

Synthesis and properties of oligonucleotides containing 5-formyl-2'-deoxycytidine: *in vitro* DNA polymerase reactions on DNA templates containing 5-formyl-2'-deoxycytidine

Naoko Karino, Yoshihito Ueno and Akira Matsuda*

Graduate School of Pharmaceutical Sciences, Hokkaido University, Kita-12, Nishi-6, Kita-ku, Sapporo 060-0812, Japan

Received March 26, 2001; Revised and Accepted April 26, 2001

ABSTRACT

Oligodeoxynucleotides (ODNs) containing 5-formyl-2'-deoxycytidine (fC) were synthesized by the phosphoramidite method and subsequent oxidation with sodium periodate. The stabilities of duplexes containing A, G, C or T opposite fC were studied by thermal denaturation. It was found that fC:A, fC:C or fC:T base pairs significantly reduce the thermal stabilities of duplexes. Next, single nucleotide insertion reactions were performed using ODNs containing fC as templates and the Klenow fragment of *Escherichia coli* DNA polymerase I. It was found that: (i) insertion of dGMP opposite fC appears to be less efficient relative to insertion opposite 5-methyl-2'-deoxycytidine (mC); (ii) dAMP is misincorporated more frequently opposite fC than mC, although the frequency of misincorporation seems to be dependent on the sequence; (iii) TMP is misincorporated more frequently opposite fC than mC. These results suggest that fC may induce the transition mutation C-G→T-A and the transversion mutation C-G→A-T during DNA synthesis.

INTRODUCTION

2'-Deoxycytidine (C) in the DNA strand can be methylated by DNA-(cytosine-5) methyltransferase, yielding 5-methyl-2'-deoxycytidine (mC). C methylation takes place only after DNA synthesis, primarily in CpG sequences (1,2). C methylation is involved in the regulation of gene expression (3–6) and silencing of invading viral genomes (7,8). Interestingly, CpG sequences are hot-spots for mutations: the predominant mutation in the human p53 gene is a C→T transition in a CpG sequence (9,10). This event has been thought to be the result of spontaneous or methylase-catalyzed hydrolytic deamination of mC to T (11–13). However, this C→T transition could occur by other mechanisms as well as by deamination of mC.

Many types of DNA damage have been identified in cells exposed to ionizing radiation and oxidizing agents. 5-Formyl-2'-deoxyuridine (fdU) is the major product when T is treated

with Fenton-type reagents (14,15). Formation of fdU was also observed when DNA was γ -irradiated *in vitro* (16,17) and when cultured human cells were exposed to hydrogen peroxide (18). It is known that fdU can miscode and is mutagenic (19–21). Primer extension analysis *in vitro* showed that DNA polymerases insert 5-formyl-2'-deoxyuridine 5'-triphosphate (fdUTP) opposite both G and A in a DNA template (19). Furthermore, Fujiwara *et al.* reported mutational properties of fdUTP in an *in vivo* system (20). They found that fdUTP induces the transition mutations G·C→A·T and A·T→G·C and transversion mutation G·C→T·A.

On the other hand, it was also reported that UVA irradiation of mC in the presence of a photosensitizer leads to formation of various compounds, including 5-formyl-2'-deoxycytidine (fC, 1) (Fig. 1; 22). Additionally, formation of 5-formylcytosine was observed upon exposure of 5-methylcytosine to UV light (23). Recently we also reported that fC is formed by γ -irradiation of mC or double-stranded DNA fragments containing mC and by aerobic treatment of DNA fragments with Fenton-type reagents (24).

Based on this background information we conjectured that fC formation might be one cause of the C→T transition mutations frequently found in cytosine methylation sites. It was reported that an intramolecular hydrogen bond between the carbonyl of the 5-formyl group and the 4-amino function is present in the 5-formylcytosine moiety of its ribonucleoside (25,26). The equilibrium between the amino and imino tautomers may be affected by this hydrogen bonding and may allow formation of a base pair with adenine.

In this paper we report chemical synthesis of oligodeoxynucleotides (ODNs) containing fC at specific sites. To study

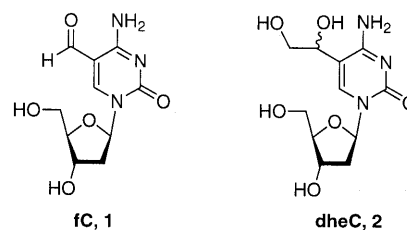


Figure 1. Structures of the modified nucleosides.

*To whom correspondence should be addressed. Tel: +81 11 706 3228; Fax: +81 11 706 4980; Email: matuda@pharm.hokudai.ac.jp

the miscoding properties of fC, primer extension analysis *in vitro* was performed using these ODNs.

MATERIALS AND METHODS

General remarks

NMR spectra were recorded at 400 (^1H) and 202 MHz (^{31}P) and are reported in p.p.m. downfield from TMS or 85% H_3PO_4 . J values are given in Hz. Mass spectra were obtained by the fast atom bombardment (FAB) method. Thin layer chromatography was done on Merck 60F₂₅₄ coated plates. The silica gel and the neutralized silica gel used for column chromatography were Merck silica gel 5715 and ICN silica 60A, respectively.

3'-O-(*t*-Butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-5-vinyl-2'-deoxycytidine (5)

A mixture of 3'-O-(*t*-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine (6.37 g, 8.26 mmol), $(\text{Ph}_3\text{P})_2\text{PdCl}_2$ (580 mg, 0.826 mmol) and tributyl(vinyl)tin (3.14 ml, 10.7 mmol) in DMF (50 ml) was stirred at 80°C for 1.5 h. The reaction mixture was filtered through a Celite pad and the filtrate was taken in EtOAc (200 ml). The organic layer was washed with diluted NaHCO_3 (50 ml \times 2), H_2O (50 ml) and brine (50 ml), dried (Na_2SO_4) and evaporated under reduced pressure. The residue was roughly purified by column chromatography (SiO_2 , 5–33% EtOAc in hexane) to give solids (4.89 g) as a mixture of **4** and tributyltin derivatives. A mixture of the above solids (400 mg), Et_3N (237 μl , 1.70 mmol), DMAP (208 mg, 1.70 mmol), 2,4,6-triisopropylbenzenesulfonyl chloride (515 mg, 1.70 mmol) in CH_3CN (5 ml) was stirred at room temperature for 28 h. The mixture was cooled in an ice bath, concentrated NH_4OH (25%, 10 ml) was added, and then stirred at room temperature for 1 h. The mixture was concentrated under reduced pressure and extracted with EtOAc (30 ml \times 2). The organic layer was washed with H_2O (10 ml \times 2) and brine (10 ml), dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography (SiO_2 , 0–4% MeOH in CHCl_3) to give **5** (305 mg, 67% from **3**) as a foam. FABMS m/z 670 (MH^+); ^1H NMR (CDCl_3) δ 7.99 (s, 1H, H-6), 7.48–7.28 (m, 9H, Ar-H), 6.90–6.88 (m, 4H, Ar-H), 6.37 (t, 1H, H-1', $J = 6.1$), 6.02 (dd, 1H, H-1'', $J = 17.6$, 11.2), 5.21 [d, 1H, H-2''(Z), $J = 17.6$], 5.02 (d, 1H, H-2''(E), $J = 11.2$), 4.47 (m, 1H, H-3'), 4.06 (m, 1H, H-4'), 3.86 (s, 6H, CH_3O), 3.55 (dd, 1H, H-5'a, $J = 10.5$, 2.9), 3.30 (dd, 1H, H-5'b, $J = 10.5$, 3.2), 2.57 (ddd, 1H, H-2'a, $J = 13.4$, 6.1, 4.6), 2.23 (ddd, 1H, H-2'b, $J = 13.4$, 6.4, 6.4), 0.88 [s, 9H, $(\text{CH}_3)_3\text{C-Si}$], 0.07–0.00 (m, 6H, $\text{CH}_3\text{-Si}$). Analysis calculated for $\text{C}_{38}\text{H}_{47}\text{N}_3\text{O}_6\text{Si}\cdot 1.2\text{H}_2\text{O}$, C, 66.00, H, 7.20, N, 6.08; found, C, 66.00, H, 6.96, N, 5.98.

3'-O-(*t*-Butyldimethylsilyl)-5-(1,2-dihydroxyethyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxycytidine (6)

A solution of OsO_4 in *t*-BuOH (5.0 mg/ml, 1.05 ml, 20.7 μmol) was added to a solution of **5** (277 mg, 0.414 mmol) and *N*-methylmorpholine *N*-oxide (97 mg, 0.828 mmol) in acetone/ $\text{H}_2\text{O}/t$ -BuOH (4:1:1, 6 ml) and the resulting mixture stirred at room temperature for 4 h. After the reaction was quenched with saturated aqueous $\text{Na}_2\text{S}_2\text{O}_3$ (10 ml) the mixture was extracted with EtOAc (20 ml \times 3). The organic layer was

washed with saturated aqueous $\text{Na}_2\text{S}_2\text{O}_3$ (20 ml) and brine (20 ml), dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography (SiO_2 , 0–10% MeOH in CHCl_3) to give **6** (225 mg, 77%) as a foam. FAB-MS m/z 704 (MH^+); ^1H NMR (CDCl_3) δ 7.89 (s, 0.33H, H-6 isomer A), 7.75 (s, 0.67H, H-6 isomer B), 7.49–7.28 (m, 9H, Ar-H), 6.94–6.91 (m, 4H, Ar-H), 6.34–6.29 (m, 1H, H-1'), 4.50 (m, 0.33H, H-3' isomer A), 4.43 (m, 0.67H, H-3' isomer B), 4.21 (m, 0.67H, *CHOH* isomer B), 4.05–4.02 (m, 1H, H-4'), 3.92 (m, 0.33H, *CHOH* isomer A), 3.86–3.85 (m, 6H, CH_3O), 3.67–3.46 (m, 3H, CH_2O , H-5'a), 3.34–3.30 (m, 1H, H-5'b), 2.52–2.44 (m, 1H, H-2'a), 2.24 (m, 0.33H, H-2'b isomer A), 2.16 (m, 0.67H, H-2'b isomer B), 0.87 [m, 9H, $(\text{CH}_3)_3\text{C-Si}$], 0.08–0.00 (m, 6H, $\text{CH}_3\text{-Si}$). HRMS (FAB) calculated for $\text{C}_{38}\text{H}_{50}\text{N}_3\text{O}_8\text{Si}$, 704.3367; found, 704.3380.

5-(1,2-Diacetoxyethyl)-*N*⁴-acetyl-3'-O-(*t*-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxycytidine (7)

A mixture of **6** (106 mg, 0.151 mmol), DMAP (2 mg, 16 μmol), Ac_2O (63.9 μl , 0.677 mmol) in pyridine (2 ml) was stirred at room temperature for 20 h. After EtOH (2 ml) was added the mixture was evaporated under reduced pressure. After successive co-evaporation with toluene the residue was dissolved in EtOAc (40 ml) and washed with H_2O (10 ml \times 3) and brine (10 ml), dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography (SiO_2 , 0–2% MeOH in CHCl_3) to give **7** (116 mg, 93%) as a foam. FAB-MS m/z 830 (MH^+); ^1H NMR (CDCl_3) δ 13.16, 13.11, 8.75, 8.62 (each br s, 1H, NH, exchanged with D_2O), 8.25, 8.14, 7.91, 7.80 (each s, 1H, H-6), 7.53–7.29 (m, 9H, Ar-H), 6.94–6.91 (m, 4H, Ar-H), 6.33, 6.18, 5.56, 5.45 (each m, 1H, *CHOAc*), 6.25 (m, 1H, H-1'), 4.50–4.05 (m, 4H, CH_2OAc , H-3' and H-4'), 3.87 (s, 6H, CH_3O), 3.61–3.51 (m, 1H, H-5'a), 3.40–3.26 (m, 1H, H-5'b), 2.50–2.42 (m, 1H, H-2'a), 2.23–2.14 (m, 1H, H-2'b), 2.34–1.49 (m, 9H, CH_3CO), 0.90 [s, 9H, $(\text{CH}_3)_3\text{C-Si}$], 0.09–0.00 (m, 6H, $\text{CH}_3\text{-Si}$). HRMS (FAB) calculated for $\text{C}_{44}\text{H}_{55}\text{N}_3\text{O}_{11}\text{Si}$, 830.3684; found, 830.3707.

5-(1,2-Diacetoxyethyl)-*N*⁴-acetyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxycytidine (8)

A solution of TBAF in THF (1.0 M, 216 μl , 0.22 mmol) was added to a solution of **7** (90 mg, 0.11 mmol) in THF (1.5 ml) in an ice bath and stirred at room temperature for 6 h. The reaction mixture was taken in EtOAc (25 ml), washed with H_2O (8 ml \times 3) and brine (10 ml), dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography (SiO_2 , 0–2% MeOH in CHCl_3) to give **8** (67 mg, 87%) as a foam. FAB-MS m/z 716 (MH^+); ^1H NMR (CDCl_3) δ 13.06, 13.00, 8.69, 8.56 (each br s, 1H, NH, exchanged with D_2O), 8.13, 7.98, 7.75, 7.63 (each s, 1H, H-6), 7.42–7.20 (m, 9H, Ar-H), 6.85–6.82 (m, 4H, Ar-H), 6.27–6.16 (m, 1H, H-1'), 6.14, 6.07, 5.49, 5.41 (each m, 1H, *CHOAc*), 4.45–3.96 (m, 4H, CH_2OAc , H-3' and H-4'), 3.78 (s, 6H, CH_3O), 3.53–3.34 (m, 2H, H-5'), 2.51–2.41 (m, 1H, H-2'a), 2.21–2.08 (m, 1H, H-2'b), 2.22–1.45 (m, 9H, CH_3CO). Analysis calculated for $\text{C}_{38}\text{H}_{41}\text{N}_3\text{O}_{11}$, C, 63.77, H, 5.77, N, 5.87; found, C, 63.58, H, 5.62, N, 5.66.

3'-O-[(Cyanoethyl)(*N,N*-diisopropylamino)phosphinyl]-5-(1,2-diacetoxyethyl)-*N*⁴-acetyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxycytidine (9**)**

2-Cyanoethyl diisopropylchlorophosphoramidite (187 μ l, 0.838 mmol) was added to a solution of **8** (300 mg, 0.419 mmol) and *N,N*-diisopropylethylamine (168 μ l, 0.964 mmol) in CH_2Cl_2 (4 ml) in an ice bath and stirred at room temperature for 0.5 h. The mixture was taken in CHCl_3 (50 ml) and washed with H_2O (20 ml \times 3) and brine (20 ml), dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography (neutral SiO_2 , 33–60% EtOAc in hexane) to give **9** (290 mg, 76%) as a foam. FAB-MS *m/z* 916 (MH^+); ^{31}P NMR (CDCl_3) δ 149.99, 149.53, 149.27. HRMS (FAB) calculated for $\text{C}_{47}\text{H}_{59}\text{N}_5\text{O}_{12}\text{P}$, 916.3897; found, 916.3888.

Synthesis of ODNs

ODNs containing 5-(1,2-dihydroxyethyl)cytidine (dheC, **2**) were synthesized on a DNA synthesizer (Applied Biosystem model 381A) by the phosphoramidite method. The fully protected ODNs were then deblocked and purified by the same procedure as for the purification of normal ODNs (27), i.e. each ODN linked to the resin was treated with concentrated NH_4OH at 55°C for 12 h and the released ODN protected by a DMTr group at the 5'-end was chromatographed on a C-18 silica gel column (1 \times 10 cm) (Waters) with a linear gradient of MeCN from 0 to 35% in 0.1 M TEAA buffer, pH 7.0. The fractions were concentrated and the residue was treated with aqueous 80% AcOH at room temperature for 20 min, then the solution was concentrated and the residue was co-evaporated with H_2O . The residue was dissolved in H_2O and the solution was washed with Et_2O , then the H_2O layer was concentrated to give the 5-(1,2-dihydroxyethyl)cytosine-containing ODNs **5** (34), **9** (18) and **13** (58). The yields are indicated in parentheses as $\text{OD}_{260\text{nm}}$ units, starting from the 1 μ mol scale.

Each ODN (5.0 OD_{260}) was treated with 50 mM aqueous NaIO_4 (50–100 equiv.) at 4°C for 30 min and then 1 M glycerol was added to the mixture. The ODN was purified by C-18 HPLC, using a J'sphere ODS M80 column (4.6 \times 150 mm) (YMC) with a linear gradient of MeCN from 7 to 21% in 0.1 M TEAA buffer, pH 7.0. The ODN was de-salted using a C-18 silica gel column (dispo SPE; YMC) to yield the 5-formylcytosine-containing ODNs **6** (3.4), **10** (1.4) and **14** (1.6). The yields are indicated in parentheses as $\text{OD}_{260\text{nm}}$ units.

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

Spectra were obtained on a Voyager Elite reflection time-of-flight mass spectrometer (PerSeptive Biosystems, Framingham, MA) equipped with a nitrogen laser (337 nm, 3 ns pulse). 2,4,6-Trihydroxyacetophenone was used as the matrix, which was prepared according to the Voyager Biospectrometry Workstation Use's Guide. Spectra were recorded in linear negative ion mode. T_9 and T_{14} were used as external standards. ODN **6**: calculated mass, 5649.7; observed mass, 5651.3. ODN **10**: calculated mass, 5609.7; observed mass, 5609.9. ODN **14**: calculated mass, 7648.0; observed mass, 7649.3.

Thermal denaturation

Each solution, which contained each ODN (3 μ M) and a complementary DNA (3 μ M) in a buffer of 10 mM sodium cacodylate (pH 7.0) and 10 mM NaCl, was heated at 95°C for 5 min. Then the solution was cooled gradually to an appropriate temperature and used for the thermal denaturation studies. Thermally induced transitions of each mixture were monitored at 260 nm on a Beckman DU 650 spectrophotometer. Sample temperature was increased at 0.5°C/min.

Single nucleotide insertion reactions

Primer 5'-termini were labeled using $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and T4 polynucleotide kinase. The single nucleotide insertion reaction was carried out in a buffer (total 10 μ l) containing an ODN template (2.0 pmol) annealed with the primer (1.0 pmol), 10 mM Tris-HCl (pH 7.5), 5 mM MgCl_2 , 7.5 mM DTT, 2.0 μ g BSA, 10 μ M dNTP and 0.5 U *Escherichia coli* Klenow fragment (KF) (either exo^+ or exo^-). After being incubated at 25°C for 20 min the reaction mixture was added to loading buffer (9 M urea, 10 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue, 40 μ l), then the mixture was heated at 95°C for 5 min. An aliquot (9 μ l) of the mixture was analyzed by electrophoresis on a 20% polyacrylamide gel containing 8 M urea (28). Densities of radioactivity of the gel were visualized with a Bas 2000 bio-imaging analyzer (Fuji Co.).

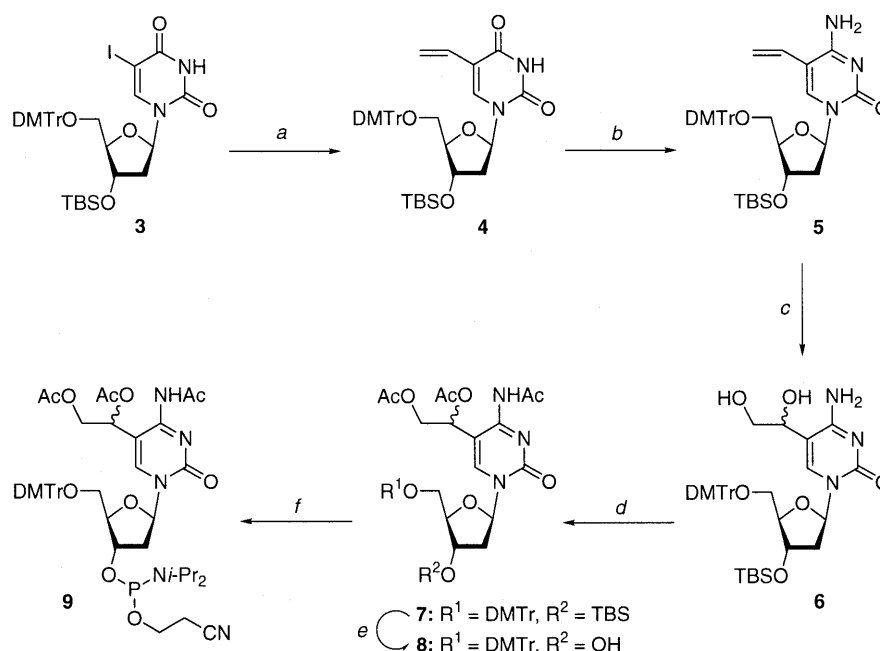
Kinetic analysis of the single nucleotide insertion reactions

Kinetic studies of incorporation of nucleotides opposite C, mC or fC during DNA synthesis were carried out under the conditions described above using 0.001–800 μ M dNTP and 0.02 U KF (exo^-). Reaction times were adjusted to give extents of reaction of 17% or less. The Michaelis constant (K_m) and the maximum rate of reaction (V_{max}) were obtained by a least squares fit of the kinetic data to the Michaelis–Menten curve using Kaleida-Graph v.3.0.

RESULTS AND DISCUSSION**Synthesis**

Previously we reported the first chemical synthesis of ODNs containing fdU using *N,N*-di(3,5-dichlorophenyl)ethylene-diamine (diCPhEDA) to protect the formyl group of fdU (29). However, this method required rather drastic conditions for deprotection of the formyl group, e.g. 80% $\text{CH}_3\text{CO}_2\text{H}$ overnight, which is not suitable for synthesis of longer ODNs. Recently another synthetic method was reported by Sugiyama *et al.* (30), who synthesized 5-(1,2-dihydroxyethyl)uracil-containing ODNs as precursors of fdU-containing ODNs. These were then converted to fdU-containing ODNs using sodium periodate. Since Sugiyama's method is more efficient than ours, we therefore applied it to the synthesis of fC-containing ODNs.

The phosphoramidite block of 5-(1,2-dihydroxyethyl)-2'-deoxycytidine **9** was prepared as shown in Scheme 1. Pd(II)-catalyzed vinylation of the 5-iodo-2'-deoxyuridine derivative **3** followed by transformation of the 4-oxo group to an amino group using a known method gave the 5-vinyl-2'-deoxycytidine derivative **5** in 67% yield in three steps. Compound **5** was then converted to the 5-(1,2-dihydroxyethyl)-2'-deoxycytidine derivative **6** in 77% yield with OsO_4 in the presence of *N*-methylmorpholine



Scheme 1. Conditions: (a) $\text{Bu}_3\text{SnCH}=\text{CH}_2$, $(\text{Ph}_3\text{P})_2\text{PdCl}_2$, DMF, 80°C , 1.5 h; (b) (1) 2,4,6-triisopropylbenzenesulfonyl chloride, DMAP, Et_3N , CH_3CN , rt, 28 h; (2) conc. NH_4OH , 0°C –rt, 1 h, 67% (three steps); (c) OsO_4 , 4-methylmorpholine-*N*-oxide, acetone- H_2O -*t*-BuOH (4:1:1), rt, 4 h, 77%; (d) Ac_2O , DMAP, py, rt, 20 h, 93%; (e) TBAF, THF, 0°C –rt, 6 h, 87%; (f) *i*- $\text{Pr}_2\text{N}(\text{Cl})\text{O}(\text{CH}_2)_2\text{CN}$, *i*- Pr_2NEt , CH_2Cl_2 , 0°C –rt, 0.5 h, 76%.

N-oxide. Acetylation of **6** with Ac_2O afforded the triacetyl derivative **7** in 93% yield. After deprotection of the TBS groups on the sugar moiety with tetrabutylammonium fluoride (TBAF) in THF the resulting nucleoside **8** was phosphitylated by a standard procedure (27) to produce the nucleoside 3'-phosphoramidite **9** in 76% yield.

The ODNs containing 5-(1,2-dihydroxyethyl)-2'-deoxycytidine (dheC, **2**) were synthesized on a DNA synthesizer. The average coupling yield of **9** was 95% using a 0.12 M solution of the amidite derivative in CH_3CN and 300 s as the coupling time. The fully protected ODNs (each 1 μmol scale) linked to a solid support were treated with concentrated NH_4OH at 55°C for 12 h. This was followed by C-18 column chromatography. De-tritylation gave the 5-(1,2-dihydroxyethyl)cytosine-containing ODNs **5**, **9** and **13**, which were oxidized with sodium periodate to produce the fC-containing ODNs **6**, **10** and **14** (Fig. 2). Conversion of the dheC-containing ODNs to the fC-containing ODNs was quantitative on HPLC analysis.

Enzymatic digestion of the isolated peak by snake venom phosphodiesterase and calf intestine alkaline phosphatase provided C, G, fC, T and A in a 2.9:7.1:1.0:0.94:5.9 ratio, confirming the structure of ODN **10**. Furthermore, ODNs **6**, **10** and **14** were analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and the observed molecular weights supported their structures (see Materials and Methods).

pK_a measurement of fC

We determined the pK_a of the base moiety of fC since the pK_a is thought to be an important parameter for understanding the base pairing properties of fC. The pK_a values of the base moieties of C, mC and fC are listed in Table 1. The pK_a of fC was 2.4. On the other hand, those of C and mC were 4.4 and

5'-CAGCTTTCGCTCX-3'
ODN **1** (primer **1**): X = C
ODN **2** (primer **2**): X = G

5'-GAGAXGGAGCGAAAGCTG-3'
ODN **3** (template **1**): X = C
ODN **4** (template **2**): X = mC
ODN **5**: X = dheC
ODN **6** (template **3**): X = fC

5'-GAGAXCGAGCGAAAGCTG-3'
ODN **7** (template **4**): X = C
ODN **8** (template **5**): X = mC
ODN **9**: X = dheC
ODN **10** (template **6**): X = fC

5'-CATCGATAGCATCCGXCACAGGCAG-3'
ODN **11**: X = T, ODN **12**: X = mC,
ODN **13**: X = dheC, ODN **14**: X = fC

5'-CTGCCTGTGXCGGATGCTATCGATG-3'
ODN **15**: X = G, ODN **16**: X = A,
ODN **17**: X = C, ODN **18**: X = T

Figure 2. Sequences of ODNs.

4.5, respectively. It was found that the pK_a of fC is 2 units lower than those of C and mC.

UV melting studies of duplexes

The stabilities of duplexes containing A, G, C or T opposite fC were studied by thermal denaturation in a buffer of 0.01 M sodium cacodylate (pH 7.0) containing 0.01 M NaCl. One

Table 1. pK_a values of cytidine and its derivatives^a

Nucleoside	pK_a
C	4.4 (4.3 ^b)
mC	4.5 (4.4 ^b)
fC	2.4

^aA 0.2 M NaCl solution was used for all pH ranges instead of buffer.^bData taken from Dawson *et al.* (32).**Table 2.** Hybridization data^a

Duplexes (base pairs)	T_m (°C)	ΔT_m (°C) ^b
ODN 11/ODN 16 (T:A)	60	
ODN 12/ODN 15 (mC:G)	61	
ODN 14/ODN 15 (fC:G)	61	
ODN 14/ODN 16 (fC:A)	55	-6
ODN 14/ODN 17 (fC:C)	53	-8
ODN 14/ODN 18 (fC:t)	55	-6

^aExperimental conditions are described in Materials and Methods.^b $\Delta T_m = [T_m(\text{each ODN}) - T_m(\text{ODN 14/ODN 15})]$.

transition was observed in the melting profile of each duplex. Melting temperatures (T_m) and ΔT_m values [T_m (mismatched DNA containing fC:A, fC:C or fC:T base pairs) - T_m (matched DNA containing a fC:G base pair)] are summarized in Table 2. T_m values of DNAs containing T:A or mC:G base pairs were 60 and 61°C, respectively. T_m of DNA containing a fC:G base pair (61°C) was the same as that of DNA containing a mC:G base pair. On the other hand, T_m values of DNAs containing fC:A, fC:C or fC:T base pairs were significantly smaller than that of DNA containing a fC:G base pair. ΔT_m values of the matched DNAs were from -6 to -8°C.

Incorporation of a single dNMP opposite the site of fC

Next we identified the nucleotides incorporated opposite the site of fC in the template during DNA synthesis. Two kinds of DNA polymerase, the Klenow fragment of *Escherichia coli* DNA polymerase I with and without 3'→5' exonuclease [KF (exo⁺) and KF (exo⁻)], were used in this study. Primer 1 (ODN 1), labeled at its 5'-end with ³²P, was annealed to templates 1-3 (ODNs 3, 4 and 6) containing C, mC or fC and extended by KF (exo⁺) or KF (exo⁻) in the presence of a single dNTP. The reaction products were analyzed by 20% PAGE containing 8 M urea (Fig. 3). Extension ratios (%) are summarized in Table 3. KF (exo⁺) and KF (exo⁻) inserted dGMP opposite fC (exo⁻, 94%; exo⁺, 97%) as efficiently as dGMP opposite C and mC. On the other hand, dCMP was virtually unincorporated opposite C, mC and fC. When KF (exo⁺) was used, the incorporation ratio for TMP opposite fC was four times greater than those of TMP opposite C and mC (C, 11%; mC, 11%; fC, 43%). When KF (exo⁻) was used, the incorporation ratio for dAMP opposite fC was slightly smaller than those of dAMP opposite C and mC (C, 89%; mC, 81%; fC, 66%). The order of incorporation ratios of dNMPs opposite fC was as follows:

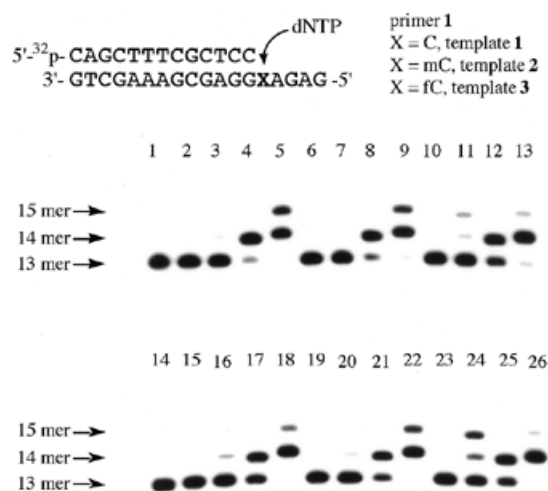


Figure 3. Primer extension assay to identify the nucleotide inserted opposite fC by KF (exo⁻) and KF (exo⁺). Lanes 2-13, KF (exo⁻); lanes 15-26, KF (exo⁺). Lanes 2-5 and 15-18, template 1; lanes 6-9 and 19-22, template 2; lanes 10-13 and 23-26, template 3. Lanes 2, 6, 10, 15, 19 and 23, dCTP; lanes 3, 7, 11, 16, 20 and 24, TTP; lanes 4, 8, 12, 17, 21 and 25, dATP; lanes 5, 9, 13, 18, 22 and 26, dGTP. Lanes 1 and 14, primer 1. Experimental conditions are described in Materials and Methods.

Table 3. Extension ratios (%)

		dCTP	TTP	dATP	dGTP
Template 1 (X = C)	KF (exo ⁻)	1	6	89	96
	KF (exo ⁺)	2	11	58	96
Template 2 (X = mC)	KF (exo ⁻)	2	5	81	95
	KF (exo ⁺)	3	11	64	96
Template 3 (X = fC)	KF (exo ⁻)	2	13	66	94
	KF (exo ⁺)	3	43	66	97
Template 4 (X = C)	KF (exo ⁻)	3	11	46	92
	KF (exo ⁺)	7	15	26	95
Template 5 (X = mC)	KF (exo ⁻)	2	5	46	93
	KF (exo ⁺)	7	14	37	94
Template 6 (X = fC)	KF (exo ⁻)	3	23	71	93
	KF (exo ⁺)	7	51	53	94

dGMP > dAMP > TMP > dCMP. From these results it was noted that TMP was more frequently incorporated opposite fC than C and mC.

To investigate whether the incorporation frequency of dNMPs opposite fC was affected by the nearest neighbouring base, a primer extension study was performed using primer 2 (ODN 2) and templates 4-6 (ODNs 7, 8 and 10) (Fig. 4). Extension ratios (%) are listed in Table 3. It was found that dGMP was likewise incorporated opposite fC (exo⁻, 93%; exo⁺, 94%) as efficiently as dGMP opposite C and mC. Furthermore, TMP was also incorporated more frequently opposite fC than C and mC (exo⁻: C, 15%; mC, 14%; fC, 51%). However, the incorporation ratio for dAMP opposite fC was significantly greater than those of dAMP opposite C and mC

Table 4. Kinetic parameters of the nucleotide insertion reaction by exonuclease-deficient Klenow fragment^a

A						B					
5'- ³² P-CAGCTTTCGCTCC 3'-GTCGAAAGCGAGGXAGAG-5'						5'- ³² P-CAGCTTTCGCTCG 3'-GTCGAAAGCGAGCXAGAG-5'					
primer 1 template 1-3						primer 2 template 4-6					
X	dNTP	K _m (μM)	V _{max} (%·min ⁻¹)	V _{max} /K _m (%·min ⁻¹ ·μM ⁻¹)	Misinsertion ratio (f) ^b	X	dNTP	K _m (μM)	V _{max} (%·min ⁻¹)	V _{max} /K _m (%·min ⁻¹ ·μM ⁻¹)	Misinsertion ratio (f) ^b
C ← G		0.012	12	1.0 × 10 ³	1	C ← G		0.010	14	1.4 × 10 ³	1
C ← A		49	5.2	1.1 × 10 ⁻¹	1.1 × 10 ⁻⁴	C ← A		22	0.79	3.6 × 10 ⁻²	2.6 × 10 ⁻⁵
C ← T		260	3.7	1.4 × 10 ⁻²	1.4 × 10 ⁻⁵	C ← T		80	1.2	1.5 × 10 ⁻²	1.1 × 10 ⁻⁵
mC ← G		0.009	9.5	1.1 × 10 ³	1	mC ← G		0.013	13	1.0 × 10 ³	1
mC ← A		34	5.4	1.6 × 10 ⁻¹	1.5 × 10 ⁻⁴	mC ← A		18	1.0	5.6 × 10 ⁻²	5.6 × 10 ⁻⁵
mC ← T		270	4.6	1.7 × 10 ⁻²	1.6 × 10 ⁻⁵	mC ← T		71	1.1	1.5 × 10 ⁻²	1.5 × 10 ⁻⁵
fC ← G		0.006	3.9	6.5 × 10 ²	1	fC ← G		0.011	6.6	6.0 × 10 ²	1
fC ← A		43	4.0	9.3 × 10 ⁻²	1.4 × 10 ⁻⁴	fC ← A		12	1.2	1.0 × 10 ⁻¹	1.7 × 10 ⁻⁴
fC ← T		110	2.5	2.3 × 10 ⁻²	3.5 × 10 ⁻⁵	fC ← T		55	1.3	2.4 × 10 ⁻²	3.9 × 10 ⁻⁵

^aExperimental conditions are described in Materials and Methods.
^bf, misinsertion ratio; $f = (V_{max}/K_m)_{dNTP} / (V_{max}/K_m)_{dGTP}$.

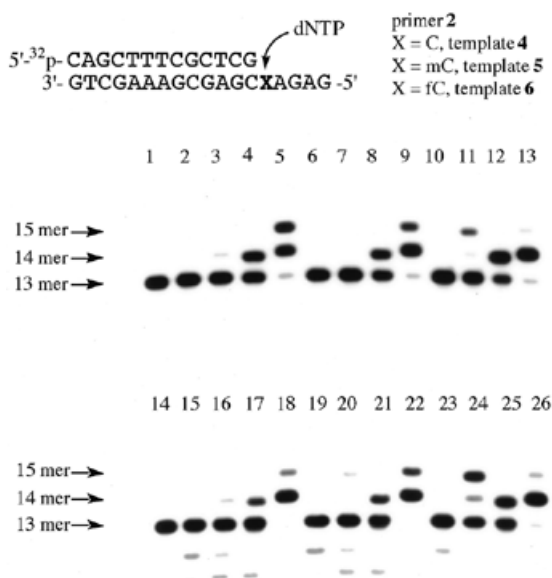


Figure 4. Primer extension assay to identify the nucleotide inserted opposite fC by KF (exo⁻) and KF (exo⁺). Lanes 2–13, KF (exo⁻); lanes 15–26, KF (exo⁺). Lanes 2–5 and 15–18, template 4; lanes 6–9 and 19–22, template 5; lanes 10–13 and 23–26, template 6. Lanes 2, 6, 10, 15, 19 and 23, dCTP; lanes 3, 7, 11, 16, 20 and 24, TTP; lanes 4, 8, 12, 17, 21 and 25, dATP; lanes 5, 9, 13, 18, 22 and 26, dGTP. Lanes 1 and 14, primer 2. Experimental conditions are described in Materials and Methods.

(exo⁻: C, 46%; mC, 46%; fC, 71%. exo⁺: C, 26%; mC, 37%; fC, 53%).

To study misinsertion of dNMPs opposite fC in detail the kinetic parameters of dNTP insertion catalyzed by KF (exo⁻) were measured during DNA synthesis. The Michaelis constant (K_m) and the maximum rate of reaction (V_{max}) were determined experimentally: the misinsertion ratio (f) was calculated from V_{max}/K_m ratios according to the equation $f = (V_{max}/K_m)_{dNTP} / (V_{max}/K_m)_{dGTP}$. The kinetic parameters are summarized in Table 4. The V_{max}/K_m values for dGMP incorporation opposite fC was

about half that opposite mC in both sequences. Thus, insertion of dGMP opposite fC seemed to be less efficient relative to that opposite mC. The f value for dAMP incorporation opposite mC and fC was almost identical when primer 1 was used, whereas that of dAMP incorporation opposite fC was ~3-fold larger than that opposite mC when primer 2 was used. The f value for TMP incorporation opposite fC was ~2- and 2.5-fold greater than those opposite mC in primers 1 and 2, respectively. Thus, TMP was found to be misincorporated more frequently opposite fC than opposite mC.

DISCUSSION

In this study we had hoped to find answers to the hypothesis ‘fC formation might be one cause of the C→T transition mutations frequently found in cytosine methylation sites’. So far C→T transition mutations have been thought to be the result of spontaneous or methylase-catalyzed hydrolytic deamination of mC to T (11–13). Indeed, it is known that cytosine methylation in vertebrates takes place exclusively in CpG sequences, which are hot-spots for mutations (1,2). Furthermore, mC residues in denatured DNA were shown to undergo alkali- or heat-induced deamination at rates ~1.5–3 times greater than those of C residues (12). However, we have recently found that fC is formed by γ-irradiation of mC or double-stranded DNA fragments containing mC and by aerobic treatment of DNA with Fenton-type reagents (24). Thus, if fC is as miscoding and mutagenic as fU, fC formation might be one cause of the C→T transition mutations frequently found in cytosine methylation sites.

It is known that an intramolecular hydrogen bond between the carbonyl of the 5-formyl group and the 4-amino function is present in the 5-formylcytosine moiety of its ribonucleoside (25,26). Thus, the equilibrium between the amino and imino tautomers may be affected by this hydrogen bonding and may allow the formation of a base pair with adenine. To investigate this hypothesis, a thermal denaturation study was carried out using duplexes containing fC. However, it was found that

fC:A, fC:C and fC:T base pairs significantly reduced the thermal stability of the duplexes.

Next, primer extension reactions were performed using fC-containing ODNs as templates. The kinetic parameters of dNTP insertion catalyzed by KF (exo⁻) were measured during DNA synthesis. It was found that: (i) insertion of dGMP opposite fC appears to be less efficient relative to insertion opposite mC; (ii) dAMP is misincorporated more frequently opposite fC than mC, although the frequency of misincorporation seems to be dependent on the sequence; (iii) TMP is misincorporated more frequently opposite fC than opposite mC. These results suggest that fC may induce the transition mutation C·G→T·A and transversion mutation C·G→A·T during DNA synthesis.

The pK_a of fC was found to be 2.4, whereas those of C and mC were 4.4 and 4.5, respectively. These results indicate that substitution of the electron withdrawing formyl group (fC) for the electron donating methyl group (mC) clearly alters the electronic structure of the pyrimidine ring. Recently Morales and Kool reported that in the majority of cases no hydrogen bonds between the incoming nucleotide and the opposite base in the template appear to be needed during the nucleotide insertion step with DNA polymerases if the requirements of correct Watson-Crick geometry and specific minor groove interactions between the enzyme and DNA are fulfilled (31). Thus, the differences in shape and electronic structure between fC and mC may influence single nucleotide insertion with KF (exo⁻).

In conclusion, we efficiently synthesized ODNs containing fC by the phosphoramidite method and subsequent oxidation with sodium periodate. Single nucleotide insertion reactions were carried out with KF (exo⁻). It was found that fC can be used to induce the transition mutation C·G→T·A and transversion mutation C·G→A·T during DNA synthesis. Thus fC formation may be one cause of the C→T transition mutations frequently found in cytosine methylation sites.

ACKNOWLEDGEMENTS

This paper constitutes part 210 of Nucleosides and Nucleotides; for part 209 in this series see Shuto *et al.* (33). We thank Dr H. Kamiya (Hokkaido University) for valuable comments and Y. Misawa (Hokkaido University) for technical assistance. This work was supported in part by a Grant-in-Aid for Encouragement of Young Scientists from the Ministry of Education, Science, Sports and Culture of Japan and a Grant from the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF97I00301).

REFERENCES

- Razin, A. and Riggs, A.D. (1980) DNA methylation and gene function. *Science*, **210**, 604–610.
- Verdine, G.L. (1994) The flip side of DNA methylation. *Cell*, **76**, 197–200.
- Cedar, H. (1988) DNA methylation and gene activity. *Cell*, **53**, 3–4.
- Hergersberg, M. (1991) Biological aspects of cytosine methylation in eukaryotic cells. *Experientia*, **47**, 1171–1185.
- Boyes, J. and Bird, A. (1991) DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*, **64**, 1123–1134.
- Barlow, D.P. (1993) Methylation and imprinting: from host defense to gene regulation? *Science*, **260**, 309–310.
- Groudine, M., Eisenman, R. and Weintraub, H. (1981) Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. *Nature*, **292**, 311–317.
- Doerfler, W. (1991) Patterns of DNA methylation—evolutionary vestiges of foreign DNA inactivation as a host defense mechanism. A proposal. *Biol. Chem. Hoppe Seyler*, **372**, 557–564.
- Rideout, W.M., Coetsee, G.A., Olumi, A.F. and Jones, P.A. (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science*, **249**, 1288–1290.
- Magewu, A.N. and Jones, P.A. (1994) Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. *Mol. Cell. Biol.*, **14**, 4225–4232.
- Shen, J.C., Rideout, W.M. and Jones, P.A. (1994) The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. *Nucleic Acids Res.*, **22**, 972–976.
- Wang, R.Y., Kuo, K.C., Gehrke, C.W., Huang, L.H. and Ehrlich, M. (1982) Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. *Biochim. Biophys. Acta*, **697**, 371–377.
- Yebrá, M.J. and Bhagwat, A.S. (1995) A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine. *Biochemistry*, **34**, 14752–14757.
- Murata-Kamiya, N., Kamiya, H., Muraoka, M., Kaji, H. and Kasai, H. (1997) Comparison of oxidation products from DNA components by γ -irradiation and Fenton-type reactions. *J. Radiat. Res.*, **38**, 121–131.
- Wakizaka, A., Aiba, N., Okuhara, E. and Kawazoe, Y. (1987) Production of 5-formyluracil from thymine in an *in vitro* active oxygen-generating system. *Biochem. Int.*, **14**, 289–295.
- Kasai, H., Iida, A., Yamaizumi, Z., Nishimura, S. and Tanooka, H. (1990) 5-Formyldeoxyuridine: a new type of DNA damage induced by ionizing radiation and its mutagenicity to salmonella strain TA102. *Mutat. Res.*, **243**, 249–253.
- Wagner, J.R., Hu, C.C. and Ames, B.N. (1992) Endogenous oxidative damage of deoxycytidine in DNA. *Proc. Natl. Acad. Sci. USA*, **89**, 3380–3384.
- Jaruga, P. and Dizdaroglu, M. (1996) Repair of products of oxidative DNA base damage in human cells. *Nucleic Acids Res.*, **24**, 1389–1394.
- Yoshida, M., Makino, K., Morita, H., Terato, H., Ohyama, Y. and Ide, H. (1997) Substrate and mispairing properties of 5-formyl-2'-deoxyuridine 5'-triphosphate assessed by *in vitro* DNA polymerase reactions. *Nucleic Acids Res.*, **25**, 1570–1577.
- Fujiwara, K., Kamiya, H. and Kasai, H. (1998) The mutations induced by oxidatively damaged nucleotides, 5-formyl-dUTP and 5-hydroxy-dCTP, in *Escherichia coli*. *Nucleic Acids Res.*, **26**, 4582–4587.
- Klungland, A., Paulsen, R., Rolseth, V., Yamada, Y., Ueno, Y., Wiik, P., Matsuda, A., Seeberg, E. and Bjelland, S. (2001) 5-Formyluracil and its nucleoside derivatives confer toxicity and mutagenicity to mammalian cells by interfering with normal RNA and DNA metabolism. *Toxicol. Lett.*, **119**, 71–78.
- Bienvendu, C., Wagner, J.R. and Cadet, J. (1996) Photosensitized oxidation of 5-methyl-2'-deoxycytidine by 2-methyl-1,4-naphthoquinone: characterization of 5-(hydroperoxymethyl)-2'-deoxycytidine and stable methyl group oxidation product. *J. Am. Chem. Soc.*, **118**, 11406–11411.
- Privat, E. and Sowers, L.C. (1996) Photochemical deamination and demethylation of 5-methylcytosine. *Chem. Res. Toxicol.*, **9**, 745–750.
- Murata-Kamiya, N., Kamiya, H., Karino, N., Ueno, Y., Kaji, H., Matsuda, A. and Kasai, H. (1999) Formation of 5-formyl-2'-deoxycytidine from 5-methyl-2'-deoxycytidine in duplex DNA by fenton-type reactions and γ -irradiation. *Nucleic Acids Res.*, **27**, 4385–4390.
- Kawai, G., Yokogawa, T., Nishikawa, K., Ueda, T., Hashizume, T., McCloskey, J.A., Yokoyama, S. and Watanabe, K. (1994) Conformational properties of a novel modified nucleoside, 5-formylcytidine, found at the first position of the anticodon of bovine mitochondrial tRNA^{Met}. *Nucl. Nucl.*, **13**, 1189–1199.
- LaFrancosi, C.J., Fujimoto, J. and Sowers, L.C. (1998) Synthesis and characterization of isotopically enriched pyrimidine deoxynucleoside oxidation damage products. *Chem. Res. Toxicol.*, **11**, 75–83.
- Gait, M.J. (ed.) (1984) *Oligonucleotides Synthesis: A Practical Approach*. IRL Press, Oxford, UK.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Ono, A., Okamoto, T., Inada, M., Nara, H. and Matsuda, A. (1994) Synthesis and properties of oligonucleotides containing 5-formyl-2'-deoxyuridine. *Chem. Pharm. Bull.*, **42**, 2231–2237.
- Sugiyama, H., Matsuda, S., Kino, K., Zhang, Q.-M., Yonei, S. and Saito, I. (1996) New synthetic method of 5-formyluracil-containing oligonucleotides and their melting behavior. *Tetrahedron Lett.*, **37**, 9067–9070.

31. Morales, J.C. and Kool, E.T. (2000) Varied molecular interactions at the active sites of several DNA polymerases: nonpolar nucleoside isosteres as probes. *J. Am. Chem. Soc.*, **122**, 1001–1007.
32. Dawson, R.M.C., Elliot, D.C., Elliot, W.H. and Jones, K.M. (1986) *Data for Biochemical Research*, 3rd Edn. Clarendon Press, Oxford, UK.
33. Shuto, S., Fukuoka, M., Manikowsky, A., Ueno, Y., Nakano, T., Kuroda, R., Kuroda, H. and Matsuda, A. (2001) Total synthesis of cyclic ADP-carbocyclic-ribose, a stable mimic of Ca²⁺-mobilizing second messenger cyclic ADP-ribose. *J. Am. Chem. Soc.*, in press.