Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 7327

www.rsc.org/obc

PAPER

Solution structure of S-DNA formed by covalent base pairing involving a disulfide bond†

Akihiko Hatano,*[a] Munehiro Okada[a] and Gota Kawai[b]

Received 18th February 2012, Accepted 12th July 2012
DOI: 10.1039/c2ob25346a

Here, we present the solution structure of a DNA duplex containing a disulfide base pair (S-DNA). The unnatural nucleoside "**S**" possessing a thiophenyl group as base was incorporated into a self-complementary singled-stranded oligonucleotide. Crosslinking of the disulfide base pair was analyzed by non-denaturing polyacrylamide gel electrophoresis. Under oxidizing conditions a high molecular weight band as 18 mer, corresponding to the double-stranded molecule (5′-GCGA**S**TCGC: 3′-CGCT**S**AGCG), was found, whereas single-stranded self-complementary 9 mer oligonucleotide GCGA**S**TCGC was detected in the presence of a reducing agent. These results suggest that the oligonucleotide is covalently linked by disulfide bonding under oxidizing conditions, which can be reversibly reduced to two thiol groups under reducing conditions. CD spectrum of S-DNA (CGA**S**TCG) under oxidizing conditions suggested that the duplex had a right-handed double-stranded structure similar to that of natural DNA (B-form, CGATCG). NMR studies confirmed that this CGA**S**TCG resembled natural B-DNA and that the two phenyl rings derived from the disulfide base pairing intercalated into the duplex. However, these two phenyl rings were not positioned in the same plane as the other base pairs. Specifically, NOEs suggest that although CGA**S**TCG adopts a structure similar to B-type DNA, the S-DNA duplex is bent at the point of disulfide base pairing to face the major groove.

Introduction

The high fidelity of molecular recognition displayed by DNA is derived from the assembly of complementary hydrogen bonding of the bases between single stranded oligonucleotides. Because hydrogen bonding and base stacking are weak interactions the DNA double strand can reversibly dissociate and associate according to thermodynamics. This reversible dissociation/association of DNA facilitates replication and gene transcription. Recently, many research groups have created functional materials by utilizing the specific properties and structural characteristics of DNA.[1,2] For example, these functional molecules can be used as genetic diagnostic agents,[3] antisense reagents,[4,5] sensor devices[6,7] as well as for the fabrication of nanostructures[8,9] and various other applications.[10,11] Many research groups are currently developing artificial base pairs that rely on both hydrogen bonding[12] and non-hydrogen bonding *e.g.*, hydrophobic interaction,[13,14] shape recognition into a duplex,[15] metal coordination,[16–18] reversible bonding[19,20] and expanded-size base pairing.[21,22] These efforts have led to the discovery of a number of alternative base pairings, which could prove very useful in terms of generating functionalized DNA[23] or expanding the genetic code.[24–26]

Oligonucleotides that form interstrand crosslinks have been extensively investigated in the field of cancer therapy.[27] The covalent bond between the two single-stranded oligonucleotides prohibits the separation of the duplex and thereby blocks vital aspects of DNA metabolism. Although there are many examples of covalently crosslinked base pairings,[28–31] duplexes linked by base pairs with disulfide bonding are unique in terms of reversibility.[32–34] Disulfide bonds, which are formed between two cysteine residues in natural proteins, play an important role in stabilizing the tertiary structure of many proteins *in vivo*. The disulfide bond is covalent and can be readily converted to two SH groups in the presence of a reducing agent. We recently described the characteristics of S-DNA in which base pairing involves disulfide bonding.[20] Although S-DNA displays a high melting temperature its stability can be controlled by the use of redox reagents. We have previously studied the thermodynamic parameters for the formation of a duplex possessing disulfide base pair(s).[35] The formation of a disulfide-linked base pair in the DNA resulted in a gain in enthalpy (ΔH) but a drop in entropy (ΔS) owing to structural restrictions. The high ΔH value

[a]*Department of Chemistry, Shibaura Institute of Technology, 307 Fukasaku, Minuma-ku, Saitama, 337-8570, Japan. E-mail: a-hatano@sic.shibaura-it.ac.jp; Fax: +81-48-687-5013; Tel: +81-48-687-5035*
[b]*Department of Life and Environmental Sciences, Faculty of Engineering, Chiba Institute of Technology, 2-17-1 Tsudanuma, Narashino, Chiba 275-0016, Japan*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c2ob25346a

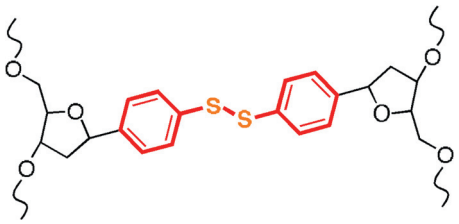

Fig. 1 Schematic representation of the disulfide base pair incorporated into the DNA.

was derived from the strong interaction as a result of covalent disulfide bonding, which considerably offset the less favorable ΔS due to structural constraints. We believe that S-DNA has inherent structural stress as a result of formation of the disulfide bond. Specifically, we were interested in investigating the overall structure of S-DNA, including the shape of the major and minor groove as well as the stacking pattern.

In this report, we have used NMR to examine the solution structure of S-DNA possessing a disulfide base pairing (Fig. 1). Specifically, we investigated the crosslinked structure of a seven-base self-complementary DNA molecule containing an unnatural nucleoside possessing a thiophenyl group located at the central position of the oligonucleotide. Under suitable redox conditions each strand of S-DNA was connected *via* a covalent bond. Our studies revealed the overall structure of the duplex resembled that of B-form DNA although it was slightly distorted by the presence of the disulfide base pair.

Results

1. Urea-containing DNA-PAGE

The crosslinking of the disulfide base pair was analyzed by urea-containing polyacrylamide gel electrophoresis (Fig. 2). Lanes 2 and 5 of Fig. 2 show normal 15-mer oligonucleotide in the presence of iodine as oxidizing reagent or in the presence of 2-mercaptoethanol as reducing reagent, respectively. These lanes confirm the presence of a 15-mer single stranded oligonucleotide. We also tested the formation of disulfide base pairing of a non-complementary single stranded oligonucleotide, TACAACA**S**TAATGTG under oxidizing conditions (Fig. 2, lane 3). Two bands are visible in lane 3 of Fig. 2, which are derived from a higher molecular weight band corresponding to a 30-mer oligonucleotide and a low molecular weight band corresponding to a 15-mer oligonucleotide. The main upper band indicates the presence of a homo complex formed from non-specific intermolecular disulfide bonding. The lower minor band indicates the presence of a single stranded oligonucleotide that has not formed a disulfide bond. Under reducing conditions, urea-containing PAGE of TACAACA**S**TAATGTG gave a band corresponding to single stranded oligonucleotide (Fig. 2, lane 6). Interestingly, the self-complementary 9-mer oligonucleotide CGAA**S**TTCG formed an 18-mer duplex due to disulfide bonding under oxidizing conditions (Fig. 2, lane 4). In the presence of reducing reagent the CGAA**S**TTCG oligonucleotide did not complex at room temperature and gave a low molecular weight band corresponding to a 9-mer (Fig. 2, lane 7). The intensity of the band in

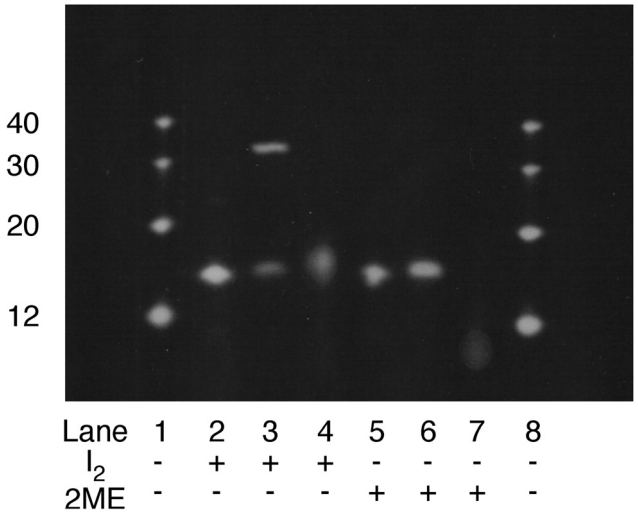

Fig. 2 7 M urea-containing DNA-PAGE to monitor formation of base pairing and disulfide bonding. Lane 1, size markers (12-, 20-, 30-, 40-mer ODN); lane 2, TACAACATTAATGTG 15-mer under oxidizing conditions; Lane 3, TACAACA**S**TAATGTG 15-mer under oxidizing conditions; lane 4, GCGA**S**TCGC 9-mer (self-pairing sequence) run under oxidizing conditions; lanes 5–7 are the same as lanes 2–4, respectively, except the samples were prepared under reducing conditions; lane 8, same as lane 1. All samples analyzed under oxidizing conditions were prepared in the presence of 0.5 mM of I_2. All samples analyzed under reducing conditions were prepared in the presence of 10 mM 2-mercaptoethanol.

lane 7 was relatively faint due the single-stranded nature of the short oligonucleotide (9-mer). Importantly, the sequence of the 9-mer possesses a mispair under these reducing conditions because the reduced disulfide match-pair gives a thiophenyl group on each strand.

2. Thermal denaturation

Fig. 3a shows the thermal denaturation profiles of self-pair DNAs, CGATCG, CGA**S**TCG in the presence and absence of 2-mercaptoethanol (100 μM). The T_m value at 30 mM bp^{-1} of 6-mer CGATCG was 25.4 °C. For duplex CGA**S**TCG containing the unnatural base **S** under oxidized conditions, the value of T_m was 65.8 °C. The unnatural S–S base pair formed a matching disulfide bond under oxidizing conditions, which made the duplex considerably more stable compared to the control DNA, CGATCG (T_m = 25.4 °C). In the presence of reducing agent (100 μM of 2-mercaptoethanol), the T_m of CGA**S**TCG could not be determined. This observation suggests that the T_m of CGA**S**TCG was very low under these conditions. Indeed, in the presence of a reducing agent each thiophenyl group on the two strands generates a mismatching base pair.

3. Circular dichroism

Circular dichroism spectroscopy was used to study the macroscopic helical geometry of the DNA duplexes. All CD spectra were analyzed at 0 °C and data are shown in Fig. 3b. Under oxidizing conditions, the CD spectra (red) of CGA**S**TCG, which

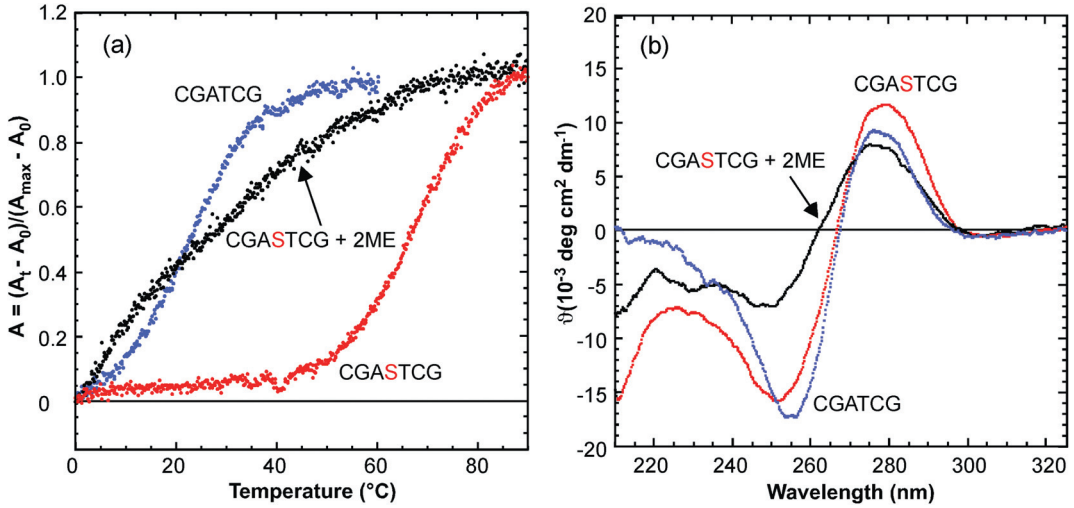

Fig. 3 Analysis of the stability of the DNA. Melting curves at 260 nm from 0 to 90 °C (a) and CD spectra (b) for the duplexes induced by the formation of a disulfide bond. Red: CGA**S**TCG, black: CGA**S**TCG in the presence of 100 μM of 2-mercaptoethanol, blue: CGATCG. These studies were performed at an oligomer concentration of 30 μM for the base-pairing in a buffer containing 1.0 M NaCl and 10 mM sodium phosphate, pH 6.8. 2-Mercaptoethanol (final concentration 100 μM) was added to the sample solution under reducing conditions.

contains a disulfide bond, revealed specific features of B-DNA that are similar to those of the control natural duplex (blue, CGATCG). The overall spectrum of CGA**S**TCG showed it to comprise a right-handed helix structure. The duplex CGA**S**TCG, which had a disulfide base pair at the central position, displayed a slight increase in positive Cotton effect at 276 nm and a corresponding decrease in negative Cotton effect at 254 nm relative to the control duplex CGATCG, which does not possess the **S** nucleoside. Under these reducing conditions, CGA**S**TCG could not generate a duplex that possessed a central mispair –SH HS– group on each strand. Hence, the T_m of CGA**S**TCG could not be observed between 0 °C and 90 °C in the presence of 2-mercaptoethanol (Fig. 3a).

4. NMR spectra

The symmetry of the self-complementary sequence results in four unique base pairs of duplex with seven chemically distinct nucleotides ($C^1G^2A^3\mathbf{S}^4T^5C^6G^7$). The ^{1}H NMR spectra of the imino proton gave further insight into several structural features of the DNA duplex. Three lowfield imino proton signals were observed for the duplex form of S-DNA (Fig. 4). These signals were relatively sharp and intense, implying that the base pairs were stabilized by the formation of the disulfide bond. Signals at 12.8 and 13.0 ppm were assigned to the imino protons of G^2:C^{13} (G^9:C^6) and G^7:C^8 (G^{14}:C^1), respectively. Probably because the base pairs of G^7:C^8 and G^{14}:C^1 are located on the terminal nucleotides of duplex, the self-correlation peak of G^7–H1 could not be observed at 13.0 ppm and a strong exchange peak with protons of solvent H_2O could be observed (ESI 5†). Next, we identified the signal of imino proton at 12.8 ppm as G^2–H1. This signal indicated the interstrand NOE of C^{13}–NH_2 at 8.50 ppm. The amino protons of deoxycytidine were exchangeable with protons of solvent H_2O. Thus, if this nucleotide is located to terminal in duplex, the amino protons of deoxycytidine cannot be detectable. The peak of 13.4 ppm could be assigned to the

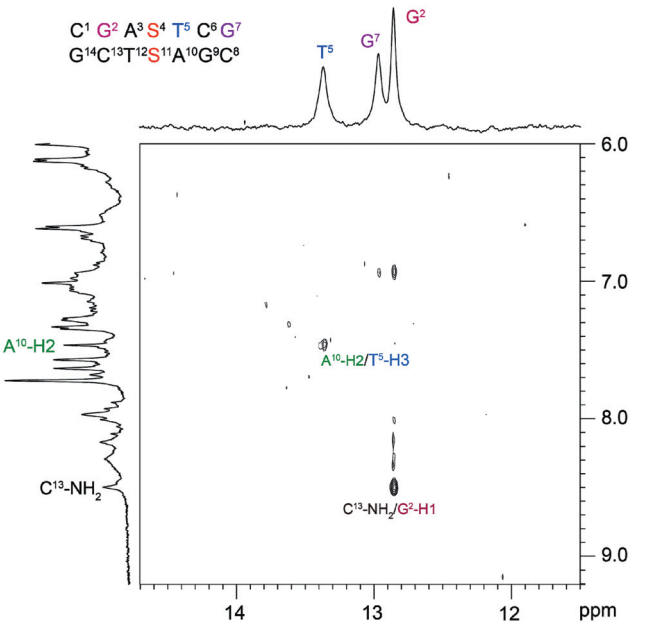

Fig. 4 NOESY spectrum between the imino protons and the base protons for $C^1G^2A^3\mathbf{S}^4T^5C^6G^7$. 0.5 mM per base of the duplex was analyzed in 10 mM Na-phosphate buffer (pH 6.8), 200 mM of NaCl dissolved in 10% D_2O (D_2O–H_2O/1 : 9) at 7 °C. The water signal was suppressed with the jump-and-return pulse sequence.[36]

imino proton of T^5 (A^{10}). We identified the interstrand NOEs between T^5–H3 and A^{10}–H2 (7.45 ppm) as well as G^2–H1 and C^{13}–NH_2 (8.50 ppm) due to the formation of interstrand base pairs (Fig. 4 and ESI 5†). Although T^5–H3 adjacently locates to G^2–H1 (G^9), the NOE correlation peak between these two imino protons seems to be too weak to be observed. According to previous report,[35] the presence of the imino proton signals at high-temperature indicated the formation of a disulfide base-pair bonded duplex, which stabilized the neighboring hydrogen

bonding network of the DNA molecule. All non-exchangeable proton resonances were assigned to specific protons in the DNA by using a combination of 2D NMR techniques including ^{1}H/^{1}H-COSY, HOHAHA and NOESY. NOESY spectra acquired in D_2O were used to determine ^{1}H/^{1}H distance constraints for the duplex. Most NOEs observed in normal B type DNA were also observed in this modified S-DNA duplex. The assigned proton resonances are tabulated in the ESI 2.†

To confirm intercalation of the thiophenyl group into the duplex, we checked the NOEs between the aromatic protons of the **S**^{4} nucleoside and the adenine protons of A^3 as well as the NOEs between the aromatic protons of **S**^{4} nucleoside and the thymine protons of T^5. NOE signals of A^3–H2 were observed for the two protons **S**^{4}–H5 and **S**^{4}–H6 (Fig. 5). However, no NOEs were observed between A^3–H8 and the aromatic protons of **S**^{4} (H5 and H6). In addition, the T^5–H6 did not interact with the two protons of the thiophenyl group of **S**^{4} in the duplex, although the NOE between T^5–H6 and the protons of C^6 (C^6–H5 and C^6–H6) were observed. Nonetheless, the T^5–5Me protons showed NOEs for **S**^{4}–H5 and **S**^{4}–H6. These results demonstrate that the disulfide base pair composed of the complementary thiophenyl groups as an unnatural bases of **S**^{4} and **S**^{11} intercalates between the base pairs of A^3:T^{12} and T^5:A^{10}.

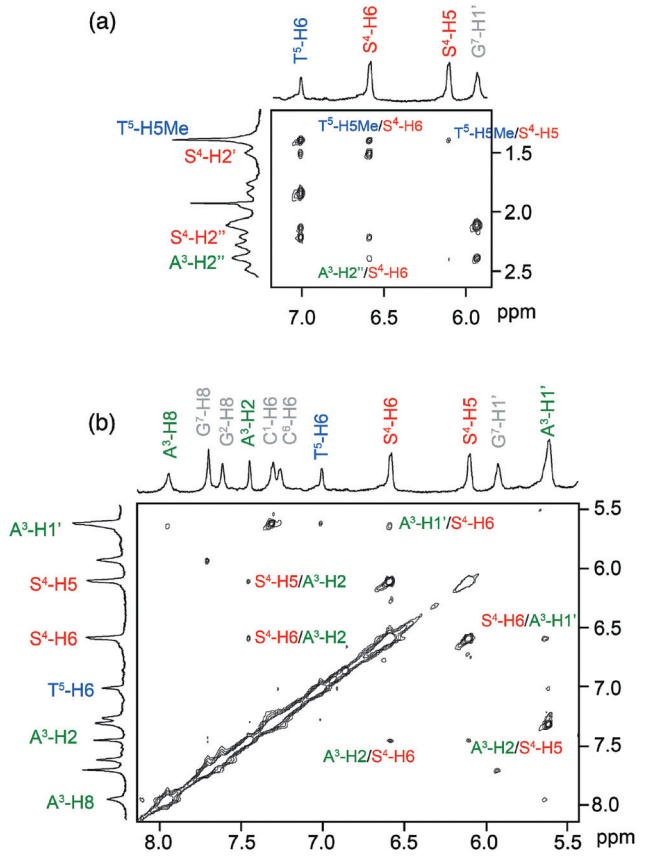

Fig. 5 NOESY spectra between the aromatic proton signals (5.5–8.0 ppm) of the **S**^{4} duplex and the protons of A^3 and T^5 acquired at 7 °C (150 ms of mixing time). (a) NOEs between T^5–H5Me and **S**^{4} aromatic proton (b) NOEs between base protons of A^3 and **S**^{4}, that of **S**^{4} and T^5.

5. Structure determination

We confirmed the formation of the interstrand Watson–Crick base pairs for 5′-C^{1}G^{2}A^3**S**^{4}T^{5}C^{6}G^7 and 3′-G^{14}C^{13}T^{12}**S**^{11}-A^{10}G^{9}C^8 from the imino proton NMR with the jump-and-return pulse sequence[36] in 10% D_2O (Fig. 4). Thus, the formation of the B-type DNA structure for the corresponding region can be assumed. We prepared 8 distance constraints for hydrogen bonds, 56 dihedral restraints for backbone and glycosidic bonds (about 5 restraints per residue), and 6 plane constraints of base pairs in this region. These three types of constraints and a total of 155 distances (102 intra-residue and 53 inter-residue distance restraints) were obtained using restrained molecular dynamic calculations with a simulated annealing protocol.[37] The structure is well defined, having a heavy atom r.m.s. deviation of 0.49 Å for 45 converged structures (Fig. 6). The r.m.s. deviations of bonds (Å) and angles (°) from the idealized geometry were 0.0023 ± 0.00011 and 0.5244 ± 0.0113 (structural statistics are summarized in Table 1). These results indicated a right-handed helix that is globally similar to that of B type DNA. The structure of

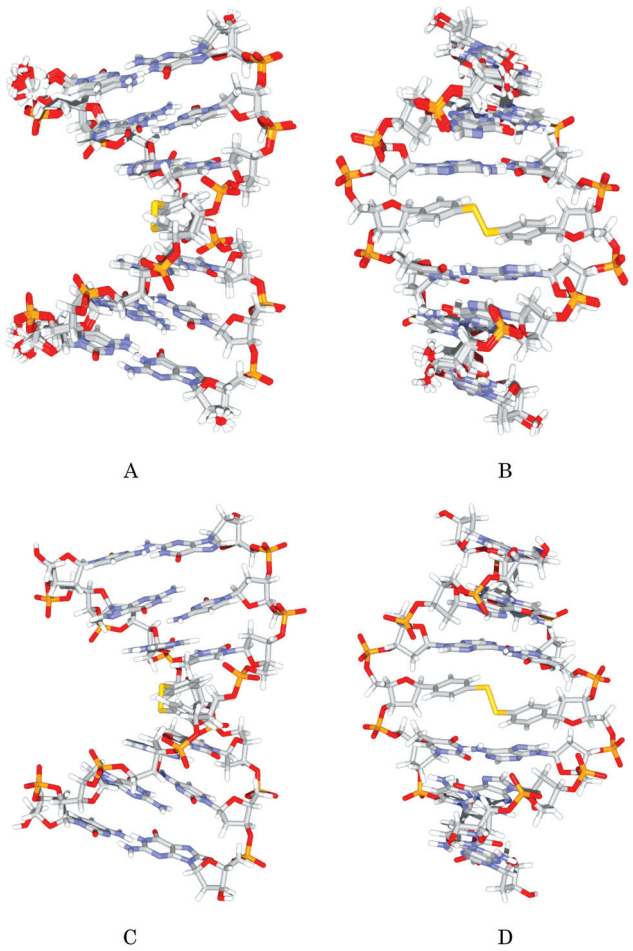

Fig. 6 Superposition of 12 S-DNA (CGA**S**TCG) structures after restrained molecular dynamics and energy minimization (A, B). The 45 final structures were accepted and we showed the lowest twelve structures to A and B. B is the view of A rotated by 90° to the right. Mean structure of 45 structures of S-DNA (C, D). D is a view of C rotated by 90° to the right.

natural B type DNA and modified S-DNA containing a disulfide base pair are different in the conformation of the backbone. We assumed that the disulfide base pair exerts an influence on the conformation of the backbone composed of phosphate diester bonding. The hydrogen bonding and base pair planarity restraints were only added for the regular stem region (C^1–A^3, T^5–G^7). The disulfide base pair ($\mathbf{S}^4$:$\mathbf{S}^{11}$) is stabilized by a strong covalent bond, but to accommodate this artificial base pair into the duplex the backbone of the whole structure needs to be twisted.

Discussion

The T_m value of the short duplex was drastically increased upon formation of the unnatural base pair with a disulfide bond.

Table 1 NMR restraints and structural statistics

Number of restraints[a]	
Intra-residue distance restraints	102
Strong	50
Medium	51
Weak	1
Inter-residue distance restraints	53
Medium	45
Weak	8
Dihedral restraints	56
Hydrogen bonding distance restraints	8
r.m.s. deviations from the idealized geometry	
Bonds (Å)	0.0023 ± 0.00011
Angle (°)	0.5244 ± 0.0113
Impropers (°)	0.2612 ± 0.0156
Heavy-atom r.m.s. deviations (Å)[b]	
All	0.490

[a] Because the calculation was performed with a dimer, actual numbers of restraints are double. [b] Averaged r.m.s. deviations between an average structure and the 45 converged structures were calculated. The converged structures did not contain any experimental distance violation of >0.5 Å or a dihedral violation >5°. The r.m.s. deviation from experimental distance and dihedral restraints were 0.0167 ± 0.0011 Å and 0.694° ± 0.038°, respectively. The converged structures did not contain any violation of hydrogen bonding distance restraints.

Where there was a mismatch at the central position of a 7-mer DNA, the T_m curve could not be observed. Moreover, the T_m value for this 7-mer could not be determined under reducing conditions. In the presence of a reducing agent the disulfide base pair was reduced to generate two thiophenyl groups on each strand, which constituted mismatched bases. The urea-containing polyacrylamide gel electrophoresis analysis indicated the formation of a disulfide bond was nonspecific and did not depend on the DNA sequence (Fig. 2, lane 3). Under oxidizing conditions each thiol group of the **S** nucleoside immediately formed a disulfide bond, whereas under reducing conditions the disulfide bond was reduced to two thiol groups. These results indicate that formation of the disulfide bond drives the assembly of the duplex. Although this driving force is strong it is also reversible. Thus, the thermal stability of the oligonucleotide to possess **S** nucleoside can be switched by redox reagents acting as an external stimulus. The disulfide bridge is a strong interaction due to covalent bonding between two sulfur atoms. The bond angle of ∠C–S–S is 103° and the bond length between each sulfur atom of the disulfide bond is 2.05 Å. NOEs between $\mathbf{S}^4$–H5, $\mathbf{S}^4$–H6 and the protons of the neighboring bases such as A^3–H2, T^5–Me were observed, whereas NOEs between $\mathbf{S}^4$–H5, $\mathbf{S}^4$–H6 and T^5–H6 could not be observed (Fig. 5). It is found that the disulfide base pair is intercalated into the sequence of the duplex. However, each phenyl ring of the disulfide base pair is not positioned in the same plane in the duplex due to rotational constraints of this bond (Fig. 7). Thus, our results suggest a conformational change for the nucleic acid induced by modified base pair formation. We analysed the conformation of S-DNA using the Curves+ application, which provided the helical and backbone parameters including a curvilinear axis and the position of bases relative to this axis.[38] The total helical axis was bent to 29° for the major groove. Moreover, the average distance of each base pair was expanded to 3.9 Å per base pair (B-DNA: 3.4 Å; see ESI 8†). The propeller twists of the base pairs in S-DNA were drastically changed from those of B type DNA. This is because the disulfide base pair formed to an opposite propeller twist with respect to a natural base pair.

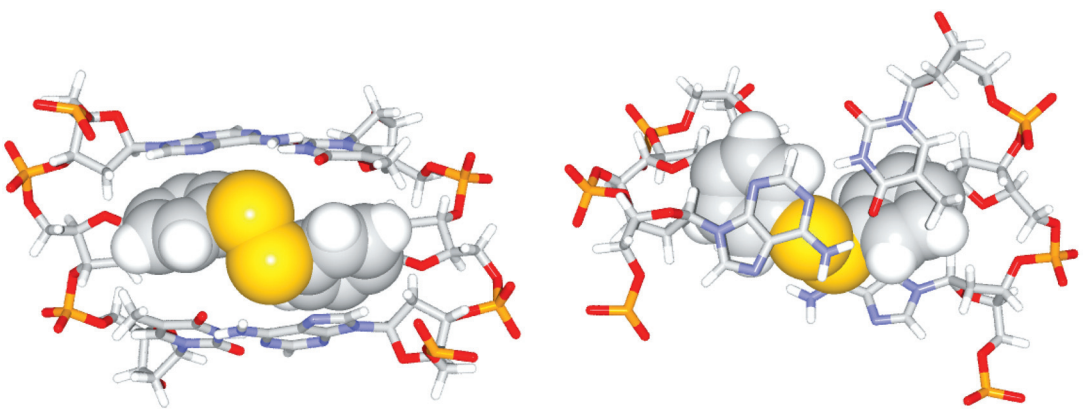

Fig. 7 Stacking structure of S-DNA ($A^3\mathbf{S}^4T^5$:$T^{12}\mathbf{S}^{11}A^{10}$) with disulfide base pair and neighboring base pairs. For clarity, only the aromatic regions of $\mathbf{S}^4$ and $\mathbf{S}^{11}$ nucleotides are shown as a space-filling model to highlight the base stacking. (A) Side view to face the major groove and (B) top view. The sulfur atoms were indicated in yellow.

We had studied the thermodynamic parameters for formation of a duplex possessing a disulfide base pair in previous work.[35] The formation of a disulfide-linked base pair in the S-DNA resulted in a gain in enthalpy (ΔH) but a drop in entropy (ΔS). The increase in enthalpy value was derived from the strong interaction due to covalent disulfide base pairing, whereas the decrease in entropy is derived from the structural constraint of the whole S-DNA molecule. The resulting thermodynamic parameters are reflected in the corresponding CD spectra. The CD spectrum of S-DNA indicated a right-handed helix that was globally similar to B-DNA under oxidizing conditions. However, the Cotton effects of S-DNA were a little different from that of the natural control DNA.

Conclusion

We determined the solution structure of the S-DNA containing a disulfide bond at the central position in the duplex by NMR. It was found that the disulfide base pair was intercalated into the sequence of the duplex. This was because the disulfide bond is covalent and the resulting base pairing unnatural. The two phenyl rings of the corresponding base pair generate a strong propeller twist. Hence, the corresponding phenyl rings are not positioned in the same plane in the duplex (Fig. 7). In addition, the disulfide base pair can give rise to inter-base pair translation between $\mathbf{S}^4$:$\mathbf{S}^{11}$ and T^5:A^{10}. As a result, S-DNA is bent at the point of the disulfide base pair to face the major groove.

Materials and methods

Melting temperature analysis

Melting studies were performed in Teflon-stoppered 1 cm path-length quartz cells under an atmosphere of nitrogen using a Shimadzu UV 2100 UV-vis recording spectrophotometer (Shimadzu, Kyoto, Japan) equipped with a thermoprogrammer. Absorbance was monitored at 260 nm. The temperature was raised from 0 to 90 °C at a rate of 0.5 °C min^{-1}. The solution for the thermal denaturation studies was prepared at an oligomer concentration of 30 μM for the base-pairing studies in a buffer containing 1.0 M NaCl and 10 mM sodium phosphate, pH 6.8. For experiments performed under reducing conditions 2-mercaptoethanol (final concentration 100 μM) was added to the sample solution.

Circular dichroism

Circular dichroism spectra were measured on a JASCO J-715 spectropolarimeter between 350 nm and 210 nm in standard buffer containing 1.0 M NaCl, 10 mM sodium phosphate, pH 6.8, at 0 °C. The duplex concentrations were 30 μM per base. Spectra were acquired every 1 nm with a bandwidth setting of 1 nm at a speed of 50 nm min^{-1}, averaging over 5 scans. For experiments performed under reducing conditions 2-mercaptoethanol (final concentration 100 μM) was added to the sample solution.

7 M urea-containing DNA-PAGE

Samples of oligonucleotide (50 μM per base) were dissolved in 10 mM phosphate buffer (pH 6.8, total volume was 50 μL). 2-Mercaptoethanol was added to the samples at a final concentration of 10 mM for experiments performed under reducing conditions. For oxidizing conditions, iodine was added to samples at a final concentration 0.5 mM. Loading buffer that included bromophenol blue was added to each sample (7.5 μL), which was then heated to 90 °C for 5 min and subsequently cooled to 4 °C over 1 h. Samples were analyzed on a 20% polyacrylamide (5% crosslinked) gel containing 7 M urea and run at 150 V using an AnaTech model 3800 power supply for 1.5 h. After staining the gel with GelRed™ (Biotium, Hayward, CA) bands were visualized on a UV transilluminator (312 nm).

NMR sample preparation

The phosphoramidite monomer carrying a thiophenol moiety was synthesized as previously described.[35] CGA**S**TCG oligonucleotide was synthesized using standard phosphoramidite chemistry and purified on an OPC column, reversed phase HPLC and then freeze-dried immediately. The buffer used consisted of 10 mM Na-phosphate buffer, 200 mM NaCl at pH 6.8. A mixture of 90% H_2O and 10% D_2O was used for experiments examining exchangeable protons. The final concentration was 2.0 mM per base for the 7-mer oligonucleotide in 400 μL of solution. For experiments carried out in D_2O, the DNA sample was lyophilized three times from D_2O and then redissolved in D_2O. Chemical shift reference was used by employing the signal of H_2O with temperature correction (4.7 ppm at 7 °C).

NMR experiments

NMR spectra were measured on a DRX-500 spectrometer (Bruker, Karlsruhe, Germany) at a probe temperature of 7 °C (280 K). 2D NOESY (150 ms of mixing time) spectra[36] in 90% H_2O were recorded using the States-TPPI method[39] and the WATERGATE 3919 pulse sequence[40] and jump-and-return pulse sequence.[36] Resonances were assigned with standard methods using a combination of COSY, HOHAHA (3919 pulse,[41] 50 ms of mixing time) and NOESY experiments (jump-and-return, 150 ms of mixing time; 3919 pulse, 200 ms of mixing time). NOESY experiments in D_2O were acquired with mixing times of 60, 100, 150 ms. Spectra were processed and analyzed using Felix software (Accelrys, San Diego, CA).

Structure calculation

A total of 155 NOE distance restraints, 56 dihedral restraints, 8 hydrogen bonding restraints and 6 base pair planarity restraints were used. Because the calculation was done for a dimer, the actual numbers of restraints were double. The program Felix (Accelrys) was used for the distance interpretation. The NOESY spectrum observed with a mixing time of 100 ms was used to produce the distance restraints. NOE intensities from non-exchangeable protons were interpreted as distances by using the fixed distance between H5 and H6 of cytidine residues, then

they were categorized into three groups, weak (>3.5 Å), medium (>2.5 Å, <3.5 Å) and strong (<2.5 Å), based on the calculated distances. Margins from the calculated distance were set for each NOE distance with −0.5 to +1.0 Å for weak and medium and −0.5 to +1.5 Å for strong. NOEs of H3/H5 or H2/H6 of the S residue were treated as r^{-6} averaged distances between the involved protons. The B-DNA conformation for backbone and the C2′– *endo–anti* conformation in the regular stem region (C^1–A^3, T^5–G^7) were added as the dihedral torsion restraints. For backbone restraints, dihedral angle values for C4′–C3′–O3′–P, C3′–O3′–P–O5′, O3′–P–O5′–C5′, P–O5′–C5′–C4′, O5′–C5′–C4′–C3′ were assumed to be 155, −96, −46, −147, 36 with margins of ±20, ±30, ±30, ±20, ±20, respectively and were applied for C^1–G^2, G^2–A^3, T^5–C^6 and C^6–G^7 (20 restraints). For the C2′–*endo* conformation, dihedral angle values for C4′–O4′–C1′–C2′, O4′–C1′–C2′–C3′, C1′–C2′–C3′–C4′, C2′–C3′–C4′–O4′, C3′–C4′–O4′–C1′ were assumed to be −4.22, 29.42, −34.92, 33.25, −18.33 with a margin of ±15 (30 restraints). For the anti conformation, a dihedral angle value for O4′–C1′–N1–C2 for pyrimidine or O4′–C1′–N9–C4 for purine was assumed to be −98 with a margin of ±30 (6 restraints). The hydrogen bonding and base pair planarity restraints were also added for the regular stem region (C^1–A^3:T^5–G^7). A distance restraint of 1.9 ± 0.1 Å was used for each hydrogen bond. The initial structure was prepared based on the regular DNA duplex. Structural parameters for the thiophenol residue were prepared based on the structure optimization with AM1 method by Gaussian98 (Gaussian, Inc., Pittsburgh, PA). The simulated annealing calculations were done with the standard protocol of CNS 1.3[37] with the script anneal.inp. Default values were used for most of the force constants as well as calculating conditions except for md.hot.step = 10 000, md.hot.noe = 50, md.cool.step = 50 000, md.cool.noe = 50, md.cart.temp = 1000, md.cart.step = 20 000, md.pow.cdih = 200, md.pow.step = 5000, md.pow.cycl = 100 and nmr.plan.scale = 50. The time step for the molecular dynamics calculation was 0.001. A set of 100 structures was calculated and the 45 final structures were accepted and a minimized average structure was calculated.

Notes and references

1 N. C. Seeman, *Nature*, 2003, **421**, 427–431.
2 Y. Lu and J. Liu, *Curr. Opin. Biotechnol.*, 2006, **17**, 580–588.
3 S. Tyagi and F. R. Kramer, *Nat. Biotechnol.*, 1996, **14**, 303–308.
4 W. J. Stec, G. Zon and W. Egan, *J. Am. Chem. Soc.*, 1984, **106**, 6077–6079.
5 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500.
6 S. Sando and E. T. KooL, *J. Am. Chem. Soc.*, 2002, **124**, 2096–2097.
7 A. Ono and H. Togashi, *Angew. Chem., Int. Ed.*, 2004, **43**, 4300–4302.
8 E. Winfree, F. Liu, L. A. Wenzler and N. C. Seeman, *Nature*, 1998, **394**, 539–544.
9 Y. He, T. Ye, M. Su, C. Zhang, A. E. Ribbe, W. Jiang and C. Mao, *Nature*, 2008, **452**, 198–202.
10 S. W. Santoro, G. F. Joyce, K. Sakthivel, S. Gramatikova and C. F. Barbas III., *J. Am. Chem. Soc.*, 2000, **122**, 2433–2439.
11 G. F. Joyce, *Curr. Opin. Struct. Biol.*, 1994, **4**, 331–336.
12 C. Switzer, S. E. Moroney and S. A. Benner, *J. Am. Chem. Soc.*, 1989, **111**, 8322–8323.
13 B. A. Schweitzer and E. T. Kool, *J. Org. Chem.*, 1994, **117**, 7238–7242.
14 D. L. McMinn, A. K. Ogawa, Y. Wu, J. Liu, P. G. Schultz and F. E. Romesberg, *J. Am. Chem. Soc.*, 1999, **121**, 11585–11586.
15 T. Mitsui, A. Kitamura, M. Kimoto, T. To, A. Sato, I. Hirao and S. Yokoyama, *J. Am. Chem. Soc.*, 2003, **125**, 5298–5307.
16 K. Tanaka and M. Shionoya, *J. Org. Chem.*, 1999, **64**, 5002–5003.
17 K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, M. Shiro and M. Shionoya, *Science*, 2003, **299**, 1212–1213.
18 E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romesberg and P. G. Schultz, *J. Am. Chem. Soc.*, 2000, **122**, 10714–10715.
19 C. Dohno, A. Okamoto and I. Saito, *J. Am. Chem. Soc.*, 2005, **127**, 16681–16684.
20 A. Hatano, S. Makita and M. Kirihara, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2459–2462.
21 H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard and E. T. Kool, *Science*, 2003, **302**, 868–871.
22 S. R. Lynch, H. Liu, J. Gao and E. T. Kool, *J. Am. Chem. Soc.*, 2006, **128**, 14704–14711.
23 K. Tanaka, G. H. Clever, Y. Takezawa, Y. Yamada, C. Kaul, M. Shionoya and T. Carell, *Nat. Nanotechnol.*, 2006, **1**, 190–194.
24 I. Hirao, M. Kimoto, T. Mitsui, T. Fujiwara, R. Kawai, A. Sato, Y. Harada and S. Yokoyama, *Nat. Methods*, 2006, **3**, 729–735.
25 Z. Yang, A. M. Sismour, P. Sheng, N. L. Puskar and S. A. Benner, *Nucleic Acids Res.*, 2007, **35**, 4238–4249.
26 D. A. Malyshev, Y. J. Seo, P. Ordoukhanian and F. E. Romesberg, *J. Am. Chem. Soc.*, 2009, **131**, 14620–14621.
27 S. R. Rajski and R. M. Williams, *Chem. Rev.*, 1998, **98**, 2723–2796.
28 B. Devadas and N. J. Leonard, *J. Am. Chem. Soc.*, 1986, **108**, 5012–5014.
29 F. Nagatsugi, T. Kawasaki, D. Usui, M. Maeda and S. Sasaki, *J. Am. Chem. Soc.*, 1999, **121**, 6753–6754.
30 X. Qiao and Y. Kishi, *Angew. Chem., Int. Ed.*, 1999, **38**, 928–931.
31 J. M. Kean, A. Murakami, K. R. Blake, C. D. Cushman and P. S. Miller, *Biochemistry*, 1988, **27**, 9113–9121.
32 A. E. Ferentz and G. L. Verdine, *J. Am. Chem. Soc.*, 1991, **113**, 4000–4002.
33 H. Wang, S. E. Osborne, E. R. P. Zuiderweg and G. D. Glick, *J. Am. Chem. Soc.*, 1994, **116**, 5021–5022.
34 R. S. Coleman, J. L. MacCary and R. J. Perez, *Tetrahedron*, 1999, **55**, 12009–12022.
35 A. Hatano, S. Makita and M. Kirihara, *Tetrahedron*, 2005, **61**, 1723–1730.
36 P. Plateau and M. Gueron, *J. Am. Chem. Soc.*, 1982, **104**, 7310–7311.
37 A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, R. J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1998, **D54**, 905–921.
38 R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute and K. Zakrzewska, *Nucleic Acids Res.*, 2009, **37**, 5917–5929.
39 J. Jeener, B. H. Meier, P. Bachmann and R. R. Ernst, *J. Chem. Phys.*, 1979, **71**, 4546–4553.
40 M. Piotto, V. Saudek and V. Sklenar, *J. Biomol. NMR*, 1992, **2**, 661–665.
41 V. Sklenar, M. Piotto, R. Leppik and V. Saudek, *J. Magn. Reson.*, 1993, **A102**, 241–245.