Gary D. Glick
Department of Chemistry,
University of Michigan,
Ann Arbor, MI 48109

Received and accepted 28 July 1998

Design, Synthesis, and Analysis of Conformationally Constrained Nucleic Acids

Abstract: *In this review I discuss straightforward and general methods to modify nucleic acid structure with disulfide cross-links. A motivating factor in developing this chemistry was the notion that disulfide bonds would be excellent tools to probe the structure, dynamics, thermodynamics, folding, and function of DNA and RNA, much in the way that cystine cross-links have been used to study proteins. The chemistry described has been used to synthesize disulfide cross-linked hairpins and duplexes, higher order structures like triplexes, nonground-state conformations, and tRNAs. Since the cross-links form quantitatively by mild air oxidation and do not perturb either secondary or tertiary structure, this modification should prove quite useful for the study of nucleic acids.* © 1998 John Wiley & Sons, Inc. Biopoly 48: 83–96, 1998

Keywords: *nucleic acid; disulfide cross-link; structure; dynamics; stability*

INTRODUCTION

Characterization of synthetic oligonucleotides shapes much of our understanding of "native," higher molecular weight DNA and RNA molecules. Although of immense utility, short oligonucleotides usually possess lower structural and thermal stability and have greater end effects than the larger nucleic acid constructs they are intended to model. Hence, the physiochemical properties of oligonucleotides may not always compare favorably to those of larger nucleic acids.[1–9] Arguably, the most successful approach to stabilize oligonucleotides is to connect the strands that comprise helical structure with a cross-link.[10–15] Methods to cross-link nucleic acids can generally be divided into two categories. In the first group, cross-links are formed using an exogenous reagent, like a (bis)electrophile. Because (bis)electrophiles can react with nearly all of the nucleophilic sites on the bases, these agents often have little or no sequence specificity and form complex mixtures of cross-linked aducts.[16–18] Although some natural products like mitomycin C and synthetic compounds such as cisplatin and psoralen can form lesions at unique sits, these reagents require the presence of specific recognition sites within a target sequence to generate the cross-link.[19–22] In some cases, cross-link formation by alkylating agents also can have undesirable effects such as disrupting base stacking.[16,21]

Cross-links can also be placed into oligonucleotides through solid-phase synthesis of oligomers site specifically labeled with modified nucleosides that present reactive groups. Positioning the reactive groups in proximity on opposing strands of a helix allows for formation of a cross-link. This second general strategy requires selection of both the loci to be bridged as well as a chemistry to form the cross-link. In one of the most effective examples of this approach, Webb and Matteucci demonstrated that

Correspondence to: Gary D. Glick; email: gglick@umich.edu
Contract grant sponsor: NIH, National Arthritis Foundation, American Cancer Society, National Science Foundation, Camille and Henry Dreyfus Foundation, Alfred P. Sloan Foundation
Contract grant number: GM-52831 and GM-43861 (NIH)
Biopolymers (Nucleic Acid Sciences), Vol. 48, 83–96 (1998)

CCC 0006-3525/98/010083-14

DNA oligomers containing cytosines bearing an aziridine group on the N4 position form cross-links after annealing to a complementary oligomer, and selective opening of the aziridine by the exocyclic amine of an opposing dC or dA.[23] These cross-links have proved useful for some experiments, but this chemistry in particular, and related methods in general, have several limitations. For example, ethano-bridged cross-links are not formed in high yield, the internal mismatch required to form the cross-link can give rise to (local) disruption of helical geometry, and these cross-links potentially can interfere with protein and drug binding.[24–26]

An alternate location for incorporation of an intrastrand cross-link is at the terminus of helical structure. For example, one end of a duplex can be covalently linked by bridging the 3′- and 5′-terminal hydroxyl groups with either oligonucleotide, (oligo)glycol, or alkyl linkers to form stem–loop structures ("hairpins").[27–34] Both ends of a duplex can also be covalently capped with linkers to generate double hairpins or "dumbbells."[35–43] Because dumbbells denature in a monomolecular fashion and do not suffer from end effects, they are particularly useful in thermodynamic experiments.[35,36] However, the synthesis of DNA dumbbells has in some cases proved to be quite challenging, which has limited their utility.[35] In addition, because the linkers are tethered to the terminal hydroxyl groups, some standard enzymatic manipulations such as end labeling are not possible.

Design of an Ideal Cross-Link

Several years ago we set out to develop a simple and general approach to cross-link oligonucleotides that would overcome the drawbacks of other methods. In developing our chemistry, we focused on addressing eight criteria (Table I). At the outset of this work, it was clearly desirable to incorporate our cross-links by solid-phase synthesis to permit the modifications to be placed site specifically and in close proximity within any given target sequence. Yet it was critical to find chemistry that would be specific and efficient for cross-link formation. We selected thiols as the reactive functional groups because the mild redox chemistry to form disulfide bonds from thiols is specific for sulfur. In addition, disulfide bonds are formed in high yield, often quantitatively, and are stable to a wide variety of solvents and reagents yet are cleaved by reduction. Finally, the use of disulfide bonds to constrain macromolecular architecture has already been demonstrated in the peptide/protein literature.[44]

The next step in developing our chemistry was to select the loci for modification. Because cross-links residing in either the major or minor groove could interfere with ligand recognition and hydration, we chose to place our cross-links at the terminus of helical structure. To avoid using the 3′- and 5′-hydroxyl groups, we decided to modify the terminal bases themselves. This strategy exploited the fact that the terminal residues in duplexes have a reduced stability due to end effects, often referred to as "end fraying."[45] Therefore, altering the hydrogen bonding groups of the two opposing bases at the terminus of a duplex should not adversely affect helical stability, provided the thiol modifications do not disrupt base stacking.[46,47,*] While this energetic penalty could be compensated for upon formation of the cross-link, it necessarily limits the overall (thermodynamic) stabilization the cross-link can provide.

When two thymidine residues directly oppose each other in duplex DNA, their N3 positions project toward the center of the helix and converge to about 4.5 Å. Based on this observation, we reasoned that replacing the terminal bases on one or both ends of a duplex with an N3-(alkylthiol)thymidine should be an ideal way to form interstrand disulfide cross-links. Molecular modeling studies suggest that disulfide cross-links can form when the terminal bases of a duplex are replaced with either N3-(methylthiol)thymidine or N3-(ethylthiol)thymidine. Since N3-(methylthiol)thymidine is expected to decompose under the acidic conditions used to detritylate synthetic oligonucleotides, N3-(ethylthiol)thymidine **(T^{*}_{SH})** was chosen (Figure 1). In the following sections we describe key aspects pertaining to the synthesis, structure, dynamics, thermodynamic stability, and in vitro biological properties of our disulfide cross-linked DNA oligonucleotides. We then discuss two applications of our cross-linking chemistry and describe recent experiments using different thiol modified nuclosides to generate disulfide cross-links within RNA secondary and tertiary structure. Last, we provide a brief outline of other methods that have been developed to incorporate disulfide cross-links into nucleic acids.

DISULFIDE CROSS-LINKED DNA DUPLEXES

Synthesis

We explored the properties of disulfide cross-linked DNA duplexes by synthesizing analogues of two pre-

* The free energy of a dT · dT mismatch located at the terminus of a duplex is approximately 1.0 kcal/mol more stable than a dT · dT mismatch within a duplex.

Table I Requirements for an "Ideal" Nucleic Acid Cross-Link

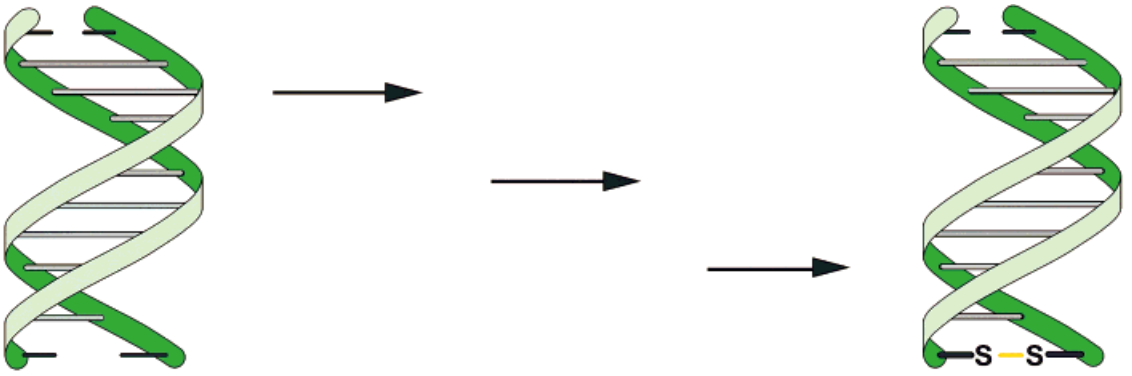
Requirement[a]	Rationale	Potential Solution
1. Increase the conformational stability toward thermal-, ionic-, pH-, and concentration-induced conformational changes	Conformationally stable, nucleic acids that are structurally homogeneous will facilitate their structural and thermodynamic analysis	Covalent interstrand cross-links confer a large degree of conformational stability
2. Form interstrand cross-link site specifically	Limits undesired adduct formation resulting in high yields of cross-linked product	Position functional groups proximal to one another on opposing strands of a helix
3. Cross-link does not alter native geometry	Necessary if the structural aspects of cross-linked constructs are to be compared to their unmodified nucleic acid precursors	Position cross-link within a sterically accessible space
4. Grooves of the helices, counterion binding, and hydration must remain unaltered	Allows for protein and ligand · nucleic acid interactions to be studied, maximizes (thermal) stability	Linker at terminus of helix
5. 3′- and 5′-Hydroxyl groups are free to [^{32}P] end label	Needed for wide range of biochemical assays (e.g., footprinting/sequencing)	Cross-link located on the base rather than terminal hydroxyl groups
6. Prepared and isolated in large quantities	Allows for high-resolution physical measurements such as nmr, DSC, and x-ray diffraction studies	Efficient cross-linking chemistry
7. Reversible	Assess structural and thermodynamic effects of base modification before and after cross-linking	Reversible disulfide cross-link
8. Flexible to a wide variety of nucleic acid structures	Broad applicability will be necessary if this methodology is to be used uniformly in nucleic acids	Position functional tether at a base or sugar site; cross-link within a sterically accessible space

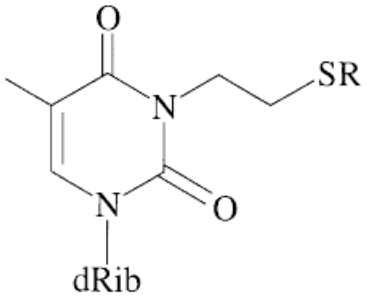
[a] Relative to the unmodified nucleic acid.

viously well-studied oligonucleotides. The first is a hairpin whose sequence is derived from the ColE1 cruciform,[48] and the second is the Dickerson/Drew dodecamer (Figure 2).[49] Synthesis of **2b, 4b,** and **5b** was conducted using standard phosphoramidite chemistry using N-benzoyl and N-isobuturyl protected bases.[52–54] After deprotection and purification of each thiol-modified DNA by reversed-phase high performance liquid chromatography (HPLC), the *t*-butyl mixed disulfide protecting groups on each sequence were cleaved with dithiothreitol (DTT; $\sim$ 20 equivalent per thiol, pH 8, 4°C, overnight). Reduced DNA samples were separated from the DTT by reversed-phase HPLC. Disulfide bond formation was performed by dissolving each sample in phosphate buffer (pH 8.3, [DNA] $\sim$ 2 μM) and stirring vigorously at 25°C exposed to air. The reactions were usually complete in 8 h as judged by a negative Ellman's test. The cross-linked DNAs were then isolated by either polyacrylamide gel electrophoresis (PAGE) or HPLC and incorporation of the modified bases was confirmed by enzymatic digestion. Importantly, disulfide bond formation was quantitative.

Structural Studies

The effects of introducing the thiol modifications were first assessed by comparing uv thermal denaturation profiles of **2a** to that of the corresponding wild-type sequence. We observed that the T_m values are within $\pm$1°C and the transitions are nearly superimposable, which suggests that the thiol linkers do not alter the stability or thermal denaturation pathways of the modified hairpin relative to its unmodified counterpart.[51] These data also suggest that all of the stability conferred by the disulfide bond will be reflected in the cross-linked structure, since there is not a





T^*_{SH}; R = H

T^*_{tBu}; R = StBu

T^*_2 ; R = ⟩ 2

FIGURE 1 (Top) Synthetic strategy. To place cross-links within a duplex requires synthesizing thiol-modified nucleosides, incorporating the modified bases within a target by solid-phase synthesis, and following removal of all protecting groups, air oxidation to form the cross-link. (Bottom left) Chemical structure of cross-link. During synthesis, the thiol group on $\mathbf{T^*_{SH}}$ is protected as a *tert*-butyl mixed disulfide **($\mathbf{T^*_{tBu}}$),** and this protecting group is stable to all conditions of solid-phase synthesis, deprotection, and purification. The coupling efficiency of this base is indistinguishable from thymidine (>99% per cycle). To place $\mathbf{T^*_{tBu}}$ at the 3′-terminus, the 3′-hydroxyl group is attached to controlled-pore glass.

significant energetic penalty for replacing the terminal bases with our N3 thiol modified thymidines. Introducing a disulfide cross-link in **1** increases the T_m from 65 to 81°C in low ionic strength buffers. If the $[Na^+] > 100$ m*M*, the cross-linked hairpin denatures above 96°C and T_m values cannot be measured by uv spectroscopy.[51]

To investigate further the structural effects of the disulfide modification, the solution-phase conformations of **1** and **2b** were determined by nmr spectroscopy.[54] The eight base-pair long stem of these sequences adopts a B-form helix whereas the five base long single-stranded loop appears to be flexible and cannot be represented by a unique static conformation. Nuclear Overhauser effect spectroscopy (NOESY) cross-peak volumes, proton (both labile and nonlabile) and phosphorus chemical shifts, as well as both homo- and heteronuclear coupling constants for the cross-linked hairpin are virtually identical to those measured for the unmodified sequence, even for the residues that are proximal to the cross-link. Thus, within the resolution of nmr spectroscopy, the two hairpins are structurally isomorphous.

Similar to the results described above for the ColE1 hairpin, replacing one or both of the terminal dG · dC bases in $d(CGCGAATTCGCG)_2$ with our modified thymidine does not significantly alter the thermal stability of these duplexes relative to the parent sequence.[53,55] The exchangeable and nonexchangeable protons along with the ^{31}P resonances of **4b** and **5b** were assigned and compared to those data for the parent duplex. We find that the spectra of **5b** are virtually identical to those obtained for **3** which strongly indicates that these sequences also adopt very similar structures (because **4b** is not symmetric, a direct comparison with either **3** or **5b** is not possible). To provide further evidence of this point, two-dimensional (2D) NOESY spectra of **5b** as a function of mixing time were measured and interproton distances were obtained from initial buildup plots. These data for **5b** are nearly identical to those for the wild-type dodecamer. Based on these findings, along with the results of CD and nuclease-susceptibility experiments, we conclude that our alkylthiol modifications, either on the disulfide cross-linked or reduced and protected forms, do not alter native structure.[53]

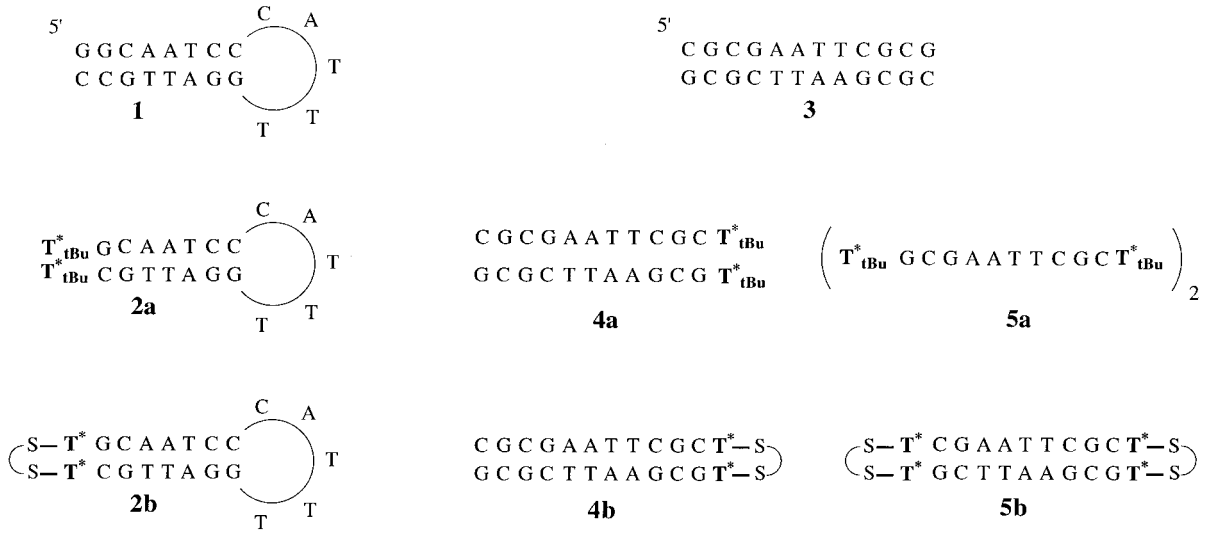


FIGURE 2 Duplex sequences and cross-linked analogs.

Thermodynamic Measurements

Differential scanning calorimetry (DSC) measurements were conducted in collaboration with K.J. Breslauer and co-workers to characterize the thermally induced denaturation of duplexes **3–5.** Our goal was to elucidate the thermal and thermodynamic consequences of modifying and constraining DNA via our disulfide chemistry.[53] These experiments represented the first use of calorimetry to investigate the thermodynamic consequences of constraining DNA with disulfide cross-links. The use of calorimetry is critical here because **3** denatures in a non-two-state fashion and cannot be analyzed using a van't Hoff treatment. Consistent with our nmr data, a large energetic penalty is not incurred by replacing the terminal base pair(s) with N3-(ethylthiol)thymidine: the melting temperature (T_m) of **3, 4a,** and **5a** are ± 1°C of each other. The overall free energy for **3, 4a,** and **5a** are also comparable, and this similarity arises due to compensations in both enthapy and entropy.[55]

The DSC measurements suggest that introducing cross-link(s) into **3** results in two changes. First, constraining the dodecamer results in a significant increase in thermal stability. A single disulfide cross-link changes the molecularity of the complex from bimolecular to monomolecular. However, a second disulfide cross-link results in a constrained conformation with a reduced entropy compared to **4b.** From the DSC data, we estimate that the cross-link imparts about 3 kcal/mol of stabilization. The change in entropy certainly reflects differences in the native and/or denatured states of **5b** as compared to **4b.** On first inspection, it would appear that the most likely source for the observed decrease in entropy is the denatured state, since the conformational freedom of the denatured state of **5b** is less than that of **4b.** However, based on the data, we cannot exclude differential entropic contributions from the native states, as well as influences from differential solvation in both the initial and the final states. Notwithstanding, our data show that the entropy term is responsible for the increase in thermal stability of **5b** relative to **4b.** To our knowledge, this is the first conformationally constrained nucleic acid system where the increase in melting temperature is predominantly due to a decrease in entropy.[35,36]

Dynamics Measurements

The static structure of DNA can explain many aspects of its function and properties. However, local base-pair opening is implicated in a number of important chemical, biological, and mechanical processes involving DNA.[56–60] Hence, defining the dynamics of base-pair opening is necessary to fully understand the physiochemical and biological properties of DNA. Although the opening kinetics for several different constructs have been measured, the opening kinetics for oligonucleotides constrained with cross-links (e.g., hairpin loops, glycol bridges, etc.) have not been reported.[60] Therefore, we chose to measure the base-pair lifetimes and apparent dissociation constants of **1** and **2b** (Figure 2).[61] Comparison of the lifetimes and apparent dissociation constants for corresponding base pairs of the two hairpins indicates that the cross-link neither increases the number of base pairs in-

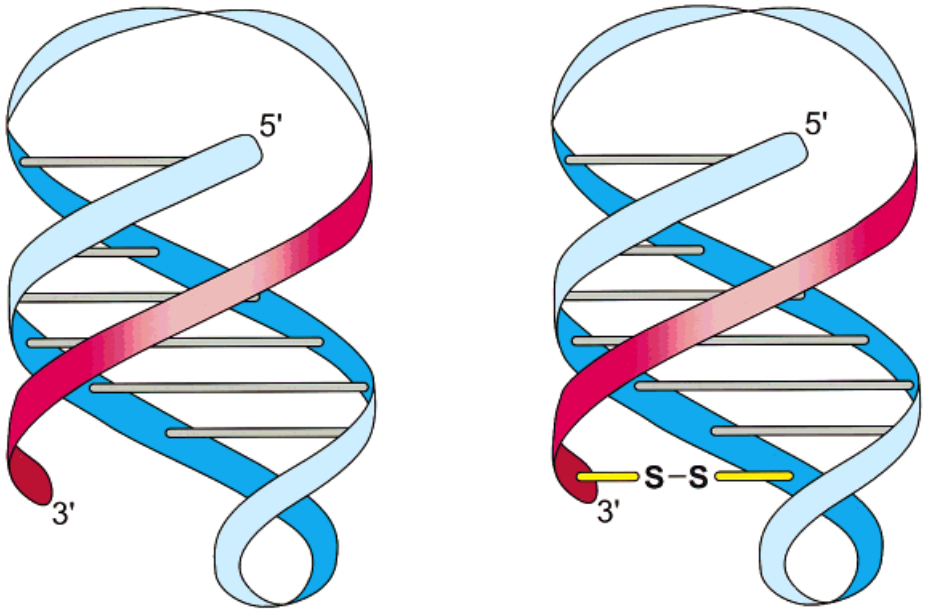


FIGURE 3 Triplex schematic.

volved in fraying nor alters the lifetime, dissociation constant, or the opened structure from which exchange occurs for the base pairs that are not frayed. The cross-link does, however, stabilize the frayed penultimate base pair of the stem duplex by increasing the closing rate of this base pair. Significantly, the disulfide cross-link is more effective at preventing fraying than the five base long hairpin loop.

DISULFIDE CROSS-LINKED DNA TRIPLEXES

Design

Folding of DNA triple helices based on the pyrimidine · purine-pyrimidine motif (py · pu–py; dot = Hoogsteen and dash = Watson–Crick base pairing) is affected by a number of factors including sequence length (i.e., the number of triplets formed), composition, the presence of base-pair mismatches, as well as solution conditions including pH and mono- and divalent counterion concentrations.[62–65] For example, a major factor in the stability of triplexes that contain C^+ · G—C triplets is the necessity of protonation of the N3 position of cytosine, which generally limits the stability of small triple helices of this type to pH values below 7.[66,67] The narrow pH range required for folding of these triplexes has hampered efforts to assess the physical properties of this motif under physiological conditions. Such data are clearly needed to design triplex sequences of higher affinity and specificity for diagonsitic purposes in vitro and for use in vivo. While methods exist to stabilize py · pu–py triple helices, they rely on modifications that may alter the native triplex structure.[68,69] Therefore, we designed experiments to determine if covalently locking the third strand of a triplex to the major groove through a structurally nonperturbing disulfide cross-link would afford constructs that are stable under physiological conditions. In other words, can the constraint provided by a disulfide cross-link alter the apparent pK_a value of a triplex?

We chose to test this hypothesis using an intramolecular triplex rather than an intermolecular construct because the former should require only one cross-link to link the Hoogsteen strand to the Watson–Crick duplex (Figure 3).[70,71] To place a cross-link between the Hoogsteen and Watson–Crick strands, sites of chemical modification on the terminal bases must first be identified. In a T · T—A triplet, which is analogous to the base-pair substitution used to cross-link the terminus of duplexes (see Figure 1), the N3 position on the Hoogsteen thymidine and the C5 position of the Watson–Crick thymidine converge. Since the cross-link used for B-form duplexes is inappropriate here, we designed a triplex cross-link using a C5-alkyl thiol modified thymidine[72] (**C_3^5S;** where the subscript three is the number of atoms in the linker, including thiol) and N3-(ethylthiol)thymidine (Figure 4).

The optimal pH for air oxidation of thiols to form disulfide bonds generally is above the pK_a of the thiol of interest (~ 8.5). At the pK_a of many thiols, triple

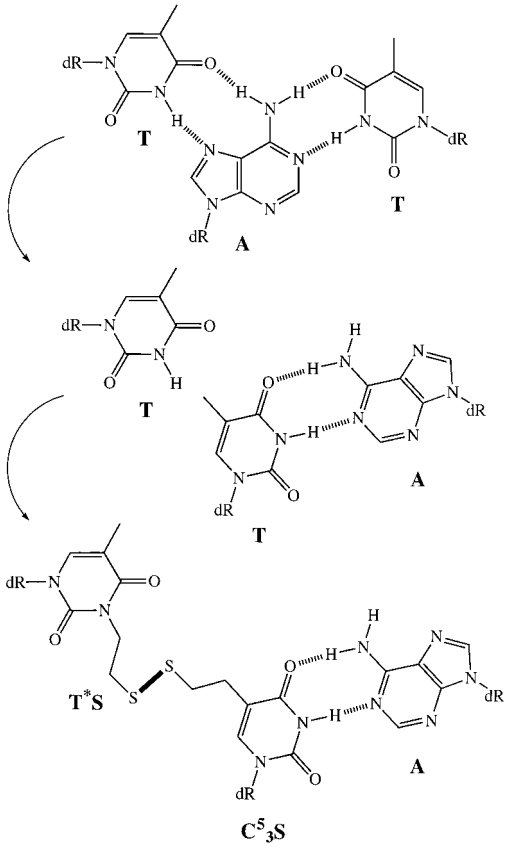


FIGURE 4 Chemical structure of triplex cross-link.

helices containing $C^+ \cdot G$—C base triplets are unfolded because the N3 position of cytosine is not protonated. To form a disulfide cross-link in a pyr · pur–pyr triple helix, it is necessary to identify a sequence that (partly) folds into a triplex at a pH where cross-link formation readily proceeds. Häner and Dervan reported a 34 base long intramolecular triple helix that remains partially folded up to pH 7.5 at 24°C in buffer containing 25 m*M* Mg^{2+}.[73] In principle, therefore, a disulfide cross-link can form under the conditions needed to fold a suitably modified variant of this oligonucleotide (Figure 5).

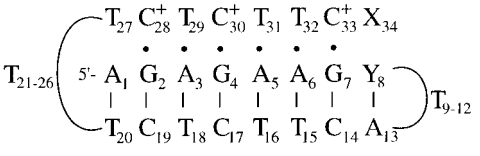
Preparation of triplex sequences containing $\mathbf{T^*_{tBu}}$ and $\mathbf{C^5_3StBu}$ was achieved via automated solid-phase DNA synthesis using standard protocols with an average stepwise coupling efficiency of > 98.5%. Removal of the phosphate and base protecting groups followed by reversed-phase HPLC purification provided **7a** in about 42% yield (based on a 1 μmol synthesis). To form the cross-link, the *tert*-butyl thiol protecting groups were removed with DTT and the reduced sample was vigorously stirred at room temperature in phosphate buffered saline (pH 7.25, 5 m*M* $MgCl_2$) while exposed to air. After 24 h, no starting material was observed by HPLC and aliquots from the reaction mixtures tested negative with Ellman's reagent. Cross-linked DNA was purified by reversed-phase HPLC to give **7b** in ~ 28% isolated overall yield (based on a 1 μmol synthesis; ~ 75% yield from **7a**). Enzymatic nucleoside composition analysis confirmed incorporation of the modified bases and formation of the cross-link, and native and denaturing PAGE indicated that the structure formed is monomeric.

Structural Studies

The imino proton spectra of **6** and **7b** were assigned using standard 2D techniques. Under conditions that favor triplex formation (pH 6, 0.5 m*M* Mg^{2+}), the imino proton spectra of both sequences are virtually identical, which indicates that our modifications do not disrupt native triplex structure.[71] When **6** or **7a** is titrated from pH 6 to 8, dissociation of the third strand is observed. By contrast, cross-linked triplex **7b** is conformationally stable over this pH range and only begins to unfold at pH ~ 9. Moreover, **7b** is stable from pH 6 to 8 in the absence of Mg^{2+}. Because of line broadening above 5°C, nmr cannot be used to study these triplexes. Therefore, we turned to CD spectroscopy. Monitoring the triplex CD signature band at 215 nm, we find that in the presence of Mg^{2+}, **7b** does not begin to melt until ~ 60°C at pH 6 (30°C for **6** and **7a**) and 40°C at pH 8 (<20°C for **6** and **7a**). Constructing titration curves from the CD data reveals that the apparent pK_a for **7b** is ~ 8.6 which is at least 1.5 pK_a units greater than **6** and **7a.** These results clearly demonstrate that our cross-link can be used to stabilize higher order DNA structures.

Thermodynamic Measurements

DSC measurements were conducted to elucidate the thermodynamic consequences of modifying and constraining DNA triplexes with our disulfide chemistry.[74] Cross-linked triplexes **6** and **7a** melt in a bipha-



6 : X, Y = T

7a: X = $\mathbf{T^*_{tBu}}$, Y = $\mathbf{C^5_3S\text{-}StBu}$ **7b**: X = $\mathbf{T^*S}$, Y = $\mathbf{C^5_3S}$ } disulfide cross-link

FIGURE 5 Triplex sequence.

sic manner above pH 6, with the initial triplex to duplex transition (Hoogsteen strand release) occurring at lower temperatures than melting of the hairpin. In contrast, cross-linking increases the thermal stability of the Hoogsteen transition such that the triplex and hairpin duplex denature simultaneously. Model-independent thermodynamic data obtained by DSC reveals the cross-link-induced increase in triplex thermal stability corresponds to a free energy stabilization of about 3 kcal/mol, with this stabilization being entirely entropic in origin. In other words, the cross-link is enthalpically neutral, but nevertheless induces a triplex stabilization of 3 kcal/mol due to a reduction in the entropy change associated with triplex melting. To deduce the origin(s) of this entropic impact, we measured the pH and ionic strength dependence of the melting transitions. From a comparison of the melting transitions at different pH values and ionic strengths, we estimate that 0.4 more protons are associated with the cross-linked triplex state than with the uncross-linked triplex, and 1.3 fewer counterions are released on melting the cross-linked triplex. Thus, the entropic stabilization is not solely a result of a reduction in conformationally entropy.

Dynamics Measurements

At the start of this project, it was not known whether the base pairs (Hoogsteen and Watson–Crick) that comprise triple-helical DNA open to any significant extent within a stable triplex. To address this question, the conformational dynamics of **6** and **7b** were studied by 2D exchange and NOE spectroscopy, and by measuring base-catalyzed imino-proton exchange rates.[75] Under conditions that promote triplex formation (pH 6.0, 1°C), **6** and **7b** exhibit a small and identical degree of conformational heterogeneity. However, at higher temperature (pH 6.0, 37°C) **6** exhibits more extensive heterogeneity than **7b.** The exchange times for Watson–Crick imino protons are $\sim$ 1 h for both triplexes. However, the Hoogsteen base-pair lifetimes of **6** could not be measured because this sequence is conformationally labile under the alkaline conditions necessary to conduct the exchange experiments.

Because of the extraordinary pH stability conferred by the cross-link, it is possible to measure the Hoogsteen lifetimes for **7b.** The lifetimes of these base pairs range from about 3–370 ms. Similar to Watson–Crick base pairs, the Hoogsteen lifetimes are highest at the central region within the triplex and taper off toward the termini, which is suggestive of imino exchange mechanisms involving end effects. In all cases, the lifetime of a Hoogsteen base pair is shorter than the Watson–Crick base pair contained in the same triplet, and is consistent with the greater stability of Watson–Crick over Hoogsteen base pairs. The rate of triplex to duplex conversion for **7b** at 1°C, pH 6.9, is low with an upper bound of $3.2 \times 10^{-4}\ s^{-1}$. The imino protons of **7b** exchange slowly in PBS with exchange times as long as 1 h, but the base-pair lifetimes are all less than 3 min, reflecting the fact that imino proton exchange is not opening-rate limited. Because the disulfide cross-link effectively prevents conformational heterogeneity associated with pyr · pur–pyr triple helices containing C^+ · G—C base triplets at neutral pH, constructs possessing this modification can serve as model systems to examine the structural and thermodynamic aspects of triplex formation in vitro, and to develop sequences that bind DNA with higher affinity and specificity.

APPLICATIONS OF DISULFIDE CROSS-LINKED DNA

Defining the biochemical, structural, dynamic, and thermodynamic impacts of constraining DNA with our disulfide chemistry as described above opens the way for applying these cross-links in a host of different experiments. Uses of this chemistry include, among others, the joining of large pieces of DNA or RNA through disulfide bonds for the construction of nanoscale architectures, the design of redox activated switches/devices, and the probing of secondary and tertiary structure. Below is described two representative applications of our chemistry. In the first, a disulfide bond was employed to stabilize or "trap" a DNA structure that would not be stable in its absence, and in the second, a disulfide bond was used to report on conformational transitions in a protein binding study.

Stabilizing Nonground-State Structures

In the absence of our disulfide cross-links, the molecules described in the second and third sections are themselves relatively stable with respect to conformational isomerization. In principle, however, it should also be possible to trap higher energy nonground-state structures with disulfide cross-links. Synthetic access to such constructs should be important in areas such as targeted drug delivery, protein–DNA recognition, and biophysical studies of alternate DNA geometries. We first demonstrated the practicality of using disulfide cross-links in such endeavors by trapping the Dickerson/Drew dodecamer premelting intermediate with a disulfide cross-link and studying the structural properties of this sequence.

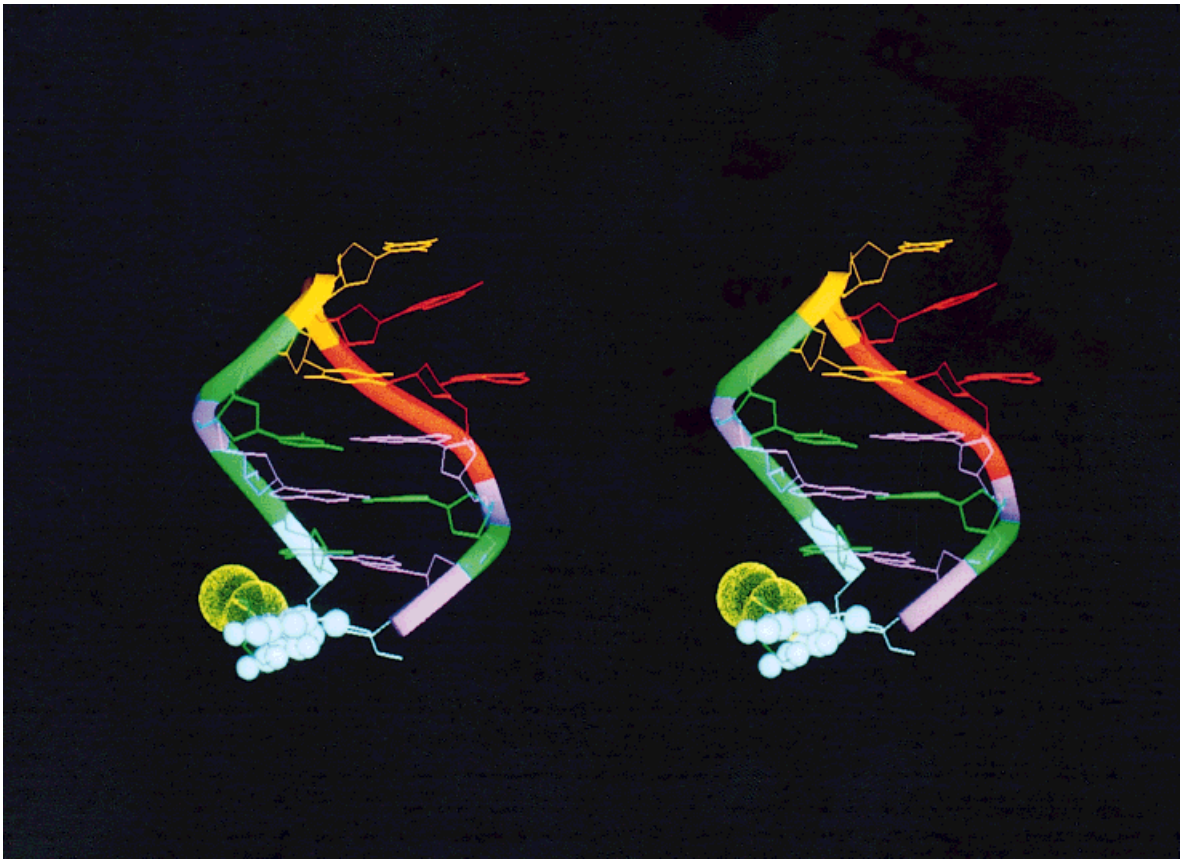


FIGURE 6 Stereoview of average structure of **8** from nmr refinement (yellow van der Waals surface: sulfur atoms; white CPK: thymidine residues that form the cross-link; green: cytosine; purple: guanosine; red: adenosine; brown: loop thymidine residues).

Both d(CGCGAATTCGCG)$_2$ and **5a** thermally denature in a biphasic manner in buffers containing [Na^+] $\leq$ 10 mM and with [DNA] $\leq$ 50 μM.[76] 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. To cross-link the premelting intermediate, we conducted air oxidation of the sulfhydryl groups in **5b** at 50°C in dilute solution in low salt buffer.[52] Unlike the stem–loop intermediate produced by initial melting of the parent dodecamer, **8** does not denature as a result of increasing temperature, [Na^+], or DNA concentration.[52,77]

To address the structural consequences of our cross-link, the solution-phase conformation of hairpin **8** was determined by nmr spectroscopy.[77,78] The stem region of this hairpin forms a B-form DNA duplex with a helical rise of 3.5 Å and a helical twist of 34.6° (Figure 6). 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 acts to close the loop. The conformation agrees with the "loop folding principle" advanced by Haasnoot, which predicts extension of the helix by three bases followed by a sharp bend in the loop.[79] The aromatic bases face into the major groove while the negatively charged backbone contacts solution. Significantly, the cross-link does not alter the geometry of the stem duplex.

Probes in Molecular Recognition Experiments

DNA ligands often undergo conformational changes upon binding to proteins. In the absence of x-ray data this "induced fit" can be difficult to study. In principle, the constraint imposed by the disulfide bond should provide a way to investigate conformational changes in DNA structure that occur upon protein recognition. To illustrate this point, we studied the binding of monoclonal antibody BV04-01 to disulfide cross-linked analogues of hairpins **9–12** (Figure 7).[80] This anti-DNA autoantibody was isolated from an autoimmune mouse that develops a syndrome related to human lupus, and BV04-01 may be involved in the

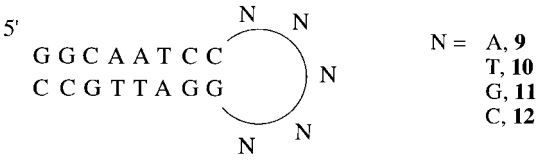


FIGURE 7 Hairpin ligands for binding study.

pathogenesis of this disorder in mice. BV04-01 binds only single-stranded DNA, which can be modeled by the loop region of DNA hairpins.[80]

If conformational reorganization of the hairpin ligands is required for binding, then BV04-01 should possess a *lower* affinity for the cross-linked sequences, since the disulfide bond renders them resistant to structural changes. However, if preorganization is important for complexation, then the more rigid oligomers should bind with equal or greater affinity than the unmodified ligands. When binding to cross-linked hairpins was measured, nearly a 100-fold increase in K_d relative to **9–12** was observed. If the weaker affinity of BV04-01 for the cross-linked molecules results from the constraint imposed by the cross-link rather than as a result of a structural perturbation introduced by the alkylthiol linkers, then removing this constraint by reduction of the disulfide bond should afford a set of ligands that bind with roughly the same affinity as the unmodified hairpins. Indeed, BV04-01 recognition of the reduced hairpins was indistinguishable from binding to **9–12.** DNA footprinting experiments showed that upon binding, residues within the duplex are recognized by single-strand specific reagents, which provides further evidence that the stem duplex of the hairpin ligands is partially denatured.[81]

DISULFIDE CROSS-LINKED RNA

Small oligoribonucleotides often exist in solution as conformationally heterogeneous mixtures, particularly at high RNA concentrations. For example, at high concentrations, even rather stable RNA hairpins can dimerize. As a means to prevent such conformational equilibria, we demonstrated that the uridine analogue of $\mathbf{T_{SH}}$, $\mathbf{U_{SH}}$, can be used to generate cross-links at the terminus of A-form duplexes (Figure 8).[82] As with DNA duplexes, cross-links with $\mathbf{U_{SH}}$ impart considerable thermal stability to small RNA constructs and effectively prevent both concentration induced aggregation and equilibration between alternate conformations. However, stabilizing small motifs for analysis is just one utility of our chemistry as it relates to RNA. Since disulfide bond formation via air oxidation is mild and highly selective, cross-links can be used to verify theoretical models of RNA tertiary structure. By varying the length of the alkyl thiol linkers, both dynamic motions and the distance between structural elements can be investigated. In addition, because of the linkage between disulfide bond formation and RNA stability, it is possible to examine the thermodynamics of RNA folding by thiol–disulfide exchange. Last, measuring the rate at which cross-linked RNAs fold and unfold should report on the structure of tertiary folding transition states.

As a first step toward using disulfide cross-links in these ways, we synthesized a set of thiol-modified nucleosides that have proven of general utility to cross-link RNA.[83–85] With these monomers, we prepared analogues of unmodified yeast $tRNA^{Phe}$ possessing a single disulfide cross-link (Figure 9).[83,85] Air oxidation to form these cross-links is quantitative, and analysis of the cross-linked products by footprinting and thermal denaturation experiments demonstrates that the disulfide bridges do not alter the structure of the thiol-modified tRNAs relative to the parent sequence. We are currently using these modified constructs to address a range questions regarding the kinetics and thermodynamics of tertiary folding.

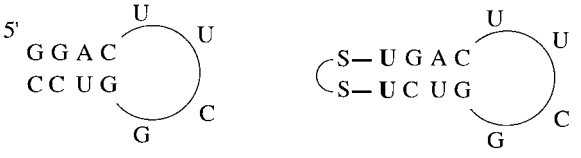


FIGURE 8 Sequence of RNA tetraloop hairpin and disulfide cross-linked analogue.

OTHER DISULFIDE CROSS-LINKED NUCLEIC ACIDS

In 1966 Lipsett showed that iodine-mediated oxidation of the two native 4-thiouridine residues in $tRNA^{Tyr}$ from *Escherichia coli* affords a unique intramolecular disulfide cross-link.[86] Although the potential of these disulfide cross-links as probes of nucleic acid structure and function was noted in this work, neither the methods to synthesize nucleic acids site specifically labeled with thiol groups, nor the biochemical and structural assays needed to analyze the cross-linked products, were available. It was not until early in 1991, several months prior to our first paper in this area, that Ferenz and Verdine reported synthesizing the first disulfide cross-linked oligonucleotide.[87,88] Using convertible nucleoside chemistry,

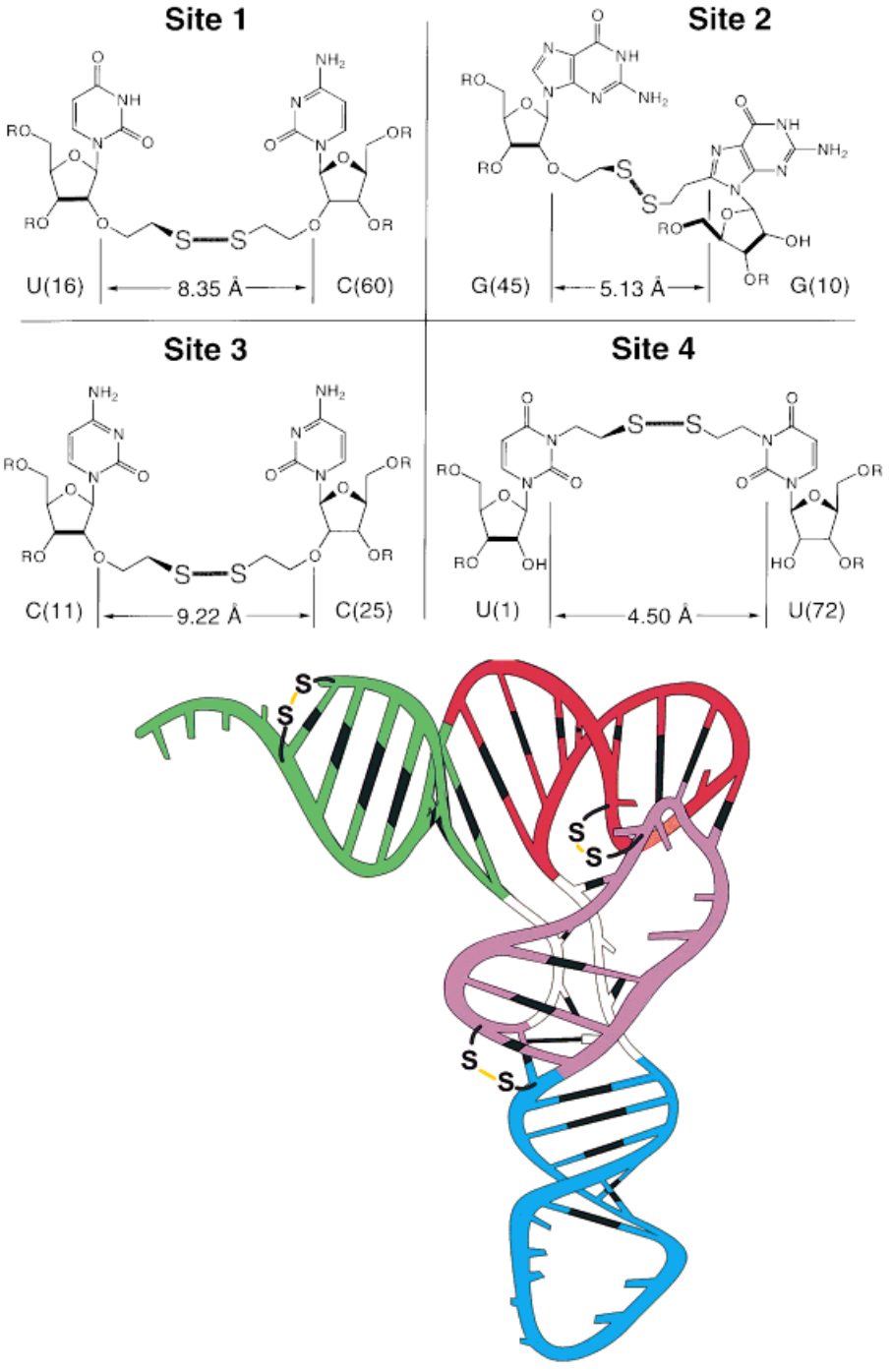


FIGURE 9 (Top) Chemical structure of RNA cross-links. (Bottom) Yeast $tRNA^{Phe}$ crystal structure showing the locations of the cross-links. Green: aminoacyl acceptor stem; purple: dihydrouridine stem and loop; blue: anticodon stem and loop; red: TΨC stem and loop. Site 1: upper right; Site 3: middle left; Site 4: upper left. The site 2 cross-link is not visible in this orientation.

they showed that placing N^6-thioalkyl derivatives of 2′-deoxyadenosine in consecutive base pairs on opposite strands of a duplex afforded disulfide cross-links upon air oxidation. These cross-links reside in the major groove and impart increased stability with minimal distortion of native DNA geometry. This

chemistry was extended several years later to the synthesis of minor groove cross-links using modified dC residues. Verdine and co-workers have used these major and minor groove cross-links with excellent effect in variety of ways, including, studies of protein · DNA interactions,[89,90] stabilization of intrinsically bent DNA,[91] the apparent induction of torsional stress in short oligonucleotides,[92] and cross-linking Z-DNA.[93] More recently, this chemistry has been adapted to synthesize disulfide cross-linked RNA molecules,[94,95] thus making it possible to conduct the type of experiments described above in the fifth section.

Several other reports of incorporating disulfide cross-links into DNA and RNA have appeared over the past few years.[32,33,96–100] Many of the chemistries developed have allowed previously unapproachable problems to be addressed. In a very interesting study on ribozyme dynamics, Cohen and Cech demonstrated that thiol groups can be incorporated into very large RNA molecules through synthesis and ligation.[100] This ability to construct large molecules modified with thiol groups at specific residues, coupled with the fact that virtually every available position on the base and sugar can be labeled with thiol groups[82–85,94,95,101–103] will facilitate the continued use of disulfide cross-links in the study of RNA, and nucleic acids in general.

I wish to thank the talented group of associates who have so skillfully conducted the research in my lab described in this article. I also wish to thank K. J. Breslauer, Jens Völker, and E. R. P. Zuiderweg, whose expertise and insight into nucleic acids had formed the basis for outstanding collaborations. The worked described here was supported by NIH Grants GM-52831 and GM-43861. Funding from the National Arthritis Foundation, the American Cancer Society, the National Science Foundation, the Camille and Henry Dreyfus Foundation, and the Alfred P. Sloan Foundation is also gratefully acknowledged.

REFERENCES

1. Baldwin, R. L. (1971) *Acc. Chem. Res.* **4,** 265–272.
2. Elson, E. L., Scheffler, I. E. & Baldwin, R. L. (1970) *J. Mol. Biol.* **54,** 401–415.
3. Scheffler, I. E., Elson, E. L. & Baldwin, R. L. (1970) *J. Mol. Biol.* **48,** 145–171.
4. Breslauer, K. J. (1986) in *Thermodynamic Data for Biochemistry and Biotechnology*, Hinz, H. J., Ed., Springer Verlag, New York, pp. 402–427.
5. Breslauer, K. J., Frank, R., Blöcker, H. & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3746–3750.
6. SantaLucia, J., Jr., Allawi, H. T. & Seneviratne, P. A. (1996) *Biochemistry* **35,** 3555–3562.
7. Record, M. T., Jr. & Lohman, T. M. (1978) *Biopolymers* **17,** 159–166.
8. Olmsted, M. C., Anderson, C. F. & Record, M. T., Jr. (1991) *Biopolymers* **31,** 1593–1604.
9. MacGregor, R. B., Jr. (1996) *Biopolymers* **38,** 321–327.
10. Pinto, A. L. & Lippard, S. (1985) *J. Biochem. Biophys. Acta* **780,** 167–180.
11. Borowy-Borowski, H., Lipman, R. & Tomasz, M. (1990) *Biochemistry* **29,** 2999–3006.
12. Kirchner, J. J. & Hopkins, P. B. (1991) *J. Am. Chem. Soc.* **113,** 4681–4682.
13. Boger, D. L., Munk, S. A. & Ishizaki, T. (1991) *J. Am. Chem. Soc.* **113,** 2779–2780.
14. Sigurdsson, S. T., Rink, S. M. & Hopkins, P. B. (1993) *J. Am. Chem. Soc.* **115,** 12633–12634.
15. Willis, M. C., Hicke, B. J., Uhlenbeck, O. C., Cech, T. R. & Koch, T. H. (1993) *Science* **262,** 1255–1257.
16. Millard, J. T., Raucher, S. & Hopkins, P. B. (1990) *J. Am. Chem. Soc.* **112,** 2459–2460.
17. Kirchner, J. J., Sigurdsson, S. T. & Hopkins, P. B. (1992) *J. Am. Chem. Soc.* **114,** 4021–4027.
18. Millard, J. T., Weidner, M. F., Kirchner, J. J., Ribeiro, S. & Hopkins, P. B. (1991) *Nucleic Acids Res.* **19,** 1885–1891.
19. Tomasz, M., Lipman, R., Chowdary, D., Pawlak, J., Verdine, G. L. & Nakanishi, K. (1987) *Science* **235,** 1204–1208.
20. Teng, S. P., Woodson, S. A. & Crothers, D. M. (1989) *Biochemistry* **28,** 3901–3907.
21. Sherman, S. E., Gibson, D., Wang, A. H. & Lippard, S. J. (1985) *Science* **230,** 412–417.
22. Lemaire, M.-A., Schwartz, A., Rahmouni, A. R. & Leng, M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1982–1985.
23. Webb, T. R. & Matteucci, M. D. (1986) *J. Am. Chem. Soc.* **108,** 2764–2765.
24. Cowart, M., Gibson, K. J., Allen, D. J. & Benkovic, S. J. (1989) *Biochemistry* **28,** 1975–1983.
25. Catalano, C. E. & Benkovic, S. J. (1989) *Biochemistry* **28,** 4374–4382.
26. Cowart, M. & Benkovic, S. J. (1991) *Biochemistry* **30,** 788–796.
27. Durand, M., Chevrie, K., Chassignol, M., Thuong, N. T. & Maurizot, J. C. (1990) *Nucleic Acids Res.* **18,** 6353–6359.
28. Kool, E. T. (1991) *J. Am. Chem. Soc.* **113,** 6265–6266.
29. Rumney, S., IV & Kool, E. T. (1992) *Angew. Chem. Int. Ed. Engl.* **31,** 1617–1619.
30. Salunkhe, M., Wu, T. & Lestinger, R. L. (1992) *J. Am. Chem. Soc.* **114,** 8768–8772.
31. Bannwarth, W., Dorn, A., Iaiza, P. & Pannekouke, X. (1994) *Helv. Chim. Acta* **77,** 182–193.
32. Gao, H., Chidambaram, N., Chen, B. C., Pelham, D. E., Patel, R., Yang, M., Zhou, L., Cook, A. & Cohen, J. S. (1994) *Bioconj. Chem.* **5,** 445–453.

33. Gao, H., Yang, M. & Cook, A. F. (1995) *Nucleic Acids Res.* **23,** 285–292.
34. Williams, D. J. & Hall, K. B. (1996) *Biochemistry* **35,** 14665–14670.
35. Erie, D. A., Jones, R. A., Olson, W. K., Jones, R. A. & Breslauer, K. J. (1987) *Biochemistry* **26,** 7150–7159.
36. Erie, D. A., Sinha, N. K., Olson, W. K., Jones, R. A. & Breslauer, K. J. (1989) *Biochemistry* **28,** 268–273.
37. Wemmer, D. E. & Benight, A. S. (1985) *Nucleic Acids Res.* **13,** 8611–8621.
38. Benight, A. S., Schurr, J. M., Flynn, P. F., Reid, B. R. & Wemmer, D. E. (1988) *J. Mol. Biol.* **200,** 377–399.
39. Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J. & Benight, A. S. (1992) *Biopolymers* **32,** 849–864.
40. Amaratunga, M., Snowden-Ifft, E., Wemmer, D. E. & Benight, A. S. (1992) *Biopolymers* **32,** 865–879.
41. Paner, T. M., Amaratunga, M. & Benight, A. S. (1992) *Biopolymers* **32,** 881–892.
42. Ashley, G. W. & Kushlan, D. M. (1991) *Biochemistry* **30,** 2927–2933.
43. Germann, M. W., Schoenwaelder, K.-H. & van de Sande, J. H. (1985) *Biochemistry* **24,** 5698–5702.
44. Clarke, J. & Fersht, A. R. (1993) *Biochemistry* **32,** 4322–4329, and references therein.
45. Patel, D. J. & Hilbers, C. W. (1975) *Biochemistry* **14,** 2651–2656.
46. Aboul-ela, F., Koh, D. & Tinoco, I., Jr. (1985) *Nucleic Acids Res.* **13,** 4811–4824.
47. Frier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986) *Biochemistry* **25,** 3209–3213.
48. Blatt, N. B., Cain, R. J., Osborne, S. E. & Glick, G. D. (1993) *Biochime* **75,** 433.
49. Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K. & Dickerson, R. E. (1980) *Nature* **287,** 755–758.
50. Drew, H. R. & Dickerson, R. E. (1981) *J. Mol. Biol.* **151,** 535–556.
51. Glick, G. D. (1991) *J. Org. Chem.* **56,** 6746–6747.
52. Glick, G. D., Osborne, S. E., Knitt, D. S. & Marino, J. P., Jr. (1992) *J. Am. Chem. Soc.* **114,** 5447–5448.
53. Osborne, S. E., Völker, J., Stevens, S. Y., Breslauer, K. J. & Glick, G. D. (1996) *J. Am. Chem. Soc.* **118,** 11993–12003.
54. Cain, R. J., Zuiderweg, E. R. P. & Glick, G. D. (1995) *Nucleic Acids Res.* **23,** 2153–2160.
55. Osborne, S. E. (1996) Ph.D. thesis, University of Michigan.
56. Ramstein, J. & Lavery, R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7231–7235.
57. Ramstein, J. & Lavery, R. (1990) *J. Biomol. Struct. Dynam.* **7,** 915–933.
58. Frank-Kamenetskii, M. D. (1985) in *Structure and Motion: Membranes, Nucleic Acids, and Proteins,* Clementi, E., Corongiu, G., Sarma, M. H. & Sarma, R. H. Eds., Adenine Press, Guilderland, NY, pp. 417–432.
59. Tari, L. W. & Secco, A. S. (1995) *Nucleic Acids Res.* **23,** 2065–2073.
60. Guéron, M. & Leroy, J. L. (1995) *Methods Enzymol.* **261,** 383–413.
61. Cain, R. J. & Glick, G. D. (1997) *Nucleic Acids Res.* **25,** 836–842.
62. Plum, G. E., Pilch, D. S., Singleton, S. F. & Breslauer, K. J. (1995) *Ann. Rev. Biophys. Biomol. Struct.* **24,** 319–350.
63. Kiessling, L. L., Griffin, L. C. & Dervan, P. B. (1992) *Biochemistry* **31,** 2829–2834.
64. Pilch, D. S., Levenson, C. & Shafer, R. H. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 1942–1946.
65. Moser, H. E. & Dervan, P. B. (1987) *Science* **238,** 645–650.
66. Singleton, S. F. & Dervan, P. B. (1992) *Biochemistry* **31,** 10995–11003.
67. Callahan, D. E., Trapane, T. L., Miller, P. S., Ts'o, P. O. P. & Kan, L.-S. (1991) *Biochemistry* **30,** 1650–1655.
68. Sun, J.-S. & Hélène, C. (1993) *Curr. Opin. Struct. Biol.* **3,** 345–356.
69. Thuong, N. T. & Hélène, C. (1993) *Angew. Chem. Int. Ed. Engl.* **32,** 666–690.
70. Goodwin, J. T., Osborne, S. E., Swanson, P. C. & Glick, G. D. (1994) *Tetrahed. Lett.* **35,** 4527–4531.
71. Osborne, S. E., Cain, R. J. & Glick, G. D. (1997) *J. Am. Chem. Soc.* **119,** 1171–1182.
72. Goodwin, J. T. & Glick, G. D. (1993) *Tetrahed. Lett.* **34,** 5549.
73. Häner, R. & Dervan, P. B. (1990) *Biochemistry* **29,** 9761–9765.
74. Völker, J., Osborne, S. E., Glick, G. D. & Breslauer, K. J. (1997) *Biochemistry* **36,** 756–767.
75. Cain, R. J. & Glick, G. D. (1998) *Biochemistry* **37,** 1456–1464.
76. Marky, L. A., Blumenfeld, K. S., Kozlowski, S. & Breslauer, K. J. (1983) *Biopolymers* **22,** 1247–1257.
77. Wang, H., Osborne, S. E., Zuiderweg, E. R. P. & Glick, G. D. (1994) *J. Am. Chem. Soc.* **116,** 5021–5022.
78. Wang, H., Zuiderweg, E. R. P. & Glick, G. D. (1995) *J. Am. Chem. Soc.* **117,** 2981–2991.
79. Hassnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N. & Kollman, P. A. (1986) *J. Biomol. Struct. Dynam.* **3,** 843–857.
80. Stevens, S. Y., Swanson, P. C., Voss, E. W., Jr. & Glick, G. D. (1993) *J. Am. Chem. Soc.* **115,** 1585–1586, and references therein.
81. Swanson, P. C., Cooper, B. C. & Glick, G. D. (1994) *J. Immunol.* **152,** 2601–2612.
82. Goodwin, J. T. & Glick, G. D. (1994) *Tetrahed. Lett.* **35,** 1647–1651.
83. Goodwin, J. T., Scholle, E., Osborne, S. E. & Glick, G. D. (1996) *J. Am. Chem. Soc.* **118,** 5207–5215.
84. Gundlach, W. C., Ryder, T. & Glick, G. D. (1997) *Tetrahed. Lett.* **38,** 4039–4042.

85. Maglott, E. J. & Glick, G. D. (1998) *Nucleic Acids Res.* **26,** 1301–1308.
86. Lipsett, M. N. (1966) *Cold Spring Harb. Symp. Quant. Biol.* **31,** 449–455.
87. Ferentz, A. E. & Verdine, G. L. (1991) *J. Am. Chem. Soc.* **113,** 4000–4002.
88. Ferentz, A. E., Keating, T. A. & Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 9006–9014.
89. Erlanson, D. A., Wolfe, S. A., Chen, L. & Verdine, G. L. (1997) *Tetrahedron* **53,** 12041–12056.
90. Erlanson, D. A., Chen, L. & Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 12583–12584.
91. Wolfe, S. A., Ferentz, A. E., Grantcharova, V., Churchill, M. E. A. & Verdine, G. L. (1995) *Chem. Biol.* **2,** 213–221.
92. Wolfe, S. A. & Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 12585–12586.
93. Erlanson, D. A., Glover, J. N. M. & Verdine, G. L. (1997) *J. Am. Chem. Soc.* **119,** 6927–6928.
94. Allerson, C. R. & Chen, S. L. (1997) *J. Am. Chem. Soc.* **119,** 7423–7433.
95. Allerson, C. R. & Verdine, G. L. (1995) *Chem. Biol.* **2,** 667–675.
96. Milton, J., Connolly, B. A., Nikforov, T. T. & Cosstick, R. (1993) *J. Chem. Soc. Chem. Commun.* 779–780.
97. Sigurdsson, S. Th., Tuschl, T. & Eckstein, F. (1995) *RNA* **1,** 575–583.
98. Chaudhuri, N. C. & Kool, E. T. (1995) *J. Am. Chem. Soc.* **117,** 10434–10442.
99. Earnshaw, D. J., Masquida, B., Muller, S., Sigurdsson, S. Th., Eckstein, F., Westhof, E. & Gait, M. J. (1997) *J. Mol. Biol.* **274,** 197–212.
100. Cohen, S. B. & Cech, T. R. (1997) *J. Am. Chem. Soc.* **119,** 6259–6268.
101. Douglas, M. E., Beijer, B. & Sproat, B. S. (1994) *Bioorg. Med. Chem. Lett.* **4,** 995–1000.
102. Manoharan, M., Johnson, L. K., Tivel, K. L., Springer, R. H. & Cook, P. D. (1993) *Bioorg. Med. Chem. Lett.* **3,** 2765–2770.
103. Sun, S., Tang, X.-Q., Merchant, A., Anjaneyulu, P. S. R. & Piccirilli, J. A. (1996) *J. Org. Chem.* **61,** 5708–5709.