

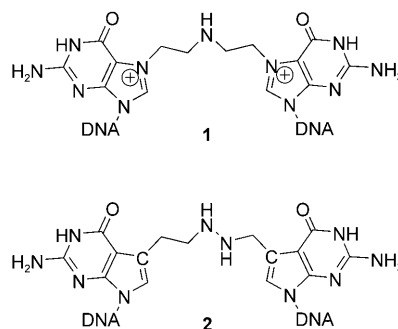
## Synthesis and Molecular Modeling of a Nitrogen Mustard DNA Interstrand Crosslink

Angelo Guainazzi,<sup>[a]</sup> Arthur J. Campbell,<sup>[b]</sup> Todor Angelov,<sup>[c]</sup> Carlos Simmerling,<sup>\*[b]</sup> and Orlando D. Schärer<sup>\*[a]</sup>

*Dedicated to Prof. Francis Johnson on the occasion of his 80th birthday*

DNA interstrand crosslinks (ICLs) are formed by bifunctional agents with the ability to covalently link two strands of duplex DNA. ICLs are extremely cytotoxic, since they block essential processes such as DNA replication and transcription. Based on these properties, agents such as nitrogen mustards (NM) or cisplatin are widely used in cancer chemotherapy.<sup>[1]</sup> ICLs induce a number of biological responses that counteract the therapeutic effects of crosslinking agents. However, the elucidation of the mechanism by which these responses cause resistance in tumor cells has been hampered by the lack of efficient methods to generate ICLs formed by antitumor agents. Current approaches toward the synthesis of site-specific ICLs have yielded mostly models of the clinically relevant lesions or ICLs formed by endogenous agents.<sup>[2]</sup> A number of recent studies have shown that ICLs with different structures are processed in distinct ways (see reference [3] for some recent reviews), emphasizing the need to generate substrates containing the ICLs formed by the clinically important drugs.

Here we report the synthesis of a stable NM ICL isostere (**2**) and use atomic detail simulations to show that it recapitulates the structural and dynamic properties of the native NM ICL **1**. To obtain the NM ICL isostere we used a strategy that we recently developed for the synthesis of major groove ICLs.<sup>[4]</sup>



This approach involves a double reductive amination reaction of two acetaldehyde functionalities (**5**) linked to the 7-position of G residues on complementary strands of dsDNA. One dG residue was substituted with 7-deaza-2'-deoxyguanosine to counteract the inherent lability of the glycosidic bond in N7-alkylated guanines.<sup>[5]</sup> This approach allowed the generation of a six-atom ICL by coupling two acetaldehyde groups with hydrazine. However, we were unsuccessful in generating ICL **3** with the five-atom bridge found in NM ICLs<sup>[6]</sup> (Scheme 1), possibly because the reductive amination with ammonia was not powerful enough to introduce the strain in the DNA caused by the  $\approx 7.5$  Å bridge of the NM.<sup>[4,7]</sup>

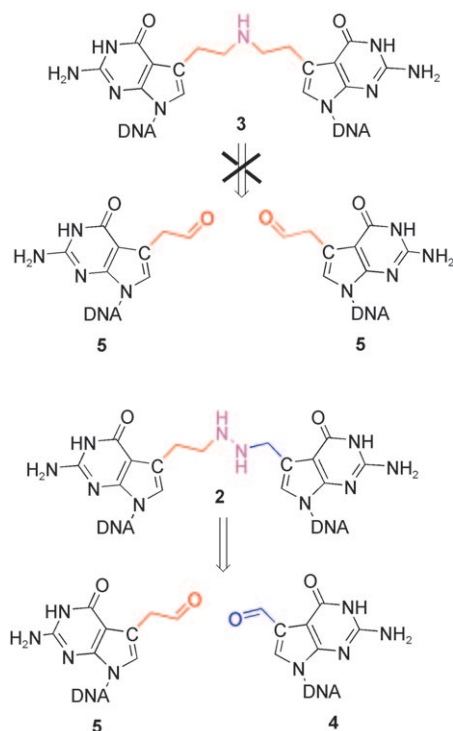
We reasoned that an ICL isosteric to those formed by NM might be formed by reaction of hydrazine with an acetaldehyde (**5**) and a formyl aldehyde derivative (**4**) of deaza-

[a] A. Guainazzi, Prof. O. D. Schärer  
Department of Pharmacological Sciences and Chemistry  
Stony Brook University, Stony Brook  
NY 11794-3400 (USA)  
Fax: (+1) 631-632-7546  
E-mail: orlando@pharm.stonybrook.edu

[b] A. J. Campbell, Prof. C. Simmerling  
Department of Chemistry and Center for Structural Biology  
Stony Brook University, Stony Brook  
NY 11794-3400 (USA)  
Fax: (+1) 631-632-1555  
E-mail: carlos.simmerling@gmail.com

[c] Dr. T. Angelov  
Institute for Molecular Cancer Research  
University of Zürich, 8057 Zürich (Switzerland)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201002041>.



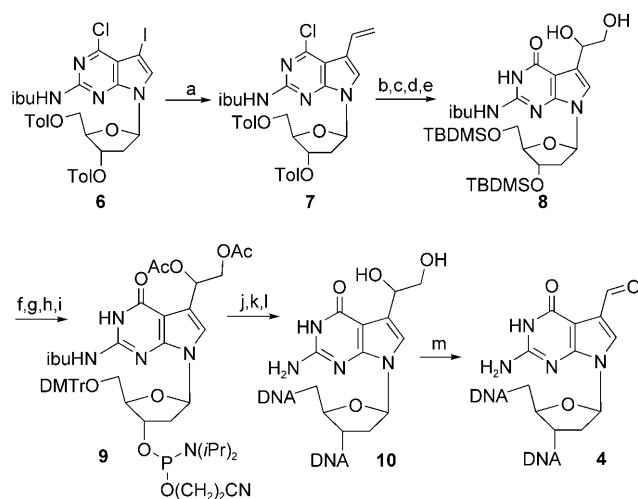
Scheme 1. Formation of NM ICLs by reductive amination was not successful using two acetaldehyde precursors **5** and  $\text{NH}_4\text{Cl}$ , prompting us to explore the generation of isostere **2** by linking precursors **4** and **5** with hydrazine.

guanine, exploiting the higher reactivity of hydrazine over ammonia (Scheme 1).<sup>[4]</sup>

We synthesized formyl aldehyde **4** in an analogous fashion to the previously reported synthesis of acetaldehyde **5**, masking the aldehyde as a protected diol during solid-phase DNA synthesis.<sup>[4,8]</sup> The synthesis started with vinylation of a protected 7-iodo-7-deazaguanine derivative (**6**) using a Stille coupling reaction (Scheme 2) to give compound **7**. Oxidation of the allyl group to the diol (**8**), protection, and functionalization yielded phosphoramidite **9**. The aldehyde precursors were incorporated into complementary 20-mer oligonucleotides in a 5'-d(GNC) sequence (the preferred sequence context for NM ICL formation<sup>[6]</sup>) by solid-phase synthesis, and the oligonucleotides were deprotected and annealed.

Following oxidation of the diols using periodate, the two aldehydes were coupled by double reductive amination with hydrazine and  $\text{NaBH}_3\text{CN}$ .<sup>[4]</sup> The reaction along with appropriate controls was analyzed by denaturing PAGE.

As already discussed,<sup>[4]</sup> ICL formation from two acetaldehyde precursors was successful with hydrazine, but not with ammonium acetate (Figure 1, lanes 2 and 3). However, reaction of a duplex containing **4** and **5** with hydrazine led to formation of the desired five-atom ICL **2**, evidenced by a band with the same mobility as the previously analyzed crosslink (Figure 1, lanes 3 and 4). Although the yield of the five-atom ICL was lower than that of the six-atom ICL



Scheme 2. Synthesis of the formyl aldehyde **4**: a) vinyl- $\text{Sn}(\text{Bu})_3$ ,  $\text{Pd}[\text{P}(\text{Ph})_3]_4$ , toluene,  $90^\circ\text{C}$ , 43 h, 70%; b) pyridine-2-carboxaldehyde,  $N,N,N',N'$ -tetramethylguanidine, dioxane, DMF, RT, 42 h, 88%; c)  $\text{NaOMe}$ , THF, RT, 5 h, 97%; d) TBDMS-Cl, imidazole, DMF, RT, 16 h, 93%; e)  $\text{OsO}_4$ , NMMO, THF,  $0^\circ\text{C}$ , 3.5 h, 63%; f)  $\text{Ac}_2\text{O}$ , pyridine, RT, 1 h, 82%; g) TBAF, AcOH, THF, RT, 20 h, 79%; h) DMTr-Cl, pyridine, RT, 1 h, 74%; i)  $i\text{Pr}_2\text{NP}(\text{Cl})\text{OC}_2\text{H}_4\text{CN}$ , DIEA,  $\text{CH}_2\text{Cl}_2$ , RT, 1 h, 72%; j) oligonucleotide synthesis; k) 33%  $\text{NH}_4\text{OH}$ ,  $50^\circ\text{C}$ , 12 h; l) annealing; m)  $\text{NaIO}_4$ , RT, 12 h. TBDMS-Cl = *tert*-butyldimethylsilyl chloride; NMMO = *N*-methylmorpholine oxide; TBAF = tetra-*n*-butylammonium fluoride; DMTr-Cl = 4,4'-dimethoxytrityl chloride; DIEA = *N,N*-diisopropylethylamine; ibu = isobutyryl.

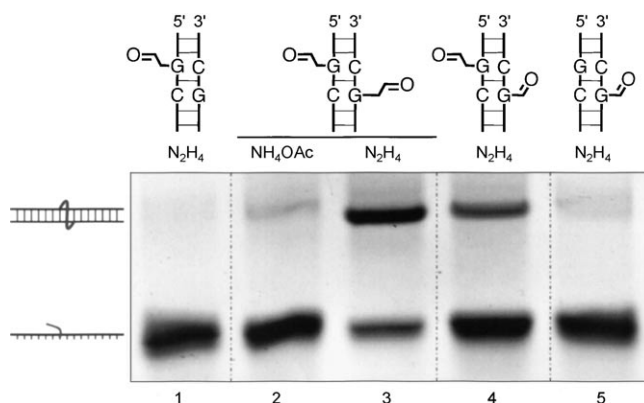


Figure 1. Analysis of ICL formation by denaturing PAGE with methylene blue staining. The duplexes, amine used and position of single-stranded or ICL-containing DNA are indicated. The sequences used were 5'-d(GTCACTGGTAXACAGCATTG) and 5'-d(CAATGCTXTCTACAGTGAC) where X represents the modified G. The small amount of bands running as duplexes in lanes 2 and 5 are likely due to residual amount of duplex that was not denatured during electrophoresis.

( $\approx 25\%$  vs.  $\approx 75\%$ ), we were able to isolate the product by gel purification and electroelution in amounts exceeding 100 nmol. For simplicity of analysis we also synthesized ICL **2** in an 11-mer duplex and unambiguously identified it as the desired product by ESI-MS ( $m/z$  calcd: 6741.23, found: 6741.6, Figure S1 in the Supporting Information). If either aldehyde precursor was present only on one strand of the

duplex, no significant amounts of the slower migrating species were observed, further demonstrating the specificity of ICL formation (Figure 1, lanes 1 and 5).

Molecular modeling was used to validate ICL **2** as a model for NM **1** and to compare the structural consequences of the two ICLs with the uncrosslinked control (**C**) in identical 11-mer sequences. We used the Amber simulation package with atomic detail and explicit water (see Supporting Information) to validate that **2** gives similar amounts of distortion when compared to the NM ICL **1** and consistent differences when both are compared to **C**.<sup>[9]</sup> Two independent simulations of 50 ns were run for each of the **1**, **2**, and **C** systems. The duplex was stable throughout all six simulations.

The N7 to N7 distances of the NM ICL **1** overlapped well with the C7 to C7 distances of **2**; both are restricted as compared to **C**, which samples a broader range (Figure 2, top).

The decreased distance between the two crosslinked bases has a direct effect on the local distortion around the ICL, and is also influenced by the sequence (Figure 2 bottom, and Table S1 in the Supporting Information). One measure of distortion is to examine the buckling and propeller

twist<sup>[11]</sup> of the two crosslinked base pairs (Table S1 and Figure S2 for sequence, in the Supporting Information). Since molecular mechanics force fields have limited accuracy, it is more important to note the trends rather than the specific values. It is well accepted that differences are more reliable in simulations than absolute values, due to cancellation of systematic error. The crosslinked guanine 16 that neighbors the large purine ring of adenine (residue 17) accommodates the crosslink by buckling of the 7:16 base pair (**C**:  $-3.5 \pm 0.2^\circ$ , **1**:  $-13.2 \pm 0.1^\circ$ , and **2**:  $-21.8 \pm 0.6^\circ$ ). The crosslinked guanine 5 that neighbors the small pyrimidine ring of thymine 6 has more room to move and accommodates the ICL by an increased propeller twist of the 5:18 base pair (**C**:  $-13.2 \pm 0.1^\circ$ , **1**:  $-22.3 \pm 0.2^\circ$ , and **2**:  $-26.9 \pm 0.8^\circ$ ). The addition of a covalent bridge between both strands of DNA nearly doubles residue to residue correlation in the central region of the duplex. Residues 5 and 16 have a correlation coefficient ( $r$ ) of  $0.29 \pm 0.04$ ,  $0.62 \pm 0.01$ , and  $0.58 \pm 0.02$  for **C**, **1**, and **2** respectively (Figure S2 in the Supporting Information). ICLs **1** and **2** have a substantial decrease in flexibility (Figure 2, top) when compared to the uncrosslinked reference CNT, resulting in the formation of a slight kink in the duplex of both **1** and **2** to accommodate the ICL (Figure 2, bottom).

To test for local duplex bending at the crosslink site we measured an angle that would encompass the tightening of the 7- to 7'-position of residues 5 and 16 (Figure 2, top). The smaller angles sampled by **1** and **2** suggest that the analogue slightly bends the duplex outside the crosslink in the same manner that the NM ICL simulations do (Figure 2, bottom). The slight bending of the central region of the ICL simulations of **1** and **2** compared to **C** is qualitatively similar to past experimental work.<sup>[7]</sup>

Circular dichroism (CD) spectra were recorded to gain experimental insight into to what extent the NM ICLs containing oligonucleotide **2** deviates from B-form DNA. The CD spectra of **2** and an unmodified duplex of the same sequence displayed the characteristic features of B-form DNA (Figure S4 in the Supporting Information), consistent with our molecular dynamics simulations, indicating the NM ICL induces only a minor bend in the DNA. It has been shown that only more dramatic distortions, such as the ones induced by cisplatin ICL result in significant changes in CD spectra.<sup>[2], [12]</sup>

In summary, we describe the synthesis of the stable NM ICL analogue **2** by post-synthetic double reductive amination. Although this ICL has three atoms substituted with respect to the NM ICL **1** (the two N at the 7-positions of dG by C and one C by N in the hydrazine-formed bridge), our molecular dynamics simulations show that the two ICLs affect DNA structure and motion in equivalent ways. The availability of stable, site-specific NM ICLs will enable studies of the structural consequences and biological responses induced by NM ICLs,<sup>[13]</sup> more than sixty years after NMs were the first agents to be used in cancer chemotherapy.<sup>[14]</sup> Such studies will provide new insights into how tumors become resistant to treatment by crosslinking agents.

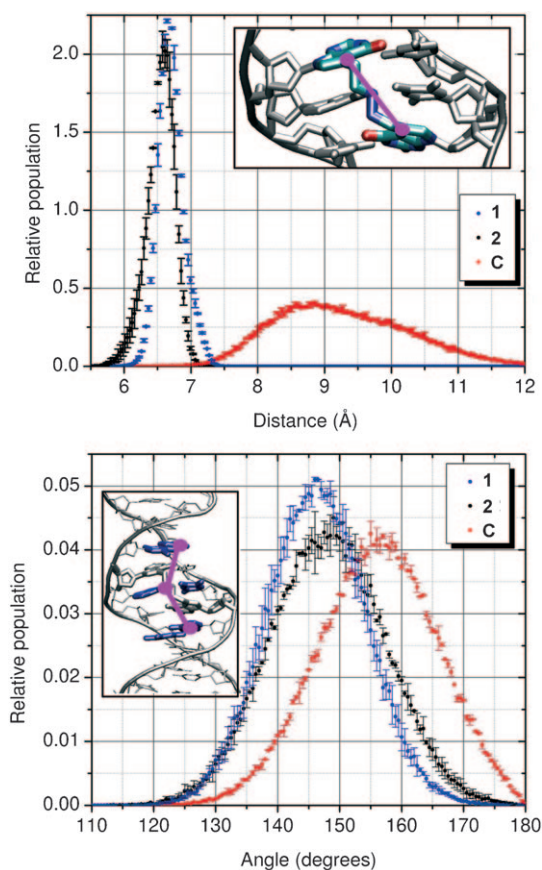


Figure 2. Data from MD simulations of **2** (black), **1** (blue) and **C** (red). Top: Distance measurement marked in pink, C/N7 to C/N7 atoms between residues 5 and 16. Bottom: Center of mass angle measurement, in which each point is defined by the heavy atoms of the base pair (highlighted in blue in inset picture): point 1 represents 4A and 19T, point 2 represents 6T and 17A and point 3 represents 8A and 15T.<sup>[10]</sup>

## Experimental Section

**Crosslink formation:** A solution of the single-strand oligonucleotides (25 nmol) in NaCl (125  $\mu$ L, 10 mM) was heated to 95°C and allowed to cool to room temperature over a period of 4 h to allow for duplex formation. After addition of sodium phosphate buffer (15  $\mu$ L, 1 M, pH 5.4) and NaIO<sub>4</sub> (10  $\mu$ L, 50 mM) the reaction mixture was kept in the dark overnight at 4°C. Excess NaIO<sub>4</sub> was removed by centrifugation through Microcon columns with a 3 K cutoff (Millipore). Then aqueous hydrazine (10  $\mu$ L, 5 mM) and NaCNBH<sub>3</sub> (10  $\mu$ L, 0.5 M) were added and the reaction mixture was left overnight at room temperature in the dark. ICL formation was assessed by electrophoresis on a denaturing 20% polyacrylamide gel. The band containing the crosslinked oligonucleotide was excised from the gel and the DNA was extracted by electroelution using D-Tube™ Dialyzer (Novagen) or the Elutrap™ (Schleicher & Schuell) device.

Full experimental details and computational methods are available in the Supporting Information.

## Acknowledgements

We are grateful to Robert Rieger and the Proteomic Center of Stony Brook University for MS spectral analyses supported by grant NIH/NCRR1 S10 RR023680-1. Support for this project was provided by NSF 0549370 (A.J.C.), NIH GM6167803 (C.S.), the New York State Office of Science and Technology and Academic Research (NYSTAR), C040069 (O.D.S.), the Swiss Cancer League OCS-01413-080-2003 (O.D.S.) and supercomputer resources through the NSF Teragrid MCA02N028 (C.S.), DOE (DE-AC02-98CH10886), and the State of New York (C.S.).

**Keywords:** antitumor agents • DNA • molecular dynamics • nitrogen mustard • oligonucleotides

[1] O. D. Schärer, *ChemBioChem* **2005**, *6*, 27.

[2] a) A. Ferentz, T. Keating, G. Verdine, *J. Am. Chem. Soc.* **1993**, *115*, 9006; b) D. A. Erlanson, L. Chen, G. L. Verdine, *J. Am. Chem. Soc.* **1993**, *115*, 12583; c) E. Harwood, S. Sigurdsson, N. Edfeldt, B. Reid, P. Hopkins, *J. Am. Chem. Soc.* **1999**, *121*, 5081; d) W. Kobertz, J. M.

- Essigmann, *J. Am. Chem. Soc.* **1997**, *119*, 5960; e) P. A. Dooley, D. Tsarouhtsis, G. A. Korbel, L. V. Nechev, J. Shearer, I. S. Zegar, C. M. Harris, M. P. Stone, T. M. Harris, *J. Am. Chem. Soc.* **2001**, *123*, 1730; f) D. M. Noll, A. M. Noronha, P. S. Miller, *J. Am. Chem. Soc.* **2001**, *123*, 3405; g) C. J. Wilds, A. M. Noronha, S. Robidoux, P. S. Miller, *J. Am. Chem. Soc.* **2004**, *126*, 9257; h) I. S. Hong, M. M. Greenberg, *J. Am. Chem. Soc.* **2005**, *127*, 3692; i) J. Alzeer, O. D. Schärer, *Nucleic Acids Res.* **2006**, *34*, 4458; j) C. J. Wilds, F. Xu, A. M. Noronha, *Chem. Res. Toxicol.* **2008**, *21*, 686.
- [3] a) P. A. Muniandy, J. Liu, A. Majumdar, S. T. Liu, M. M. Seidman, *Crit. Rev. Biochem. Mol. Biol.* **2010**, *45*, 23; b) T. V. Ho, O. D. Schärer, *Environ. Mol. Mutagen.* **2010**, *51*, 552; c) E. M. Hlavín, M. B. Smeaton, P. S. Miller, *Env. Mol. Mutagen.* **2010**, *51*, 604.
- [4] T. Angelov, A. Guainazzi, O. Schärer, *Org. Lett.* **2009**, *11*, 661.
- [5] S. Lee, B. R. Bowman, Y. Ueno, S. Wang, G. L. Verdine, *J. Am. Chem. Soc.* **2008**, *130*, 11570.
- [6] a) J. O. Ojwang, D. A. Grueneberg, E. L. Loechler, *Cancer Res.* **1989**, *49*, 6529; b) J. T. Millard, S. Raucher, P. B. Hopkins, *J. Am. Chem. Soc.* **1990**, *112*, 2459; c) S. M. Rink, M. S. Solomon, M. J. Taylor, S. B. Rajur, L. W. McLaughlin, P. B. Hopkins, *J. Am. Chem. Soc.* **1993**, *115*, 2551.
- [7] a) S. M. Rink, P. B. Hopkins, *Biochemistry* **1995**, *34*, 1439; b) Q. Dong, D. Barskt, M. E. Colvin, C. F. Melius, S. M. Ludeman, J. F. Moravek, O. M. Colvin, D. D. Bigner, P. Modrich, H. S. Friedman, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12170.
- [8] S. Khullar, C. V. Varaprasad, F. Johnson, *J. Med. Chem.* **1999**, *42*, 947.
- [9] D. A. Case, T. E. Cheatham 3rd, T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668.
- [10] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.
- [11] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *7*, 655.
- [12] C. Hofr, V. Brabec, *J. Biol. Chem.* **2001**, *276*, 9655.
- [13] a) M. Raschle, P. Knipscheer, M. Enoiu, T. Angelov, J. Sun, J. D. Griffith, T. E. Ellenberger, O. D. Schärer, J. C. Walter, *Cell* **2008**, *134*, 969; b) P. Knipscheer, M. Raschle, A. Smogorzewska, M. Enoiu, T. V. Ho, O. D. Schärer, S. J. Elledge, J. C. Walter, *Science* **2009**, *326*, 1698.
- [14] L. Goodman, M. Wintrobe, W. Dameshek, M. Goodman, A. Gilman, *JAMA J. Am. Med. Assoc.* **1946**, *132*, 126.

Received: July 19, 2010  
Published online: September 14, 2010