

Published in final edited form as:

*Bioorg Med Chem.* 2011 October 1; 19(19): 5788–5793. doi:10.1016/j.bmc.2011.08.024.

## Probing DNA interstrand cross-link formation by an oxidized abasic site using nonnative nucleotides

Jonathan T. Sczepanski<sup>a</sup>, Christine N. Hiemstra<sup>a</sup>, and Marc M. Greenberg<sup>a,\*</sup>

<sup>a</sup>Department of Chemistry, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218

### Abstract

The C4'-oxidized abasic site (C4-AP) forms two types of interstrand cross-links with the adjacent nucleotides in DNA. Previous experiments revealed that dG does not react with the lesion and that formation of one type of cross-link is catalyzed by the opposing dA. *Iso*-guanosine•dC and 2-aminopurine•dT base pairs were used to determine why dG does not cross-link with C4-AP despite its well known reactivity with other bis-electrophiles. 7-Deaza-2'-deoxyadenosine was used to probe the role of the nucleotide opposite C4-AP in the catalysis of interstrand cross-link formation.

### Keywords

DNA damage; interstrand cross-link; abasic sites; modified oligonucleotides; modified nucleotides

## 1. Introduction

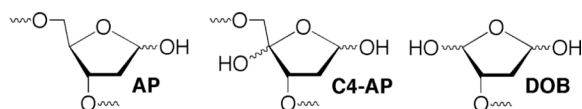
Abasic sites (AP) and related oxidized abasic lesions (e.g. C4-AP, DOB) are common forms of DNA damage.<sup>1,2</sup> While toxic in their own right, recent studies have shown these lesions can react with their surroundings to form even more deleterious types of damage.<sup>3–6</sup> This includes interstrand cross-links (ICLs), which are potent blocks to replication and transcription.<sup>7</sup> Accurate ICL repair is essential to genomic integrity and how this occurs is an active area of research.<sup>8–11</sup> ICLs are produced by a variety of cytotoxic damaging agents, as well as ionizing radiation. Therapeutic agents whose cytotoxicity is linked to ICL formation include the nitrogen mustards, mitomycin C, *cis*-platin, and the enediyne C1027.<sup>12–15</sup> A variety of bis-electrophiles, such as acrolein, and others that are produced endogenously also alkylate opposing strands of duplex DNA to produce ICLs.<sup>16,17</sup> Recently, an adduct of 2'-deoxycytidine was identified in cellular DNA that was consistent with interstrand (and/or intrastrand) cross-link formation between it and the oxidized abasic site resulting from hydrogen atom abstraction from the C4'-position of DNA (C4-AP).<sup>18</sup> C4-AP is produced by a variety of oxidizing agents, including bleomycin, the enediynes, and ionizing radiation.<sup>19–22</sup> This oxidized abasic lesion was known to be mutagenic and has been shown to irreversibly inhibit DNA polymerase  $\beta$ , but until recently had not been suggested to form DNA interstrand cross-links.<sup>18,23–25</sup>

© 2011 Elsevier Ltd. All rights reserved.

\*Corresponding author. Tel.: +0-410-516-8095; fax: +0-410-516-7044; mgreenberg@jhu.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

C4-AP was unequivocally shown to form two types of ICLs with dA and one with dC by using chemically synthesized 31 base pair oligonucleotide substrates (**1**, Figure 1A) containing the lesion at a defined site.<sup>4,26</sup> Both strands were intact in Type I ICLs (high molecular weight, **2**, Figures 1B, C), which were only observed with dA in duplexes containing this nucleotide at position 15.  $\beta$ -Elimination from C4-AP gave rise to Type II ICLs (low molecular weight, **3**, Figures 1B, D) with dA and dC at multiple positions (N<sub>14</sub>–N<sub>17</sub>) in the vicinity of the abasic site. Type I ICLs (**2**) formed reversibly, and reverted to C4-AP in several hours, but the Type II ICLs (**3**) were stable.<sup>26</sup> Type II ICLs are transformed into double strand breaks during nucleotide excision repair, indicating that a simple oxidized abasic lesion can give rise to the most deleterious form of DNA damage.<sup>27</sup>



Although the alkylating agents such as acrolein typically react with dG, cross-linking was not detected between C4-AP and this nucleotide regardless of its position in the duplex relative to the oxidized abasic lesion. Another unusual feature of C4-AP cross-link formation was that the DNA in the vicinity of the abasic site catalyzed the reaction. These aspects of this chemical process are explored further in this contribution using nonnative nucleotides as probes and by taking advantage of our ability to use chemical synthesis to prepare any desired DNA sequence containing C4-AP.

## 2. Results and Discussion

### 2.1 Substrate design

The C4'-oxidized abasic site (C4-AP) is produced by a variety of DNA oxidizing agents, but in greatest yield by bleomycin under O<sub>2</sub> deficient conditions. However, these agents typically do not produce C4-AP at a single site in substrates longer than several nucleotides. Consequently, we and other groups have developed methods for independently producing C4-AP from stable precursors in chemically synthesized oligonucleotides.<sup>4,28–31</sup> In this study, C4-AP was generated on an as needed basis from the photolabile precursor (**5**), which was incorporated into oligonucleotides from phosphoramidite **4** during solid-phase DNA synthesis (Scheme 1).<sup>4</sup> The general sequence in **1** is part of a restriction fragment from pBR322, which possesses a hot-spot for cleavage by bleomycin.<sup>4</sup>

### 2.2 Probing the absence of C4-AP cross-linking with dG

Diffusible bis-electrophiles such as acrolein, 4-hydroxynon-2-enal and but-2-ene-dial react readily with dG.<sup>16,17,32</sup> An ICL between dG and AP was also trapped under reductive conditions.<sup>3</sup> Although proximity influences C4-AP reactivity, it alone does not explain the lack of cross-linking between C4-AP and dG. Molecular models suggested that the N2-amino group of dG at position 17 (dG<sub>17</sub>) is ~2 Å closer to the lesion's reactive C1-aldehyde than is the N6-amine of dA at position 15 (dA<sub>15</sub>, Figure 2).<sup>26</sup> However, interstrand cross-links occurred exclusively with the more distant N6 of dA<sub>15</sub>. Similarly, dG does not form ICLs with dioxobutane, or the bis-electrophile derived from furan oxidation, which are also 1,4-dicarbonyl molecules.<sup>5,33,34</sup>

**2.2.1 Cross-linking in duplexes containing iso-guanosine**—The exocyclic amine and carbonyl group in guanosine and *iso*-guanosine form comparable hydrogen bonds to dC and *iso*-guanosine (iG), respectively, but with opposite orientation with respect to the double helix. Introducing an iG:iC base pair positions the amine of iG in the major groove in approximately the same position as the N6-amine of dA (Figure 3), which forms Type I and

II cross-links with C4-AP (Table 1). *Iso*-Guanosine was introduced into **1** at position 17 (**1l**, **1m**) and position 15 (**1c**, **1d**, **1g**, **1h**) in **1**. Type I ICLs were not observed in any of the duplexes. This is consistent with previous studies on duplexes containing C4-AP and dG at the corresponding positions (**1a**, **1b**, **1e**, **1f**, **1i**, **1j**).<sup>26</sup> Type II ICLs formed when iG was incorporated at N<sub>17</sub> (**1l**, **1m**), albeit not with *iso*-guanosine. The relatively higher ICL yield when C4-AP was opposite dA (**1l**) compared to dT (**1m**) was consistent with the previous reported ICL catalysis by the purine nucleotide opposite the lesion (**1i**, **1j**).<sup>4,26</sup>

The lack of reactivity with iG when it is present at dN<sub>17</sub> is not surprising, given the positioning of the exocyclic amine. For comparison, ICLs do not form with dA<sub>17</sub> either. iG is better positioned to react with C4-AP when it is incorporated at dN<sub>15</sub> (**1c**, **1d**, **1g**, **1h**) because its exocyclic amine is in a position comparable to the N6-amino group of dA<sub>15</sub>, where cross-links do form. Type II ICLs (**3**) are produced in each of these duplexes and iG<sub>15</sub> is involved in all cases (Table 1). In contrast, no reaction is observed with dG<sub>15</sub> in comparable duplexes (**1a**, **1b**, **1e**, **1f**). Type II ICL (**3**) yields increase when iG<sub>15</sub> is mispaired with dT (**1g** versus **1c**, **1h** versus **1d**) and all cross-linking occurs with the nonnative nucleotide. These data suggest that iG accessibility improves its reactivity. However, we do not know whether this is due to greater access to its Watson-Crick face in the mispair and/or greater dynamic flexibility overall.

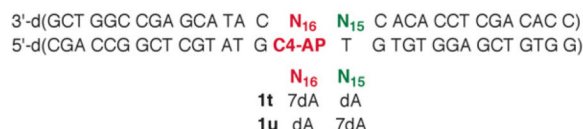
The increased iG reactivity when part of a mispair (**1g**, **1h**) led us to examine the reactivity of dG<sub>17</sub> when it too is opposite dT (**1k**). Type II ICL (**3**) formation was similar to that when a dC:dG base pair was present at this position (**1i**). The similarity included the absence of any reactivity at dG<sub>17</sub>. However, dT<sub>46</sub>:dG<sub>17</sub> containing DNA (**1k**) also yielded Type I ICLs (**2**, 6.8 ± 0.1 %). Moreover, cross-linking occurred at dG<sub>17</sub>. (Cross-linking sites were determined using hydroxyl radical cleavage.) This is the only sequence of the more than 30 examined in which a Type I ICL was observed between C4-AP and a native nucleotide other than dA. This observation is also consistent with the idea that the lack of reactivity previously observed with dG is partially attributable to relative inaccessibility to its nucleophilic functional groups, perhaps due to stronger base pairing with dC than that observed in a dA:dT base pair.<sup>26</sup>

**2.2.2 Cross-linking in duplexes containing 2-aminopurine: thymidine base pairs**—Cross-linking by iG<sub>15</sub> and dT<sub>46</sub>:dG<sub>17</sub> led us to introduce 2-aminopurine (2AP) at the comparable locations (**1p**, **1q**, **1s**, Table 2). 2AP<sub>17</sub> positions the exocyclic amine in the vicinity where the N2-amino group of dG<sub>17</sub> resides. Type I and II cross-links form in duplexes containing 2AP<sub>17</sub> when C4-AP is opposite dA or dT (**1p**, **1q**, Table 2). As in other instances involving comparable duplexes containing dA<sub>17</sub> instead of 2AP<sub>17</sub> (e.g. **1n**, **1o**), ICL yields are higher when C4-AP is opposite dA (**1p**) than dT (**1q**). The subtleties of local duplex structure are evidenced by the distribution of cross-links. Type II cross-links (**3**) form at dC<sub>14</sub> and dA<sub>16</sub> rather than at 2AP<sub>17</sub> when C4-AP is opposite dA (**1p**), even though the overall yield is more than 20%. This reactivity is consistent with an otherwise identical duplex containing dA<sub>17</sub> (**1n**, Table 2). In contrast, a low yield of Type II ICL at 2AP<sub>17</sub> is produced when dT is opposite C4-AP (**1o**) even though none form at dA<sub>17</sub> when **1o** reacts.

The formation of Type I ICLs (**2**) at 2AP<sub>17</sub> from **1p** and **1q** is most distinctive (Table 2). 2-Aminopurine is the only correctly base paired nucleotide that forms a Type I ICL at position 17. The importance of the positioning 2AP is underscored by the lack of cross-linking with this nucleotide when it is incorporated at position 15 (**1s**, Table 2). Although no cross-linking involving 2AP<sub>15</sub> (**1s**) is observed, Type I ICLs readily form with dA<sub>15</sub> (**1r**).

**2.2.3 Adenine catalysis of Type II cross-link formation in duplexes containing 2-aminopurine or iso-guanosine**—Previous C4-AP cross-linking experiments (e.g. **1b**,

**1j**, **1o**, Table 3) showed that the addition of adenine and other purines that contain exocyclic amines increased the yield of Type II ICLs when C4-AP was opposite thymidine.<sup>26</sup> The effect of adenine saturates at ~1 mM, consistent with binding of the purine by the C4-AP containing duplex.<sup>4</sup> The purines were proposed to increase the rate constant of the rate determining  $\beta$ -elimination of C4-AP to **6** (Scheme 2), which then reacts with one or more nucleotides on the opposite strand. Adenine has a similar effect on Type II ICLs (**3**) formed from duplexes containing 2AP (**1q**) or iG (**1d**, **1h**, **1m**, Table 3). In instances for which data are available from the reaction of comparable duplexes containing native nucleotides (e.g. **1o** versus **1q**, **1b** versus **1d**, and **1j** versus **1m**), the relative increases in yield of **3** upon addition of adenine are very similar. These data suggest that the formation of Type II ICLs from duplexes containing the nonnative nucleotides, 2AP and iG proceed via the same mechanism through **6** as those containing dA and dG, respectively.



## 2.2 Probing the catalysis of ICL formation via an opposing dA using 7-deaza-2'-deoxyadenosine (7dA)

The role of the dA opposing C4-AP as a catalyst for ICL formation was probed further by preparing duplexes containing 7-deaza-2'-deoxyadenosine (7dA). 7dA has a higher  $pK_a$  (5.2) than dA (3.5).<sup>35,36</sup> A greater fraction of 7dA molecules will be protonated at pH 7.2 where the reaction is carried out, suggesting that it would be a more effective acid-base catalyst. Consequently, 7dA was substituted for dA<sub>16</sub> in **1r** (**1t**). ICL formation was followed in **1t** over the course of 10 h (Figure 4). The amount of Type I ICL (**2**) reached a maximum in ~4 h and then slowly decreased. This is consistent with the reactions of other duplexes containing C4-AP, and is an indication that **2** forms reversibly.<sup>4,26</sup> Type II ICL (**3**) formed exclusively with dA<sub>15</sub> in **1t**. Its yield reached more than 70% in 10 h, which is more than twice as great as when dA is opposite C4-AP (**1r**) in an otherwise identical sequence.

Placing 7dA at position 15 (**1u**), which is where cross-linking occurs when dA is present (**1r**, Table 2), also alters ICL formation (Figure 5). The overall rate at which the starting material disappears is comparable to when 7dA is opposite C4-AP (**1t**, Figure 4). However, the distribution of ICL products is different. The Type I ICL (**2**, Figure 5) is formed more rapidly and in higher yield (at the expense of **3**) than when 7dA is opposite C4-AP. Furthermore, Type II cross-links (**3**) are distributed over dA<sub>16</sub> and dC<sub>17</sub>, as well as at 7dA<sub>15</sub> in **1u**.

## 3. Conclusions

Interstrand cross-link formation involving C4-AP occurs in cells subjected to bleomycin or ionizing radiation, and was detected at low levels in cells not exposed to oxidative stress.<sup>18</sup> Chemical studies using duplexes containing C4-AP at specific sites indicated that the formation of cross-links with the surrounding nucleotides was dependent upon the local sequence. However, reaction with dG did not occur even though this nucleotide reacts with a number of diffusible electrophiles. This surprising lack of reactivity with dG was investigated using the nonnative nucleotides *iso*-guanosine (iG) and 2-aminopurine (2AP). The data indicate that the lack of dG cross-linking to C4-AP is attributable to two factors 1) poor positioning of its N2-amino group when it is opposite the nucleotide bonded to the 3'-phosphate of C4-AP (dG<sub>15</sub>), and 2) overall reduced reactivity compared to dA and 2AP. The latter is at least partially due to stronger base pairing with dC compared to dA (and 2AP) with dT, as evidenced by the increased cross-linking when dG or iG are mispaired.

Experiments using duplexes containing 7-deaza-2'-deoxyadenosine (7dA) reinforce the hypothesis that the neighboring nucleotides affect ICL formation by catalyzing the elimination of C4-AP. How the nucleotide opposite C4-AP catalyzes ICL formation is uncertain. It may function as an acid or base catalyst, or a nucleophilic one via Schiff-base formation. However, it is unlikely that the opposing nucleotide functions as a catalyst independent of the local sequence. 2'-Deoxycytidine ( $pK_a = 4.2$ ) induces Type II ICL (**3**) formation less efficiently than dA ( $pK_a = 3.5$ ), suggesting that other properties, such as base stacking may influence reactivity.<sup>36</sup>

## 4. Experimental

### General Methods

Oligonucleotides were synthesized on an Applied Biosystems Incorporated 394 oligonucleotide synthesizer. Commercially available oligonucleotide synthesis reagents, including the phosphoramidite for 7dA, iG, iC, and 2AP were purchased from Glen Research (Sterling, VA). T4 polynucleotide kinase and terminal deoxynucleotide transferase were obtained from New England Biolabs.  $\gamma$ -<sup>32</sup>P-ATP and  $\alpha$ -<sup>32</sup>P-cordycepin 5'-triphosphate were purchased from Perkin Elmer. C18-Sep-Pak cartridges were obtained from Waters. Quantification of radiolabeled oligonucleotides was carried out using a Molecular Dynamics Phosphorimager 840 equipped with ImageQuant TL software. All photolyses of oligonucleotides were carried out in clear eppendorf tubes in a Rayonet photoreactor (RPR-100) fitted with 16 lamps having an output maximum at 350 nm. Syntheses of DNA substrates containing **5** were carried out as previously reported.<sup>4</sup>

### General procedure for ICL formation in C4-AP containing duplexes

DNA (200 nM) in PBS buffer (100 mM NaCl and 10 mM sodium phosphate, pH 7.2) was photolyzed for 30 min at room temperature. One part DNA was then diluted to 100 nM with one part PBS buffer or a 200 mM solution of adenine in PBS buffer. These reactions were typically incubated at 37 °C for 10 h. For time course experiments, a 3  $\mu$ L aliquot was removed every 2 h. Aliquots were mixed with an equal volume of formamide loading buffer and the DNA was resolved using 20% denaturing polyacrylamide gel electrophoresis (PAGE).

### Determination of ICL location using the hydroxyl-radical cleavage reaction

Duplex DNA constructed of a C4-AP containing strand and a 3'-labeled complement was incubated at 37°C overnight in PBS. Note: low yielding ICL reactions were incubated for as long as three days. The reaction was concentrated and purified by 20% analytical (0.4 mm thickness) PAGE. The bands corresponding to the ICL products were carefully cut from the gel, crushed, and eluted in 500  $\mu$ L of water for 1 h at room temperature with periodic vortexing. In some cases, Type II ICLs (**3**) were evident as multiple resolvable bands and were isolated independently. The solution was filtered using a 10 mL Polyprep Column (Biorad) and desalted using a C<sub>18</sub> column (100 mg Sep-Pak). After concentration, a portion of the DNA (30,000–60,000 cpm) was dissolved in 8  $\mu$ L of water and mixed with 10  $\mu$ L of the 2  $\times$  oxidation buffer (20 mM NaCl, 20 mM sodium phosphate, pH 7.2, 2 mM sodium ascorbate, and 1 mM hydrogen peroxide). To initiate the reaction, 2  $\mu$ L of the Fe•EDTA solution (1 mM EDTA and 0.5 mM Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>•6H<sub>2</sub>O) was added, and the reaction was allowed to proceed for no longer than 3 min. The reaction was quenched with 10  $\mu$ L of 100 mM thiourea and concentrated. The DNA was then suspended in formamide loading buffer and separated by 20% PAGE.

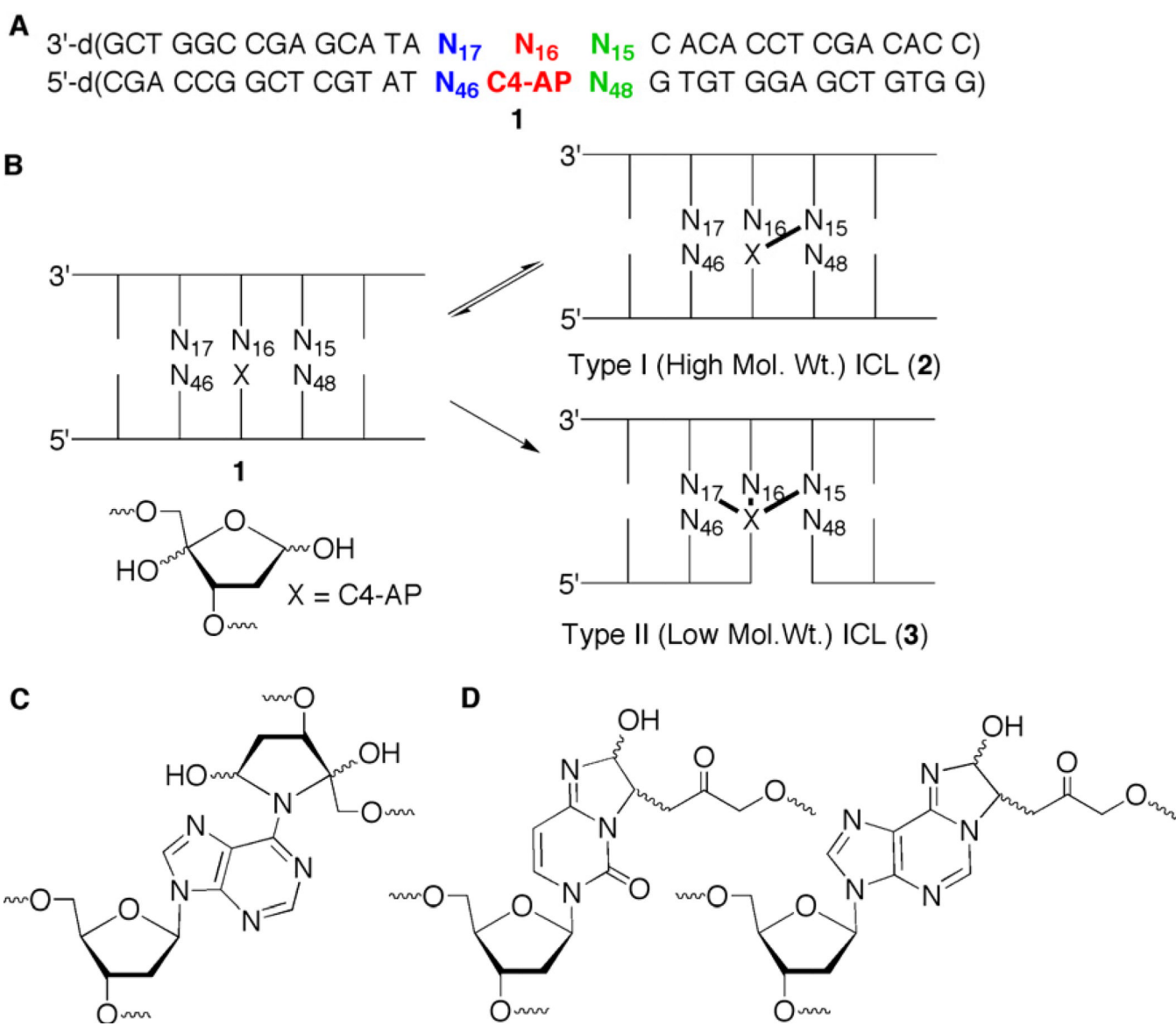
## Acknowledgments

We are grateful for support of this research from the National Institute of General Medical Sciences (GM-063028).

## References and notes

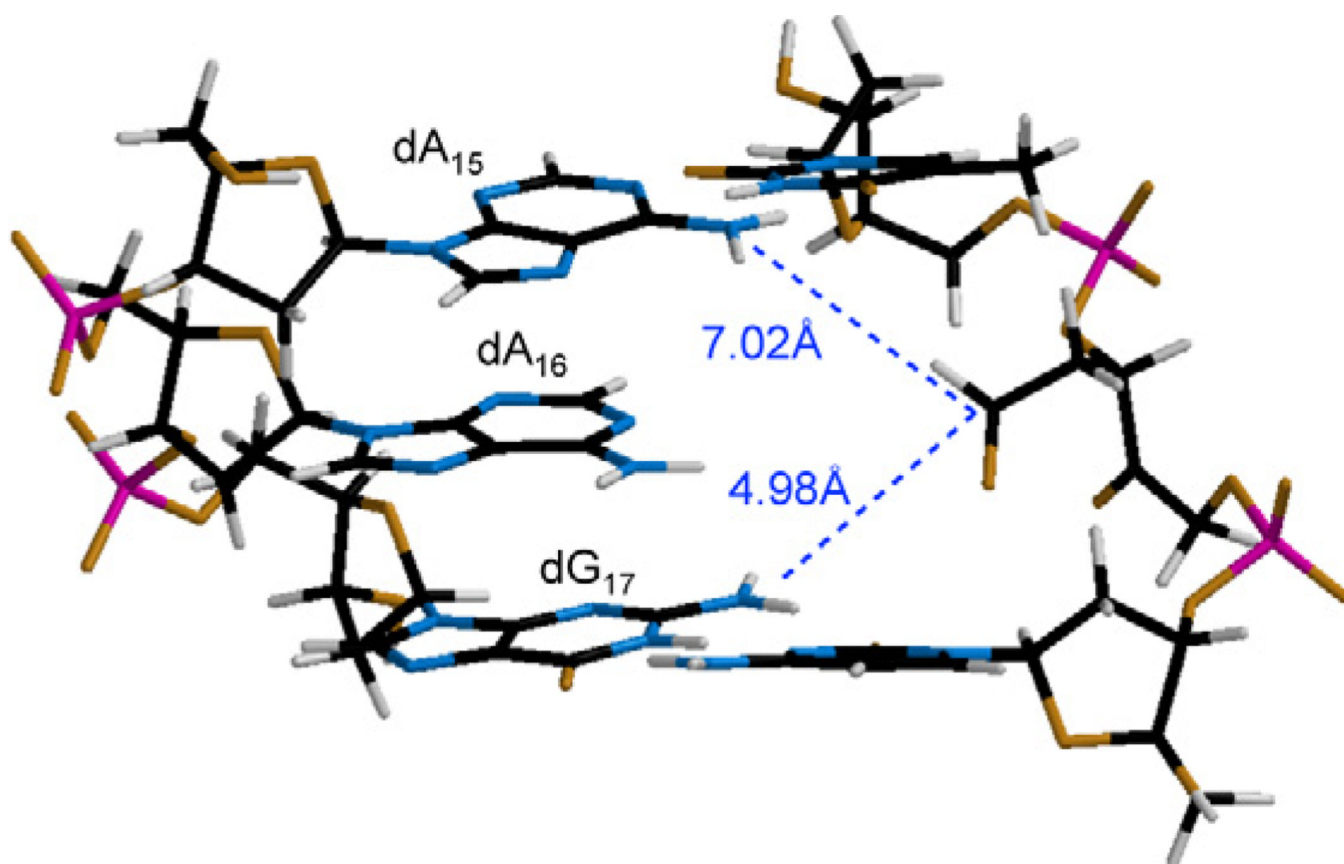
1. Gates KS. *Chem. Res. Toxicol.* 2009; 22:1747–1760. [PubMed: 19757819]
2. Chatgililoglu C. *Wiley Ser. React. Intermed. Chem. Biol.* 2009; 2:99–133.
3. Dutta S, Chowdhury G, Gates KS. *J. Am. Chem. Soc.* 2007; 129:1852–1853. [PubMed: 17253689]
4. Sczepanski JT, Jacobs AC, Greenberg MM. *J. Am. Chem. Soc.* 2008; 130:9646–9647. [PubMed: 18593126]
5. Guan L, Greenberg MM. *J. Am. Chem. Soc.* 2009; 131:15225–15231. [PubMed: 19807122]
6. Sczepanski JT, Wong RS, McKnight JN, Bowman GD, Greenberg MM. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:22475–22480. [PubMed: 21149689]
7. Kumari A, Minko IG, Harbut MB, Finkel SE, Goodman MF, Lloyd RS. *J. Biol. Chem.* 2008; 283:27433–27437. [PubMed: 18697749]
8. Kumaresan KR, Sridharan DM, McMahan LW, Lambert MW. *Biochemistry.* 2007; 46:14359–14368. [PubMed: 18020456]
9. Smeaton MB, Hlavin EM, McGregor Mason T, Noronha AM, Wilds CJ, Miller PS. *Biochemistry.* 2008; 47:9920–9930. [PubMed: 18702509]
10. Räschle M, Knipsheer P, Enoiu M, Angelov T, Sun J, Griffith JD, Ellenberger TE, Schärer OD, Walter JC. *Cell.* 2008; 134:969–980. [PubMed: 18805090]
11. Knipscheer P, Raschle M, Smogorzewska A, Enoiu M, Ho TV, Scharer OD, Elledge SJ, Walter JC. *Science.* 2009; 326:1698–1701. [PubMed: 19965384]
12. Panasci L, Xu ZY, Bello V, Aloyz R. *Anticancer Drugs.* 2002; 13:211–220. [PubMed: 11984064]
13. Zietlow L, Smith LA, Bessho M, Bessho T. *Biochemistry.* 2009; 48:11817–11824. [PubMed: 19908865]
14. Kennedy DR, Lu J, Shen B, Beerman TA. *Proc. Nat. Acad. Sci. USA.* 2007; 104:17632–17637. [PubMed: 17978180]
15. Kothandapani A, Dangeti VSMN, Brown AR, Banze LA, Wang X-H, Sobol RW, Patrick SM. *J. Biol. Chem.* 2011; 286:14564–14574. [PubMed: 21357694]
16. Kozekov ID, Turesky RJ, Alas GR, Harris CM, Harris TM, Rizzo CJ. *Chem. Res. Toxicol.* 2010; 23:1701–1713. [PubMed: 20964440]
17. Stone MP, Cho Y-J, Huang H, Kim H-Y, Kozekov ID, Kozekova A, Wang H, Minko IG, Lloyd RS, Harris TM, Rizzo CJ. *Acc. Chem. Res.* 2008; 41:793–804. [PubMed: 18500830]
18. Regulus P, Duroux B, Bayle P-A, Favier A, Cadet J, Ravanat J-L. *Proc. Nat. Acad. Sci. USA.* 2007; 104:14032–14037. [PubMed: 17715301]
19. Sugiyama H, Xu C, Murugesan N, Hecht SM. *J. Am. Chem. Soc.* 1985; 107:4104–4105.
20. Rabow LE, Stubbe J, Kozarich JW. *J. Am. Chem. Soc.* 1990; 112:3196–3203.
21. Xi, Z.; Goldberg, IH. *Comprehensive Natural Products Chemistry.* Kool, ET., editor. Vol. Vol. 7. Amsterdam: Elsevier; 1999. p. 553-592.
22. von Sonntag, C. *The Chemical Basis of Radiation Biology.* London: Taylor & Francis; 1987.
23. Kroeger KM, Kim J, Goodman MF, Greenberg MM. *Biochemistry.* 2004; 43:13621–13627. [PubMed: 15504024]
24. Guan L, Greenberg MM. *J. Am. Chem. Soc.* 2010; 132:5004–5005. [PubMed: 20334373]
25. Guan L-R, Bebenek K, Kunkel TA, Greenberg MM. *Biochemistry.* 2010; 49:9904–9910. [PubMed: 20961055]
26. Sczepanski JT, Jacobs AC, Majumdar A, Greenberg MM. *J. Am. Chem. Soc.* 2009; 131:11132–11139. [PubMed: 19722676]
27. Sczepanski JT, Jacobs AC, Van Houten B, Greenberg MM. *Biochemistry.* 2009; 48:7565–7567. [PubMed: 19606890]
28. Kim J, Gil JM, Greenberg MM. *Angew. Chem. Int. Ed.* 2003; 42:5882–5885.

29. Kim J, Weledji YN, Greenberg MM. *J. Org. Chem.* 2004; 69:6100–6104. [PubMed: 15373495]
30. Chen J, Stubbe J. *Biochemistry.* 2004; 43:5278–5286. [PubMed: 15122893]
31. Aso M, Usui K, Fukuda M, Kakihara Y, Goromau T, Suemune H. *Org. Lett.* 2006; 8:3183–3186. [PubMed: 16836361]
32. Chen B, Bohnert T, Zhou X, Dedon PC. *Chem. Res. Toxicol.* 2004; 17:1406–1413. [PubMed: 15540938]
33. Op de Beeck M, Madder A. *J. Am. Chem. Soc.* 2011; 133:796–807. [PubMed: 21162525]
34. Stevens K, Claeys DD, Catak S, Figaroli S, Hocek M, Tromp JM, Schürch S, Van Speybroeck V, Madder A. *Chem. - Eur. J.* 2011; 17:6940–6953.
35. Suydam T, Levandoski SD, Strobel SA. *Biochemistry.* 2010; 49:3723–3732. [PubMed: 20373826]
36. Saenger, W. *Principles of Nucleic Acid Structure.* New York: Springer Verlag; 1983.

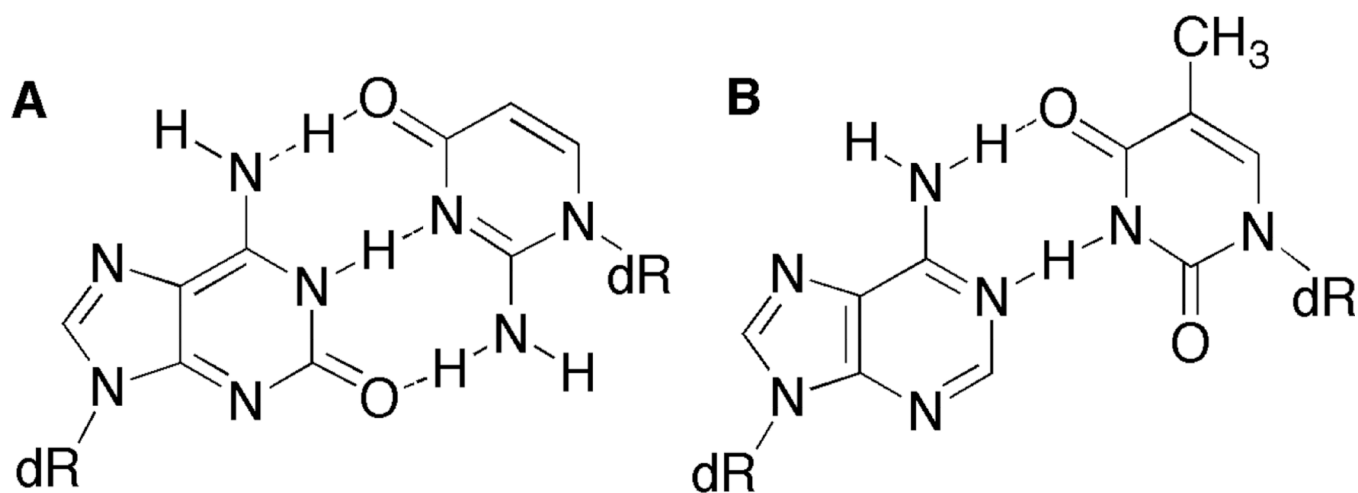


**Figure 1.** DNA interstrand cross-link formation from C4-AP. A) General sequence (1) used in cross-linking experiments. B) Cartoon illustrating the formation of two types of cross-links from C4-AP. C) Detailed structure of the Type I ICL formed between C4-AP and dA. D) Detailed structure of the Type II ICL formed between C4-AP and dC or dA.

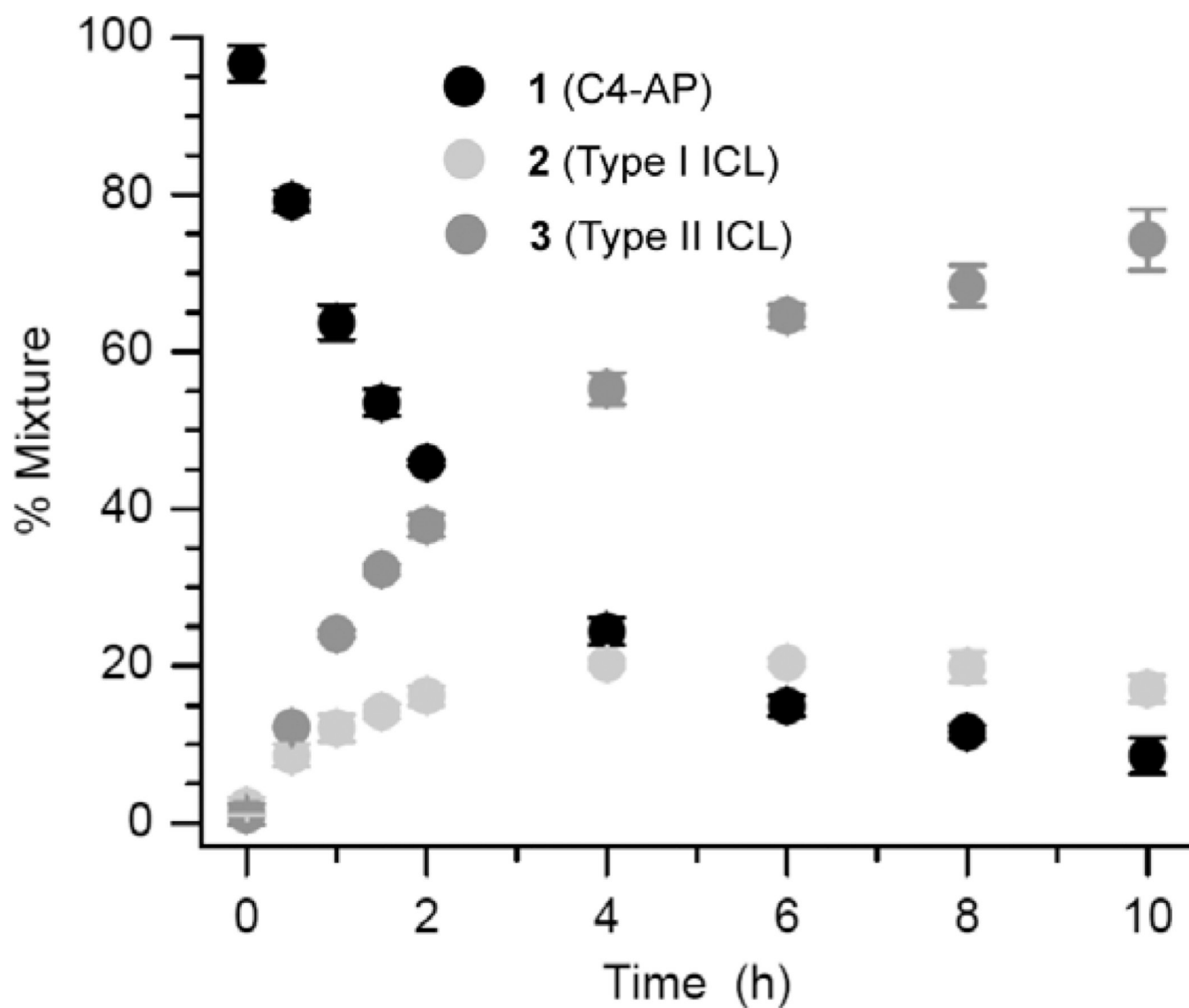




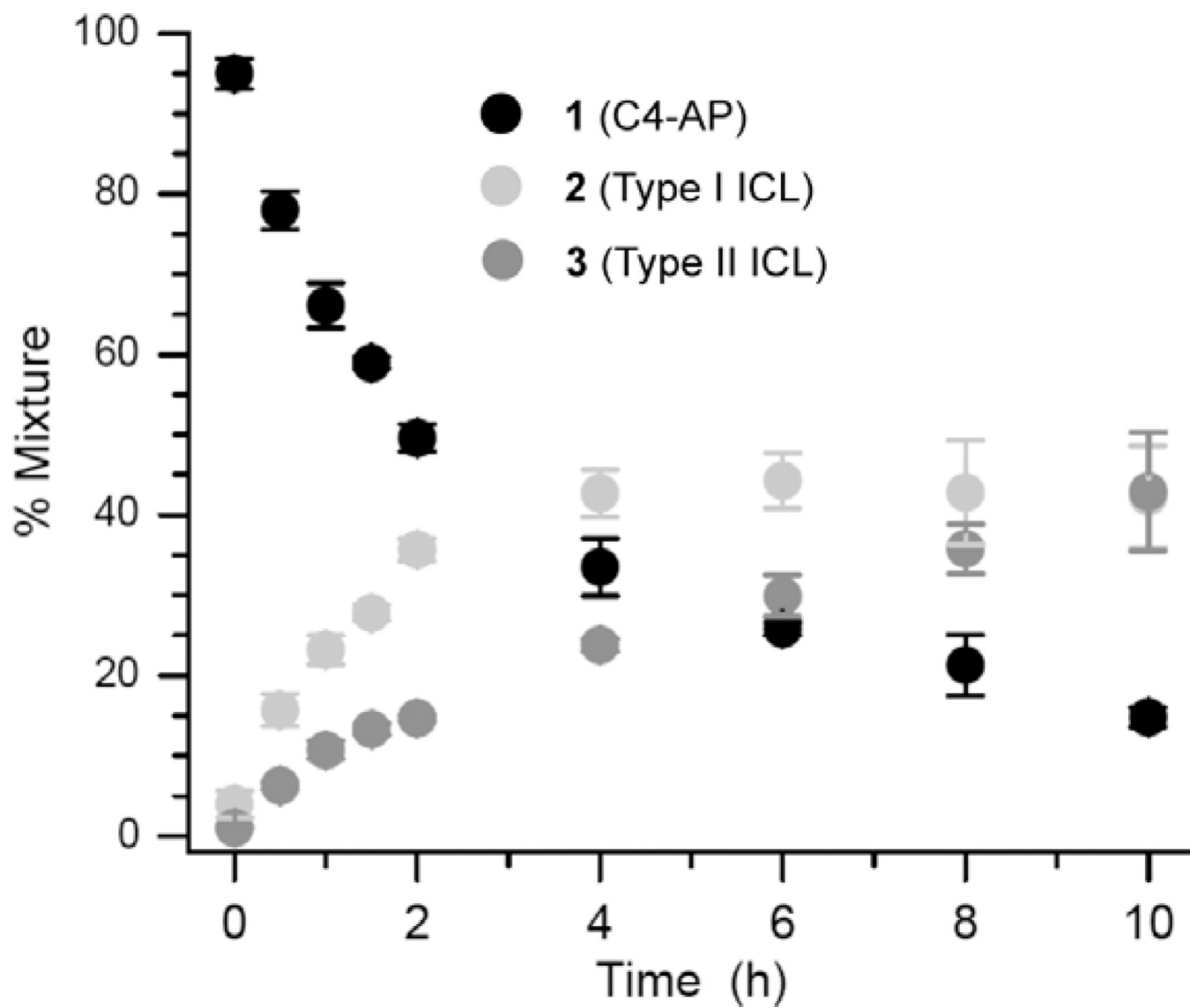
**Figure 2.** Molecular model of C4-AP showing proximity of potential nucleophiles created using Spartan<sup>®</sup>.



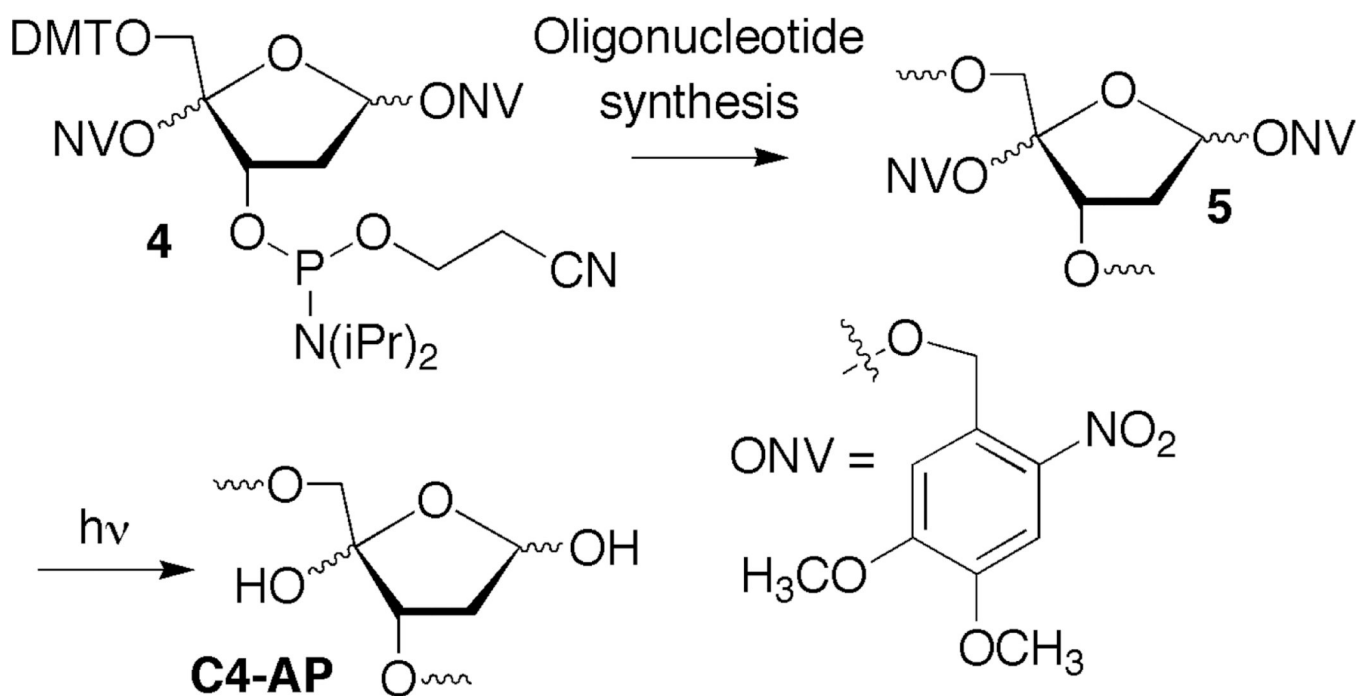
**Figure 3.** Base pairing of A) *iso*-guanosine with *iso*-cytosine and B) dA with dT showing similar positioning of exocyclic amines on *iso*-guanosine and dA in the major groove.



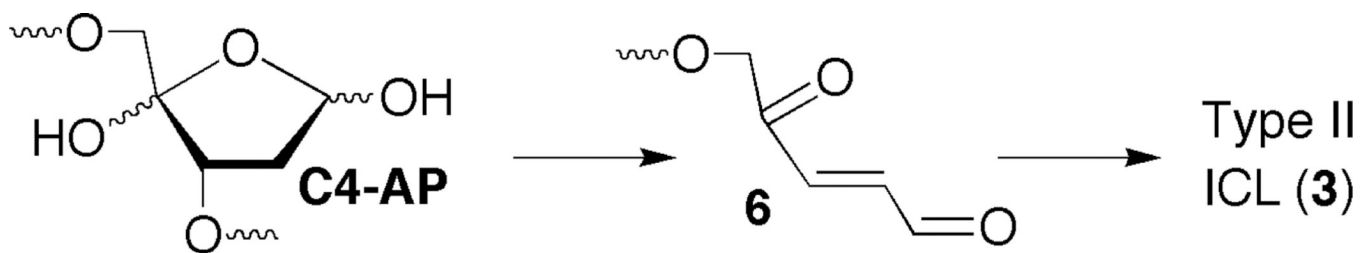
**Figure 4.**  
The effect of 7-deaza-2'-deoxyadenosine (7dA) on cross-link formation when it is opposite C4-AP (**1t**).



**Figure 5.** The effect of 7-deaza-2'-deoxyadenosine (7dA) on cross-link formation involving C4-AP when it is opposite the 3'-adjacent thymidine (**1u**).



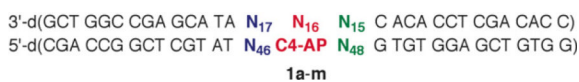
**Scheme 1.**  
Generation of C4-AP in synthetic oligonucleotides.



**Scheme 2.**  
Type II ICL (3) formation via  $\beta$ -elimination from C4-AP.<sup>26</sup>

**Table 1**

The effect of *iso*-guanosine and dG:dT mispairing on Type II on (3) interstrand cross-link formation.



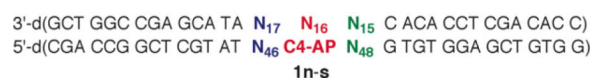
Substrate		%Yield 3 <sup>a</sup>	ICL Position
	<b>N<sub>46</sub> C4-AP N<sub>48</sub></b>		
	<b>N<sub>17</sub> N<sub>16</sub> N<sub>15</sub></b>		
<b>1a<sup>b</sup></b>	G C4-AP C C A G	5.1 ± 0.4	A <sub>16</sub> , C <sub>17</sub> , (0.9:1)
<b>1b<sup>b</sup></b>	G C4-AP C C T G	0.7 ± 0.1	C <sub>17</sub>
<b>1c</b>	G C4-AP iC C A iG	7.2 ± 1.1	iG <sub>15</sub>
<b>1d</b>	G C4-AP iC C T iG	2.0 ± 0.1	C <sub>17</sub> , iG <sub>15</sub> , (2.3:1)
<b>1e<sup>b</sup></b>	G C4-AP T C A G	28.8 ± 0.8	A <sub>16</sub> , C <sub>17</sub> , (1.3:1)
<b>1f<sup>b</sup></b>	G C4-AP T C T G	1.6 ± 0.3	C <sub>17</sub>
<b>1g</b>	G C4-AP T C A iG	18.4 ± 0.5	iG <sub>15</sub>
<b>1h</b>	G C4-AP T C T iG	3.7 ± 0.3	iG <sub>15</sub>
<b>1i<sup>b</sup></b>	C C4-AP A G A T	49.3 ± 1.6	A <sub>16</sub>
<b>1j<sup>b</sup></b>	C C4-AP A G T T	0.3 ± 0.1	Not determined
<b>1k</b>	T C4-AP A G A T	36.5 ± 1.5	A <sub>16</sub> , C <sub>14</sub> (1.1)
<b>1l</b>	iC C4-AP A iG A T	33.2 ± 2.1	A <sub>16</sub> , C <sub>14</sub> , (1.3:1)
<b>1m</b>	iC C4-AP A iG T T	0.9 ± 0.1	C <sub>14</sub>

<sup>a</sup>Yields are an average of at least 3 replicates.

<sup>b</sup>Data take from ref. 26.

**Table 2**

The effect of 2-aminopurine on interstrand cross-link formation.



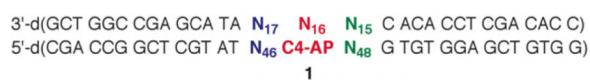
Substrate	<b>N<sub>46</sub></b> <b>C4-AP</b> <b>N<sub>48</sub></b> <b>N<sub>17</sub></b> <b>N<sub>16</sub></b> <b>N<sub>15</sub></b>	% Yield 2 (Pos.) <sup>a</sup>	% Yield 3 (Pos.) <sup>a</sup>
<b>1n<sup>b</sup></b>	T C4-AP A A A T	Not observed	21.0 ± 0.6 (A <sub>16</sub> )
<b>1o<sup>b</sup></b>	T C4-AP A A T T	Not observed	0.7 ± 0.1 (ND)
<b>1p</b>	T C4-AP A 2AP A T	14.8 ± 1.2 (2AP <sub>17</sub> )	20.5 ± 0.2 (A <sub>16</sub> , C <sub>14</sub> , 1.2:1)
<b>1q</b>	T C4-AP A 2AP T T	2.7 ± 0.4 (2AP <sub>17</sub> )	1.3 ± 0.2 (2AP <sub>17</sub> )
<b>1r<sup>b</sup></b>	G C4-AP T C A A	21.3 ± 0.9 (A <sub>15</sub> )	7.3 ± 1.0 (A <sub>15</sub> )
<b>1s</b>	G C4-AP T C A 2AP	3.9 ± 0.1 (ND)	18.8 ± 1.1 (C <sub>17</sub> , C <sub>14</sub> , 1:1)


<sup>a</sup>Yields are an average of at least 3 replicates.<sup>b</sup>Data take from ref. 26.



**Table 3**

The effect of adenine on Type II (3) interstrand cross-link formation.



Substrate		% Yield 3 <sup>a</sup>	
		No Adenine	0.1 mM Adenine
1o <sup>b</sup>	T C4-AP A A T T	0.7 ± 0.1	4.2 ± 0.8
1q	T C4-AP A 2AP T T	1.3 ± 0.2	5.2 ± 0.8
1b <sup>b</sup>	G C4-AP C C T G	0.7 ± 0.1	12.2 ± 0.8
1d	G C4-AP iC C T iG	2.0 ± 0.1	32.2 ± 1.8
1j <sup>b</sup>	C C4-AP A G T T	0.3 ± 0.1	3.3 ± 0.4
1m	iC C4-AP A iG T T	0.9 ± 0.1	9.1 ± 0.1
1h	G C4-AP T C T iG	3.7 ± 0.3	14.1 ± 1.0

<sup>a</sup>Yields are an average of at least 3 replicates.<sup>b</sup>Data take from ref. 26.