Synthesis and Structural Properties of New Oligodeoxynucleotide Analogues Containing a 2′,5′-Internucleotidic Squaryldiamide Linkage Capable of Formation of a Watson−Crick Base Pair with Adenine and a Wobble Base Pair with Guanine at the 3′-Downstream Junction Site

Kousuke Sato,[a] Ryuya Tawarada,[a] Kohji Seio,[a] and Mitsuo Sekine*[a]

Keywords: DNA / Hybridization / Hydrogen bonds / Phosphate mimic / Squaric acid / Watson−Crick base pair / Wobble base pair

A TpT dimer analogue ($U_{2'}sq_{5'}T$), in which the 3′-5′ phosphodiester linkage was replaced by a 2′-5′ squaryldiamide linkage and the 5′-upstream T was replaced by a 3′-deoxyuridine, was synthesized in almost quantitative yield from diethyl squarate. This new dimer structural motif was designed to eliminate the squaryldiamide skeleton-induced overall strain in $T_{3'}sq_{5'}T$, previously incorporated into DNA fragments as a new TpT mimic, through the change in the connection mode from the 3′-5′ linkage to a 2′-5′ linkage. Spectral analyses of $U_{2'}sq_{5'}T$ suggest that the overall structure of this dimer mimic is basically similar to that of TpT. A DNA 10mer 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ incorporating this dimer was synthesized. From the CD analysis, it turned out that the overall structure of a DNA duplex of 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ is closer to that of the unmodified duplex than the DNA duplex of 5′-d(CGCA$T_{3'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′. Interestingly, extensive *Tm* experiments suggest that d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ exhibits intriguing inherent hybridization affinity not only for the completely complementary oligodeoxynucleotide 3′-d(GCG*AA*TCGG)-5′, but also for 3′-d(GCGT*A***G**TCGG)-5′, with a mismatched dG. The unique property of the 3′-downstream dT moiety of $U_{2'}sq_{5'}T$ – the ability to recognize both dA and dG – was also supported by more detailed computational analysis of $U_{2'}sq_{5'}T$ and TpT.

Introduction

In our previous paper[1] we proposed that disubstituted derivatives **2** of squaric acid (**1**) might exhibit chemical behavior similar to that of the phosphate group[2] as far as the charge distribution is concerned,[3−7] as depicted in Figure 1.

We had synthesized a new type of modified oligodeoxynucleotide containing a neutral but highly polarized squaryldiamide group as a novel mimic of the internucleotidic phosphate group. As the result, it was found that modified oligonucleotides **3** (Figure 2) possessing a modified 3′-5′ linked dimer block ($T_{3'}sq_{5'}T$) with a squaryldiamide linkage have lower affinities for their complementary strands, but that the two thymidines at the modified dimer site can explicitly form hydrogen bonding with the consecutive deoxyadenosines at the opposite site.[1] In addition, extensive FRET experiments and computer simulations showed that duplexes containing this modified dimer have bent structures.

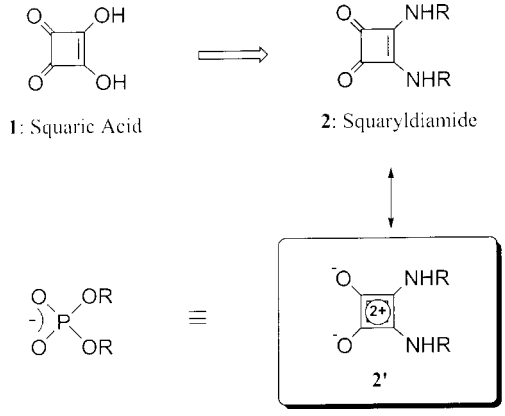

Figure 1. Structural similarity of squaryldiamides and dialkyl phosphates; the tautomeric form of the squaryldiamides resembles an internucleotidic phosphodiester linkage

[a] Department of Life Science, Graduate School of Bioscience and Biotechnology,
Nagatsuta, Midoriku, Yokohama 226-8501, Japan
[b] Frontier Collaborative Research Institute,
Nagatsuta, Midoriku, Yokohama 226-8501, Japan
[c] CREST, JST (Japan Science and Technology Corporation), Tokyo Institute of Technology, Nagatsuta, Midoriku, Yokohama 226-8501, Japan
Fax: (internat.) + 81-45-924-5772
E-mail: msekine@bio.titech.ac.jp

  DOI: 10.1002/ejoc.200300682

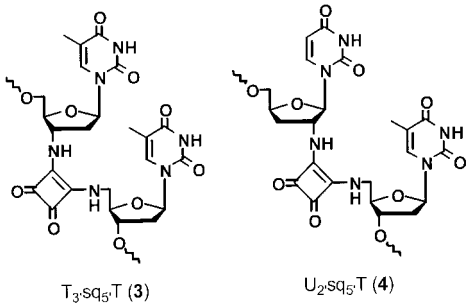

Figure 2. The structures of previously reported oligonucleotide analogue **3** and new analogue **4** containing a squaryldiamide linkage

We therefore planned a new analogue with a 2′-5′ internucleotidic squaryldiamide linkage. We chose uridine derivative as a 5′-side unit because ribothymidine is expensive and because we thought that structural difference between thymidine and uridine (5-methyl group) should not affect the backbone structures of oligonucleotides.

Our preliminary MD simulation experiments (Figure 3) suggested that this modified oligonucleotide containing a 2′-5′ internucleotidic squaryldiamide linkage has an overall structure closer to that of the natural DNA than the previous modified oligonucleotide containing a 3′-5′ internucleotidic squaryldiamide linkage.[1] In this paper we therefore report the synthesis of a 2′-5′ linked 3′-deoxyuridine-thymidine dimer derivative ($U_{2'}sq_{5'}T$) possessing a squaryldiamide linkage and an oligonucleotide derivative of 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ incorporating this dimer block, as shown in structure **4** in Figure 2, and also its interesting base recognition ability at the A−T matched site and the G−T and C−T mismatched sites upon duplex formation with the complementary DNA fragments 3′-d(GCGT*AX*TCGG)-5′ with a matched base (X = dA) and with mismatched bases (X = dG or dC). The 2′-5′ linked $U_{2'}sq_{5'}T$ dimer was designed, on the basis of computer simulations, to eliminate the overall skeleton strain induced by the squaryldiamide in the previous dimer mimic $T_{3'}sq_{5'}T$.

Results and Discussion

Synthesis of 3′-Deoxyuridine-Thymidine Dimer Derivative (9: $U_{2'}sq_{5'}T$) Containing a Squaryl Group

3′-Deoxyuridine-thymidine dimer derivative (**9**: $U_{2'}sq_{5'}T$) was synthesized by use of two amino components (2′-aminouridine[8] and 5′-aminothymidine[9−12] derivatives; Figure 4).

Treatment of 2′-amino-5′-*O*-(4,4′-dimethoxytrityl)-2′,3′-dideoxyuridine (**5**)[8] with 1.0 equiv. of diethyl squarate in ethanol gave the amidoester derivative **6** in 97% yield. The remaining ethoxy group of **6** was successively substituted by treatment with 1.0 equiv. of 5′-amino-5′-deoxythymidine (**7**) to give the squaryldiamide derivative **8** in quantitative yield. The 3′,5′-*O*-free dimer **9** was obtained in 81% yield by detritylation of **8**.

UV Spectra of $U_{2'}sq_{5'}T$ (9) and Its Model Compound 11

The UV spectra of $U_{2'}sq_{5'}T$ (**9**), TpT, and 3-isopropylamino-4-(methylamino)cyclobuten-1,2-dione (**11**), a model compound for **9**, were measured. As shown in Figure 5 (a), $U_{2'}sq_{5'}T$ (**9**) has a $\lambda_{max.}$ peak at 270 with a shoulder at

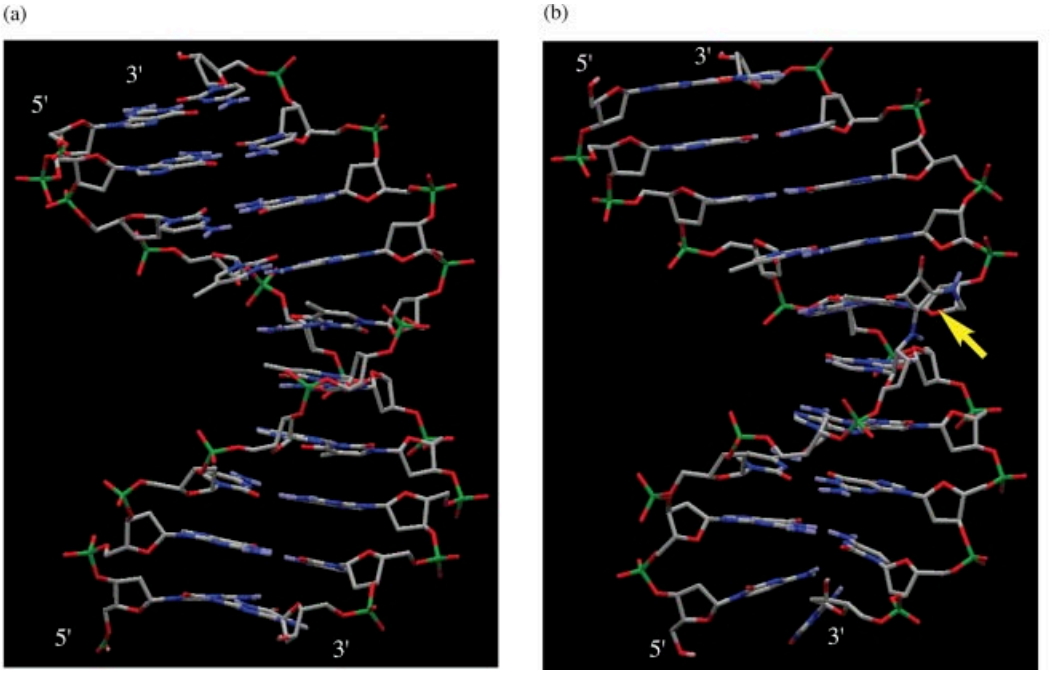

Figure 3. Structures obtained after MD simulations: (a) averaged 5′-d(CGCA*TT*AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ structure obtained after 2 ns MD simulation; (b) averaged 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ structure obtained after 2 ns MD simulation; the yellow arrow indicates the squaryldiamide group

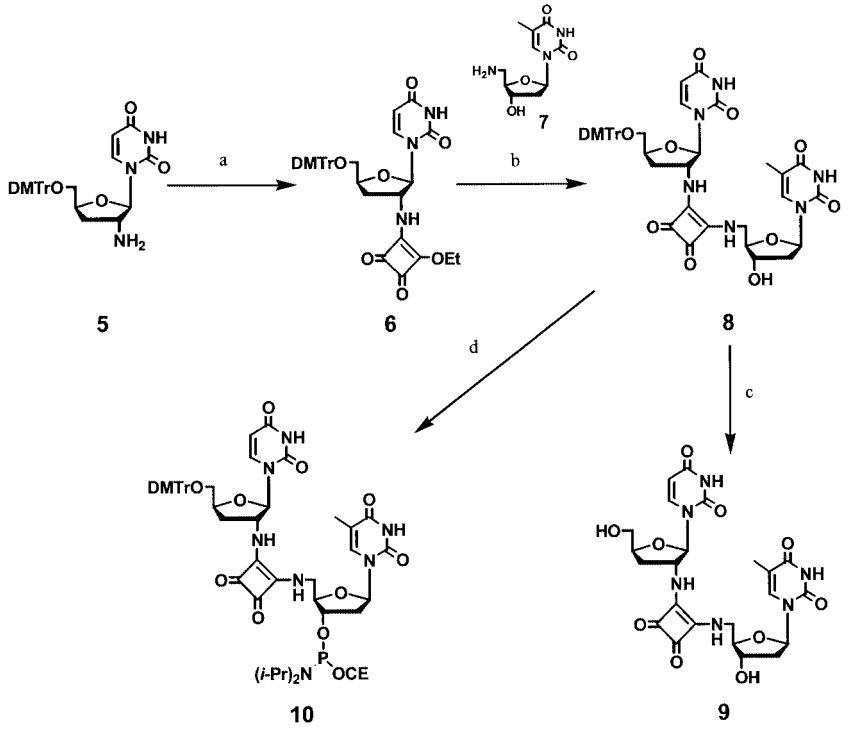

Figure 4. a) Diethyl squarate (1.0 equiv.), iPr$_2$NEt (0.5 equiv.), EtOH; 97%; b) 5′-amino-5′-deoxythymidine (**7**) (1.0 equiv.), iPr$_2$NEt (1.0 equiv.), EtOH/CH_2Cl_2 (1:1, v/v), 40 °C; quant.; c) 80% AcOH aq.; 81%; d) 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.5 equiv.), NEt_3 (5.0 equiv.), THF; 75%

292 nm. This result is similar to that of $T_{3'}sq_{5'}T$ (Figure 5, b). The $\lambda_{max.}$ of **9** (270 nm) is due to the $\lambda_{max.}$ peak of dUpT (263 nm) and the shoulder of **11** at 276 nm. The longer-wavelength peak (292 nm) of **9** is due to the $\lambda_{max.}$ peak of **11** at 293 nm. Since **11** has a large ε ($\lambda_{max.}$) value, greater than those of the canonical nucleosides, it is also surmised that the squaryldiamide group contributes substantially to the overall CD spectrum of **9**. Therefore, in order to confirm this, CD spectra of **9** were measured.

CD Spectra of $U_{2'}sq_{5'}T$ (9)

The CD spectra of $U_{2'}sq_{5'}T$, $T_{3'}sq_{5'}T$, and TpT were measured at 5 °C and these results are shown in Figure 6 (a). These data are basically similar as far as the outline shape is concerned, but the CD spectrum of $U_{2'}sq_{5'}T$ (blue) showed a positive Cotton effect somewhat different from that of $T_{3'}sq_{5'}T$, which has a wider positive Cotton effect at 285 nm (green) than TpT at 279 nm (red). These subtle differences are due to contributions not only from the squaryldiamide group, which has strong UV absorption in the 260−300 nm region, but also from the whole structure of the dimer, with a change in stacking of two pyrimidines to induce a new CD spectrum arising from the surrounding chirality. The intensity of the positive Cotton effect of $U_{2'}sq_{5'}T$ {$[\theta] = 3.2 \times 10^4$ (deg·cm^2/dmol)} is very similar to those of $T_{3'}sq_{5'}T$ and TpT, while the negative Cotton effect of $T_{3'}sq_{5'}T$ at around 260 nm is about half that of TpT, while $U_{2'}sq_{5'}T$ showed a weaker negative Cotton effect than $T_{3'}sq_{5'}T$. As shown in Figure 6 (b), the CD spectrum of $U_{2'}sq_{5'}T$ approaches that of $T_{3'}sq_{5'}T$ at 85 °C. The big difference in the CD intensity (Figure 6, b) in the region from 275 to 320 nm at 85 °C between TpT and $U_{2'}sq_{5'}T$ or $T_{3'}sq_{5'}T$ is due to the inherent contribution of the squaryldiamide skeleton arising from the nearest chirality at the 2′- and 3′-positions, respectively. The intensity of the maximum absorption (θ) value of the $U_{2'}sq_{5'}T$ dimer was reduced by 38% and those of the $T_{3'}sq_{5'}T$ and the TpT dimers were reduced by 45 and 53%, respectively. Our preliminary computer MM calculation suggested that the $U_{2'}sq_{5'}T$ dimer has an overall structure closer to that of TpT than $T_{3'}sq_{5'}T$, but the actual CD analysis of $U_{2'}sq_{5'}T$ showed a complex pattern with a shift of the positive Cotton effect to a longer wavelength than those of $T_{3'}sq_{5'}T$ and TpT. The reason for this difference is not clear.

Synthesis of Phosphoramidite Building Unit 10 and Its Incorporation into Oligodeoxynucleotides

In order to synthesize oligodeoxynucleotides containing the modified dimer block **9**, the 3′-deoxyuridine-thymidine dimer phosphoramidite unit **10** was synthesized as shown in Figure 4. Phosphitylation of **8** in the usual way gave the phosphoramidite building unit **10** in 75% yield. Standard solid-phase synthesis with **10** by the phosphoramidite method was performed to obtain 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′. The coupling time used for the building unit **10** was 15 min, and the coupling yield for the dimer phosphoramidite building unit was 94% (DMTr analysis). Since the squaryldiamide linkage proved to be rather sensitive to hydrolysis, the protecting groups on dC, dA, and dG were removed by treatment with conc. NH_3 at

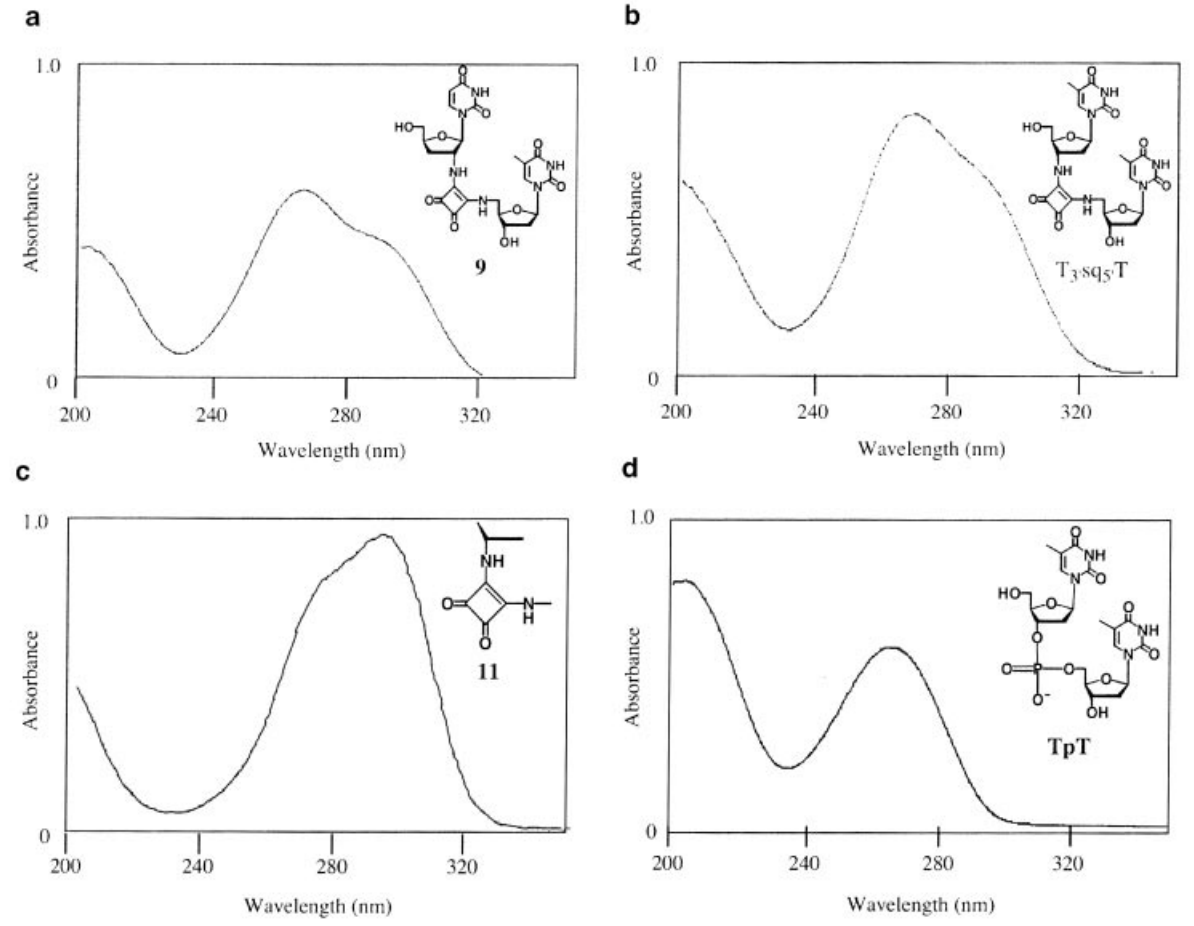

Figure 5. UV spectra of: (a) $U_{2'}sq_{5'}T$ (**9**), (b) $T_{3'}sq_{5'}T$, (c) *N*-isopropyl-*N*′-methylsquaryldiamide (**11**), and (d) TpT

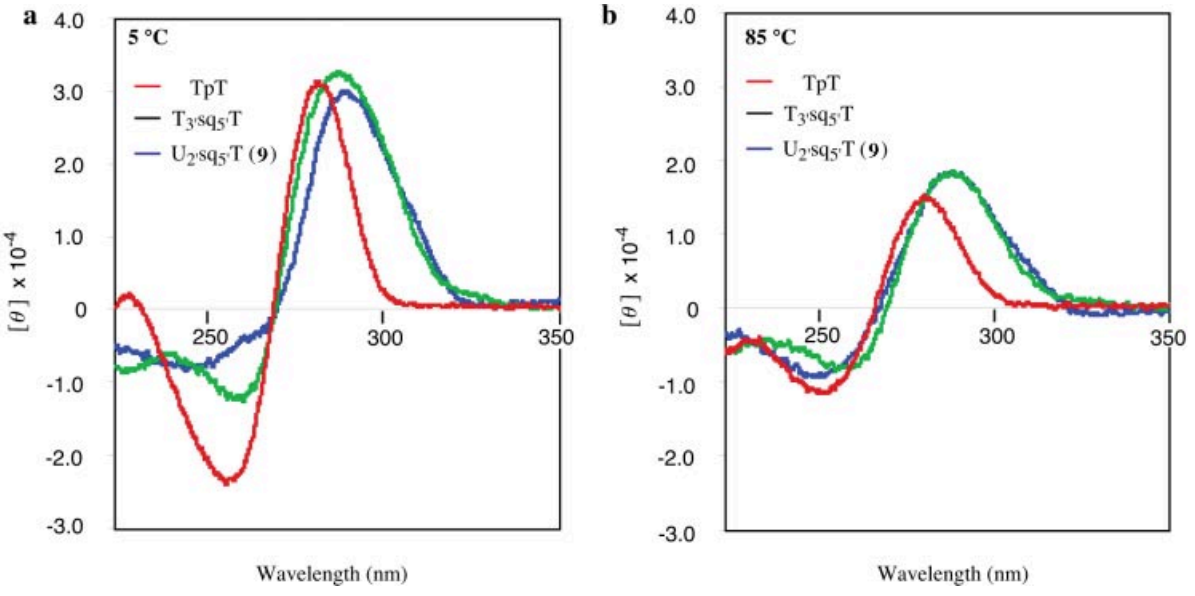

Figure 6. CD spectra of: TpT (red), $T_{3'}sq_{5'}T$ (green), and $U_{2'}sq_{5'}T$ (blue; **9**): (a) at 5 °C and (b) at 85 °C

room temperature. The modified oligodeoxynucleotide 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ was thus successfully isolated in 43% yield.

Enzyme digestion of 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ with snake venom phosphodiesterase and calf intestinal alkaline phosphatase gave three deoxynucleosides and $U_{2'}sq_{5'}T$ (**9**)

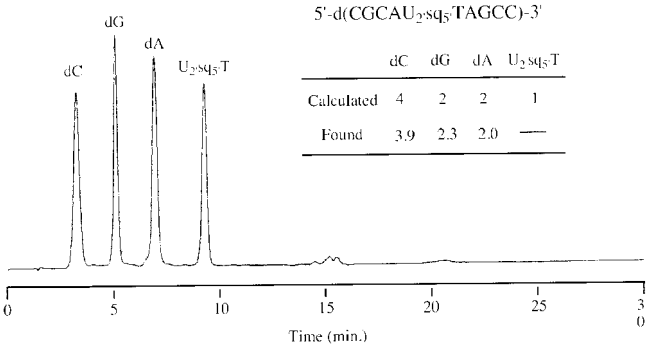

Figure 7. Reversed-phase HPLC analysis of the crude mixture obtained by enzymatic treatment of 5′-d(CGCAU$_{2'}$sq$_{5'}$TAGCC)-3′ with snake venom phosphodiesterase and calf intestinal alkaline phosphatase

in the correct ratios (Figure 7). This result also suggested that U$_{2'}$sq$_{5'}$T (**9**) is apparently resistant to snake venom phosphodiesterase and alkaline phosphatase.

Thermal Stabilities and Thermodynamic Parameters of the Duplexes Formed between Modified Oligodeoxynucleotides and Their Complementary Oligodeoxynucleotides

The thermal stability of a duplex formed between 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′ and 3′-d(GCGT*AA*TCGG)-5′ was measured (Table 1). We selected this 10mer since it is a mixed sequence composed of G−C and A−T base pairs, and one mismatch can be detected clearly within the 12.2−17.0 °C range as Δ*Tm*, as shown in Entries 2 and 3 in Table 1. This enables us to discuss a subtle change in the duplex. As shown in Table 1, the *Tm* value (26.4 °C) of the modified duplex (Entry 7) was lower than that of the control unmodified duplex (41.7 °C) by 15.3 °C (Entry 1). This result suggested the possibility that the central U$_{2'}$sq$_{5'}$T dimer block might not form a hydrogen bond with the dimer sequence of d(ApA) at the opposite site. In order to investigate whether the U−A and T−A base pairs at the 5′-upstream U and 3′-downstream T of the modified site could be formed, we also studied the thermal stabilities of the duplexes of 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′/3′-(GCGT**G***A*TCGG)-5′ and 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′/3′-(GCGT*A***G**TCGG)-5′, in which one of two dAs at the middle site of the complementary strands was replaced with dG. The result was that the *Tm* value of the former mismatched duplex of 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′/3′-(GCGT**G***A*TCGG)-5′ (Entry 8) was lower than that of the matched modified duplex of 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′/3′-d(GCGTAATCGG)-5′ (Entry 7) by 11.1 °C. This result is similar to those for 5′-d(CGCA*T$_{3'}$sq$_{5'}$T*AGCC)-3′/3′-(GCGT*AA*TCGG)-5′ (Entry 13) and 5′-d(CGCA*T$_{3'}$sq$_{5'}$T*AGCC)-3′/3′-(GCGT**G***A*TCGG)-5′ (Entry 14) reported previously by us.[1]

In contrast with these results, it was surprisingly found that the *Tm* value of 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′/3′-d(GGCGT*A***G**TCGG)-5′ (Entry 9) was only 2.5 °C lower than that of the matched duplex (Entry 7). This should be emphasized as an unexpected but intriguing result, since the *Tm* values of 5′-d(CGCA*TT*AGCC)-3′/3′-(GCGT*A***G**TCGG)-5′ (Entry 3) and 5′-d(CGCA*T$_{3'}$sq$_{5'}$T*AGCC)-3′/3′-d(GCGT*A***G**TCGG)-5′ (Entry 15) were significantly lower, by 12.2 and 10.3 °C, respectively, than those of the corresponding matched duplexes (Entries 1 and 13). There are two possible explanations for the unexpectedly moderate drop in the *Tm* value observed in Entry 9. One is that the 3′-downstream T of the core structure U$_{2'}$sq$_{5'}$T does not form hydrogen bonds with both A and G at the opposite site, so that these *Tm* values are in close proximity. The other is that the 3′-downstream T can form not only weak T−A Watson−Crick-type hydrogen bonds, but also T−G-type wobble hydrogen bonds when dA or dG is located at the opposite site. The T−G wobble base pair is less stable than the T−A Watson−Crick base pair in the natural DNA duplex. Therefore, this new modified U$_{2'}$sq$_{5'}$T dimer block could be utilizable as a new structural motif capable of both Watson−Crick and wobble base pairs of similar strengths. To ascertain whether such bifunctional base pairing was indeed produced, we examined the hybridization affinity of 5′-d(CGCA*U$_{2'}$sq$_{5'}$T*AGCC)-3′ for the mismatched strands 3′-d(GCGT**C***A*TGCG)-5′ and 3′-d(GCGT*A***C**TGCG)-5′, as shown in Entries 5, 6, 11, and 12 in Table 1. Since a T−C mismatch can form only one hydrogen bond and exhibits smaller stacking effects than a T−G mismatch, it cannot form a stable base pair in natural DNA duplexes, and so produces a significant drop in the *Tm* value in relation to matched base pairs. Actually, the *Tm* values of the DNA duplexes with TpT/CA and TpT/AC mismatches at the central positions were 15.6 and 17.6 °C, respectively (Entries 5 and 6). On the other hand, the *Tm* value of the modified duplex with a U−C mismatch could not be detected (Entry 11) and that of the modified duplex with a T−C mismatch was 14.5 °C (Entry 12). These results show an apparent sharp *Tm* drop when U$_{2'}$sq$_{5'}$T/GA (Entry 8) and U$_{2'}$sq$_{5'}$T/AG (Entry 9) were replaced by U$_{2'}$sq$_{5'}$T/CA (Entry 11) and U$_{2'}$sq$_{5'}$T /AC (Entry 12),

Table 1. Hybridization affinity of modified and unmodified oligodeoxynucleotides, (data in Entries 13−16 cited from ref.[1])

5′-CGCAXX′AGCC-3′ 3′-GCGTYY′TCGG-5′ Entry	XX′/YY′	*Tm* values *Tm* (°C)	Δ *Tm* (°C)	ΔΔ *Tm* (°C)
1	TpT/AA	41.7	−	
2	TpT/GA	24.7	−17.0	
3	TpT/AG	29.5	−12.2	
4	TpT/GG	17.1	−24.6	
5	TpT/CA	15.6	−26.1	
6	TpT/AC	17.6	−24.1	
7	U$_{2'}$sq$_{5'}$T/AA	26.4	−15.3	−
8	U$_{2'}$sq$_{5'}$T/GA	15.3	−26.4	−11.1
9	U$_{2'}$sq$_{5'}$T/AG	23.9	−17.8	−2.5
10	U$_{2'}$sq$_{5'}$T/GG	15.9	−25.8	−10.5
11	U$_{2'}$sq$_{5'}$T/CA	n.d.	n.d.	n.d.
12	U$_{2'}$sq$_{5'}$T/AC	14.5	−27.2	−11.9
13	T$_{3'}$sq$_{5'}$T/AA	30.4	−11.3	−
14	T$_{3'}$sq$_{5'}$T/GA	21.4	−20.3	−9.0
15	T$_{3'}$sq$_{5'}$T/AG	20.1	−21.6	−10.3
16	T$_{3'}$sq$_{5'}$T/GG	16.2	−25.5	−14.2

respectively, by G−C conversion. On the basis of the above, it was concluded that both U−G of $U_{2'}sq_{5'}$T/*G*A and T−G of $U_{2'}sq_{5'}$*T*/A*G* can form wobble base pairs and that the latter base pair is somewhat more stable than the former. These results showed that the mismatched duplex with a $U_{2'}sq_{5'}$T structural motif can form strong wobble-type hydrogen bonds at the central T−G base pair site, with stability similar to that of the matched T−A Watson−Crick base pair (Entries 7 and 9), much lower − by 9.1 (Entries 2 and 5) and 11.9 °C (Entries 3 and 6) − than the corresponding T−G mismatched duplexes.

CD Spectra of the Modified Duplex

It turned out that the modified oligodeoxynucleotide containing $U_{2'}sq_{5'}$T has a similar hybridization affinity for T−A matched and T−G mismatched oligonucleotides. The above reported CD spectrum of 5′-d(CGCA$T_{3'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ showed a positive Cotton effect at 288 nm, which is shifted to a longer-wavelength region than that of the control duplex (Figure 8, a). We thought that this result was caused by the global conformational change of the duplex, which can use the two A−T base pairs. In contrast with this result, the new modified duplex showed a different CD spectral pattern. At 5 °C, it is clear that the modified and unmodified control strands form duplexes, as judged from the *Tm* experiments. The positive Cotton effect at 285 nm is similar to that of the control duplex but is wider than those of the control and of the $T_{3'}sq_{5'}$T-containing duplex. When these duplexes were heated to 85 °C, the three CD spectra were now almost overlapping (Figure 8, b). At this temperature, all of the modified and the control duplexes are dissociated into single strands, as mentioned above from the results of the *Tm* experiments. The unique CD profile observed in 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ is a function not only intrinsically of the $U_{2'}sq_{5'}$T itself but also of the local conformational change in the duplex due to formation of the A−U and A−T pairs. The CD spectrum of 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AG*TCGG)-5′ at 5 °C is also different from that of the matched duplex. This spectrum showed stronger positive and negative Cotton effects at 289 nm and 256 nm, respectively. These results suggested that the whole structure of the mismatched duplex had been changed by the formation of a wobble base pair.

We next examined computational simulations at the dimer level with TpT and $U_{2'}sq_{5'}$T. It is can reasonably be expected that the local hydrogen bonds in the modified duplex should differ geometrically from those in the natural one. It is also of great importance to study the global change in the backbone structure of such modified oligonucleotides by use of computer simulations, but few such studies have been reported to date.

Computational Simulations

To ascertain whether a modified duplex possessing a 2′-5′ squaryldiamide linkage should have hydrogen bonding affinity for both dA and dG, computer simulations of $U_{2'}sq_{5'}$T and TpT were carried out with MacroModel ver 6.0.[13−15] As shown in Figure 9, the two dimer structures of $U_{2'}sq_{5'}$T and TpT extracted from structures obtained by MD simulation of the unmodified and modified DNA duplex structures (Figure 3) have different angles between the 3′-downstream and 5′-upstream bases. Figure 10 shows the bases of TpT in red, and those of $U_{2'}sq_{5'}$T in yellow. In Figure 10 (a), the 5′-upstream T and U bases are almost overlapped. The two 3′-downstream T bases (Figure 10, b), however, are not overlapped. The 3-NH of $U_{2'}sq_{5'}$T is near the 4-carbonyl group of TpT, and the 2-carbonyl group of $U_{2'}sq_{5'}$T is near the 3-NH of TpT. These relative positions are suitable for the T−G wobble base pair. These conformational changes arising from the squaryldiamide linkage

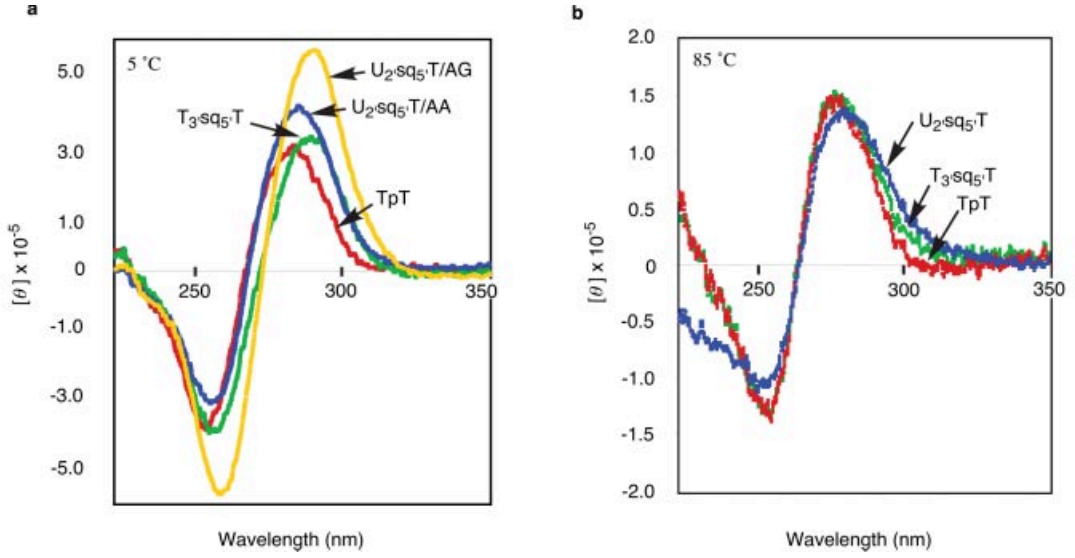

Figure 8. CD spectra of 5′-d(CGCA*TT*AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ (control; red), 5′-d(CGCA$T_{3'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ (green), 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AA*TCGG)-5′ (blue) and 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′/3′-d(GCGT*AG*TCGG)-5′ (yellow): (a) at 5 °C; (b) at 85 °C

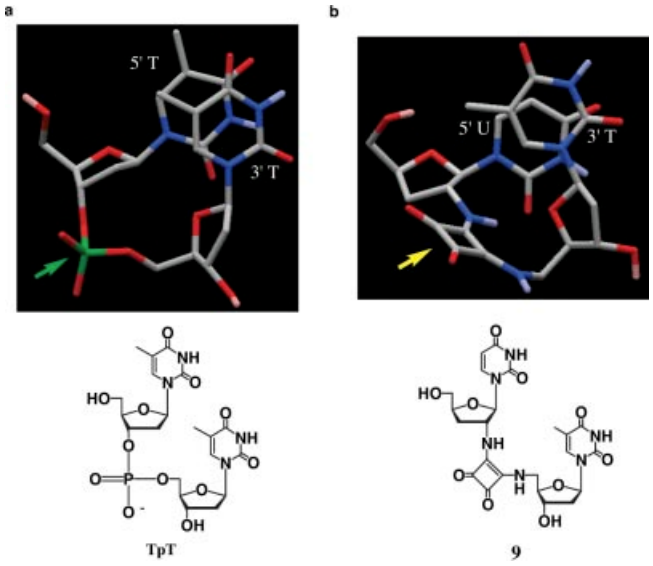

Figure 9. The dimer structures of TpT and $U_{2'}sq_{5'}T$ taken from the structures obtained by MD simulation of the unmodified and modified DNA duplexes: (a) TpT (green arrow shows phosphate group) and (b) $U_{2'}sq_{5'}T$ (**9**; yellow arrow shows squaryl group)

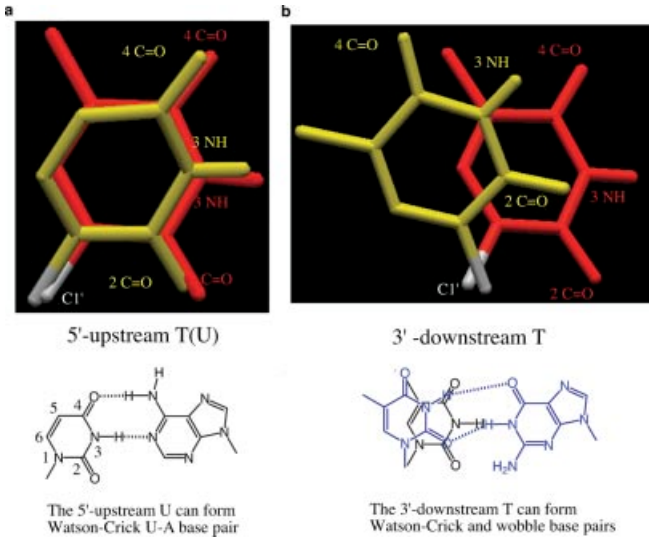

Figure 10. Superimposed structures of 5′-upstream T(U) and 3′-downstream T of TpT (red) and $U_{2'}sq_{5'}T$ (yellow; **9**) dimers: (a) 5′-upstream T(U) and (b) 3′-downstream T

may induce such flexible hydrogen bonding affinity (Watson−Crick and wobble base pairs). This analogue may therefore provide new insight into the design of the synthesis of oligonucleotides that can form Watson−Crick and wobble base pairs.

Conclusions

In our studies, the squaryldiamide linkage of $U_{2'}sq_{5'}T$ was found to decrease the base pairing affinity for complementary oligonucleotides, but to have unique base-pairing affinity for oligonucleotides with G at the site opposite to the T of $U_{2'}sq_{5'}T$. Since the change of the base recognition ability by modification of the backbone has not previously been reported, this structural motif is so interesting and important as to warrant study of the hybridization analysis with oligonucleotides having a complementary sequence and a sequence with a "wobble" base.

We were easily able to synthesize the $U_{2'}sq_{5'}T$ dimer by a two-step substitution of diethyl squarate with two different amino nucleoside derivatives in quantitative yields under mild conditions. The squaryl diamide linkage was stable under the deprotection conditions and was resistant to snake venom phosphodiesterase and alkaline phosphatase. These properties of squaryldiamide derivatives as synthetic intermediates may provide easy and effective synthesis of a wide variety of oligonucleotide analogues. Further studies in this direction are now underway.

Experimental Section

General Remarks: TLC was performed with Merck silica gel 60 (F_{254}) plates. ^1H, ^{13}C, and ^{31}P NMR spectra were obtained on a JEOL GX-270 apparatus at 270, 68, and 109 MHz, respectively. The chemical shifts were measured from tetramethylsilane (0 ppm) or [D_6]DMSO (2.49 ppm) for ^1H NMR, $CDCl_3$ (77.0 ppm), [D_6]DMSO (39.7 ppm) for ^{13}C NMR, and 85% phosphoric acid (0 ppm) for ^{31}P NMR spectroscopy. Column chromatography was performed with Wako silica gel C-200 or Kanto Kagaku Co. N60. Reversed-phase HPLC was performed on a combination of Waters 2690 and 996 systems by use of a μBondasphere C-18 column (Waters Co., Ltd., 3.9 × 150 mm) with a linear gradient of 0−15% CH_3CN/H_2O containing 0.1 M NH_4OAc (pH 7.0) at a flow rate of 1.0 mL/min. Anion-exchange HPLC was performed on a SHIMADZU 10A system with a Gen-Pak™ FAX column (Waters Co., Ltd., 4.6 × 100 mm) and use of a linear gradient of 0−0.6 M NaCl/10% CH_3CN-25 mM phosphate buffer (pH 6.0) at a flow rate of 1.0 mL/min. ESI mass spectra were measured with Mariner™ (PerSeptive Biosystems Inc.). MALDI-TOF mass spectra were measured on Voyager RP (Applied Biosystems Inc.). UV spectra were measured on a U-2000 spectrophotometer (Hitachi, Ltd.). Thymidine and uridine were purchased from Yamasa Co., Ltd. Snake venom phosphodiesterase was purchased from Boehringer Mannheim Biochemica Co., Ltd., and calf intestine alkaline phosphatase was purchased from Takara Shuzou Co., Ltd. Dry THF was purchased from Wako Pure Chemical Industries, Ltd. Triethylamine was distilled from CaH_2 and stored over molecular sieves (4Å).

5′-*O*-(4,4′-Dimethoxytrityl)-2′-(2-ethoxy-3,4-dioxocyclobuten-1-yl)-amino-2′,3′-dideoxyuridine (6): A solution of 2′-amino-5′-*O*-(4,4′-dimethoxytrityl)-2′,3′-dideoxyuridine (**5**)[8] (0.6 g, 1.13 mmol) and diisopropylethylamine (95 μL, 0.57 mmol) in 11 mL of ethanol was stirred at room temperature, and 3,4-diethoxy-3-cyclobuten-1,2-dione (0.19 g, 1.13 mmol) was added. After being stirred for 2.5 h, the solution was evaporated, diluted with $CHCl_3$, and extracted with H_2O and 5% aq. $NaHCO_3$. The organic layers were collected, and dried with Na_2SO_4. After filtration, the solvents were evaporated under reduced pressure. The residue was purified by silica gel column chromatography (C-200, 20 g, methanol/chloroform, 0 to 0.5%) to give **6** as a foam (0.72 g, 97%). ^1H NMR ($CDCl_3$): δ = 1.26−1.31 (t, J = 6.75, 7.29 Hz, 3 H, CH_2CH_3), 2.20−2.30 (m, 1 H, 3′H), 2.50−2.57 (m, 1 H, 3′′H), 3.21−3.45 (m, 2 H, 5′H), 3.66−3.71 (m, 7 H, 2′H, OCH_3), 4.41 (m, 1 H, 4′H), 4.57−4.65 (q, J = 6.75, 7.29 Hz, 2 H, CH_2CH_3), 5.33, 5.36 (d, J = 8.24 Hz,

1 H, 5 H), 6.10−6.12 (m, 1 H, 1′H), 6.73−6.79 (m, 4 H, DMTr), 7.13−7.39 (m, 9 H, DMTr), 7.79, 7.82 (d, J = 8.24 Hz, 1 H, 6 H) ppm. ^{13}C NMR ($CDCl_3$): δ = 15.97, 30.98, 34.20, 55.28, 63.67, 64.54, 70.19, 78.11, 87.05, 88.23, 103.06, 113.23, 123.68, 127.07, 127.66, 127.92, 127.96, 128.32, 128.50, 129.02, 129.97, 131.87, 132.02, 133.19, 134.97, 135.18, 135.97, 139.60, 144.13, 149.53, 150.89, 158.55, 162.87, 171.82, 178.04 ppm. ESIMS calcd. $C_{36}H_{35}N_3O_9Na$ $[M + Na]^+$ 676.2271, found 676.2219.

***N′*-[3′-Deoxy-5′-*O*-(4,4′-dimethoxytrityl)]uridin-2′-yl]-3,4-dioxo-*N*-(thymidin-5′-yl)cyclobutene-1,2-diamine (8):** To a solution of **6** (0.55 g, 0.84 mmol) and diisopropylethylamine (0.14 mL, 0.84 mmol) in 8 mL of CH_2Cl_2/ethanol (v/v, 1:1) stirred at room temperature and 5′-amino-5′-deoxythymidine (**7**)[9−12] (0.2 g, 0.84 mmol) was added and stirred at 40 °C for 20 h. The solution was evaporated under reduced pressure. The residue was purified by dry silica gel column chromatography (C-200, 20 g + N60, 3 g, methanol/chloroform, 0 to 4%) to give **8** (0.71 g, quant.). ^{1}H NMR ([D$_6$]DMSO): δ = 1.78 (s, 3 H, T3′-CH_3), 2.08−2.38 (m, 4 H, U5′-3′H, T3′-2′H), 3.20−3.38 (m, 5 H, U5′-5′H, T3′-5′H, T3′-4′H), 3.73−3.78 (m, 7 H, U5′-2′H, OCH_3), 4.18 (m, 1 H, T3′-4′H), 4.42 (m, 1 H, T3′-3′H), 5.32, 5.35 (d, J = 8.24 Hz, 1 H, U5′-5 H), 5.42−5.43 (d, J = 3.63 Hz, 1 H, T3′-3′OH), 5.84−5.86 (d, J = 3.96 Hz, 1 H, U5′-1′H), 6.16−6.21 (t, J = 6.59, 7.26 Hz, 1 H, T3′-1′H), 6.88−6.91 (m, 4 H, DMTr), 7.23−7.43 (m, 9 H, DMTr), 7.67, 7.70 (d, J = 7.91 Hz, 1 H, U5′-6 H), 7.74−7.81 (m, 1 H, T5′-6 H), 11.26 (br., 2 H, NH × 2) ppm. ^{13}C NMR ($CDCl_3$): δ = 15.84, 33.77, 55.20, 63.59, 64.35, 70.06, 78.15, 86.91, 88.41, 102.78, 113.15, 123.64, 126.96, 127.84, 127.90, 129.90, 131.93, 134.94, 135.14, 139.69, 144.08, 149.35, 150.78, 158.45, 163.19, 172.25, 177.81, 183.62, 189.02. ESIMS calcd. $C_{44}H_{44}N_6O_{12}Na$ $[M + Na]^+$ 871.2915, found 871.2996.

***N′*-(3′-Deoxyuridin-2′-yl)-3,4-dioxo-*N*-(thymidin-5′-yl)cyclobutene-1,2-diamine (9):** Compound **8** (0.16 g, 0.19 mmol) was dissolved in acetic acid (80%,5.0 mL), and the mixture was stirred at room temperature for 3 h. The solution was extracted with $CHCl_3$/H_2O, and the aqueous layer was extracted with $CHCl_3$ (2×). The aqueous layer was evaporated under reduced pressure and coevaporated three times with water to remove the last traces of acetic acid to give **9** as a white solid (83 mg, 81%). ^{1}H NMR ([D$_6$]DMSO): δ = 1.78 (s, 3 H, T3′-CH_3), 2.09−2.29 (m, 4 H, U5′-3′H, T3′-2′H), 3.47−3.89 (m, 5 H, U5′-2′H, 5′H, T3′-5′H), 4.18, 4.27 (m, 2 H, U5′-4′H, T3′-4′H), 4.62 (m, 1 H, T3′-3′H), 5.16 (br., 1 H, 5′OH), 5.41 (br., 1 H, 3′OH), 6.62, 6.65 (d, J = 7.91 Hz, 1 H, U5′-5 H), 5.82−5.84 (d, J = 4.29 Hz, 1 H, U5′-1′H), 6.15−6.20 (t, $J_{1',2'}$ = 6.27, 6.60 Hz, 1 H, T3′-1′H), 7.43 (s, 1 H, T3′-6 H), 7.70 (br., 1 H, NH of squarylamide), 7.88, 7.91 (d, J = 7.91 Hz, 1 H, U5′-6 H), 8.04 (br., 1 H, NH of squarylamide), 11.23 (br., 2 H, NH of bases) ppm. ^{13}C NMR ([D$_6$]DMSO): δ = 12.10, 32.75, 45.42, 57.78, 62.55, 70.53, 78.99, 83.49, 84.94, 88.78, 101.71, 109.83, 135.81, 140.31, 150.25, 150.37, 162.89, 163.45, 166.86, 167.72, 181.56, 182.68 ppm. ESIMS calcd. $C_{23}H_{27}N_6O_{10}$ $[M + H]^+$ 547.1789, found 547.1791.

3′-Deoxyuridine-Thymidine Dimer Building Unit 10: Chloro(2-cyanoethoxy)(diisopropylamino)phosphane (0.12 mL, 0.53 mmol) was added at room temperature to a stirred solution of **8** (0.3 g, 0.35 mmol) and triethylamine (0.29 mL, 1.77 mmol) in dry THF (3.5 mL). After the mixture had been stirred for 3.5 h, methanol (1 mL) was added, and the mixture was diluted with $CHCl_3$ and extracted with 5% aq. $NaHCO_3$. The organic layers were collected and dried with Na_2SO_4. After filtration, the solvents were evaporated under reduced pressure. The residue was purified by silica gel column chromatography (N60, 6 g, methanol/chloroform, 0 to 1.5%). The fractions containing the product were combined and evaporated under reduced pressure, and the residue was dissolved in $CHCl_3$ (1.0 mL). The $CHCl_3$ solution was poured into hexane (20 mL) to give **10** as a white solid (0.26 g, 75%). ^{1}H NMR ($CDCl_3$): δ = 1.15−1.30 (m, 12 H, CH_3 of *i*Pr), 1.87 (s, 3 H, CH_3 of thymine), 2.25−2.73 (m, 8 H, U5′-3′H, T3′-2′H, CE), 3.31−4.14 (m, 16 H, U5′-2′H, 4′H, 5′H, T3′-3′H, 4′H, 5′H, OCH_3, *i*Pr), 4.43, 4.69 (s, each 1 H, NH of squaryl amide), 5.31−5.39 (m, 1 H, U5′-5 H), 5.95 (2m, U5′-1′H, T3′-1′H), 6.80−6.83 (m, 4 H, DMTr), 7.12−7.40 (m, 10 H, DMTr, T3′-6 H), 7.86, 7.89 (d, J = 8.57 Hz, 1 H, U5′-6 H) ppm. ^{13}C NMR ($CDCl_3$): δ = 12.36, 20.53, 21.57, 24.63, 24.73, 26.20, 37.81, 43.26, 43.43, 45.08, 45.92, 55.27, 57.97, 63.90, 84.22, 86.87, 100.03, 103.25, 113.21, 118.16, 126.96, 127.90, 128.08, 130.04, 134.99, 135.19, 144.09, 149.98, 151.32, 164.45, 168.51, 176.55, 176.56, 182.24, 182.25 ppm. ^{31}P NMR ($CDCl_3$): δ = 149.37, 149.86 ppm. ESIMS calcd. $C_{53}H_{61}N_8O_{14}PNa$ $[M + Na]^+$ 1087.3943, found 1087.3945.

UV Spectra: UV spectra of $U_{2'}sq_{5'}T$ (**9**) and its model compound **11** were measured in aqueous solution at room temperature on a U-2000 spectrophotometer (Hitachi, Ltd.). The $\lambda_{max.}$ values of **9** are 270 and 292 nm; those of **11** are 276 and 293 nm.

Synthesis of Oligonucleotides: TpT and all oligodeoxynucleotides were synthesized on an ABI DNA/RNA synthesizer 392 on a 1.0 μmol scale and released from the CPG polymer support in the DMTr-on mode. The coupling time used for the uridine-thymidine dimer building unit **10** was 15 min, and the coupling with the dimer building unit proceeded in 94% yield, as estimated by DMTr cation analysis. The protecting groups used for dC, dA, and dG were acetyl, (phenoxy)acetyl, and (isopropylphenoxy)acetyl, respectively, and they were removed by treatment with concd. NH_3 (1.5 mL) at room temperature for 5 h. The ammonia solution was filtered, and the solvents were evaporated under reduced pressure. The residue was purified on a Sep-Pak™ reversed-phase column purchased from Waters Co., Ltd.. After the column was immersed with CH_3CN (5.0 mL) and 0.1 M NH_4OAc (pH 7.0) (5.0 mL) for 1 min, the solution was removed by filtration. The sample was loaded onto the column. Subsequently, the failure sequences without the DMTr group were washed out with 10−12% CH_3CN/0.1 M NH_4OAc (pH 7.0) (10 mL). After that, the column was treated with 1.0% aqueous TFA (5.0 mL) at room temperature for 5 min to remove the DMTr groups of DMTr-oligodeoxynucleotides bound to the column. Further elution was carried out with NH_4OAc (0.1 M, pH 7.0, 5.0 mL) and CH_3CN in water (15%, 10 mL) to obtain the desired oligodeoxynucleotide. After removal of the solvents, the residue was purified by anion exchange HPLC with a liner gradient of 0−60% buffer B in buffer A for 45 min at a flow rate of 1 mL/min (buffer A: 10% CH_3CN/25 mM phosphate buffer (pH 6.0), buffer B: 10% CH_3CN/25 mM phosphate buffer (pH 6.0), 1.0 M NaCl). The isolated yields were calculated by use of ε values obtained by Cantor's method.[16] The ε value (= 23314) of the $U_{2'}sq_{5'}T$ dimer was calculated by enzyme digestion of 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ with snake venom phosphodiesterase and calf intestine alkaline phosphatase. 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′: MALDI-TOF Mass calcd. for $C_{99}H_{121}N_{38}O_{56}P_8$: 2985.57; found 2985.30.

Enzyme Analysis of Modified Oligodeoxynucleotides: The enzymatic digestion was performed with a modified oligodeoxynucleotide (0.5 OD), snake venom phosphodiesterase (4 μL), and calf intestine alkaline phosphatase (2 μL) in 50 μL of alkaline phosphatase buffer (pH 9.0) at 37 °C for 2 h. After the enzymes had been deactivated by heating at 100 °C for 1 min, the solution was diluted and filtered (0.45 μm filter, Millex®-HV, MILLIPORE). This solution was analyzed by reversed-phase HPLC with a linear gradient

of 0−20% CH_3CN in 0.1 M NH_4OAc (pH 7.0) for 30 min at the flow rate of 1 mL/min.

CD Spectra of $U_{2'}sq_{5'}T$, $T_{3'}sq_{5'}T$, TpT, and Modified and Unmodified DNA Duplexes: CD spectra of $U_{2'}sq_{5'}T$, $T_{3'}sq_{5'}T$, TpT, and modified and unmodified DNA duplexes were measured with a JASCO J-725 spectropolarimeter. The plots of molar ellipticity [θ] vs. wavelength (nm) were obtained under conditions of 10 mM phosphate buffer (pH 7.0), 150 mM NaCl and 0.1 mM EDTA at 5, 25, 45, 65 and 85 °C. The blank experiments were also done under the same conditions. The concentration of each dimer was 1.0 OD/mL. The concentration of each oligonucleotide was 2.0 μM. Measurement was carried out in a 1.0-mL cell with a 0.5-cm path length. In a typical experiment, a solution of a duplex composed of two oligodeoxynucleotides was heated to 90 °C for 1 min and cooled to room temperature overnight. The CD spectra of this solution were measured eight times over the range from 200 to 350 nm at 5 to 85 °C.

Thermal Denaturation Studies: Plots of A_{260} vs. T (°C) for 5′-d(CGCA$U_{2'}sq_{5'}T$AGCC)-3′ in the presence of 3′-(GCGT*AA*TCGG)-5′, 3′-d(GCGT**G***A*TCGG)-5′, 3′-d(GCGT*A***G**TCGG)-5′, 3′-d(GCGT**C***A*TCGG)-5′, or 3′-d(GCGT*A***C**TCGG)-5′ were obtained in 10 mM phosphate buffer (pH 7.0), 150 mM NaCl and 0.1 mM EDTA with a BECKMAN DU 650 spectrophotometer. The blank experiments were done under the same conditions. A 2.0 μM solution of each oligonucleotide was put in a 400-μL cell with a 1-cm path length. In a typical experiment, a solution of a duplex composed of two oligodeoxynucleotides was heated to 60 °C and then cooled to 5 °C over the period of 55 min. Data points were then collected as the temperature was increased at the rate of 1 °C per 1 min. The first derivative of each curve was calculated by use of Igor Pro (WaveMetrics, Inc.).

Computational Simulations: The initial structure of $U_{2'}sq_{5'}T$ (**9**) was constructed by use of the uridine and thymidine residues taken from a library of MacroModel ver 6.0. The force field used was AMBER* (MacroModel ver 6.0).[14,15] The default values for the atomic partial charges were used in all the calculations. The initial structure was energy-minimized, and the optimized structure was used as the starting structure for the next simulation. The local minimum was searched by 5000-step energy minimization, and global minimum structures were searched by 5000 trials with the MCMM option. One of the low-energy structures was orientated in an *anti* conformation, and the uracil and thymine bases of this were introduced into the center of the 10mer duplex. The local minimum structure was searched for by 5000-step energy minimization, and the optimized structure was used as the starting structure for the next MD simulation. MD simulation was carried out at 300 K in H_2O for 2000 ps in which 1000 samples were collected. The averaged structure was obtained from the simulated structures in the last 1000 ps. The dimer structures of Figure 9 and 10 were obtained from the structures obtained by MD simulation of the unmodified and modified DNA duplexes.

Acknowledgments

This work was supported by a Grant from the Japanese Science and Technology Corporation (CREST) and a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

[1] K. Sato, K. Seio, M. Sekine, *J. Am. Chem. Soc.* **2002**, *124*, 12715−12724.
[2] L. M. Schwartz, L. O. Howard, *J. Phys. Chem.* **1971**, *75*, 1798−1803.
[3] H. Terao, T. Sugawara, Y. Kita, N. Sato, E. Kaho, S. Takeda, *J. Am. Chem. Soc.* **2001**, *123*, 10468−10474.
[4] A. H. Schmidt, *Synthesis* **1980**, 961−994.
[5] R. West, H. Y. Niu, *J. Am. Chem. Soc.* **1963**, *85*, 2589−2590.
[6] G. Maahs, P. Hegenberg, *Angew. Chem.* **1966**, *78*, 927−931.
[7] S. Cohen, J. R. Lacher, J. D. Park, *J. Am. Chem. Soc.* **1959**, *81*, 3480.
[8] M. Kawana, H. Kuzuhara, *J. Chem. Soc., Perkin Trans. 1* **1992**, 469−477.
[9] T. Hata, I. Yamamoto, M. Sekine, *Chem. Lett.* **1976**, 601−604.
[10] I. Yamomoto, M. Sekine, T. Hata, *J. Chem. Soc., Perkin Trans. 1* **1980**, 306−310.
[11] T.-S. Lin, W. H. Prusoff, *J. Med. Chem.* **1978**, *21*, 109−112.
[12] J. P. Horwitz, J. Chua, M. Noel, *J. Org. Chem.* **1964**, *29*, 2076−2078.
[13] For the use of MacroModel software, see: H. Tsuruoka, K. Shohda, T. Wada, M. Sekine, *J. Org. Chem.* **2000**, *65*, 7479−7494; M. Sekine, O. Kurasawa, K. Shohda, K. Seio, T. Wada, *J. Org. Chem.* **2000**, *65*, 6515−6524; M. Sekine, O. Kurasawa, K. Shohda, K. Seio, T. Wada, *J. Org. Chem.* **2000**, *65*, 3571−3578; K. Seio, T. Wada, M. Sekine, *Helv. Chim. Acta* **2000**, *83*, 162−180 and papers cited therein.
[14] MacroModel − an integrated software system for modeling organic and bioorganic molecules using molecular mechanics: F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440−467.
[15] D. Q. McDonald, W. C. Still, *Tetrahedron Lett.* **1992**, *33*, 7743−7746.
[16] Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides: C. R. Cantor, M. M. Warshaw, H. Shapiro, *Biopolymers* **1970**, *9*, 1059−1077.

Received October 28, 2003

