Three-Dimensional Structure of a Disulfide-Stabilized Non-Ground-State DNA Hairpin

Hong Wang,[†,‡] Scott E. Osborne,[†]
Erik R. P. Zuiderweg,[‡] and Gary D. Glick*[,†,‡]

Department of Chemistry and Biophysics Research Division, University of Michigan Ann Arbor, Michigan 48109-1055

Received February 14, 1994

Although nucleic acid structures possessing disulfide cross-links have been known for almost 30 years,[1-3] use of this modification to design DNA[4-12] and RNA[13] constructs for biophysical characterization and molecular recognition studies has been realized only recently. This achievement is due principally to advances in solid-phase synthesis, which facilitates the incorporation of thiols at specific loci within nucleic acids.[14-20] However, continued use of disulfide-modified nucleic acids requires an understanding of the structural consequences of incorporating such cross-links. Here we present the three-dimensional structure of a disulfide-stabilized analog of the Dickerson dodecamer hairpin[5,21] determined using NMR spectroscopy and provide evidence that the cross-link is accommodated without altering DNA geometry.

Both d(CGCGAATTCGCG)$_2$ and the corresponding thiol-modified dodecamer **2** thermally denature in a biphasic manner in buffers containing [Na$^+$] ≤ 10 mM and with [DNA] ≤ 50 μM.[5,21] The first transition defines melting of the duplex to a hairpin structure, while the second transition represents denaturation of the hairpin to a random coil. Unlike the stem-loop intermediate produced by initially denaturing d(CGCGAATTCGCG)$_2$, cross-linked hairpin **3** does not undergo further structural transitions as a result of increasing temperature, sodium ion concentration, or DNA concentration (Figure 1).[5] Thus, the disulfide cross-link effectively stabilizes this non-ground-state conformation over a wide range of conditions, including those suitable for NMR spectroscopy.

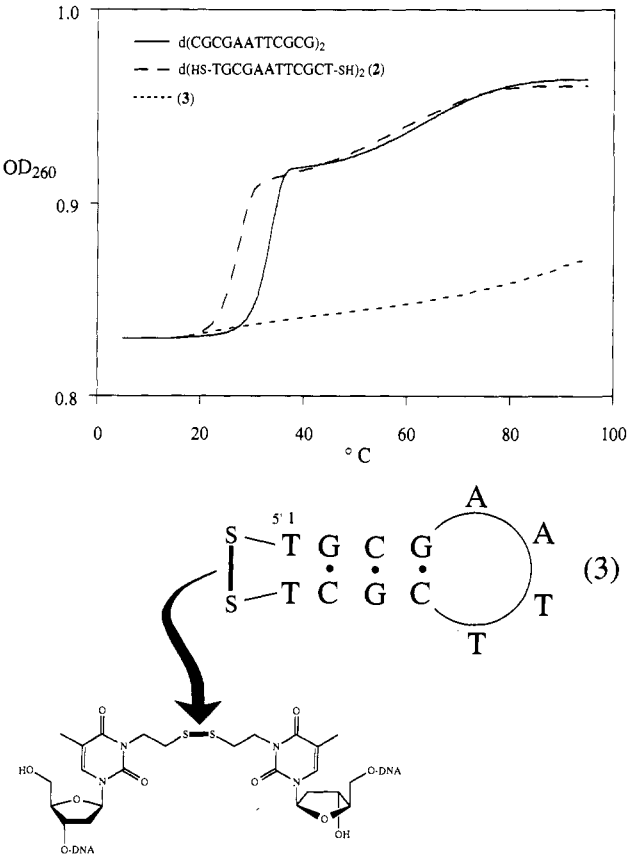

Figure 1. (A, top) Normalized UV thermal denaturation curves of d(CGCGAATTCGCG)$_2$, **2**, and **3** measured in 1 mM NaCl buffer, pH 7, 50 μM in DNA. Duplex **2** is slightly destabilized relative to the parent sequence, as indicated by the T_m of the first transition. We attribute this T_m decrease to unfavorable steric interactions between the alkanethiol linkers.[4,5,22] Higher [Na$^+$] stabilize duplex DNA, and in buffers where [Na$^+$] > 10 mM, the T_m differences bewteen **2** and d(CGCGAATTCGCG)$_2$ are <1 °C.[21,23] (B, bottom) Chemical structure of the disulfide cross-linked hairpin. Air oxidation to form the disulfide bond was conducted at 50 °C as previously described.[5]

Using COSY, TOCSY, and NOESY spectra measured in D_2O, the nonlabile protons (and phosphorus) resonances in **3** have been assigned.[24,25] The sequential connectivity network indicates that the G·C residues form a duplex which stops at positions A5 and T8 in the loop. To assign the labile protons, a NOESY spectrum was measured in 90% H_2O–10% D_2O with a jump–return read pulse for solvent suppression. Only three G·C base pairs are present in this spectrum, which indicates that there is no hydrogen bonding between the loop residues and that the presence of the cross-link does not disrupt canonical Watson–Crick base pairing in the duplex.[26]

A combined distance geometry (DG)–simulated annealing (SA)–iterative relaxation matrix analysis (IRMA) protocol was

* Address correspondence to this author at the Department of Chemistry. Phone: (313) 764-4548. FAX: (313) 764-8815. e-mail: GLICK@Chem.LSA.UMich.edu.
† Department of Chemistry.
‡ Biophysics Research Division.

(1) Lipsett, M. N. *Cold Spring Harbor Symp. Quant. Biol.* **1966**, *31*, 449–455.
(2) Lipsett, M. N. *J. Biol. Chem.* **1967**, *212*, 4067–4071.
(3) Lipsett, M. N.; Doctor, B. P. *J. Biol. Chem.* **1967**, *212*, 4072–4077.
(4) Glick, G. D. *J. Org. Chem.* **1991**, *56*, 6746–6747.
(5) Glick, G. D.; Osborne, S. E.; Knitt, D. S.; Marino, J. P., Jr. *J. Am. Chem. Soc.* **1992**, *114*, 5447–5448.
(6) Stevens, S. Y.; Swanson, P. C.; Voss, E. W., Jr.; Glick, G. D. *J. Am. Chem. Soc.* **1993**, *115*, 1585–1586.
(7) Swanson, P. C.; Glick, G. D. *BioMed. Chem. Lett.* **1993**, *3*, 2117–2118.
(8) Ferentz, A. E.; Keating, T. A.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 9006–9014 and references therein.
(9) Wolfe, S. A.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 12585–12586.
(10) Erlanson, D. A.; Chen, L.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 12583–12584.
(11) Milton, J.; Connolly, B. A.; Nikiforov, T. T.; Cosstick, R. *J. Chem. Soc., Chem. Commun.* **1993**, 779–780.
(12) Swanson, P. C.; Cooper, B. C.; Glick, G. D. *J. Immunol.* **1994**, *152*, 2601–2612.
(13) Goodwin, J. T.; Glick, G. D. *Tetrahedron Lett.* **1994**, *35*, 1647–1650.
(14) Zuckermann, R.; Corey, D.; Schultz, P. G. *Nucleic Acids Res.* **1987**, *15*, 5305–5321.
(15) Chu, B. C.; Orgel, L. E. *Nucleic Acids Res.* **1988**, *16*, 3671–3691.
(16) Gaur, R. K. *Nucleosides Nucleotides* **1991**, *10*, 895–909.
(17) Gupta, K. C.; Sharma, P.; Kumar, P.; Sathyanarayana, S. *Nucleic Acids Res.* **1991**, *19*, 3019–3025.
(18) Fidanza, J. A.; McLaughlin, L. W. *J. Org. Chem.* **1992**, *57*, 2340–2346.
(19) Bradley, D. H.; Hanna, M. M. *Tetrahedron Lett.* **1992**, *33*, 6223–6226 and references therein.
(20) Goodwin, J. T.; Glick, G. D. *Tetrahedron Lett.* **1993**, *34*, 5549–5552.
(21) Marky, L. A.; Blumenfeld, K. S.; Kozlowski, S.; Breslauer, K. J. *Biopolymers* **1983**, *22*, 1247–1257.
(22) Above 19 °C, the terminal base pairs in d(CGCGAATTCGCG)$_2$ are frayed, so introducing the thiol modification does not necessarily result in loss of Watson–Crick hydrogen bonds: Patel, D. J.; Kozlowski, S. A.; Marky, L. A.; Broka, C.; Rice, J. A.; Itakura, K.; Breslauer, K. J. *Biochemistry* **1982**, *21*, 428–436 and references therein.
(23) Osborne, S. E.; Stevens, S. Y.; Glick, G. D., unpublished observations.
(24) 1.5 mM in **3**, 50 mM NaCl, 10 mM Na_2PO_4, pH 7, 12 °C, 500 MHz.
(25) Details concerning the resonance assignments and structure calculations will be reported elsewhere: Wang, H.; Zuiderweg, E. R. P.; Glick, G. D., manuscript in preparation.
(26) Changes in the imino proton region of the ^{1}H NMR spectrum were not observed when the measurements were conducted at 1 °C.

0002-7863/94/1516-5021$04.50/0

employed to determine the structure of **3**. To begin, 60 starting conformations were embedded using the DGII program.[25,27] For this calculation we used 177 interatomic distance constraints between nonlabile protons, nine hydrogen bond constraints for the three G·C base pairs, 40 torsion angle constraints for the 12 deoxyribose sugars, and the 11 ϵ backbone angles. After SA refinement of each DGII generated structure, NOESY spectra (τ_m = 75, 150, and 300 ms) were back-calculated. For each mixing time, the experimental and back-calculated spectra were compared, and IRMA was used to compute a new constraint matrix that better represented the experimental data. The SA protocol was repeated using these new constraints, and the entire SA–IRMA refinement was iterated three times to minimize the differences between the experimental and the back-calculated spectra.

Of the 60 starting structures generated by DGII, 43 converged after repetitive SA–IRMA protocol (Figure 2A). The G·C stem of **3** forms a short helix with a rise of 3.45 Å and a rotation of 34.6°. These values, in addition to the other helical parameters and backbone angles, fall within the range of observed values for B-DNA.[25] The cross-link is not as defined as the stem or the loop because only a few experimental constraints are available for this region of the molecule (Figure 2B). However, none of the conformations of the cross-link produced by the SA–IRMA refinement distorts the duplex geometry away from B-like DNA. All of the glycosidic angles in **3** are *anti* except for residue A5 in the loop, which adopts a *syn* orientation. The first three nucleotides in the loop stack over the 5′-end of the helix and are followed by a sharp turn at residue T8 which closes the loop.[28] The aromatic bases point into the major groove, while the phosphate backbone projects toward solvent, a conformation recently predicted for DNA hairpins.[29]

To summarize, we have demonstrated the utility of our disulfide cross-link to facilitate characterization of a unique non-ground-state DNA conformation that would otherwise be inaccessible to structural analysis at high resolution. Moreover, having provided evidence that this cross-link does not appear to alter B-DNA geometry, we anticipate that it will see greater use to aid in biophysical analysis of other confomationally labile secondary structures and as a tool in protein/drug·DNA binding studies.[6,9,12,30]

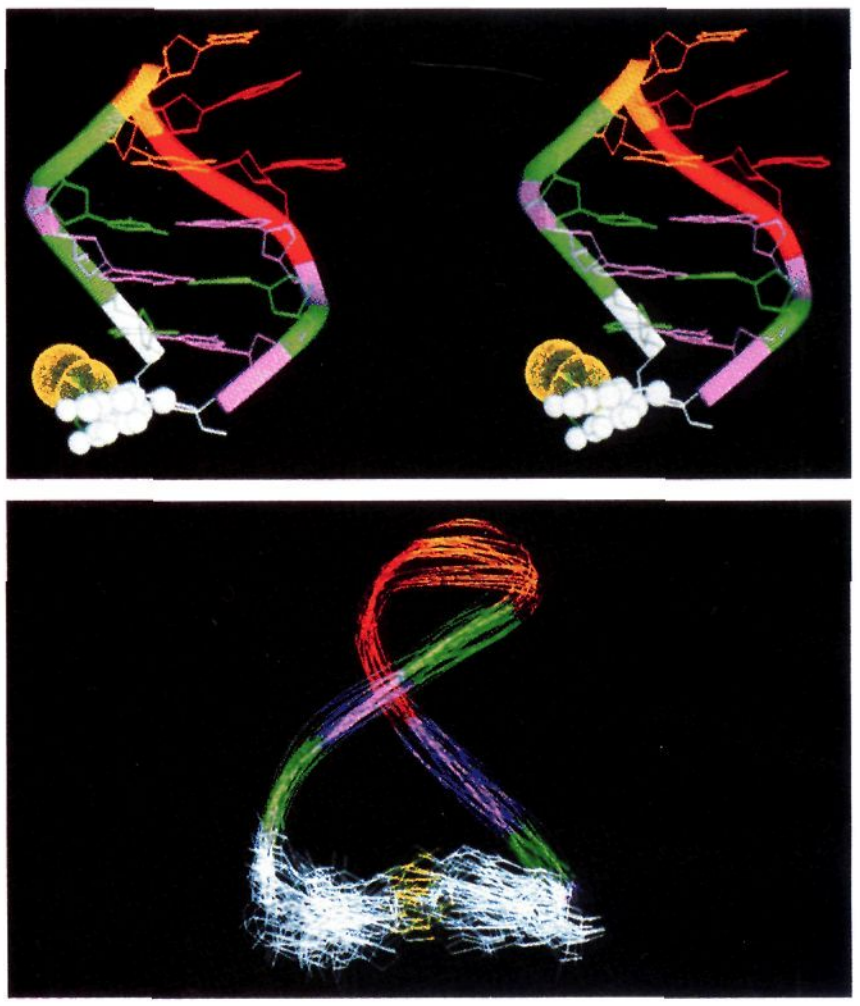

Figure 2. (A, top) Stereoview of the average structure of **3** (yellow, van der Waals surface, sulfur atoms; white, CPK, thymidine residues that form the cross-link; green, cytosine; purple, guanosine; red, adenosine; brown, loop thymidine residues). The overall heavy-atom root mean square deviation for superimposing all 43 structures is 1.1 Å. (B, bottom) Backbone trace of **3**. The backbone of each structure in the final conformational ensemble is shown in wire format, and the cross-link is depicted explicitly. In this orientation, **3** is rotated by ca. 90° along the *y*-axis relative to the view shown in A.

(27) As implemented in the NMRchitect module of the INSIGHT/DISCOVER software package (Biosym Technologies).

(28) Haasnoot, C. A. G.; Hilbers, C. W.; van der Marel, G. A.; van Boom, J. H.; Singh, U. C.; Pattabiraman, N.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1986**, *3*, 843–857.

(29) Erie, D. A.; Suri, A. K.; Breslauer, K. J.; Jones, R. A.; Olson, W. K. *Biochemistry* **1993**, *32*, 436–454.

(30) Supported in part by an NSF Young Investigator Award with matching funds from Merck & Co. Partial funding of the 500-MHz NMR spectrometer was provided by NIH Shared Instrumentation Grant RR 06739.