Protein–DNA Interactions

DOI: 10.1002/ange.200703625

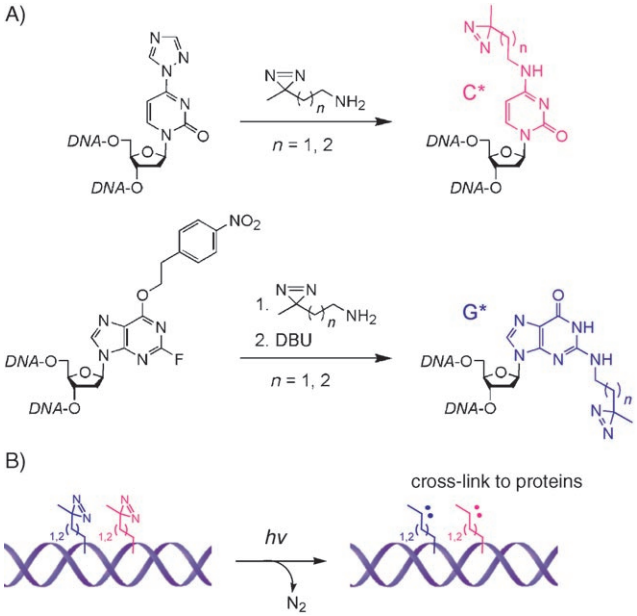
Diazirine-Based DNA Photo-Cross-Linking Probes for the Study of Protein–DNA Interactions**

*Uddhav Kumar Shigdel, Junliang Zhang, and Chuan He**

Protein–DNA interactions occur in fundamental life processes, such as replication, transcription, DNA modification, and DNA repair. Chemical and photochemical cross-linking have been used extensively in probing protein–DNA interactions.[1–3] Chemical cross-linking methods enable the trapping and characterization of various forms of protein–DNA complexes that are labile in the absence of covalent linkages.[2] Photo-cross-linking has been utilized widely to map out protein residues involved in protein–DNA interactions and to trap/identify proteins that interact with DNA.[1,3]

To obtain enough material for characterization, it is essential to use photo-cross-linkers that provide high yields of cross-linking. Various photoactive groups have been introduced into DNA for cross-linking studies;[1,3] however, most of these probes suffer from low cross-linking efficiency as a result of either low reactivity or photodecomposition, and this poor cross-linking efficiency makes biochemical studies tedious and sometimes unreliable. Aryl azides and benzophenone can be tethered to DNA or RNA through thiol-modified DNA or RNA backbones. Such probes have been utilized to obtain information on protein–DNA and protein–RNA interactions.[3c,d] However, the size of these groups may interfere with binding that is sensitive to steric hindrance.

We envisaged the possibility of incorporating the diazirine photophore into DNA as a photo-cross-linking group. Under UV irradiation, diazirine eliminates a molecule of N_2 to form a reactive carbene intermediate (Scheme 1).[1,4] This carbene species can cross-link readily with nearby protein residues with high efficiency. The use of diazirine has several advantages: 1) The photo-generated carbene intermediate has high photo-cross-linking efficiency; 2) diazirine has superb chemical stability prior to photolysis and photolyzes rapidly at wavelengths beyond those at which most biological macromolecules absorb UV light; 3) the introduction of the small diazirine unit reduces potential steric hindrance; 4) the method is nonspecific and does not require the presence of specific protein residues for cross-linking. Aryl trifluoromethyldiazirine moieties have been attached to DNA bases;[5] however, the large size of the modification still leads to the steric problem that we wish to address.



Scheme 1. A) Incorporation of diazirine with a C_2 ($n=1$) or C_3 ($n=2$) linker into the major (through the deprotection of O^4-triazolyl-dU-substituted DNA) or minor groove of DNA (through the deprotection of 2-F-dI-substituted DNA) by using the corresponding diazirine amines. B) The diazirine-modifed DNA can be photoactivated to generate a carbene intermediate for cross-linking to DNA-binding proteins. DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene.

Our strategy was to use the convertible nucleoside method[6] to introduce diazirine units into the major and minor grooves of DNA. Diazirine amines with two- and three-carbon-atom linkers (C_2: $n=1$, C_3: $n=2$ in Scheme 1) were synthesized by modifying a reported procedure.[4] Oligonucleotides with O^4-triazolyl-dU-CE phosphoramidite or 2-F-dI-CE phosphoramidite incorporated at specific positions were prepared by solid-phase synthesis, and the diazirine amines were introduced by postsynthetic modification/deprotection (Scheme 1).[6] The single-stranded DNA ssDNA-1 and ssDNA-6 were made with the modifications C* and G*, respectively (Figure 1). C* stands for an N^4-modified cytosine residue, whereas G* indicates a modification at the N^2 position of guanosine. ssDNA-1 was annealed with complementary strands to give the double-stranded DNA probes dsDNA-2, dsDNA-3, and dsDNA-4 with G, A, or T opposite the diazirine-modified C* residue, respectively (Figure 1).

[*] U. K. Shigdel, Dr. J. Zhang,[+] Prof. C. He
Department of Chemistry, University of Chicago
5735 South Ellis Avenue Chicago, IL 60637 (USA)
Fax: (+1) 773-702-0805
E-mail: chuanhe@uchicago.edu
[+] Current address:
Department of Chemistry
East China Normal University, Shanghai (P.R. China)

[**] This research was supported by the W. M. Keck Foundation, the Arnold and Mabel Beckman Foundation, the Camille & Henry Dreyfus Foundation, the Research Corporation, and a Concept Award from the U.S. Army Medical Research Office.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

Figure 1. DNA sequences used in the photo-cross-linking studies.

The annealing of a complementary strand with an abasic site opposite the C* residue gave dsDNA-5. DNA probes 7–15 were prepared with the minor-groove modification G*. Complementary strands were annealed to ssDNA-6 to yield dsDNA-7 and dsDNA-8 with C or A opposite G*, respectively (Figure 1). DNA probes 9–11 were prepared by annealing ssDNA-6 to complementary strands containing an abasic site. DNA probes 12–15 were prepared with G* positioned in matched or mismatched base pairs adjacent to an abasic site (Figure 1). The abasic site and mismatched DNA base pairs were chosen because of their preferences in interacting with some known base-flipping proteins.[2a,d,7]

To determine whether the diazirine-tethered duplex DNA was any different from the analogous alkylamine- or disulfide-tethered dsDNA, we measured the melting temperatures (T_m) of the modified duplex oligonucleotides by using a differential scanning calorimeter.[8] The T_m values of dsDNA-2 and dsDNA-7 with the C_2 linker were determined to be 55.0 and 61.2 °C, respectively (see the Supporting Information). The differences in the T_m values of the diazirine-tethered DNA and normal DNA correlate with the differences observed previously between the T_m values of normal dsDNA and those of alkylamine- or disulfide-tethered dsDNA.[6b,d] Thus, as expected, the diazirine tether does not introduce additional destabilization relative to that observed with simple alkylamine tethers.

As carbenes can cross-link indiscriminately with organic groups nearby, we also investigated whether the carbene generated from diazirine would cause any interstrand cross-linking in dsDNA. We did not observe any products of interstrand cross-linking with the various major-groove and minor-groove probes that we tested (data not shown). In analogy with the model presented for disulfide-tethered dsDNA, the diazirine moiety on dsDNA may also point away from the DNA helix and thus generate no products of interstrand cross-linking.[6b,d]

To showcase the utility of these DNA probes, we subjected them to photo-cross-linking with DNA-binding proteins. *E. coli* DNA adenine methyltransferase (EcoDam) methylates the N6 position of adenine in GATC sequences. Structural information on how this protein interacts with the sequence-specific DNA was limited until a recent structural report.[9a] We were interested in how this protein interacts with DNA, and in particular in its sequence-nonspecific DNA-binding mode. We expressed and purified EcoDam and studied its photo-cross-linking to various DNA probes containing a diazirine moiety in either the major or the minor groove of duplex DNA. We mixed EcoDam (1 equiv) with various DNA probes (3 equiv) on ice for 16 h before irradiating the samples with a mercury lamp for 10 min. The samples were then analyzed on Coomassie Blue stained SDS-PAGE gel. The appearance of a low-mobility band indicated the formation of cross-linked products. To our delight, very good yields (20–50 %) of cross-linking were observed with most of the DNA probes (notably with probes 1, 2, 3, 6, 10, and 11; Figure 2). To ensure maximum cross-linking, all experiments were performed with 3 equivalents of DNA and 1 equivalent of EcoDam. However, the use of an excess of the DNA probe may not be necessary (Figure 2 A; compare lanes 2, 4, and 7). A UV irradiation time of 10 min was found to be optimal for the initiation of cross-linking (Figure 2 A, compare lanes 6, 7, and 8).

We surveyed a range of the probes shown in Figure 1 to test their versatility. In the first group of probes tested, the diazirine-containing base was positioned opposite an abasic site or in a matched or mismatched base pair. Probes 2, 3, 10, and 11 gave the best results with cross-linking yields of approximately 40–50 % (Figure 2 A, lane 7; Figure 2 C, lane 6; Figure 2 D, lanes 10, 12, 14, and 16). If the reaction mixture was preincubated with a 10-fold excess of unmodified dsDNA, no significant cross-linking was observed between EcoDam and one of the best probes, dsDNA-10 (Figure 2 A, lane 10). This DNA probe does not contain the sequence recognized by EcoDam and can be titrated away with the excess unmodified DNA, thus further demonstrating that the protein–DNA interaction is critical for photo-cross-linking.

Next, we tested probes that contain G* adjacent to the abasic site (probes 12–15, Figure 1). We chose G* because the minor-groove modification led to the best cross-linking yield among the DNA probes that contain an abasic site. We observed moderate cross-linking of EcoDam with probes 12, 13, and 14 with both C_2 and C_3 linkers, whereas minimal cross-linking was observed when probe 15 was used (Figure 2 B,C). The minor-groove modification in probes 7 and 8 also generated cross-linked products with EcoDam, although not

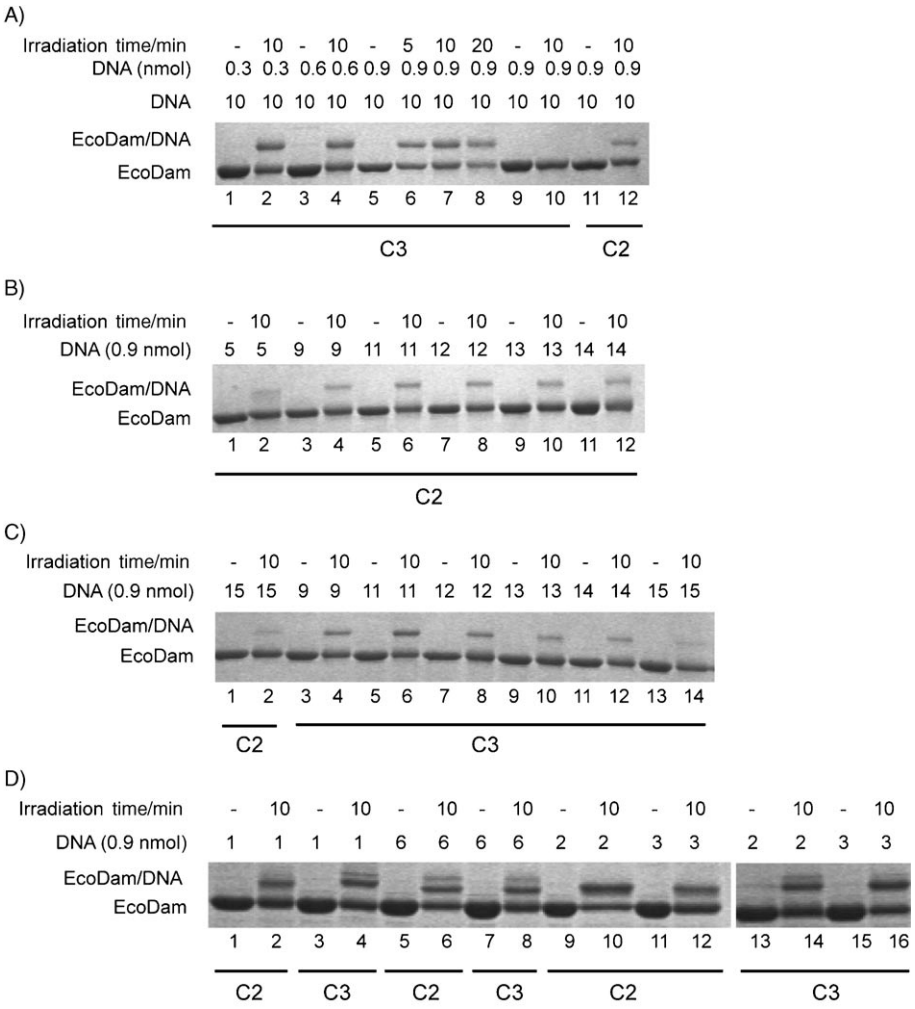


Figure 2. SDS-PAGE analysis of the photo-cross-linking reactions between *E. coli* DNA adenine methyltransferase (EcoDam; 0.3 nmol) and various DNA probes (0.9 nmol). C_2 and C_3 are the lengths of the carbon tethers between the diazirine ring and N2 of guanosine or N4 of cytosine. A) The photo-cross-linking between probe 10 and EcoDam was studied for its dependence on the irradiation time with UV light and on the DNA equivalence. A 10-fold excess of a competing unmodified DNA molecule with the same sequence as that of probe 10 was used in the experiments shown in lanes 9 and 10. No cross-linking was observed in these two experiments. B–D) Different probes were tested for cross-linking to EcoDam with and without UV irradiation.

as efficiently as probes 1, 2, 3, 6, 10, and 11 (data not shown), but the best yields (ca. 50%) were observed with probes 10 and 11. Base-flipping proteins, such as EcoDam, are known to recognize unstable sites in a specific manner,[7] which may account for the higher yields of cross-linking observed for probes 10 and 11. When the single-stranded DNA probes 1 and 6 were used, we observed 1:2 protein–DNA complexes on the gel as well as the 1:1 protein–DNA complexes (Figure 2D). These additional complexes might arise from the binding of a small fraction of EcoDam to two ssDNA molecules simultaneously. Our results show that one can identify several DNA probes that cross-link efficiently with the target protein from a quick screen (see the Supporting Information).

Given that a nonspecific sequence was used in the study with EcoDam, the results of the experiments with probes 2 and 3 are excellent for photo-cross-linking reactions between a protein and DNA. These DNA probes appear to react with DNA-binding residues on the surface of EcoDam, as comparable photo-cross-linking yields were observed with matched and mismatched probes. EcoDam methylates adenine in a sequence-specific manner.[9] It must scan the DNA base pairs in its search for a specific sequence. The close contacts with DNA base pairs may account for cross-linking with diazirine moities on matched DNA probes. The structural characterization of this "nonspecific" interaction, which we can now trap in high yields, would provide valuable information about how EcoDam binds to nonspecific DNA.

We also studied human O^6-alkylguanine-DNA alkyltransferase (hAGT), a protein that searches DNA base damage sequence-non-specifically.[10] When hAGT was incubated with our DNA probes, very good yields of cross-linked products were observed (Figure 3A, lanes 5, 7, 9–11, and 13; Figure 3B, lanes 6 and 10). Experiments with dsDNA-3 and hAGT in various ratios indicated that almost all of the hAGT in solution was consumed with 3 equivalents of DNA (Figure 3A, compare lanes 5, 7, and 10). An optimal UV irradiation time of 10 min was determined from a time-dependence study (Figure 3A, compare lanes 9, 10, and 11).

Low cross-linking was observed when the DNA probe 2 with either a C_2 or a C_3 linker was used (Figure 3A, lane 3; Figure 3B, lane 8), whereas the mismatched probe 3 with a C_2 or C_3 linker gave the photo-cross-linked product in very good yield (50–75%; Figure 3A,B, lane 10). The DNA probe 4, with a C*:T mismatch, underwent less cross-linking than that observed with probe 3 (C*:A; Figure 3B). The use of ssDNA-1 with a C_2 or C_3 linker led to little cross-linking (Figure 3B, lanes 2 and 4). When a 10-fold excess of unmodified, mismatched dsDNA was added to the reaction mixture, less of the covalently linked complex was formed (Figure 3A, lane 13), which indicates that hAGT can be titrated away from dsDNA-3 with mismatched dsDNA. We observed minimal photo-cross-linking of hAGT with dsDNA-5 and other DNA probes containing minor-groove modifications (results not shown), most likely as a result of a lack of contact of the tethered diazirine with the protein.[10b,c] Overall,

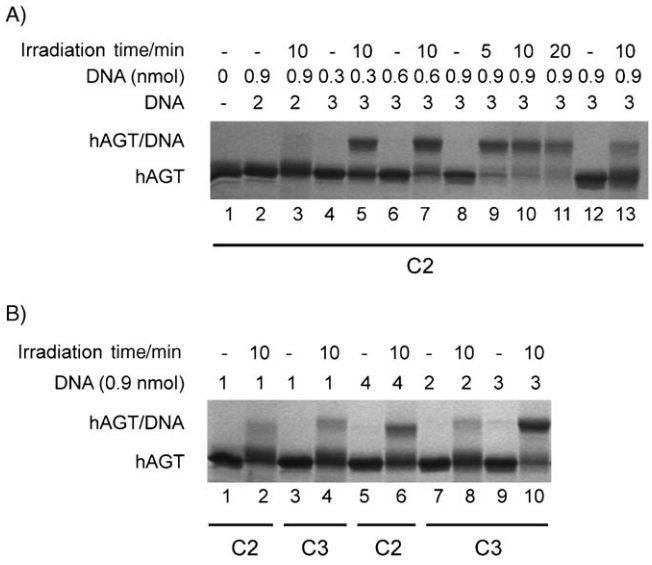


Figure 3. SDS-PAGE analysis of the photo-cross-linking reactions between human AGT (0.3 nmol) and DNA probes 1–4 (0.9 nmol). C_2 and C_3 are the lengths of the carbon tethers between the diazirine ring and N4 of cytosine. A) Comparison of the photo-cross-linking efficiency of matched (probe 2) and mismatched probes (probe 3) with hAGT. Lanes 12 and 13 show the extent of competition from a 10-fold excess of an unmodified mismatched DNA with the same sequence as that of probe 3. B) Photo-cross-linking with different DNA probes and tether lengths.

the most efficient cross-linking with hAGT was observed with dsDNA-3 with a C_2 linker.

In summary, we have described a simple and quick method for the incorporation of diazirine into the major and minor grooves of DNA for efficient photo-cross-linking to proteins. Good to excellent yields of cross-linking with two different proteins were observed for the resulting DNA probes with small diazirine modifications. This method appears to be suitable for mapping out protein–DNA interactions, in particular those that may be sensitive to steric hindrance. The high cross-linking efficiency observed in several cases would allow direct purification of the photo-cross-linked products for structural characterization. Alternatively, protein residues engaged in photo-cross-linking can be identified readily by mass spectrometry and mutated to Cys residues in preparation for chemical disulfide cross-linking with disulfide-tethered oligonucleotides. The resulting disulfide-linked complex can be subjected to structural studies. Furthermore, the relatively high cross-linking efficiency of these diazirine-based probes makes them promising tools for the covalent trapping of proteins from cell extracts.

Received: August 9, 2007
Published online: November 12, 2007

Keywords: cross-linking · diazirine · DNA probes · photochemistry · proteins

[1] a) Y. Hatanaka, Y. Sadakane, *Curr. Top. Med. Chem.* **2002**, *2*, 271–288; b) K. M. Meisenheimer, T. H. Koch, *Crit. Rev. Biochem. Mol. Biol.* **1997**, *32*, 101–140; c) F. Kotzyba-Hibert, I. Kapfer, M. Goeldner, *Angew. Chem.* **1995**, *107*, 1391–1408; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1296–1312; d) S. N. Khodyreva, O. I. Lavrik, *Curr. Med. Chem.* **2005**, *12*, 641–655; e) J. Brunner, *Annu. Rev. Biochem.* **1993**, *62*, 483–514.

[2] a) G. L. Verdine, D. P. G. Norman, *Annu. Rev. Biochem.* **2003**, *72*, 337–366; b) A. Banerjee, W. L. Santos, G. L. Verdine, *Science* **2006**, *311*, 1153–1157; c) Y. Mishina, C. He, *J. Am. Chem. Soc.* **2003**, *125*, 8730–8731; d) E. M. Duguid, Y. Mishina, C. He, *Chem. Biol.* **2003**, *10*, 827–835; e) D. G. Osterman, G. D. DePillis, J. C. Wu, A. Matsuda, D. V. Santi, *Biochemistry* **1988**, *27*, 5204–5210; f) S. Klimasauskas, S. Kumar, R. J. Roberts, X. Cheng, *Cell* **1994**, *76*, 357–369; g) K. M. Reinisch, L. Chen, G. L. Verdine, W. N. Lipscomb, *Cell* **1995**, *82*, 143–153.

[3] a) M. M. Coman, M. Jin, R. Ceapa, J. Finkelstein, M. O'Donnell, B. T. Chait, M. M. Hingorani, *J. Mol. Biol.* **2004**, *342*, 1457–1469; b) C. X. Zhang, P. V. Chang, S. J. Lippard, *J. Am. Chem. Soc.* **2004**, *126*, 6536–6537; c) T. Lagrange, T.-K. Kim, G. Orphanides, Y. W. Ebright, R. H. Ebright, D. Reinberg, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10620–10625; d) K. Musier-Forsyth, P. Schimmel, *Biochemistry* **1994**, *33*, 773–779; e) V. Pingoud, H. Thole, F. Christ, W. Grindl, W. Wende, A. Pingoud, *J. Biol. Chem.* **1999**, *274*, 10235–10243.

[4] a) R. F. R. Church, M. J. Weiss, *J. Org. Chem.* **1970**, *35*, 2465–2471; see also the Supporting Information; b) M. Suchanek, A. Radzikowska, C. Thiele, *Nat. Methods* **2005**, *2*, 261–267.

[5] a) T. Yamaguchi, K. Suyama, K. Narita, S. Kohgo, A. Tomikawa, M. Saneyoshi, *Nucleic Acids Res.* **1997**, *25*, 2352–2358; b) J. J. Tate, J. Persinger, B. Bartholomew, *Nucleic Acids Res.* **1998**, *26*, 1421–1426; c) G. A. Korshunova, N. V. Sumbatyan, A. N. Topin, M. T. Mtchedlidze, *Mol. Biol.* **2000**, *34*, 823–839.

[6] a) Y.-Z. Xu, Q. Zheng, P. F. Swann, *J. Org. Chem.* **1992**, *57*, 3839–3845; b) A. M. MacMillan, G. L. Verdine, *J. Org. Chem.* **1990**, *55*, 5931–5933; c) A. M. MacMillan, G. L. Verdine, *Tetrahedron* **1991**, *47*, 2603–2616; d) D. A. Erlanson, L. Chen, G. L. Verdine, *J. Am. Chem. Soc.* **1993**, *115*, 12583–12584.

[7] a) X. Cheng, R. J. Roberts, *Nucleic Acids Res.* **2001**, *29*, 3784–3795; b) R. J. Roberts, X. Cheng, *Annu. Rev. Biochem.* **1998**, *67*, 181–198; c) G. L. Verdine, *Cell* **1994**, *76*, 197–200.

[8] a) J. G. Duguid, V. A. Bloomfield, J. M. Benevides, G. J. Thomas, Jr., *Biophys. J.* **1996**, *71*, 3350–3360; b) J. Völker, H. H. Klump, K. J. Breslauer, *J. Am. Chem. Soc.* **2007**, *129*, 5272–5280.

[9] a) J. R. Horton, K. Liebert, M. Bekes, A. Jeltsch, X. Cheng, *J. Mol. Biol.* **2006**, *358*, 559–570; b) A. Løbner-Olesen, O. Skovgaard, M. G. Marinus, *Curr. Opin. MicroBiol.* **2005**, *8*, 154–160.

[10] a) A. E. Pegg, *Cancer Res.* **1990**, *50*, 6119–6129; b) E. M. Duguid, P. A. Rice, C. He, *J. Mol. Biol.* **2005**, *350*, 657–666; c) D. S. Daniels, T. T. Woo, K. X. Luu, D. M. Noll, N. D. Clarke, A. E. Pegg, J. A. Tainer, *Nat. Struct. Mol. Biol.* **2004**, *11*, 714–720.

