Characterization of the Low Molecular Weight, Thermodynamic ICL. Hydroxyl radical cleavage showed that the low molecular weight cross-link obtained from **2d** in the presence of adenine (0.1 mM) resulted from formal reaction of the cleaved C4-AP containing a strand at the site of the lesion with the opposing oligonucleotide. Greater structural detail was obtained following enzymatic digestion of **4d**. This material was compared by HPLC and ESI-MS/MS to the products obtained from the reaction of dA with independently prepared **8** and subsequent deacetylation (Scheme 4). Sufficient quantities of the cross-linked material were obtained by using dT opposite

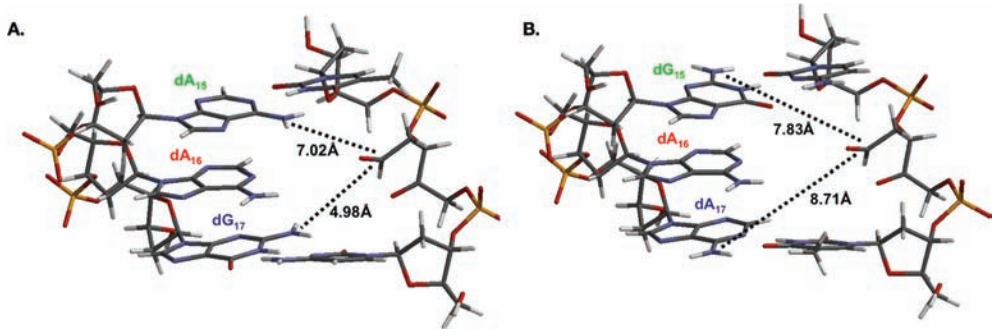


Figure 2. Proximity of potential nucleophilic partners for C4-AP when opposed by dA and flanked by (A) 5′-dC and 3′-dT (B) 5′-dT and 3′-dC. Structures were obtained using *Spartan*.

Table 3. Opposing Nucleotide Effect on the Yield and Location of the Low Molecular Weight Interstrand Cross-Link (**4**)

5'-d(CGA CCG GCT CGT AT N_{46} C4-AP N_{48} G TGT GGA GCT GTG G)
3'-d(GCT GGC CGA GCA TA N_{17} N_{16} N_{15} C ACA CCT CGA CAC C)

Substrate	N_{46} C4-AP N_{48} / N_{17} N_{16} N_{15}	Yield **4** (%)[a]	ICL Location
2a[b]	G C4-AP T / C A A	7.3 ± 1.0	A_{15}
2b[b]	G C4-AP T / C C A	3.1 ± 0.3	C_{16}
2c	G C4-AP T / C G A	2.1 ± 0.5	A_{15}, C_{17} (2.2 : 1.0)[c]
2d[b]	G C4-AP T / C T A	2.6 ± 0.3	A_{15}, C_{17} (1.9 : 1.0)[c]
2e[b]	G C4-AP T / C Ne A	5.5 ± 0.4	A_{15}, C_{17} (2.0 : 1.0)[c]

[a] Yields are the average of at least 3 replicates. [b] See ref 21. [c] The ratio of reaction at the nucleotides is in parentheses.

Table 4. Flanking Sequence Effect on the Yield of the Low Molecular Weight Interstrand Cross-Link (**4**)

5'-d(CGA CCG GCT CGT AT N_{46} C4-AP N_{48} G TGT GGA GCT GTG G)
3'-d(GCT GGC CGA GCA TA N_{17} N_{16} N_{15} C ACA CCT CGA CAC C)

Substrate	N_{46} C4-AP N_{48} / N_{17} N_{16} N_{15}	Yield **4** (%)[a]	ICL Location
2f	G C4-AP T / C A Ne	8.0 ± 0.8	A_{16}, C_{17} (1.7 : 1.0)[c]
2g	G C4-AP T / C T Ne	2.0 ± 0.3	C_{17}[b]
2h	G C4-AP G / C A C	7.3 ± 0.7	C_{15}
2i	G C4-AP G / C T C	2.8 ± 0.8	C_{15}, C_{17} (1.9 : 1.0)[c]
2j	G C4-AP G / C C C	1.5 ± 0.2	C_{16}
2k	G C4-AP C / C A G	5.1 ± 0.4	A_{16}, C_{17} (0.8 : 1.0)[c]
2l	G C4-AP C / C T G	0.7 ± 0.1	C_{17}[b]
2m	G C4-AP A / C A T	33.9 ± 0.5	A_{16}, C_{17} (0.9 : 1.0)[c]
2n	G C4-AP A / C T T	0.8 ± 0.1	C_{17}[b]
2o	C C4-AP A / G A T	49.3 ± 1.6	A_{16}
2p	C C4-AP A / G T T	0.3 ± 0.1	n.d.
2q	T C4-AP A / A A T	21.0 ± 0.6	A_{16}
2r	T C4-AP A / A T T	0.7 ± 0.1	n.d.
2s	A C4-AP A / T A T	10.9 ± 0.4	A_{16}

[a] Yields are the average of at least 3 replicates. [b] Determined by reaction in the presence of 0.1 mM adenine. [c] The ratio of reaction at the nucleotides is in parentheses.

C4-AP in conjunction with exogenous adenine (0.1 mM). On the basis of the structure of **1** and the similarities in reactivity of dA and dC with other bis-electrophiles, we expected that adducts between dA and the elimination product of C4-AP consist of stereoisomers of **9** (and/or **10**). In the case of **1**, four diastereomers of a single regioisomer were formed when dC was reacted with **8** (Scheme 4). However, when we reacted dA under similar conditions only two stereoisomers were observed

Table 5. Effect of Adenine on the Yield of the Low Molecular Weight Interstrand Cross-Link (**4**)

5'-d(CGA CCG GCT CGT AT N_{46} C4-AP N_{48} G TGT GGA GCT GTG G)
3'-d(GCT GGC CGA GCA TA N_{17} N_{16} N_{15} C ACA CCT CGA CAC C)

Substrate	N_{46} C4-AP N_{48} / N_{17} N_{16} N_{15}	Yield **4** (%)[a]	
		No adenine	0.1 mM Adenine
2b	G C4-AP T / C C A	3.1 ± 0.3	13.4 ± 1.2
2c	G C4-AP T / C G A	2.1 ± 0.5	2.9 ± 0.6
2d	G C4-AP T / C T A	2.6 ± 0.3	12.5 ± 0.6
2i	G C4-AP G / C T C	2.8 ± 0.8	22.2 ± 1.6
2j	G C4-AP G / C C C	1.5 ± 0.2	5.4 ± 0.3
2l	G C4-AP C / C T G	0.7 ± 0.1	12.2 ± 0.8
2n	G C4-AP A / C T T	0.8 ± 0.1	14.4 ± 0.5
2p	C C4-AP A / G T T	0.3 ± 0.1	3.3 ± 0.4
2r	T C4-AP A / A T T	0.7 ± 0.1	4.2 ± 0.8
2t	G C4-AP G / C G C	2.0 ± 0.2	3.2 ± 0.4

[a] Yields are the average of at least 3 replicates.

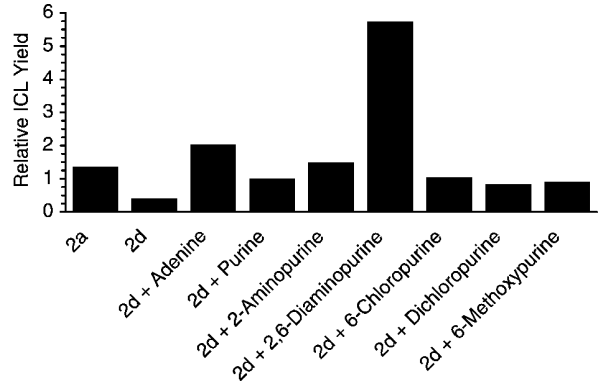
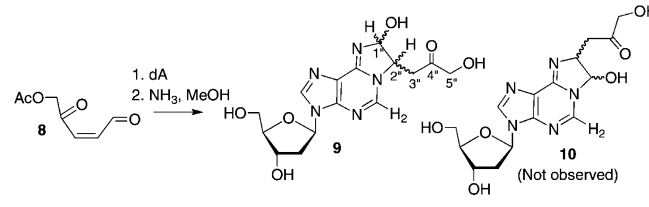


Figure 3. Effects of purine additives on the yield of the low molecular weight ICL (**4**). Yields are expressed relative to the yield of **4d** obtained from **2d** in the presence of purine (0.1 mM).

Scheme 4



by HPLC. The yield of the dA adducts were low (~15% of both isomers), even after incubation with **8** for 3 days at 37 °C. The adducts were unstable at alkaline pH where they reverted to dA and unknown compounds. This was not surprising because reactions of dA with other bis-electrophiles such as 1,4 butenedial have been reported to be slow and the resulting products unstable, especially when compared to analogous reactions with dC.[33,34] Each isomer had a corresponding mass of 365 and a fragmentation of 365 → 249 in agreement with

glycosidic bond cleavage of **9** (and/or **10**).[36] As was the case for **1**, the isolated isomers did not interconvert.[27,36] This behavior differs from the respective adducts of dA and 1,4-butenedial, which rapidly isomerize.[34] Each isomer was characterized separately by 2D COSY and ROESY NMR experiments and assigned as diastereomers of **9** (Scheme 4).[36] The COSY spectra enabled us to establish the *J*-coupling connectivity of the atoms in the adduct. Correlation between a single proton doublet at 6.23 ppm and a single proton multiplet at 5.23 ppm led to their assignments as the 1″ and 2″ protons, respectively. The correlation between the proton at 5.23 ppm and protons (2) at 2.62 ppm resulted in their assignment as the 3″ methylene group. Adduct regiochemistry was more firmly established in ROESY experiments that correlate protons, which are in spatial proximity (<5 Å).[36] In both isomers, the H_2 proton showed a weak cross-peak to $H_{2''}$ but not to $H_{1''}$, as expected from **9** (but not **10**) in Scheme 4. This is confirmed by the presence of additional, medium to strong intensity cross peaks between $H_{2''}$ and the ($H_{1''}$, $H_{3''}$) groups of protons. The NMR data are also in good agreement with assignments made for **1**, again consistent with the proposal that the two isolated molecules are diastereomers of **9**. These assignments are also consistent with the reactions of other aldehydes with the exocyclic amines of the nucleobases.[14,27,33,34]

To verify the presence of **9** in cross-linked DNA, isolated cross-links from **2d** (reacted in the presence of 0.1 mM adenine) were subjected to enzymatic digestion. Alkaline conditions were avoided through the use of antarctic phosphatase, replacing the more commonly used alkaline phosphatase.[38] HPLC analysis of the digested material revealed the four common nucleosides along with two additional peaks with retention times corresponding to the two isomers of **9**.[36] Coinjections with authentic **9** suggested that the compounds were identical. In addition, MS/MS spectra of **9** obtained from ICL that was digested by enzymes showed the same fragmentation pattern as independently synthesized material. These data lead us to conclude that the thermodynamic ICLs (**4**) derived from C4-AP and 2′-deoxyadenosine consist of two diastereomers of **9** whose structures are analogous to **1**. In sequences in which cross-linking was observed to dC, the actual adduct was not isolated but was assumed to be **1** based upon the previous work of Cadet as well as the similarities noted above in the reaction between dA and **8**.[27]

Mechanism for Formation of the Low Molecular Weight, Thermodynamic ICL Product (4). Because of the reversible formation of the kinetically preferred, high molecular weight ICL (**3**), formation of the low molecular weight ICL (**4**) does not follow simple first-order kinetics. The disappearance of isolated high molecular weight ICL (**3a**) and the appearance of C4-AP (**2a**) and low molecular weight ICL (**4a**) were measured as a function of time (Figure 4). The formation of elimination

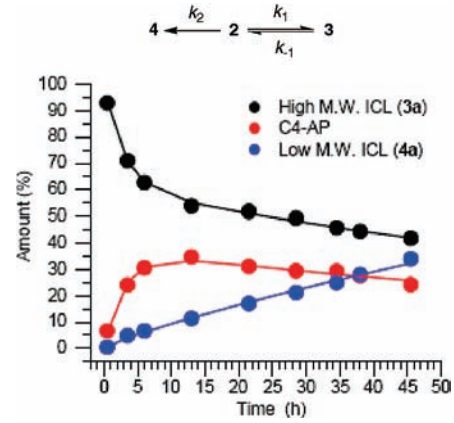
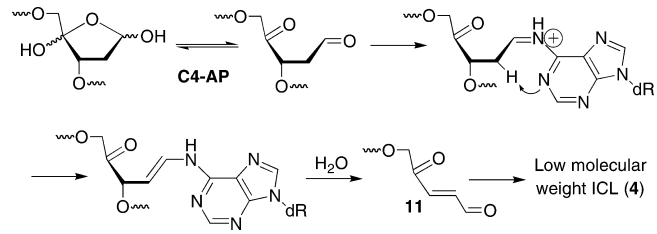


Figure 4. Kinetic fitting of the interconversion of C4-AP (**2a**), **3a**, and **4a**.

Scheme 5



product from **2a** was negligible under these conditions. The product profiles were fit to the kinetic scheme described by these reactivity parameters, and rate constants for each step were determined using an iterative fitting procedure.[39] The data affirm the designation of the high molecular weight cross-link as the kinetic product. The rate constant for its formation ($k_1 = 3.0 \pm 0.1 \times 10^{-5}$ s^{-1}) is almost 10 times faster than that which describes production of the low molecular weight ICL ($k_2 = 3.4 \pm 0.3 \times 10^{-6}$ s^{-1}). The data also reveal that the equilibrium between C4-AP and the high molecular weight cross-link lies on the side of the oxidized abasic lesion ($k_{-1} = 6.0 \pm 0.2 \times 10^{-5}$ s^{-1}).

Whereas the kinetic fitting places the competitive formation of the two ICLs in a quantitative context, it does not address the questions of how the low molecular product is formed or the role of the surrounding sequence in its formation. Previous studies revealed that independent generation of the α,β-unsaturated aldehyde produced by elimination from C4-AP (**11**, Scheme 5) using endonuclease III rapidly formed the low molecular weight ICL.[21,40] The rate constant for ICL formation from the α,β-unsaturated aldehyde was estimated to be $>10^{-4}$ s^{-1}. The data obtained from the fitting indicate that this potential intermediate reacts much more rapidly than the rate constant with which C4-AP (**2**) yields the low molecular weight cross-link (**4**) and that its formation could be the rate-determining step in the overall process. We originally did not consider the free unsaturated aldehyde (**11**) to be a precursor to the low molecular weight ICL because **2u**, which cannot form a similar cross-link, also did not produce a significant yield of the respective elimination product, even in the presence of 0.1 mM adenine.[21] This proposal was flawed in that it assumed that only the opposing nucleotide affected the ICL formation rate.

(31) Chen, B.; Vu, C. C.; Byrns, M. C.; Dedon, P. C.; Peterson, L. A. *Chem. Res. Toxicol.* **2006**, *19*, 982–985.
(32) Byrns, M. C.; Vu, C. C.; Peterson, L. A. *Chem. Res. Toxicol.* **2004**, *17*, 1607–1613.
(33) Gingipalli, L.; Dedon, P. C. *J. Am. Chem. Soc.* **2001**, *123*, 2664–2665.
(34) Byrns, M. C.; Predecki, D. P.; Peterson, L. A. *Chem. Res. Toxicol.* **2002**, *15*, 373–379.
(35) Sodum, R. S.; Shapiro, R. *Bioorg. Chem.* **1988**, *16*, 272–82.
(36) See Supporting Information.
(37) Lhomme, J.; Constant, J. F.; Demeunynck, M. *Biopolymers* **1999**, *52*, 65–83.
(38) Byrns, M. C.; Vu, C. C.; Neidigh, J. W.; Abad, J.-L.; Jones, R. A.; Peterson, L. A. *Chem. Res. Toxicol.* **2006**, *19*, 414–420.
(39) Bamford, C. H.; Tipper, C. F. H. The Theory of Kinetics; Elsevier: Amsterdam, 1969; Vol. 2.
(40) Greenberg, M. M.; Weledji, Y. N.; Kim, J.; Bales, B. C. *Biochemistry* **2004**, *43*, 8178–8183.

Table 6. Dithiothreitol Competition with Elimination (**11**) and Interstrand Cross-Link (**4**) Products

5′-d(CGA CCG GCT CGT AT N_{46} C4-AP N_{48} G TGT GGA GCT GTG G)
3′-d(GCT GGC CGA GCA TA N_{17} N_{16} N_{15} C ACA CCT CGA CAC C)

Substrate	N_{46} C4-AP N_{48} / N_{17} N_{16} N_{15}	Yield **4** (%)[a] ICL (**4**) + **11**[b]	DTT adduct[c]
2a	G C4-AP T / C A A	22.5 ± 2.2	29.9 ± 1.4
2m	G C4-AP A / C A T	63.2 ± 1.0	67.6 ± 1.5
2n	G C4-AP A / C T T	5.8 ± 1.5	7.5 ± 1.0
2n[d]	G C4-AP A / C T T	38.8 ± 1.0	39.0 ± 1.3
2o	C C4-AP A / G A T	70.6 ± 1.1	83.8 ± 0.6
2u	A C4-AP A / T T T	2.5 ± 0.7	3.3 ± 1.0

[a] Yields are the average after 16 h of at least 3 replicates. [b] Percent yield of low molecular weight ICL + C4-AP β-elimination product in the absence of DTT. [c] Percent yield of DTT adduct with **11**. [d] Reaction was carried out in the presence of 0.1 mM adenine.

Consequently, the possibility that C4-AP elimination to **11** was the rate-determining step in low molecular weight ICL (**4**) formation was investigated further. First, the nucleotides at which cross-linking occurs when the α,β-unsaturated aldehyde was independently generated by endonuclease III treatment were determined by hydroxyl radical cleavage in several duplexes.[36] In every instance, the positions of the ICL were identical to those when the cross-links were produced from the same duplexes from which **11** was not preformed by the enzyme (Tables 3 and 4). Furthermore, in those instances in which multiple positions were cross-linked, if the duplex was treated with endonuclease III the ratios of reaction at the nucleotides were the same (**2d**, **2i**, **2m**, Tables 3 and 4).[36] These observations are consistent with the intermediacy of **11** in ICL formation without endonuclease digestion. More direct evidence for **11**′s role in low molecular weight ICL (**3**) formation was obtained by trapping it with dithiothreitol (DTT).[41−43] Several C4-AP containing duplexes were incubated at 37 °C for 16 h with or without 1 mM DTT (Table 6). The thiol had no effect on the yield of high molecular weight ICLs (**3**), indicating that it did not react with C4-AP itself.[36] The thiol effect was quite different on the other products. Although the yields varied widely depending upon sequence, in the absence of DTT each duplex yielded a mixture of low molecular weight ICL (**4**) and a cleavage product, which was believed to be the β-elimination product (**11**) or one derived from it (Figure 5). No low molecular weight ICL was observed in the presence of DTT. In addition, **11** was replaced by two new products. The electrophoretic mobility of the new products was similar to that of **11** and were attributed to the respective DTT adducts but were not independently synthesized. The agreement between the yields of the DTT adducts and the respective products formed in the absence of thiol were very good (Table 6). In addition, when the C4-AP duplex was incubated in the presence of adenine (**2n**, Table 6) the greater DTT adduct yield over the same reaction time showed good correlation with the increased amounts of ICL

Figure 5. Dithiothreitol (DTT, 1 mM) intercepts low molecular weight ICL formation (**4o**) from **2o** and forms adducts with independently generated **11** by Endo III treatment of **2o**. * indicate proposed DTT adducts.

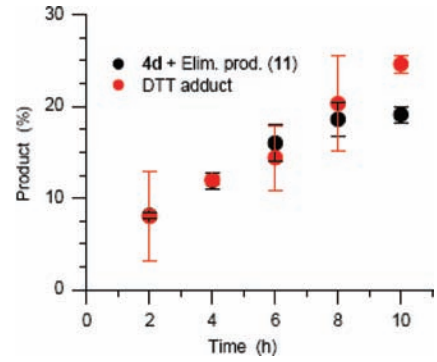


Figure 6. Correlation between yields of DTT (1 mM) adducts with **11** with the sum total of C4-AP cleavage product(s) (**11**) and low molecular weight ICLs (**4d**) in the absence of DTT from **2d** + 0.1 mM adenine.

and **11**. This suggested that the role of the additive was to accelerate C4-AP elimination. Additional correlation between DTT adducts and products formed in the thiol's absence were obtained by measuring their yields as a function of time from **2b** incubated in the presence of 0.1 mM adenine (Figure 6). The good agreement between the yield of thiol adducts and the sum total of ICL and elimination product formed in the absence of the trap suggests that DTT intercepts a C4-AP cleavage intermediate en route to the cross-linked product. Mechanistic economy leads us to propose that the β-elimination product of C4-AP is the immediate precursor to the low molecular weight ICL and its formation is the rate determining step. These data also indicate that the role of the purines is to promote C4-AP elimination. The greater proclivity of amino-substituted purines to induce elimination may be due to stronger binding, which increases their effective molarity. However, we cannot rule out the possibility that the amino-substituted purines form intermediate Schiff bases and that the ring nitrogens act as general bases (Scheme 5). The purines act as acid−base catalysts in

(41) Sato, K.; Greenberg, M. M. *J. Am. Chem. Soc.* **2005**, *127*, 2806–2807.
(42) Hwang, J.-T.; Tallman, K. A.; Greenberg, M. M. *Nucleic Acids Res.* **1999**, *27*, 3805–3810.
(43) Zheng, Y.; Sheppard, T. L. *Chem. Res. Toxicol.* **2004**, *17*, 197–207.

reactions mediated by ribozymes.[44−47] The role of the opposing (and possibly flanking) adenines in the rate determining step for ICL formation from C4-AP also explains why the yields of cross-links with dC_{15-17} in some sequences are not as high as one would expect based upon the expected reactivity of the pyrimidine with other bis-electrophiles (Tables 3 and 4).[14,27,33,34] This is borne out by the large increase in ICLs with dC when adenine is added (Table 5).

Conclusions

The surrounding sequence has a large effect on the chemoselectivity, rate, and nucleotide location of interstrand cross-links produced from the C4′-oxidized abasic site (C4-AP). Although proximity plays a role in the ICLs formed, the nucleophilicity of the nucleobase and/or adduct stability are believed to be the dominant factors. This is evidenced by the ineffectiveness of dG at reacting with C4-AP, even when the N2-amino group is well positioned to form interstrand cross-links with the lesion. The sequences most favored for C4-AP formation by bleomycin, the antitumor agent that is commonly associated with this lesion, are also some of the most proficient for forming the thermodynamic ICL. Despite this, to our knowledge ICLs were not previously described in the reactions of bleomycin with DNA. One possible reason for this is that ICL formation requires incubation of the damaged DNA. In addition, carrying damaging reactions out in Tris buffer could inhibit cross-link formation due to competing cyclic adduct formation with C4-AP.[48,49] Interstrand cross-links have also been detected in the reactions of enediynes with DNA under anaerobic conditions, although these are believed to involve reaction of the bound drug with the DNA radicals that they generate.[50−54] However, the enediyne antitumor agents also generate C4-AP and like bleomycin may produce DNA interstrand cross-links indirectly via the formation of this and other metastable DNA lesions.[27]

Recently, a C4-AP adduct of dC was detected in cellular DNA following enzyme digestion.[27] The experiments described here suggest that these may be derived from interstrand cross-links. However, they may also be derived from intrastrand cross-links, a type of DNA lesion for which a general assay has not yet been reported. Both of these types of complex lesions could prove challenging to repair and therefore may have deleterious consequences on cells.

Acknowledgment. We are grateful for support of this research by the National Institute of General Medical Sciences (GM-063028). We thank Professor Jack Norton for assistance with analyzing the kinetics for cross-link formation.

Supporting Information Available: Complete description of experimental procedures, ESI-MS of oligonucleotides containing **6**, sample autoradiograms, NMR spectra of **9**. This material is available free of charge via the Internet at http://pubs.acs.org.

JA903404V

(44) Ditzler, M. A.; Sponer, J.; Walter, N. G. *RNA* **2009**, *15*, 560–575.
(45) Nam, K.; Gao, J.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 4680–4691.
(46) Wilson, T. J.; Ouellet, J.; Zhao, Z.-Y.; Harusawa, S.; Araki, L.; Kurihara, T.; Lilley, D. M. J. *RNA* **2006**, *12*, 980–987.
(47) Meli, M.; Vergne, J.; Maurel, M.-C. *J. Biol. Chem.* **2003**, *278*, 9835–9842.
(48) Dhar, S.; Kodama, T.; Greenberg, M. M. *J. Am. Chem. Soc.* **2007**, *129*, 8702–8703.
(49) Kodama, T.; Greenberg, M. M. *J. Org. Chem.* **2005**, *70*, 9916–9924.
(50) Xu, Y.-j.; Xi, Z.; Zhen, Y.-s.; Goldberg, I. H. *Biochemistry* **1997**, *36*, 14975–14984.
(51) Xu, Y.; Zhen, Y.; Goldberg, I. H. *J. Am. Chem. Soc.* **1997**, *119*, 1133–1134.
(52) Gates, K. S. In *Comprehensive Natural Products Chemistry, DNA and Aspects of Molecular Biology*; Kool, E. T., Ed.; Elsevier: Amsterdam, 1999; Vol. 7.
(53) Beerman, T. A.; Gawron, L. S.; Shin, S.; Shen, B.; McHught, M. M. *Cancer Res.* **2009**, *69*, 593–598.
(54) Kennedy, D. R.; Lu, J.; Shen, B.; Beerman, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17632–17637.