Preferential formation of (5*S*,6*R*)-thymine glycol for oligodeoxyribonucleotide synthesis and analysis of drug binding to thymine glycol-containing DNA

Tatsuhiko Shimizu, Koichiro Manabe, Shinya Yoshikawa, Yusuke Kawasaki and Shigenori Iwai*

Division of Chemistry, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan

ABSTRACT

We previously reported the chemical synthesis of oligonucleotides containing thymine glycol, a major form of oxidative DNA damage. In the preparation of the phosphoramidite building block, the predominant product of the osmium tetroxide oxidation of protected thymidine was (5*R*,6*S*)-thymidine glycol. To obtain the building block of the other isomer, (5*S*,6*R*)-thymidine glycol, in an amount sufficient for oligonucleotide synthesis, the Sharpless asymmetric dihydroxylation (AD) reaction was examined. Although the reaction was very slow, (5*S*,6*R*)-thymidine glycol was obtained in preference to the (5*R*,6*S*) isomer. The ratio of (5*S*,6*R*)- and (5*R*,6*S*)-thymidine glycols was 2:1, and a *trans* isomer was also formed. When an ionic liquid, 1-butyl-3-methylimidazolium hexafluorophosphate, was used as a co-solvent, the reaction became faster, and the yield was improved without changing the preference. The phosphoramidite building block of (5*S*,6*R*)-thymidine glycol was prepared, and oligonucleotides containing 5*S*-thymine glycol were synthesized. One of the oligonucleotides was used to analyze the binding of distamycin A to thymine glycol-containing DNA by Circular dichroism (CD) spectroscopy and surface plasmon resonance (SPR) measurements. Distamycin A bound to a duplex containing either isomer of thymine glycol within the AATT target site, and its binding was observed even when the thymine glycol was placed opposite cytosine.

INTRODUCTION

Thymine glycol (5,6-dihydro-5,6-dihydroxythymine) is a major type of oxidative DNA damage that results from the reaction of a thymine base with reactive oxygen species generated by ionizing radiation (1) or as a consequence of aerobic metabolism (2). While this damage causes mutations only at a low frequency (3), it effectively blocks DNA replication (4) and must be repaired by the base excision repair pathway (5). There are four diastereomers of thymine glycol, i.e. (5*R*,6*S*), (5*R*,6*R*), (5*S*,6*R*) and (5*S*,6*S*), but thymine glycol exists as either the 5*R* *cis–trans* pair [(5*R*,6*S*) and (5*R*,6*R*)] or the 5*S* *cis–trans* pair [(5*S*,6*R*) and (5*S*,6*S*)] in solution, due to epimerization at the C6 position (6). It has been reported that the 5*R* and 5*S* isomers are formed in equal amounts in γ-irradiated DNA (7), but the oxidation of thymidine or thymidine-containing oligonucleotides preferentially yields (5*R*,6*S*)-thymine glycol (8,9). We observed the same preference [the ratio of the (5*R*,6*S*) and (5*S*,6*R*) isomers was 6:1] in our previous study on the chemical synthesis of thymine glycol-containing oligonucleotides using a phosphoramidite building block (10), and a large-scale preparation was required to obtain (5*S*,6*R*)-thymidine glycol in a yield sufficient for its incorporation into oligonucleotides (11). Biochemical studies on translesion replication (12,13) and base excision repair (14–16) revealed that the stereochemistry of thymine glycol is recognized by proteins responsible for these biological processes. Therefore, it is becoming more important to synthesize oligonucleotides containing each isomer of thymine glycol, which can be used for comparative experiments. The problem to be solved is the development of a practical method to obtain the building block of (5*S*,6*R*)-thymidine glycol. Here we report the synthesis of (5*S*,6*R*)-thymidine glycol, using Sharpless asymmetric

dihydroxylation (AD) (17), and its incorporation into oligo-nucleotides. This reaction was previously employed by Barvian and Greenberg (18) to synthesize (5S,6R)-thymidine glycol, but their study was only at the nucleoside level, because the 5′- and 3′-hydroxyl functions of the sugar moiety were both protected by the *tert*-butyldimethylsilyl (TBDMS) group, which is unsuitable for oligonucleotide synthesis. In our current study, the reaction was slow due to the bulky protecting group, but the yield was improved using an ionic liquid (19).

One of the oligonucleotides synthesized in this study was used to analyze drug binding to thymine glycol-containing DNA. Previously, we found that distamycin A, a natural antibiotic, bound to DNA duplexes containing the (6-4) photoproduct, one of the major types of ultraviolet (UV)-induced lesions, although the photoproduct formation alters both the chemical structure of the base moiety and the local tertiary structure of the duplex (20). This compound did not bind to a duplex containing the *cis–syn* thymine dimer, another UV lesion, and the binding stoichiometry for the AATT target site was changed to 2:1 upon the formation of the (6-4) photoproduct. In the chemical structure, thymine glycol resembles the 5′ component of the (6-4) photo-product, and the thermodynamic properties of the duplexes containing these lesions are similar to each other (11,21). Therefore, we expected that distamycin A might be able to bind to the duplexes containing thymine glycol in a similar manner.

MATERIALS AND METHODS

For the chemical synthesis, the general methods were basically the same as those reported previously (11). ^{1}H-NMR spectra were measured on a JEOL AL-400 or Varian INOVA 600 spectrometer, and mass spectra were obtained on a Micromass LCT spectrometer. High-performance liquid chromatography (HPLC) analyses were carried out on a Gilson gradient-type analytical system equipped with a Waters 2996 photodiode array detector. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra of oligo-nucleotides were measured in the negative ion mode on an Applied Biosystems Voyager DE PRO spectrometer, using 3-hydroxypicolinic acid as a matrix.

The sharpless AD reaction for the preparation of 2a

To a mixture containing $K_2OsO_4 \cdot 2H_2O$ (28.0 mg, 76 μmol), (DHQ)$_2$PHAL (284 mg, 0.36 mmol), $K_3Fe(CN)_6$ (1.45 g, 4.41 mmol), K_2CO_3 (609 mg, 4.41 mmol) and methane-sulfonamide (140 mg, 1.47 mmol) in *tert*-butanol and water (1:1, v/v, 36ml), 5′-O-(4,4′-dimethoxytrityl)-3′-O-benzoyl-thymidine (**1**, 953 mg, 1.47 mmol) was added, and the resultant mixture was stirred for 7 days. The mixture was cooled in an ice bath, and Na_2SO_3 (2.42 g, 19.2 mmol) was added. After stirring for 30 min on ice, the mixture was poured into water (100 ml), and the products were extracted three times with chloroform (200 ml in total). The organic layer was dried with Na_2SO_4 and concentrated *in vacuo*, and the products were purified on silica gel (0–0.5% methanol in chloroform). The starting material (**1**, 151 mg, 232 μmol) was recovered. Yield:

2a, 254 mg (372 μmol, 25%); **2b**, 141 mg (207 μmol, 14%); *trans* isomer, 175 mg (256 μmol, 17%).

For the synthesis using an ionic liquid, **1** (973 mg, 1.50 mmol) was added to a solution of $K_2OsO_4 \cdot 2H_2O$ (27.6 mg, 75 μmol), (DHQ)$_2$PHAL (292 mg, 375 μmol) and 4-methylmorpholine N-oxide (900 mg, 7.68 mmol) in a mixture of water, 1-butyl-3-methylimidazolium hexafluoro-phosphate and *tert*-butanol (1:1:2, v/v/v, 12 ml), and this solu-tion was stirred for 2 days. After the reaction mixture was poured into water (100 ml), the products were extracted three times with chloroform (200 ml in total) and were purified by column chromatography. Wakogel C-200 (15 g, Wako Pure Chemical Industries) was first used to obtain all of the products together, and after the ratio of the products was determined by NMR spectroscopy, the *cis* isomers were separated on Wako-gel C-300 (30 g) eluted with chloroform. Yield: **2a**, 410 mg (600 μmol, 40%); **2b**, 179 mg (262 μmol, 17%).

2a: TLC (chloroform–methanol, 10:1), R_f 0.57; ^{1}H-NMR (CDCl$_3$, δ, p.p.m.), 8.03 (dd, 2H, Bz, $J = 7.2$, 1.2 Hz), 7.60 (t, 1H, Bz, $J = 7.4$ Hz), 7.51–7.41 (m, 4H, Bz, DMT), 7.35 (s, 1H, NH), 7.34–7.21 (m, 7H, DMT), 6.83 (d, 4H, DMT, $J = 8.8$ Hz), 6.48 (dd, 1H, H1′, $J = 8.3$, 6.7 Hz), 5.65–5.61 (m, 1H, H3′), 5.01 (d, 1H, H6, $J = 2.4$ Hz), 4.23 (dd, 1H, H4′, $J = 6.6$, 2.9 Hz), 3.77 (s, 6H, OCH$_3$), 3.59 (d, 1H, 6-OH, $J = 2.0$ Hz), 3.50 (dd, 1H, H5′, $J = 10.8$, 4.0 Hz), 3.39 (dd, 1H, H5′, $J = 10.8$, 2.8 Hz), 3.10 (s, 1H, 5-OH), 2.46–2.36 (m, 2H, H2′), 1.39 (s, 3H, CH$_3$); ESI-TOF-MS, calcd [M + Na]$^+$ = 705.2419, found 705.2406.

2b: TLC (chloroform–methanol, 10:1), R_f 0.54; ^{1}H-NMR (CDCl$_3$, δ, p.p.m.), 8.05 (dd, 2H, Bz, $J = 7.2$, 1.2 Hz), 7.59 (t, 1H, Bz, $J = 7.4$ Hz), 7.57–7.44 (m, 4H, Bz, DMT), 7.41 (s, 1H, NH), 7.37–7.28 (m, 7H, DMT), 6.85 (d, 4H, DMT, $J = 8.8$ Hz), 6.33 (dd, 1H, H1′, $J = 9.2$, 5.2 Hz), 5.72 (d, 1H, H3′, $J = 6.4$ Hz), 5.27 (d, 1H, H6, $J = 1.2$ Hz), 4.23–4.20 (m, 1H, H4′), 3.78 (s, 6H, OCH$_3$), 3.57 (dd, 1H, H5′, $J = 10.4$, 3.6 Hz), 3.46 (s, 1H, 5-OH), 3.32 (dd, 1H, H5′, $J = 10.4$, 2.8 Hz), 2.83 (d, 1H, 6-OH, $J = 1.6$ Hz), 2.79–2.71 (m, 1H, H2′), 2.53–2.46 (m, 1H, H2′), 1.36 (s, 3H, CH$_3$); ESI-TOF-MS, calcd [M + Na]$^+$ = 705.2419, found 705.2410.

Trans isomer: TLC (chloroform–methanol, 10:1), R_f 0.51; ^{1}H-NMR (CDCl$_3$, δ, p.p.m.), 8.03 (dd, 2H, Bz, $J = 7.2$, 1.2 Hz), 7.59 (t, 1H, Bz, $J = 7.2$ Hz), 7.47–7.39 (m, 4H, Bz, DMT), 7.33 (s, 1H, NH), 7.34–7.27 (m, 7H, DMT), 6.86 (d, 4H, DMT, $J = 8.4$ Hz), 6.36 (dd, 1H, H1′, $J = 9.6$, 5.2 Hz), 5.72 (d, 1H, H3′, $J = 6.0$ Hz), 5.30 (d, 1H, H6, $J = 5.2$ Hz), 4.22–4.20 (m, 1H, H4′), 3.79 (s, 6H, OCH$_3$), 3.61 (dd, 1H, H5′, $J = 10.4$, 2.4 Hz), 3.50 (dd, 1H, H5′, $J = 10.6$, 2.6 Hz), 2.75–2.68 (m, 1H, H2′), 2.65 (s, 1H, 5-OH), 2.55–2.50 (m, 1H, H2′), 2.46 (d, 1H, 6-OH, $J = 5.2$ Hz), 1.08 (s, 3H, CH$_3$); ESI-TOF-MS, calcd [M + Na]$^+$ = 705.2419, found 705.2412.

Deprotection of the isomers of thymidine glycol

Compound **2a**, synthesized in the present study, or compound **2b**, prepared by the previous method (10), (1.4 mg) was dissolved in 0.5 M K_2CO_3/MeOH (0.7 ml). After 2.5 h, the solution was mixed with 0.5 M sodium phosphate buffer (pH 5.0, 0.8 ml), and the product was extracted with chloroform. The organic layer was concentrated *in vacuo*, and 80% acetic acid (2.0 ml) was added. After 1 h, the acetic

acid was removed by evaporation and co-evaporation with water. Water and diethyl ether were added to the residue, and after concentration, the aqueous layer was analyzed by HPLC, using an Inertsil ODS-3 column (4.6 × 250 mm, GL Science) eluted with 0.1 M triethylammonium acetate (pH 7.0).

Oligodeoxyribonucleotide synthesis

The phosphoramidite building block of (5S,6R)-thymidine glycol (**3**), prepared from **2a** following the previous report (11), was dissolved in anhydrous acetonitrile to a concentration of 0.1 M and was installed on an Applied Biosystems Model 394 DNA/RNA synthesizer. Nucleoside phosphoramidites for ultramild DNA synthesis (Glen Research), as well as the base-unprotected thymidine phosphoramidite, were also dissolved in acetonitrile to make 0.1 M solutions and were installed on the synthesizer. Oligonucleotides were synthesized on a 1.0 µmol scale, and the reaction time for the coupling of **3** was prolonged to 5 min. After chain assembly and removal of the 4,4′-dimethoxytrityl (DMT) group at the 5′ end on the synthesizer, the solid supports containing the oligonucleotides were treated with 28% aqueous ammonia (2 ml) at room temperature for 2 h. The resulting ammoniac solutions were concentrated to dryness on a rotary evaporator equipped with a vacuum pump. The residues were dissolved in triethylamine trihydrofluoride (500 µl, Aldrich), and the mixtures were kept at 40°C overnight. After desalting on a NAP-10 column (Amersham Biosciences), the oligonucleotides were analyzed and purified by HPLC, using a µBondasphere C18 5 µm 300 Å column (3.9 × 150 mm, Waters) with a linear gradient of acetonitrile (5–11% for the 13 and 14mers and 7–13% for the 30mer, 20 min) in 0.1 M triethylammonium acetate (pH 7.0).

Circular dichroism (CD) spectroscopy

The oligonucleotide duplexes shown in Figure 4 were prepared by heating the mixtures of the two strands (5.0 nmol) in water (50 µl) at 80°C for 3 min and cooling them to room temperature. The CD spectra of the distamycin A–DNA complexes were measured at 15°C on a JASCO J-805 spectrophotometer. The sample solutions (600 µl) contained 2.5 µM each duplex, 500 mM NaCl and 10 mM sodium phosphate (pH 7.0), and titrations were conducted by adding 1.5 µl aliquots of a distamycin solution to increase the distamycin/duplex molar ratio by 0.5.

Surface plasmon resonance (SPR)

Oligonucleotides bearing biotin at the 5′ end were synthesized using 5′-biotin phosphoramidite (Glen Research), following the manufacturer's instructions. Oligonucleotide duplexes were prepared by heating the mixtures of the lesion-containing strand (3.0 nmol) and the biotin-linked strand (2.5 nmol) in water (50 µl) at 80°C for 3 min and cooling them to room temperarure. Experiments were performed on a Biacore 2000 system, using streptavidin sensor chips (Biacore Sensor Chip SA). The duplexes were immobilized at a flow rate of 5 µl/min with a DNA concentration of 100 nM, using a buffer containing 10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA and 0.005% surfactant P20, and the amount of the immobilized duplex was 1000 resonance units (RUs). One

flow cell was left intact and was used as a blank for reference. The sensorgrams were collected at a flow rate of 20 µl/min, using a buffer containing 10 mM sodium phosphate (pH 7.0), 500 mM NaCl, 3 mM EDTA and 0.005% Tween-20. All of the experiments were performed at 15°C to avoid dissociation of the duplexes.

RESULTS

AD of protected thymidine

We chose the DMT and benzoyl groups for the protection of the 5′- and 3′-hydroxyl functions, respectively, because phosphoramidite building blocks of thymidine glycol for the oligonucleotide synthesis had been successfully prepared using these protecting groups (10,11). The AD reaction using 5′-*O*-DMT-3′-*O*-benzoylthymidine (**1**) as the starting material (Scheme 1) was very slow. When 5 mol% $K_2OsO_4 \cdot 2H_2O$ and 25 mol% hydroquinine 1,4-phthalazinediyl diether $(DHQ)_2PHAL$ were used, following the previous report (18), three products were detected, but the starting material still remained after 7 days at room temperature. After purification of the products on silica gel, the configuration of the base moiety was determined by NMR spectroscopy. As described previously (10,11), none of the obtained products showed a NOESY crosspeak between H6 and H1′, and the configurations of thymine glycol in the products with and without an NOE between H6 and H2′ were assigned to (5S,6R) (**2a**) and (5R,6S) (**2b**), respectively. A *trans* isomer, which had 5-CH₃–6-OH and 5-OH–H6 crosspeaks in the NOESY spectra, was also obtained, as reported previously (18), and the NMR spectra suggested that its configuration was (5S,6S). To confirm the stereochemistry, the *cis* products were deprotected and characterized by reversed-phase HPLC. Typical results are shown in Figure 1. The peaks of (5R,6S)-thymidine glycol prepared by the previous method (10) (Figure 1a) and the (5S,6R) isomer synthesized in the present study (Figure 1b) were separated clearly in the co-injection experiment (Figure 1c). It should be noted that the peak separation of the isomers was much improved, as compared with that in the previous report (11), by using a different column. The ratio of the (5S,6R)-, (5R,6S)-, and *trans*-thymine glycols produced in the AD reaction was determined from the ^{1}H-NMR spectra of a crude mixture before the separation of the products. This ratio was 4:2:3, respectively, and the isolation yield of the (5S,6R) isomer was 25%. The lower stereoselectivity, as well as the slow reaction and the low yield, could be attributed to the bulky DMT group at the 5′ position, which would prevent the complex formation. The reaction rate and the yield were similar when the benzoyl group at the 3′ position was changed to the TBDMS group (data not shown).

Improvement using an ionic liquid as a co-solvent

We tried to improve the reaction rate. First, the concentrations of the reagents were increased. When 10 mol% $K_2OsO_4 \cdot 2H_2O$ and 50 mol% $(DHQ)_2PHAL$ were used, all of the starting material (**1**) was consumed within 3 days. However, it was disadvantageous, because the ratio of (5S,6R)- and (5R,6S)-thymine glycols became 3:2. As an alternative method, we
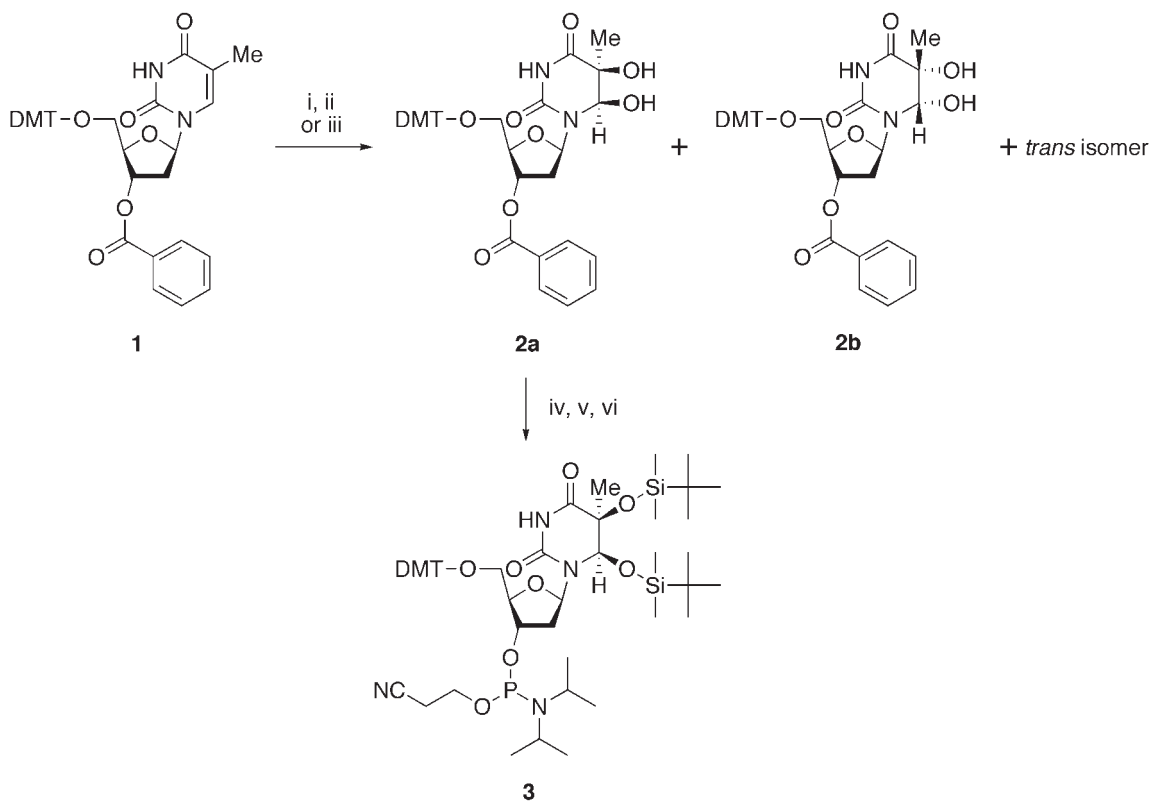
Scheme 1. Synthesis of the phosphoramidite building block of (5S,6R)-thymidine glycol. Reagents and conditions: (i) $K_2OsO_4 \cdot 2H_2O$, $(DHQ)_2PHAL$, $K_3Fe(CN)_6$, K_2CO_3, $CH_3SO_2NH_2$, $(CH_3)_3COH$, H_2O, room temperature, 7 days; (ii) Na_2SO_3, 0°C, 30 min; (iii) $K_2OsO_4 \cdot 2H_2O$, $(DHQ)_2PHAL$, 4-methylmorpholine N-oxide, $(CH_3)_3COH$, H_2O, 1-butyl-3-methylimidazolium hexafluorophosphate, room temperature, 2 days; (iv) TBDMS-Cl (5 equiv), imidazole, DMF, 37°C, 24 h; (v) K_2CO_3, MeOH, room temperature, 2 h; (vi) $[(CH_3)_2CH]_2NP(Cl)OCH_2CH_2CN$, $[(CH_3)_2CH]_2NC_2H_5$, THF, room temperature, 30 min.

tried to use a room temperature ionic liquid as a co-solvent. Branco and Afonso (19) reported that 1-butyl- or 1-octyl-3-methylimidazolium hexafluorophosphate could facilitate the AD reaction and improve the enantiomeric excess. Among the combinations they tested, we chose 1-butyl-3-methylimidazolium hexafluorophosphate, 4-methylmorpholine N-oxide and PHAL as the co-solvent, the co-oxidant, and the ligand, respectively, because using this combination they obtained a good result in the AD reaction of 1-methylcyclohexene, which was most similar to the thymine base. In the same way as the previous experiment, 5 mol% $K_2OsO_4 \cdot 2H_2O$ and 25 mol% $(DHQ)_2PHAL$ were used. The reaction was monophasic, and the starting material (**1**) was consumed within 2 days. TLC analysis showed that the products were the same as those obtained with the conventional solvent system. These products were roughly purified on silica gel, and the putative mixture of (5S,6R)-, (5R,6S)- and *trans*-thymine glycols was analyzed by ^{1}H-NMR spectroscopy. It was found from the integration values of the methyl signals that the ratio of these thymine glycols was 4:2:3, respectively, which was identical with that obtained without the ionic liquid. Since the starting material was completely consumed in this case, the products were easily isolated by chromatography on fine-particle silica gel, and their chemical structures were identified by ^{1}H-NMR spectroscopy. The yield of the (5S,6R) isomer (**2a**) was improved to 40%. Recycling and re-use of the catalyst described in the original report (19) was not tested.

Synthesis of 5S-thymine glycol-containing oligonucleotides

The phosphoramidite building block of (5S,6R)-thymidine glycol (**3**) was prepared from the oxidized nucleoside (**2a**) in three steps, as described previously (11). The yields of the TBDMS protection, the 3′-deprotection and the phosphitylation were 78, 100 and 76%, respectively. Although the two stereoisomers of the phosphoramidite (**3**), due to the chirality of the phosphorus atom, were separated on silica gel, as observed in the previous study (11), they were mixed together before the oligonucleotide synthesis. Using this building block (**3**) and nucleoside phosphoramidites bearing N^6-phenoxyacetyladenine, N^4-acetylcytosine, N^2-(4-isopropylphenoxy)acetylguanine, and thymine as normal bases, three oligonucleotides containing 5S-thymine glycol, d(ACGCGATg*ACGCCA), d(CGCGAATg*TGCGCCC) and d(CTCGTCAGCATCTTg*CATCATACAGTCAGTG), in which Tg* represents 5S-thymine glycol, were synthesized on a DNA synthesizer. The reaction time for the coupling of the thymidine glycol building block (**3**) was prolonged to 5 min, and the 1 μmol scale synthesis was carried out. After chain assembly and removal of the 5′-terminal DMT group on the synthesizer, the oligonucleotides were cleaved from the solid support by a treatment with 28% aqueous ammonia at room temperature for 2 h. The protecting groups at the base moieties and the internucleotide phosphates were removed simultaneously by this treatment. The ammoniac
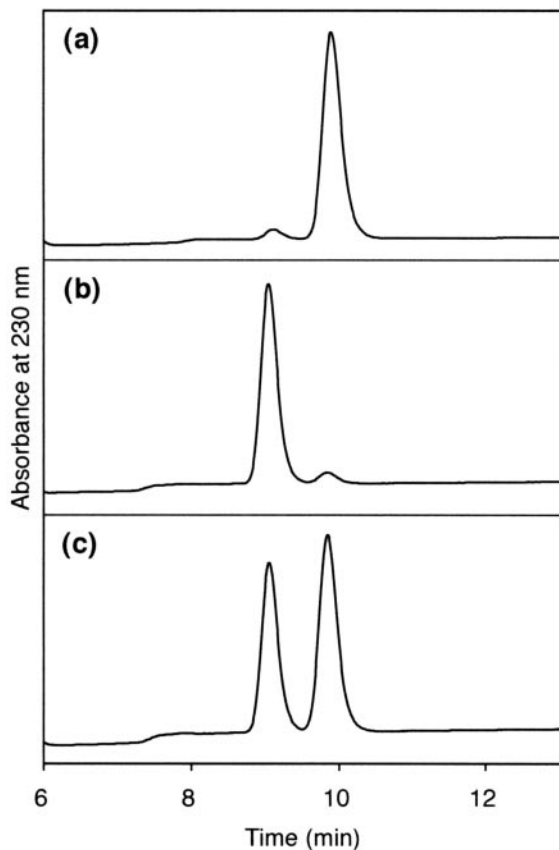
Figure 1. HPLC analysis of deprotected thymidine glycol. (**a**) (5*R*,6*S*)-thymidine glycol obtained by the previously reported OsO$_4$ oxidation; (**b**) (5*S*,6*R*)-thymidine glycol obtained in the present study; (**c**) a mixture of (a) and (b).

Figure 2. HPLC analysis of the crude samples of the 13- (**a**), 14- (**b**), and 30mers (**c**) containing 5*S*-thymine glycol.

solutions were evaporated to dryness, and the TBDMS group was removed by treating the oligonucleotides with triethylamine trihydrofluoride at 40°C for 24 h (22). After desalting on a disposable gel-filtration column, the oligonucleotides were analyzed by reversed-phase HPLC, as shown in Figure 2, and only a single major peak was detected in each case. In our previous study (11), the TBDMS group was removed with tetrabutylammonium fluoride, and a relatively large peak of a byproduct was detected. Analysis of the byproduct by mass spectrometry suggested that the thymine glycol was converted to 5-hydroxy-5-methylhydantoin during deprotection, but this side reaction was prevented in the present study by using triethylamine trihydrofluoride for the removal of the TBDMS group. The oligonucleotides were purified by reversed-phase HPLC, and their molecular weights were confirmed by MALDI-TOF mass spectrometry, as shown in Figure 3.

Analysis of distamycin A binding to thymine glycol-containing DNA

One of the oligonucleotides containing the 5*S*-thymine glycol was used for a study on the DNA recognition of distamycin A, which is a natural antibiotic known as a minor groove binder and can bind to DNA containing the (6-4) photoproduct caused by UV-irradiation (20). The DNA duplexes used in the present study contained either of the thymine glycol isomers within the target site for distamycin A, AATT·AATT (Figure 4).
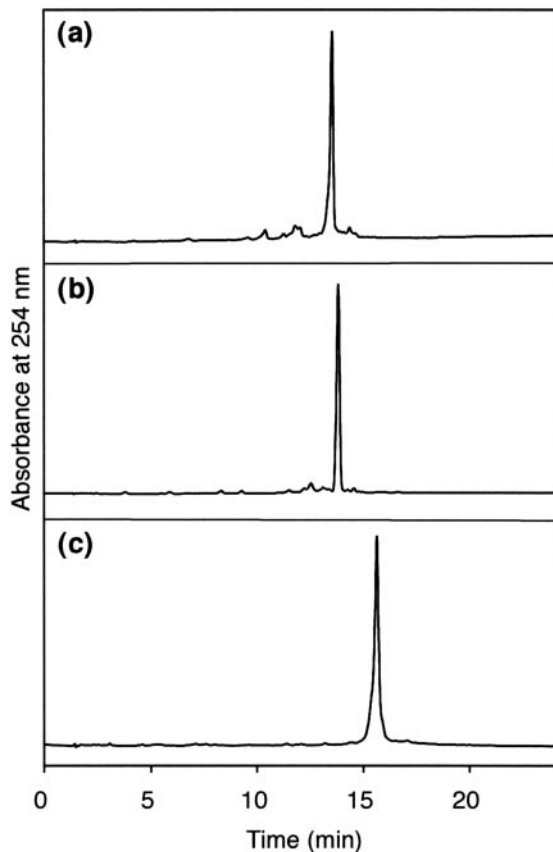
Mismatched duplexes, designated as T·C-14 and Tg*·C-14, were also tested. The binding of distamycin A was analyzed by CD spectroscopy, as reported previously (20), and the positive band induced at wavelengths between 300 and 360 nm upon the addition of distamycin A indicated the drug binding. Figure 5A shows the authentic binding of distamycin A to its target sequence. An induced CD signal was observed when the target sequence was disrupted by changing one of the T·A base pairs into G·C (Figure 5B). However, its intensity was about one half of that observed for T·A-14, and there was no isoelliptic point at 305 nm in the spectra for G·C-14. This pattern of CD spectra was observed for nonspecific binding in our previous study (20). In contrast, the CD spectra obtained for the 5*R*- and 5*S*-thymine glycol-containing duplexes were very similar to those for T·A-14 (Figure 5, C and D). The difference between the thymine glycol isomers was that the second isoelliptic points from the long-wavelength side were observed at 275 and 265 nm for the duplexes containing 5*R*- and 5*S*-thymine glycols, respectively. We also analyzed the distamycin A binding to mismatched duplexes. A duplex containing a T·C mismatch within the target sequence (T·C-14) showed only weak, nonspecific binding (Figure 5E), but when thymine glycol was placed opposite cytosine (Tg*·C-14), the CD spectra resembled the specific binding pattern more closely (Figure 5F).

The binding of distamycin A to the thymine glycol-containing duplexes was confirmed by SPR measurements. For this experiment, the thymine glycol-containing
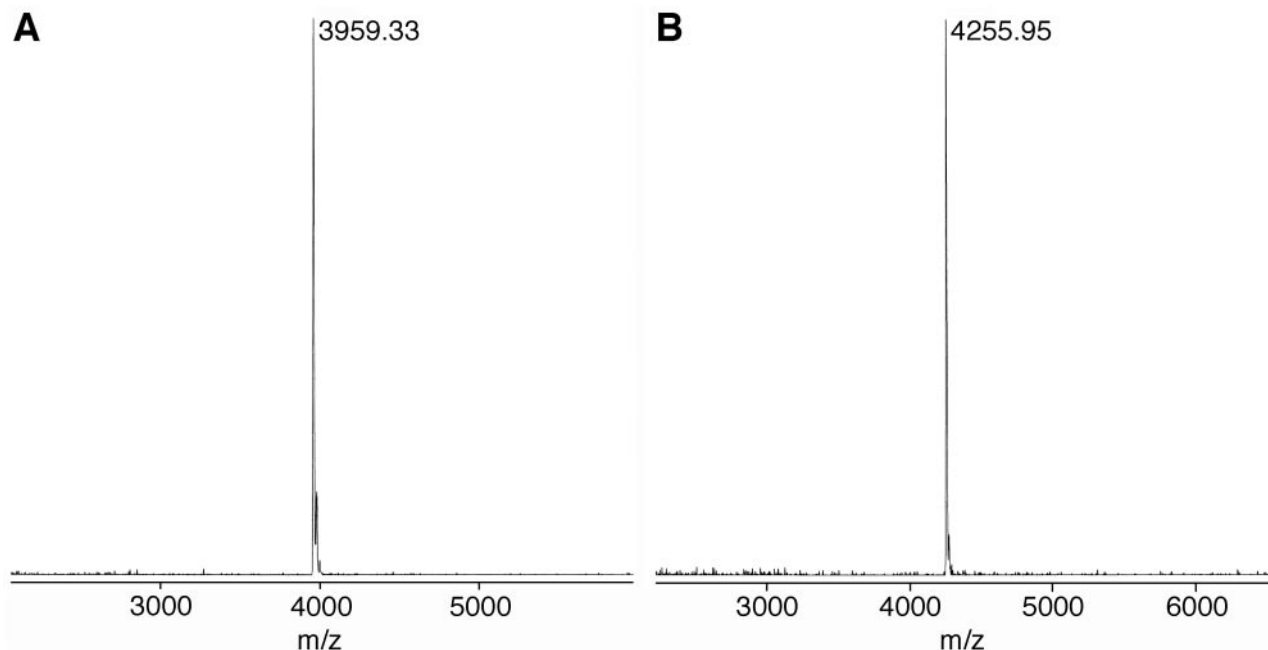
Figure 3. Analysis of the 13- (**A**) and 14mers (**B**) by MALDI-TOF mass spectrometry. Their calculated molecular weights are 3959.70 and 4255.73, respectively.
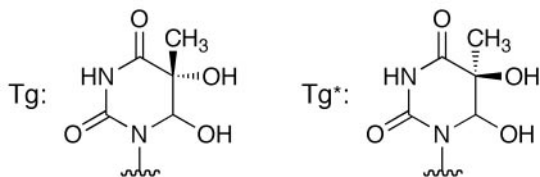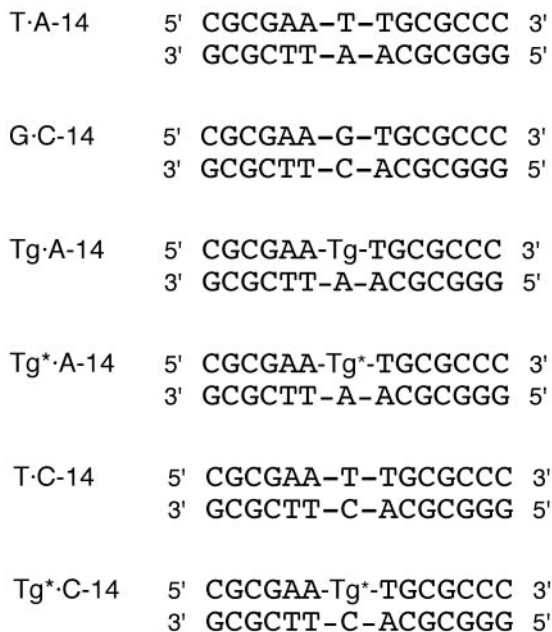

Figure 4. Oligonucleotide duplexes used for the analysis of distamycin A binding. Tg and Tg* represent 5R- and 5S-thymine glycols, respectively. For the SPR measurement, oligonucleotides bearing biotin at the 5′ end were used as the bottom strand.

oligonucleotides were hybridized to a complementary strand bearing biotin via a linker at the 5′ end, and the resulting duplexes were immobilized on streptavidin-linked sensor chips. As shown in Figure 6, sensorgrams indicating the binding of distamycin A were obtained for the thymine glycol-containing duplexes, regardless of the configuration at the C5 position, whereas G·C-14 showed no binding in this SPR experiment. All of the flow cells had essentially the same amount of immobilized duplexes (1000 RU), and saturation occurred at 45–50 RU for both of the thymine glycol-containing duplexes. The dissociation constants calculated from the SPR data were 2.8×10^{-8} and 2.6×10^{-8} M for Tg·A-14 and Tg*·A-14, respectively. These values are slightly smaller than that obtained previously for the (6-4) photo-product-containing duplex by CD spectroscopy (20).

DISCUSSION

Oxidation of thymine produces thymine glycol, which has two chiral carbon atoms, C5 and C6. Since epimerization occurs at the C6 position in an aqueous solution, thymine glycol exists as either the 5R *cis–trans* pair (5R-thymine glycol) or the 5S *cis–trans* pair (5S-thymine glycol), in which the *cis* isomers predominate (6). Recently, several groups reported that the stereochemistry of thymine glycol could be distinguished by proteins. Human DNA polymerase η, which catalyzes efficient and accurate translesion synthesis (TLS) past a thymine dimer produced by UV, could efficiently replicate DNA past 5R-thymine glycol, whereas the primer extension was considerably inhibited at 5S-thymine glycol (12). By contrast, human DNA polymerase κ, another Y-family DNA polymerase responsible for TLS, showed efficient and accurate incorporation opposite 5S-thymine glycol, but it exhibited reduced fidelity for nucleotide incorporation opposite the 5R
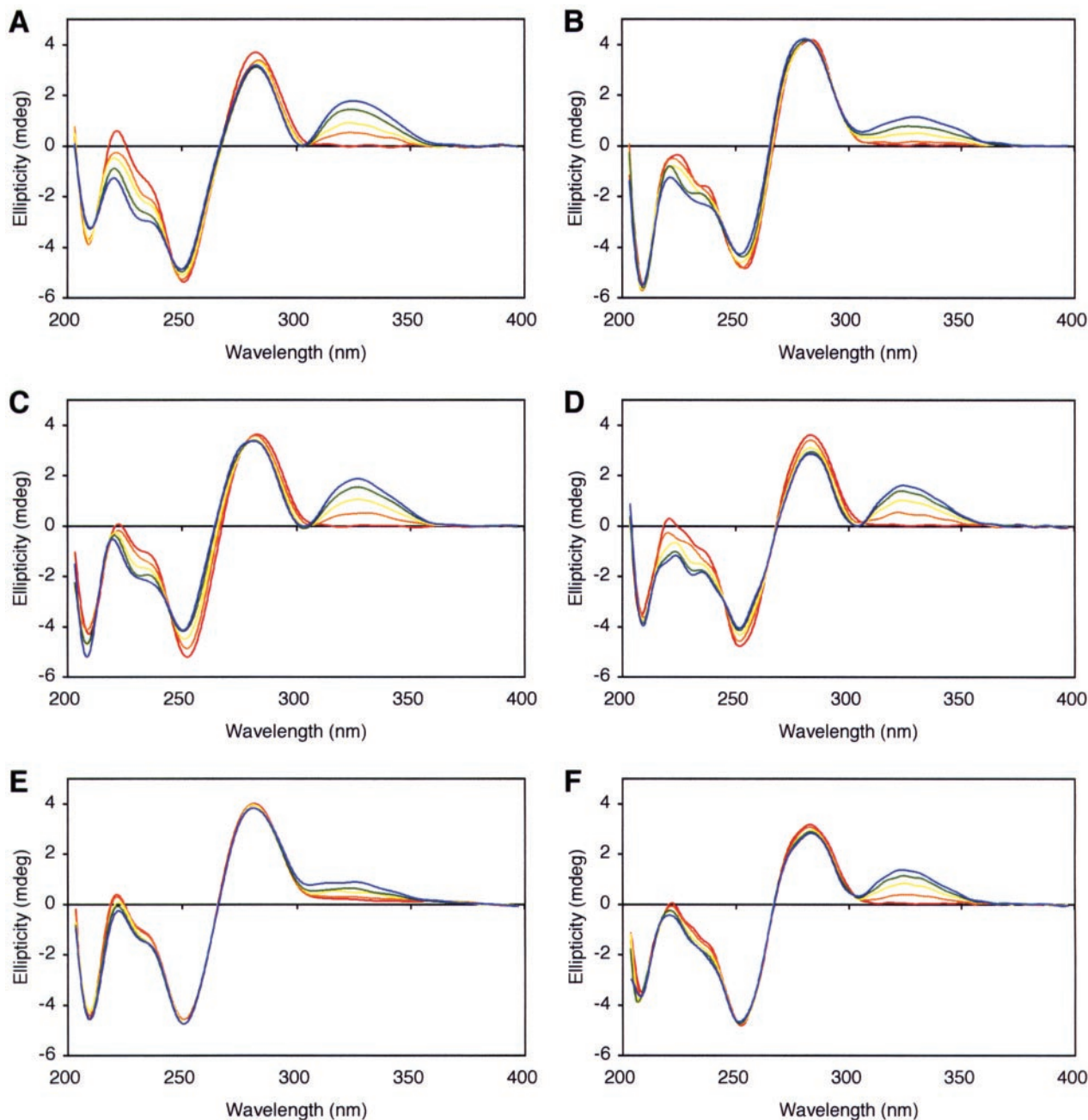
Figure 5. CD spectra of distamycin A complexed with T·A-14 (**A**), G·C-14 (**B**), Tg·A-14 (**C**), Tg*·A-14 (**D**), T·C-14 (**E**) and Tg*·C-14 (**F**). The distamycin/duplex ratios are 0 (red), 0.5 (orange), 1.0 (yellow), 1.5 (green) and 2.0 (blue).

counterpart (13). Among the base excision repair enzymes for thymine glycol, *Escherichia coli* endonuclease VIII, mouse and human NTH1, and human NEIL1 incised DNA containing 5*R*-thymine glycol more efficiently than that containing the 5*S* isomer, whereas *E.coli* endonuclease III preferred 5*S*-thymine glycol (14–16).

Oligonucleotides containing a single *cis–trans* pair of thymine glycol are important materials in these studies. One of the authors of the present study developed a method for incorporating the thymine glycol isomers separately into oligonucleotides (10,11). However, the synthesis of 5*S*-thymine glycol-containing oligonucleotides was not practical,

because the amount of (5*S*,6*R*)-thymidine glycol, which was formed by the OsO_4 oxidation of protected thymidine and was used for the preparation of the phosphoramidite building block, was much smaller than that of the (5*R*,6*S*) isomer. In this study, we tried to obtain (5*S*,6*R*)-thymidine glycol in a yield sufficient for oligonucleotide synthesis. The Sharpless AD reaction (17) was examined for this purpose, because this reaction was previously employed to obtain (5*S*,6*R*)-thymidine glycol (18). The major problem in its application to the oligonucleotide synthesis was that the bulky DMT group was attached to the 5′-hydroxyl function of thymidine as a protecting group. Using 5 mol% osmate and 25 mol% ligand,
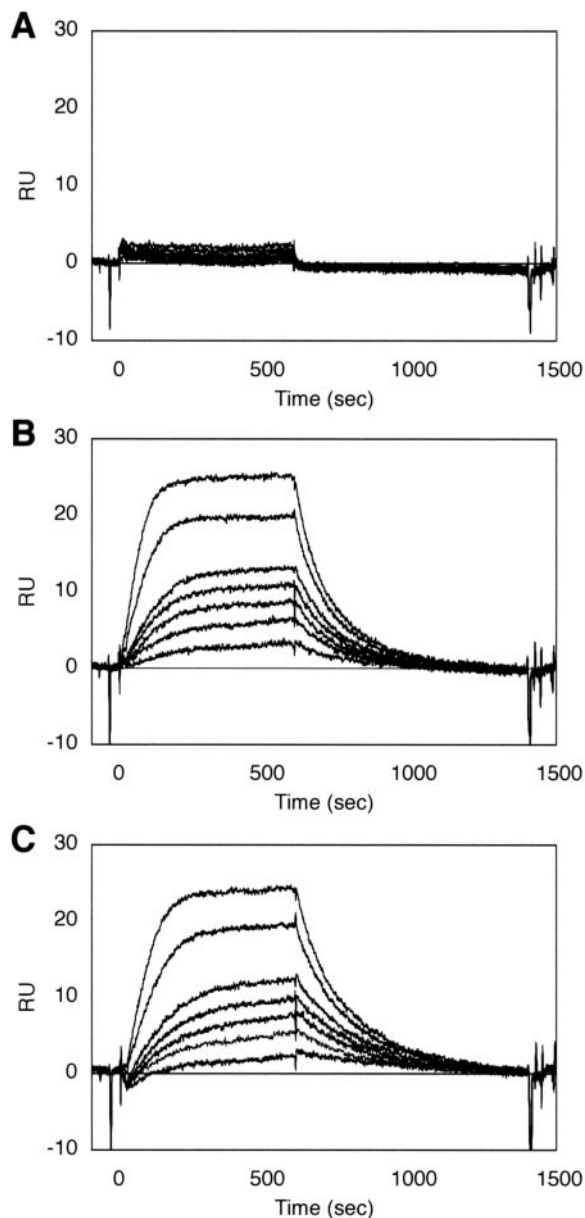
Figure 6. SPR sensorgrams for the interaction of distamycin A with G·C-14 (**A**), Tg·A-14 (**B**) and Tg*·A-14 (**C**). The concentrations of distamycin A are 2, 4, 6, 8, 10, 20 and 30 nM.

the starting material remained after 7 days in the present study, although the previous report (18) stated that the 5′-TBDMS-protected thymidine was converted to the products within 72 h, under the same conditions. Since the increased concentrations of the reagents resulted in a loss of the preference for the (5S,6R) isomer, we tried another method, i.e. the addition of a room temperature ionic liquid as a co-solvent (19). This procedure accelerated the reaction rate, probably because the ionic liquid made the reaction monophasic. The starting material was consumed within 2 days without changing the product ratio, and the desired product, (5S,6R)-thymidine glycol, was isolated easily by column chromatography on silica gel. Thus, the AD reaction using the ionic liquid as a co-solvent is a practical, good method to obtain (5S,6R)-thymidine glycol for oligonucleotide synthesis.

Oligonucleotides containing 5S-thymine glycol were synthesized in the same way as described previously (11), and in the deprotection process, we confirmed that the use of triethylamine trihydrofluoride for the removal of the TBDMS group prevented the byproduct formation that had been found in our previous study (11), as reported by Glen Research [*The Glen Report*, Vol. 16, No. 1, page 4 (2003)].

Using the 14mer synthesized in this study, d(CGCGAATg*TGCGCCC), together with a 14mer containing 5R-thymine glycol in the same sequence context, the binding of distamycin A to thymine glycol-containing DNA was analyzed. In our previous study, we found that this compound could bind to duplexes in which one of the TT sites within its target sequence, AATT·AATT, was changed to the (6-4) photoproduct, a major UV lesion, although photoproduct formation alters both the chemical structure of the base moiety and the local tertiary structure of the duplex (20). We expected that distamycin A would be able to bind to duplexes containing thymine glycol for the following reasons: (i) the chemical structure of (5R,6S)-thymine glycol resembles that of the 5′ component of the (6-4) photoproduct; (ii) NMR studies revealed that thymine glycol is extrahelical in a duplex (23), and the base stacking is also lost at the (6-4) photoproduct because the two base rings are linked perpendicularly; and (iii) the thermodynamic properties of the thymine glycol-containing duplexes (11) are similar to those of the (6-4) photoproduct-containing ones (21). If distamycin A recognizes the chemical structure of the damaged base, then there might be a difference in the binding properties between 5R- and 5S-thymine glycols.

As shown in Figure 5, distamycin A was able to bind to the duplexes containing thymine glycol. Since a relatively large signal was induced for G·C-14, used as a negative control in the analysis by CD spectroscopy, we measured the SPR, which has been used in recent studies on minor groove binders (24–26), to distinguish between specific and nonspecific binding. In this experiment, the duplexes containing 5R- and 5S-thymine glycols yielded binding sensorgrams, whereas no obvious binding was detected for G·C-14, as shown in Figure 6. The discrepancy between the CD and SPR methods may have been caused by the different concentrations of distamycin A. The dilute solution (up to 30 nM) used in the SPR measurement may have well excluded nonspecific binding. Although one of the isoelliptic points in the CD spectra was slightly different between the 5R- and 5S-thymine glycols, the sensorgrams obtained for the 5R-thymine glycol-containing duplex were indistinguishable from those for the 5S counterpart, and we suppose that the stereochemistry at the C5 position was not recognized by distamycin A. Besides the duplexes in which the base opposite thymine glycol was adenine (Tg·A-14 and Tg*·A-14), distamycin A bound to a duplex containing cytosine opposite thymine glycol (Tg*·C-14), whereas specific binding was not observed for the duplex containing a T·C mismatch (T·C-14). These results suggest that some property of damage-containing duplexes, rather than the chemical structure of the damaged base or the sequence context, is recognized by distamycin A. Based on the results obtained in this study, as well as those described previously (20), studies aimed toward elucidating the mechanism of damaged DNA recognition by the minor groove binders are in progress.

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

REFERENCES

1. Teoule,R., Bert,C. and Bonicel,A. (1977) Thymine fragment damage retained in the DNA polynucleotide chain after gamma irradiation in aerated solutions. II. *Radiat. Res.*, **72**, 190–200.
2. Adelman,R., Saul,R.L. and Ames,B.N. (1988) Oxidative damage to DNA: relation to species metabolic rate and life span. *Proc. Natl Acad. Sci. USA*, **85**, 2706–2708.
3. Basu,A.K., Loechler,E.L., Leadon,S.A. and Essigmann,J.M. (1989) Genetic effects of thymine glycol: site-specific mutagenesis and molecular modeling studies. *Proc. Natl Acad. Sci. USA*, **86**, 7677–7681.
4. McNulty,J.M., Jerkovic,B., Bolton,P.H. and Basu,A.K. (1998) Replication inhibition and miscoding properties of DNA templates containing a site-specific *cis*-thymine glycol or urea residue. *Chem. Res. Toxicol.*, **11**, 666–673.
5. Slupphaug,G., Kavli,B. and Krokan,H.E. (2003) The interacting pathways for prevention and repair of oxidative DNA damage. *Mutation Res.*, **531**, 231–251.
6. Lustig,M.J., Cadet,J., Boorstein,R.J. and Teebor,G.W. (1992) Synthesis of the diastereomers of thymidine glycol, determination of concentrations and rates of interconversion of their *cis-trans* epimers at equilibrium and demonstration of differential alkali lability within DNA. *Nucleic Acids Res.*, **20**, 4839–4845.
7. Teebor,G., Cummings,A., Frenkel,K., Shaw,A., Voituriez,L. and Cadet,J. (1987) Quantitative measurement of the diastereoisomers of *cis* thymidine glycol in gamma-irradiated DNA. *Free Rad. Res. Commun.*, **2**, 303–309.
8. Vaishnav,Y., Holwitt,E., Swenberg,C., Lee,H.-C. and Kan,L.-S. (1991) Synthesis and characterization of stereoisomers of 5,6-dihydro-5, 6-dihydroxythymidine. *J. Biomol. Struct. Dyn.*, **8**, 935–951.
9. Kao,J.Y., Goljer,I., Phan,T.A. and Bolton,P.H. (1993) Characterization of the effects of a thymine glycol residue on the structure, dynamics, and stability of duplex DNA by NMR. *J. Biol. Chem.*, **268**, 17787–17793.
10. Iwai,S. (2000) Synthesis of thymine glycol containing oligonucleotides from a building block with the oxidized base. *Angew. Chem. Int. Ed.*, **39**, 3874–3876.
11. Iwai,S. (2001) Synthesis and thermodynamic studies of oligonucleotides containing the two isomers of thymine glycol. *Chem. Eur. J.*, **7**, 4343–4351.
12. Kusumoto,R., Masutani,C., Iwai,S. and Hanaoka,F. (2002) Translesion synthesis by human DNA polymerase η across thymine glycol lesions. *Biochemistry*, **41**, 6090–6099.
13. Fischhaber,P.L., Gerlach,V.L., Feaver,W.J., Hatahet,Z., Wallace,S.S. and Friedberg,E.C. (2002) Human DNA polymerase κ bypasses and extends beyond thymine glycols during translesion synthesis *in vitro*, preferentially incorporating correct nucleotides. *J. Biol. Chem.*, **277**, 37604–37611.
14. Miller,H., Fernandes,A.S., Zaika,E., McTigue,M.M., Torres,M.C., Wente,M., Iden,C.R. and Grollman,A.P. (2004) Stereoselective excision of thymine glycol from oxidatively damaged DNA. *Nucleic Acids Res.*, **32**, 338–345.
15. McTigue,M.M., Rieger,R.A., Rosenquist,T.A., Iden,C.R. and de los Santos,C.R. (2004) Stereoselective excision of thymine glycol lesions by mammalian cell extracts. *DNA Repair*, **3**, 313–322.
16. Katafuchi,A., Nakano,T., Masaoka,A., Terato,H., Iwai,S., Hanaoka,F. and Ide,H. (2004) Differential specificity of human and *Escherichia coli* endonuclease III and VIII homologues for oxidative base lesions. *J. Biol. Chem.*, **279**, 14464–14471.
17. Sharpless,K.B., Amberg,W., Bennani,Y.L., Crispino,G.A., Hartung,J., Jeong,K.-S., Kwong,H.-L., Morikawa,K., Wang,Z.-M., Xu,D. *et al.* (1992) The osmium-catalyzed asymmetric dihydroxylation: a new ligand class and a process improvement. *J. Org. Chem.*, **57**, 2768–2771.
18. Barvian,M.R. and Greenberg,M.M. (1993) Diastereoselective synthesis of hydroxylated dihydrothymidines resulting from oxidative stress. *J. Org. Chem.*, **58**, 6151–6154.
19. Branco,L.C. and Afonso,C.A.M. (2004) Ionic liquids as a convenient new medium for the catalytic asymmetric dihydroxylation of olefins using a recoverable and reusable osmium/ligand. *J. Org. Chem.*, **69**, 4381–4389.
20. Inase,A., Kodama,T.S., Sharif,J., Xu,Y., Ayame,H., Sugiyama,H. and Iwai,S. (2004) Binding of distamycin A to UV-damaged DNA. *J. Am. Chem. Soc.*, **126**, 11017–11023.
21. Jing,Y., Kao,J.F. and Taylor,J.-S. (1998) Thermodynamic and base-pairing studies of matched and mismatched DNA dodecamer duplexes containing *cis-syn*, (6-4) and Dewar photoproducts of TT. *Nucleic Acids Res.*, **26**, 3845–3853.
22. Westman,E. and Strömberg,R. (1994) Removal of *t*-butyldimethylsilyl protection in RNA-synthesis. Triethylamine trihydrofluoride (TEA, 3HF) is a more reliable alternative to tetrabutylammonium fluoride (TBAF). *Nucleic Acids Res.*, **22**, 2430–2431.
23. Kung,H.C. and Bolton,P.H. (1997) Structure of a duplex DNA containing a thymine glycol residue in solution. *J. Biol. Chem.*, **272**, 9227–9236.
24. Tanious,F.A., Hamelberg,D., Bailly,C., Czarny,A., Boykin,D.W. and Wilson,W.D. (2004) DNA sequence dependent monomer–dimer binding modulation of asymmetric benzimidazole derivatives. *J. Am. Chem. Soc.*, **126**, 143–153.
25. Mallena,S., Lee,M.P.H., Bailly,C., Neidle,S., Kumar,A., Boykin,D.W. and Wilson,W.D. (2004) Thiophene-based diamidine forms a 'super' AT binding minor groove agent. *J. Am. Chem. Soc.*, **126**, 13659–13669.
26. Buchmueller,K.L., Staples,A.M., Howard,C.M., Horick,S.M., Uthe,P.B., Le,N.M., Cox,K.K., Nguyen,B., Pacheco,K.A.O., Wilson,W.D. and Lee,M. (2005) Extending the language of DNA molecular recognition by polyamides: unexpected influence of imidazole and pyrrole arrangement on binding affinity and specificity. *J. Am. Chem. Soc.*, **127**, 742–750.