Neighbouring bases in template influence base-pairing of isoguanine

Agnieszka M. MACIEJEWSKA, Katarzyna D. LICHOTA and Jarosław T. KUŚMIEREK¹

Department of Molecular Biology, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 5A Pawińskiego Str., 02 106 Warsaw, Poland

Assuming that the efficiency of the incorporation of 5-methyl-2'-deoxyisocytosine-5' triphosphate (dMiCTP) and dTTP opposite isoguanine (iG) is a measure of the proportion of the keto and enol tautomers of iG in the *Thermus aquaticus* DNA polymerase active centre, we studied the influence of temperature and iG-neighbouring bases in the template on base-pairing of iG. On the basis of experiments with four sequences (3'-TXT-5', 3'-GXX-5', 3'-CXC-5', 3'-CXT-5', where X = iG) at 37, 50, 65 and 80 °C, we found that 3'-neighbours decrease the fraction of the keto tautomer in the order C > G ≥ T, whereas temperature apparently does not influence the tautomeric equilibrium of iG. The variability of the ratio of incorporation of dMiCTP versus dTTP (5–20) primarily reflects the variability of K_m values, since

V_{max} values are roughly similar, which indicates that the iG · MiC and iG · T pairs fit the polymerase active centre equally well. The altering of the base-pairing of iG by sequence context is discussed in relation to tautomerism and miscoding of this oxidized adenine derivative. A key derivative for preparation oligodeoxynucleotides, *O*²-diphenylcarbamoyl-*N*⁶-[(dimethylamino)ethylidene]-2'-deoxyisoguanosine, is extremely labile ($t_{1/2}$ = 3.5 min) in 3 % trichloroacetic acid/dichloromethane, i.e. under the conditions of automated DNA synthesis, which results in low yield and length heterogeneity of templates.

Key words: isoguanine–isocytosine pair, isoguanine–thymine pair, sequence context, tautomerism.

INTRODUCTION

Isoguanosine (also called crotonoside) was discovered 70 years ago as a constituent of the croton bean, *Croton tiglium* L. [1]. Although isoguanine (iG), as a ribonucleoside, and its methyl derivatives occurs in Nature [1–4], it has not been established that this base represents a natural constituent of nucleic acids. However, iG (1,2-dihydro-2-oxoadenine or 2-hydroxyadenine) is formed in DNA, among other oxidized bases, in the reaction of reactive oxygen species with adenine [5]. The miscoding, mutagenicity and repair of this oxidized adenine derivative have been extensively studied [6–12].

Benner and co-workers, in their search for 'new letters of the genetic alphabet', have found that iG can form a pair with another odd base, isocytosine (iC), and also with thymine (or uracil) [13–16]. They proposed that iC pairs with iG in its keto form, whereas thymine (or uracil) pairs with iG in its enol form (Figure 1). In fact, it had been shown previously by Shugar and co-workers that isoguanosine exists in solution as a mixture of these tautomeric forms, and their ratio depends on solvent polarity [17]. Of course, the ambiguity of pairing caused by tautomerism limits the use of iG as a 'letter in the genetic alphabet'; on the other hand, this fact makes it possible to study the influence of various factors on the tautomeric equilibrium of iG and its coding properties. Benner and co-workers [15,16] observed significant differences in the acceptance by various polymerases of iG · iC versus iG · T (or iG · U) pairs, studied in manifold combinations, i.e. in the template or in NTP, as ribo- or deoxyribo-derivatives. They postulated that one of the reasons for this phenomenon could be the differences in polarity of the microenvironment within active sites of various polymerases, which in turn influences the tautomeric status of iG.

Since the keto–enol equilibrium of iG depends on the polarity of its microenvironment, one can expect that the nature of neighbouring bases in the template would also influence this

equilibrium and could result in alteration of the coding properties of iG. Our former studies [8], with the use of randomly composed polyribonucleotides containing iG and avian myeloblastosis virus reverse transcriptase as the enzyme catalysing the template-directed synthesis, did not give evidence for such an influence. However, because of experimental limitations, we were able to test only two different neighbourhoods, namely iG situated between cytosines or between hypoxanthines. In the present study, we report the results of our studies on the influence of template neighbourhood and temperature on the coding properties of iG, with the use of synthetic oligodeoxynucleotides with iG located at pre-selected sites and *Thermus aquaticus* DNA (*Taq*) polymerase. We assume that the efficiency of incorporation of 2'-deoxy-5-methylisocytosine 5'-triphosphate (dMiCTP) and that of dTTP opposite iG determined in a kinetic assay is a measure

Figure 1 Base-pairing of iG

(A) iG in its keto tautomeric form (N1-H, 2-oxo, 6-amino) forms a pair with MiC. (B) iG in its enol tautomeric form (2-OH (*cis*), 6-amino) forms a pair with thymine (T). MiC and T can be replaced by iC and uracil respectively, without changing the pattern of hydrogen bonds in these Watson–Crick-type pairs. See also Figure 4 for some other tautomeric forms of iG.

Abbreviations used: AF, ammonium formate; AM, (dimethylamino)ethylidene group; DMT, 4,4-dimethoxytriphenyl group; DPC, diphenylcarbamoyl group; iC, isocytosine; iG, isoguanine; MiC, 5-methylisocytosine; RT, retention time; *Taq*, *Thermus aquaticus* DNA polymerase; TCA, trichloroacetic acid.

¹ To whom correspondence should be addressed (e-mail jareq@ibb.waw.pl).

of the proportion of tautomers of iG within the polymerase active site.

EXPERIMENTAL

Commercial enzymes and nucleotides

T4 polynucleotide kinase and *Thermus aquaticus* (Taq; strain YT1) DNA polymerase were purchased from Promega. Deoxythymidine 5'-triphosphate was from Pharmacia and [γ - 32 P]ATP (3000 Ci/mmol) was from Amersham.

Synthesis and stability of 5-methyl-2'-deoxyisocytidine 5'-triphosphate (dMiCTP)

The 5'-triphosphate was prepared according to the methods described in [8]. In contrast with the reported instability of 2'-deoxyisocytidine 5'-triphosphate [15], the 5-methyl derivative appears to be relatively stable. Samples of dMiCTP stored for years at -20°C in a water solution show some degradation, the extent of which depends on the number of its thawing and freezing cycles. A sample containing approx. 70% of non-degraded 5'-triphosphate, after complete dephosphorylation by bacterial alkaline phosphatase, showed the presence of 20% of the deprimidination product, MiC, in addition to the expected 2'-deoxy-5-methylisocytidine, whereas in a sample in which practically all material was in the 5'-triphosphate form, only 2% of free base was detected. The deamination product, 2'-deoxythymidine, was detected in both samples at a level of 0.1% using a Waters Nova-Pak[®] C₁₈, 4 μm , 4.6 mm $\times$ 250 mm HPLC column, with a linear gradient of Buffer A [20 mM ammonium formate (AF; pH 9.0)] and Buffer B [aq. 30% (v/v) methanol], using a flow rate of 1 ml/min for 30 min and a wavelength of 250 nm for detection. Retention times (RT) were as follows: MiC, 8.9 min; 2'-deoxy-5-methylisocytidine, 11.4 min; 2'-deoxythymidine, 14.1 min.

Synthesis of fully protected phosphoramidite of 2'-deoxyisoguanosine (diG)

2'-Deoxyisoguanosine (diG) was a gift from Dr Z. Kazimierzczuk [18]. The preparation of protected derivatives of diG [*N*⁶AM-diG, where AM is (dimethylamino)ethylidene, *O*²DPC-diG, where DPC is diphenylcarbonyl, and *O*²DPC-*N*⁶AM-diG] and its 5'-*O*-(4,4'-DMT-methyl) derivative, where DMT is 4,4'-dimethoxytriphenyl (5'-DMT-*O*²DPC-*N*⁶AM-diG), was accomplished using the methods reviewed in [19]. The derivatives were identified on the basis of UV spectra recorded in aq. 50% (v/v) methanol (*O*²DPC-diG, λ_{max} = 260 nm, *N*⁶AM-diG, λ_{max} = 335 nm, and *O*²DPC-*N*⁶AM-diG, λ_{max} = 310 nm) which were in conformity with published UV data [20]. The preparation of the 3'-*O*-[di-isopropylamino(2-cyanoethoxy)phosphino] derivative of 5'-DMT-*O*²DPC-*N*⁶AM-diG was accomplished following the procedure for adducts of 2'-deoxynucleosides [21].

The stability of protected diG derivatives under acidic conditions

The stability of *O*²DPC-diG, *N*⁶AM-diG and *O*²DPC-*N*⁶AM-diG in 0.1 M HCl in aq. 50% (v/v) methanol, and of *O*²DPC-diG and *O*²DPC-*N*⁶AM-diG in 3% (w/v) trichloroacetic acid (TCA) in CH₂Cl₂, was studied at 20–22 $^{\circ}\text{C}$. The concentration of the compounds tested was approx. 50 A_{max} units/ml. Reactions were quenched by addition of an excess of 1 M triethylammonium bicarbonate. Samples were evaporated at 37 $^{\circ}\text{C}$, redissolved in aq. 50% (v/v) methanol and analysed by HPLC using a Waters Symmetry C₈, 3.9 mm $\times$ 150 mm column. A linear gradient of aq. 50% and 80% (v/v) methanol over 10 min, then isocratic

Table 1 Acid instability of protected 2'-deoxyisoguanosine derivatives

Instability in 3% (v/v) TCA in CH₂Cl₂ and 0.1 M HCl in aq. 50% (v/v) methanol was measured at 20–22 $^{\circ}\text{C}$ as a disappearance of substrate followed by HPLC (the present study). Data for 0.1 M HCl, obtained by measuring at 25 $^{\circ}\text{C}$ changes of UV absorbance at selected wavelength of the whole reaction mixture, were taken from [20]. ND, not determined.

	$t_{1/2}$ (min)		
	3% TCA in CH ₂ Cl ₂	0.1 M HCl in aq. 50% MeOH	0.1 M HCl
<i>O</i> ² DPC-diG	100	160	125
<i>N</i> ⁶ AM-diG	ND	30	13
<i>O</i> ² DPC- <i>N</i> ⁶ AM-diG	3.5	72	284

Figure 2 Exemplary sequence of template with annealed complementary primer

The variable parts of sequences are shown in bold. The variable regions (positions 20–21) of the four studied templates are listed in Table 2.

elution with aq. 80% (v/v) methanol at a flow rate of 0.5 ml/min, was applied for separation of reaction products of *O*²DPC-diG (RT, 11.7 min; base RT, 12.2 min) and *O*²DPC-*N*⁶AM-diG (RT, 12.7 min; base RT, 13.1 min). Separations were monitored at 260 or 310 nm. For separation of reaction products of *N*⁶AM-diG (RT, 13.8 min; base RT, 12.3 min) a linear gradient of H₂O and aq. 30% (v/v) methanol over 10 min, then isocratic elution with aq. 30% (v/v) methanol at a flow rate of 0.5 ml/min, and detection at 335 or 260 nm was applied. The half-life values ($t_{1/2}$) of the compounds studied were calculated on the basis of disappearance of substrate followed by HPLC (Table 1).

Preparation of 25-mers containing diG

The fully protected phosphoramidite of diG was used for preparation of 25-mer oligodeoxynucleotides containing diG at position 21 relative to the 3'-termini (Figure 2), with the aid of a Beckman Oligo 1000M DNA Synthesizer under standard conditions. Deprotection of the oligonucleotides was performed in aq. 25% (v/v) NH₃ at 60 $^{\circ}\text{C}$ for 18 h. The 5'-DMT-oligomers were purified by HPLC with a Waters Prep Nova-Pak HR C₁₈, 3.9 mm $\times$ 300 mm column, using a linear gradient of Buffer A (20 mM AF, pH 7) and Buffer B [aq. 30% (v/v) methanol] over 30 min, with a flow rate of 1 ml/min and detection at 260 nm. The failure sequences (RT, 22–26 min) constituted 50–60% of the whole oligonucleotide material. The 5'-DMT-oligomer fractions (RT, 26–30 min) were collected and evaporated on a Speed Vac. The 5'-DMT residues were removed from the oligomers adsorbed on Waters Sep-Pak⁺ C₁₈ cartridges with 2% (v/v) trifluoroacetic acid, as described in [21]. The oligomers that eluted from the cartridges, after evaporation, were again HPLC purified under the above conditions. The peaks of oligomers (RT, 23–26 min) were evaporated, dissolved in TE buffer (10 mM Tris and 1 mM EDTA, pH 8) and stored at -20°C . The 25-mers were 5'- ^{32}P -labelled and subjected to PAGE

in order to estimate their integrity (see below for procedures). The amount of abortive products was 20–30 %, and 20-mers constituted about half of this amount. The 25-mers were used in experiments at this stage of purification, since additional purification did not improve their integrity.

Taq polymerase reactions and determination of kinetic parameters

Kinetic parameters for site-specific incorporation of dNTPs were determined using the methodology previously described [22,23]. Primers were labelled with T4 polynucleotide kinase and [γ - 32 P]ATP, usually 50 μ Ci per 0.0007 A_{260} of primer was used, and annealed to complementary templates. The polymerase reaction mixtures contained (in 6 μ l of buffer provided by the enzyme supplier): 0.0035 A_{260} /ml of 32 P-labelled primer, 0.0060 A_{260} /ml of template, 0.1 unit of enzyme and 0.1–100 μ M dMiCTP or 1–500 μ M dTTP. The mixtures were incubated at 37, 50, 65 or 80 °C for 90 s, and the reaction stopped by addition of 12 μ l of stop buffer [formamide/0.5 M Na₂-EDTA, 9:1 by vol.]. Samples (5–10 μ l) of the reaction mixtures were loaded on a 12 % polyacrylamide gel (20 cm $\times$ 40 cm $\times$ 0.04 cm) prepared in TBE buffer (89 mM Tris, 89 mM boric acid and 10 mM EDTA, pH 8) containing 8 M urea, and electrophoresed at 2500–3000 V (preset temperature 50 °C) using Bio-Rad Sequi-Gen apparatus, for approx. 1.5 h, until the Bromophenol Blue dye front had run to the bottom of the gel. After electrophoresis, gels were vacuum dried and overlaid with X-ray film for 1–24 h exposures. For determination of kinetic parameters, the gel autoradiographs were scanned on an Ultrosan XL LKB Bromma Densitometer and band intensities were evaluated with the Gelscan XL Pharmacia Data System. The K_m and V_{max} values were calculated using the Eisenthal and Cornish-Bowden direct linear plot algorithm [24,25].

The procedures for determination of kinetic parameters of incorporation of dTTP and dMiCTP were always performed in parallel, using the same dilutions of the enzyme and substrates, and the reactions were run on the same gel, in order to compensate and minimize possible errors. The relatively high heterogeneity of length of the deoxynucleotide templates (20–30 % of oligodeoxynucleotides were shorter than 25-mers) should not influence the final results, since in each experiment, both dNTPs interacted with the same population of templates.

Statistics

Statistical calculations were performed using the In Stat for Windows (GraphPad Software) computer program. For the analysis of statistical significance, the one-way analysis of variance (ANOVA) test and the Student–Newman–Keuls post-test were used.

RESULTS

Preparation of oligodeoxynucleotide templates containing diG

Incorporation of diG into DNA in solid-phase synthesis requires the protection of O^2 and N^6 of iG [15,19,26,31]. In our preparation of 25-mer templates containing diG, we followed the strategy elaborated by Seela and co-workers [19], i.e. AM protection of N^6 and DPC protection of O^2 of diG, and we used 5'-DMT- O^2 DPC- N^6 AM-diG phosphoramidite as a building block in automated DNA synthesis. We found, however, during purification of 5'-DMT-oligomers on a reversed-phase column, that failure sequences constituted more than half of the whole oligonucleotide material. HPLC analysis of the failure and full-

length oligonucleotide fractions digested to deoxynucleosides showed that the former contained approx. $5 \times$ less diG than the latter (results not shown). This indicates that the problem lies in the coupling of the diG building block or in its stability under the conditions of solid-phase synthesis. In order to check the second possibility, we re-investigated acid stability of the protected derivatives of diG [20]. In contrast with the quoted work, we studied this stability not only in aqueous solution but also in dichloromethane containing 3 % TCA, i.e. under conditions which would remove 5'-DMT groups in a DNA synthesizer. Another difference was that we determined the kinetics following the disappearance of the substrate peak in a HPLC chromatogram, monitored at various wavelengths, whereas the quoted authors measured UV changes at a chosen wavelength for the whole reaction mixture. Our results (together with relevant data from [20]) are gathered in Table 1. The enormous instability of O^2 DPC- N^6 AM-diG in TCA/CH₂Cl₂ ($t_{1/2}$ = 3.5 min) in comparison with 0.1 M HCl in aq. 50 % (v/v) methanol ($t_{1/2}$ = 72 min) was the most striking result of our studies. Another observed difference between non-aqueous and aqueous conditions was the composition of reaction mixtures. The HPLC profiles of aqueous hydrolysis of bis-protected diG monitored at 260 nm (λ_{max} for O^2 DPC-diG) showed a fast formation of two peaks with RT identical to the RT of peaks observed in aqueous and non-aqueous reactions of O^2 DPC-diG (11.7 and 12.2 min) in addition to the substrate (12.7 min) and its base (13.1 min) peaks. The same reaction monitored at 310 nm (λ_{max} for O^2 DPC- N^6 AM-diG) showed peaks with RTs of 12.7 and 13.1 min, whereas the pair of more polar compounds was barely detectable. This suggests that in aqueous solution, in addition to the relatively slow hydrolysis of the glycosyl bond in O^2 DPC- N^6 AM-diG, fast hydrolysis of the dimethylaminoethylidene group occurs in the bis-protected nucleoside and base. In contrast, the HPLC profiles of the TCA/CH₂Cl₂ reaction monitored at either wavelength, 260 or 310 nm, are consistent with the rationale that scission of the glycosyl bond in O^2 DPC- N^6 AM-diG is very fast, whereas N^6 -deprotection is slow. In reactions of mono-protected derivatives, O^2 DPC-diG and N^6 AM-diG, only scission of the glycosyl bond was observed. The HPLC profiles did not show the presence of polar compounds which would correspond to diG or iG. This indicates that the protecting groups, O^2 DPC and N^6 AM, in mono-protected derivatives are acid-stable under the conditions tested.

We did not study the mechanistic aspects of the enhanced acid-lability of the glycosyl bond of O^2 DPC- N^6 AM-diG in non-aqueous solution. However, practical consequences of this fact are very serious. In each cycle of solid-phase synthesis, the removal of acid-labile 5'-DMT protecting group causes substantial depurination at the iG site. The extent of entire depurination increases with the number of cycles after coupling of the diG building block and the duration of the 5'-DMT de-blocking step (four cycles, 75 s each, in the present study). The depurinated oligonucleotide undergoes chain scission at the apurinic site during deprotection of alkali-labile groups in aq. NH₃ at 60 °C. We suppose that the depurination at the iG site under the conditions of solid-phase synthesis is mainly responsible for the observed high content of failure sequences. Most likely, the lability of the glycosyl bond of diG (which is approx. 10-fold less stable than that of common purine deoxynucleosides [20]) is also responsible for the fragility of deprotected oligodeoxynucleotides, since we observed some degradation at the iG site during purification procedures.

The problem of stability of oligodeoxynucleotides containing diG has been left unaddressed by many investigators. It was, however, considered by Seela and Wei [27], who used oligomers

Table 2 Kinetic parameters of dMiCTP and dTTP incorporation opposite isoguanine by *Taq* polymerase

The K_m and V_{max} are the mean values ($\pm$ S.D.) of 3–6 independent experiments. The relative velocity of dNTP incorporation was measured as the ratio of the elongated primer (I_2) and primer (I_1) band intensities observed on a gel autoradiograph after 1.5 min reaction time. The f_{MIC} and f_t represent the incorporation efficiency of dMiCTP and dTTP respectively. They were calculated as the mean values ($\pm$ S.D.) of three most converged V_{max}/K_m determined in independent experiments. They may differ from the ratios directly calculated on the basis of K_m and V_{max} values presented in the Table. $F = f_{MIC}/f_t$ represents the relative incorporation efficiency of dMiCTP versus dTTP. In the sequences listed, the closest neighbourhood of isoguanine in template is shown.

Sequence (3' $\rightarrow$ 5') temperature	dMiCTP			dTTP			<i>F</i>
	K_m (μ M)	V_{max} (I_2/I_1) _{max}	f_{MIC} (μ M ⁻¹)	K_m (μ M)	V_{max} (I_2/I_1) _{max}	f_t (μ M ⁻¹)	
TiGT							
37 °C	2.7 $\pm$ 1.25	0.67 $\pm$ 0.18	0.263 $\pm$ 0.044	35.6 $\pm$ 10.9	0.55 $\pm$ 0.29	0.016 $\pm$ 0.003	16.4 $\pm$ 4.1
50 °C	1.35 $\pm$ 0.29	0.71 $\pm$ 0.07	0.540 $\pm$ 0.075	19.59 $\pm$ 5.22	0.75 $\pm$ 0.12	0.041 $\pm$ 0.004	13.2 $\pm$ 2.2
65 °C	4.35 $\pm$ 3.80	0.90 $\pm$ 0.49	0.321 $\pm$ 0.049	40.85 $\pm$ 3.43	0.54 $\pm$ 0.18	0.016 $\pm$ 0.002	20.1 $\pm$ 3.9
80 °C	1.56 $\pm$ 0.48	0.67 $\pm$ 0.11	0.419 $\pm$ 0.031	29.46 $\pm$ 7.43	0.78 $\pm$ 0.17	0.029 $\pm$ 0.002	14.4 $\pm$ 1.5
GiGG							
37 °C	5.11 $\pm$ 2.76	0.50 $\pm$ 0.22	0.115 $\pm$ 0.014	36.31 $\pm$ 25.35	0.33 $\pm$ 0.20	0.010 $\pm$ 0.002	11.5 $\pm$ 2.7
50 °C	5.22 $\pm$ 0.77	1.37 $\pm$ 0.12	0.250 $\pm$ 0.016	63.31 $\pm$ 20.71	1.18 $\pm$ 0.20	0.019 $\pm$ 0.003	13.2 $\pm$ 2.2
65 °C	5.39 $\pm$ 4.50	0.78 $\pm$ 0.66	0.162 $\pm$ 0.017	53.13 $\pm$ 49.77	0.65 $\pm$ 0.34	0.008 $\pm$ 0.001	20.3 $\pm$ 3.3
80 °C	4.63 $\pm$ 1.00	1.12 $\pm$ 0.21	0.246 $\pm$ 0.044	47.23 $\pm$ 14.31	1.00 $\pm$ 0.23	0.027 $\pm$ 0.006	9.1 $\pm$ 2.6
CiGC							
37 °C	5.37 $\pm$ 1.05	0.25 $\pm$ 0.14	0.054 $\pm$ 0.008	47.04 $\pm$ 24.45	0.26 $\pm$ 0.11	0.006 $\pm$ 0.0004	9.0 $\pm$ 1.5
50 °C	3.12 $\pm$ 1.96	0.19 $\pm$ 0.04	0.064 $\pm$ 0.022	63.31 $\pm$ 43.55	0.48 $\pm$ 0.27	0.008 $\pm$ 0.0008	8.0 $\pm$ 2.9
65 °C	5.98 $\pm$ 4.78	0.67 $\pm$ 0.36	0.063 $\pm$ 0.019	42.76 $\pm$ 15.67	0.41 $\pm$ 0.16	0.008 $\pm$ 0.0021	7.9 $\pm$ 3.1
80 °C	6.62 $\pm$ 2.40	0.35 $\pm$ 0.01	0.049 $\pm$ 0.008	76.43 $\pm$ 56.91	0.59 $\pm$ 0.45	0.008 $\pm$ 0.001	6.1 $\pm$ 1.3
CiGT							
37 °C	10.05 $\pm$ 8.89	0.75 $\pm$ 0.80	0.077 $\pm$ 0.012	102.24 $\pm$ 15.22	0.82 $\pm$ 0.35	0.009 $\pm$ 0.002	8.6 $\pm$ 2.3
50 °C	17.11 $\pm$ 6.80	1.12 $\pm$ 0.46	0.075 $\pm$ 0.012	79.11 $\pm$ 52.61	1.03 $\pm$ 0.45	0.015 $\pm$ 0.004	5.0 $\pm$ 1.6
65 °C	13.61 $\pm$ 6.53	0.58 $\pm$ 0.17	0.048 $\pm$ 0.008	77.91 $\pm$ 13.56	0.57 $\pm$ 0.16	0.007 $\pm$ 0.0007	6.9 $\pm$ 1.3
80 °C	16.85 $\pm$ 2.52	1.14 $\pm$ 0.29	0.067 $\pm$ 0.007	72.20 $\pm$ 18.94	1.11 $\pm$ 0.35	0.014 $\pm$ 0.001	4.8 $\pm$ 0.6

Figure 3 Influence of template context on the relative incorporation efficiency of dMiCTP versus dTTP opposite iG (parameter *F*) determined at various temperatures

Asterisks denote statistical significance: **P* < 0.05 (significant), ***P* < 0.01 (very significant) and ****P* < 0.001 (extremely significant). See Table 2 and text for details.

containing 7-deaza-2'-deoxyisoguanosine in their studies, instead of the unstable oligomers containing diG. The deaza analogue has similar base-pairing properties and is much more stable than diG.

Four 25-mer templates with an iG residue at position 21 relative to the 3'-terminus and three complementary 20-mer primers were used to study the incorporation of dMiCTP and dTTP with the aid of *Taq* polymerase. An example sequence of the template with annealed primer is shown in Figure 2, and the variable regions of the templates (positions 20–22 relative to the 3'-terminus) are listed in Table 2.

The influence of neighbouring bases in template and temperature on incorporation of dTTP and dMiCTP opposite iG by *Taq* polymerase

Kinetic parameters (K_m and V_{max}) for the incorporation of dTTP and dMiCTP opposite iG located in four different template neighbourhoods were determined at four different temperatures. The parameters f_T and f_{MiC} , representing the efficiency of incorporation of dTTP and dMiCTP respectively, as well as the parameter $F = f_{MiC}/f_T$, were calculated and are presented in Table 2. The ratio of efficiency of incorporation of dMiCTP versus dTTP (parameter F) varies within the range 5–20. A statistical analysis of the influence of iG flanking bases on the variance of F obtained for each temperature is shown in Figure 3. The differences between TiGT and GiGG versus CiGC and CiCT sequences vary from significant ($P < 0.05$) to extremely significant ($P < 0.001$), depending on the temperature. The difference between TiGT and GiGG was significant ($P < 0.01$) only at 80 °C, whereas no significant differences were found between CiGC and CiGT sequences at all temperatures tested.

An analogous analysis of the influence of temperature on the variance of F was carried out for all four sequences. With one exception, where for the sequence GiGG the F value determined at 65 °C was different ($P < 0.05$) from the values determined at 37, 50 and 80 °C, it does not appear that F is significantly influenced by temperature (results not shown).

DISCUSSION

The accuracy of DNA replication results from selective forces at the active site of the replicating enzyme. These include not only the formation of hydrogen-bonded Watson–Crick base pairs between the primer–template and the incoming nucleotide, but also specific interactions between all these elements and amino acid residues of the DNA polymerase. The elucidation of the crystal structures of complexes of DNA polymerases with primer–template and dNTPs has brought about a significant progress in our knowledge about specific interactions within the polymerase active site (see [28] for a recent review). Since this knowledge is fairly new and limited to the relatively few examples of oligonucleotides and dNTPs containing natural bases, the discussion of the coding properties of modified bases has to be restricted primarily to such topics as tautomeric equilibrium of these bases, energy of base–base interactions and conformity with the geometry of the DNA double helix.

The tautomerism of 9-substituted iGs was studied by spectroscopic methods using a set of reference derivatives, corresponding to the various fixed tautomeric forms, i.e. 6-NH₂, N1-H, N3-H and 2-OH. The main conclusion from these studies was that iG exists in solution in the 6-amino form and that the N1-H tautomer (structure I, Figure 4) is predominant in aqueous solution, whereas formation of the 2-OH tautomer (structures II and III, Figure 4) is favoured by a decrease in solvent polarity

[17,29]. Quantum mechanical calculations, particularly calculations accomplished for the dielectric constant of 40, which approximates to the interior of the DNA double helix, revealed that amongst those studied, the stability of two tautomers, 2-oxo, N3-H, 6-amino (structure IV, Figure 4) and N1-H, 2-oxo, N3-H, 6-imino (*trans*) (structure V, Figure 4), falls into the same energy range as the stability of two tautomers whose existence has been proven experimentally [30,31]. In our opinion, the possible participation of these tautomers in the interactions of iG should also be taken into consideration. The participation of the 2-oxo, N3-H, 6-amino tautomeric form (structure IV, Figure 4) in various base pairs involving iG was postulated by Eschenmoser and co-workers in order to explain the formation of complexes of various unnatural polynucleotides studied in their laboratory [32,33]. Isocytosine can exist in various tautomeric forms, however, the experimental and theoretical data are consistent with the fact that 1-substituted iC adopts the 2-amino-4-oxo form (Figure 1A) in environments of intermediate to high polarity [31] and references cited therein). 2'-Deoxy-5-methylisocytidine exists in the 2-amino-4-oxo form in the crystalline state [34].

Quantum mechanical calculations have shown that the iG · iC pair, in which complementary bases are in their main tautomeric forms (Figure 1A), is comparable in stability to the G · C pair [35]. No such calculations have been performed for other tautomeric forms of iG and iC. In fact, data obtained from oligonucleotide duplex thermal denaturation studies showed that the iG · iC pair is even more stable than the G · C pair [26,27,31]. Calculations for the iG · T pair were performed for the 2-OH, 6-amino and 2-oxo, N3-H, 6-amino tautomeric forms of iG. The Watson–Crick type pairing, involving the enol form of iG with the OH group in *cis* arrangement (structure II, Figure 4), was considerably more stable than that formed with the OH group in *trans* arrangement (structure III, Figure 4) and the pairing involving the N3-H form of iG (structure IV, Figure 4) [30]. X-ray crystallography and NMR studies of an oligodeoxynucleotide duplex containing iG · T pairs showed that, depending on conditions, the iG · T pair can adopt Watson–Crick (probably involving the enol form of iG) or wobble conformation [36].

The formation of hydrogen bonds between the lone electron pairs of N3 of purines and O2 of pyrimidines at the minor groove side of the base pair and the amino acid side chains was postulated as the mechanism controlling the selectivity of DNA polymerases. However, recent work on the structure of the ternary complex of *Taq* polymerase has shown that these specific interactions occur in the base pair formed by the already incorporated dNMP and the corresponding template base, but not in the nascent base pair [37]. This means that the possible discrimination of the putative N3-H tautomer by *Taq* polymerase itself cannot take place at the moment of incorporation of the dNTP opposite iG.

We cannot exclude the participation of tautomers of iG, other than structures I and II (Figure 4), in its possible plenitude of interactions with natural and non-natural bases considered by the authors quoted above [30–33]. It is worth emphasizing that relevant experimental data on the structure of iG base-pairing are rather scarce [36], whereas the evaluations of the stability and geometry of base pairs are often just of a qualitative nature [31].

Taking all these considerations together, we assume that in the *Taq* polymerase active centre the N1-H, 2-oxo, 6-amino tautomer of iG forms a base pair with MiC, whereas the 2-OH, 6-amino forms one with T during incorporation of dMiCTP or dTTP respectively, as shown in Figure 1. The proposed structures are compatible with Watson–Crick geometry and structural requirements within the polymerase active centre. The presence of three hydrogen bonds in each pair implicates strong interactions

Figure 4 Tautomeric forms of iG

Structures I and II are referred to as 'keto' and 'enol' forms respectively in the text (see also Figure 1).

between partners. The participation of other tautomeric forms, as well as wobble arrangements, appears to be highly improbable.

Our results show that *Taq* polymerase catalyses incorporation of both dTTP and dMiCTP opposite iG. Since various polymerases accept non-standard base pairs with their own characteristic idiosyncrasies, the acceptance of iG · T and iG · MiC by *Taq* was not obvious before the experiments were performed. For example, T4 DNA polymerase does not incorporate diGTP opposite iC, whereas T7 RNA polymerase incorporates UTP but not iCTP opposite iG [15]. An isosteric analogue of iG, 2-chloroadenine, completely stops *Taq* polymerase when present in the template, which is rather unexpected, since many prokaryotic and eukaryotic polymerases accept the 2-chloroadenine-thymine pair, albeit to various extents [38].

In our experimental system, the efficiency of incorporation of dMiCTP and dTTP can be regarded as a probe for the presence of the keto and enol tautomeric forms of iG, structures I and II respectively (Figure 4). Hence the variability of the parameter *F* can be a measure of the influence of neighbouring bases or temperature on the proportion of the tautomeric forms of iG in the *Taq* polymerase active centre. On the basis of the observed *F* values, the estimated ratio of the keto to the enol form varies from 5:1 to 20:1, depending on the conditions (Table 2). In our former experiments with avian myeloblastosis virus reverse transcriptase, we found this ratio to be approx. 10:1 [8], and the same value was reported for monomeric isoguanosine in water solution [17]. The variability of the *F* parameter primarily reflects the variability of K_m values, since the V_{max} values are roughly similar, especially when these values are compared for the same sequences and temperatures (Table 2). This indicates that the steric constraints during incorporation of dMiCTP and dTTP

are similar, in other words, that the iG · MiC and iG · T pairs fit the polymerase active centre equally well.

In order to evaluate the significance of variance of *F* depending on iG neighbouring bases or on temperature, we employed the one-way analysis of variance (ANOVA) test and Student–Neuman–Keuls post-test. The analysis has shown that *F* does not depend on the temperature. In only one case (out of 16, see Results section) the difference was significant ($P < 0.05$) and it is very likely that this was an accidental result. This would suggest that temperature does not influence the tautomeric equilibrium of iG in the *Taq* active centre. It has been found, on the basis of UV spectra, that the proportion of the enol tautomer increases with temperature in chloroform, whereas no changes in UV spectra were observed in ethanol solution over the temperature range 15–70 °C [17]. Similarly, in our studies, we did not notice any changes in UV spectrum of isoguanosine in water solution over the temperature range 20–80 °C (results not shown). On the other hand, the observation based on the temperature-dependence of NMR spectra of an oligodeoxynucleotide duplex containing iG · T pairs suggests that the population of the enol tautomer increases with temperature over the range 2–40 °C [36]. The reported examples do not give conclusive evidence regarding the type of influence that temperature has on the tautomeric equilibrium of iG, but only demonstrate that this influence is markedly dependent on the nature of the microenvironment. Apparently, temperature does not influence this equilibrium in the *Taq* active centre over the studied range (37–80 °C) and, as a consequence, it does not influence the ratio of efficiency of incorporation of dMiCTP versus dTTP (parameter *F*).

In contrast with temperature, bases neighbouring iG in the template significantly change the relative efficiency of incor-

poration of dMiCTP and dTTP (Figure 3). Cytosine as the 3'-neighbour of iG causes a decrease of incorporation of dMiCTP versus dTTP in comparison with thymine or guanine at all temperatures studied. However, the differences between thymine and guanine are less marked; the ability of the 3'-neighbour of iG to increase the proportion of the enol form decreases in the order $C > G \geq T$. It is interesting to note that, at higher temperatures (65 and 80 °C), the discrimination among sequences is more pronounced than at lower temperatures (37 and 50 °C). This may be related to the fact that the optimum activity of *Taq* polymerase lies within higher ranges of temperature.

Apparently, the 5'-neighbour does not influence the tautomeric equilibrium of iG, which is not unexpected, since the 5'-neighbour of iG is flipped out of the stacking arrangement with the template-primer duplex in the *Taq* polymerase active site at the moment of incorporation of a nucleotide opposite iG [37]. This indicates that the nature of stacking interactions of the base pair preceding iG (i.e. 3'-neighbour of iG paired with 3'-terminal base of primer) with iG is responsible for tautomeric shift of this base. We assume that the thermodynamic contributions to duplex formation of single-base dangling-ends determined in solution in the absence of polymerase may be a good approximation of the state where the primer-template duplex is interacting with polymerase. The thermodynamic parameters of dangling-ends reflect the stacking ability of the dangling base on the duplex closing base pair.

The thermodynamic contributions to duplex formation of all 32 possible single-base dangling-ends on a Watson-Crick pair were determined and reported by SantaLucia and co-workers [39]. Out of all the structures studied, we have chosen for comparison three which are relevant to our work, namely $C \cdot G/X$, $G \cdot C/X$ and $T \cdot A/X$, where X stands for 5'-dangling A, G, C or T. It has been found for each dangling base, irrespective of its nature, that the free-energy contribution of closing base pair to duplex stability follows the order $C \cdot G > G \cdot C > T \cdot A$. Since the order of stacking ability of the discussed base pairs does not depend on the nature of the dangling base, we can assume that their stacking with dangling iG will follow the same order. In conclusion, we postulate that stronger stacking interactions which occur in $C \cdot G/iG$ structure would favour enolization of iG to a higher extent than weaker interactions, which occur in $G \cdot C/iG$ and in $T \cdot A/iG$ structures.

To our knowledge, this is the first experimental evidence to show that the change of pairing of a base can be the result of a tautomeric shift caused by stacking interactions. The tautomeric shift caused by stacking interactions was postulated to explain different miscoding of $N^2,3$ -ethenoguanine present in poly(A) or poly(C) templates [40]; however, the formation of wobble pairs and looping-out of the modified base were not excluded in that study. N^4 -Methoxydeoxycytidine 5'-triphosphate, a nucleotide containing a base which exhibits appreciable tautomerism, can replace either dCTP or dTTP during DNA replication. The ease of formation of pairs between N^4 -methoxycytosine and template G or A exhibits a distinct sequence dependence [41]. This would suggest that stacking between the last base in the primer and N^4 -methoxycytosine influences the tautomeric equilibrium of the latter, but no specific studies were undertaken in this direction.

An *in vitro* kinetic assay revealed an influence of sequence context on base-pairing for 8-oxoguanine located at different template neighbourhoods [42]. However, in that case, it was not caused by a tautomeric shift, since it appears that this oxidized base exists, under physiological conditions, in a single tautomeric N1-H, N7-H, 6,8-dioxo form [43–45]. The formation of the non-mutagenic Watson-Crick-type 8-oxoG · C pair with 8-oxodG in *anti* conformation and the mutagenic Hoogsteen-type 8-oxoG · A

pair with 8-oxodG in *syn* conformation was proposed on the basis of NMR spectra of designed oligodeoxynucleotide complexes [46,47]. Hence, it is very likely that neighbouring bases shift the *syn-anti* equilibrium of 8-oxodG in the template, resulting in a change of efficiency for incorporation of dCTP versus dATP [42].

We showed in the present study that sequence context affects the relative efficiency of polymerase-catalysed formation of $iG \cdot MiC$ and $iG \cdot T$ pairs, which is a result of shifts in the proportion of the keto and enol tautomeric forms of iG. This finding by itself does not give a direct view on the mutagenic potential of iG, since the formation of the $iG \cdot T$ pair does not lead to mutations and MiC is not a natural base. It demonstrates, however, that sequence context exerts an influence on base-pairing in general, which in turn can alter the mutagenic potential of iG. In fact, studies performed by Kasai and co-workers [7,48] gave experimental evidence that incorporation of four natural dNTPs opposite iG is also sequence dependent. In addition to incorporation of dTTP occurring in each case, preferred incorporation of dCTP and dGTP opposite iG was observed along several contexts tested, with the exception of the 3'-AiGT-5' sequence, in which iG preferentially directs dATP incorporation by α and β mammalian DNA polymerases and by the exonuclease-deficient Klenow fragment of *Escherichia coli* DNA polymerase I. This group also compared thermodynamic stability of duplexes containing the 3'-AiGT-5' and 3'-GiGC-5' sequences. The authors found a good correlation between the stability of duplexes with the $iG \cdot N$ pair at the end of the duplex, which mimics the nucleotide incorporation step, and the efficiency of incorporation of a given dNTP opposite iG in each of the two sequences studied [49].

The structures of the pairs formed by iG with A, C and G during DNA replication remain unknown in general. The participation of various tautomeric forms of iG as well as the *syn* or *anti* conformers of each deoxynucleoside residues composing a pair can not be excluded. Since the tautomeric and the *syn-anti* equilibrium can be altered by stacking interactions, a detailed analysis of the influence of sequence context on the miscoding properties of iG is not a simple matter.

This work was supported by the State Committee for Scientific Research (KBN grant No. 6 P203 01906) and the Centre of Excellence in Molecular Biotechnology, Institute of Biochemistry and Biophysics, Polish Academy of Sciences (grant WP10). We thank Professor Zygmunt Kazimierzczuk (Department of Chemistry, Warsaw Agricultural University, Poland) for the sample of 2'-deoxyisoguanosine. A.M.M. is funded by the Institute of Biochemistry and Biophysics, Polish Academy of Sciences.

REFERENCES

- Cherbuliez, E. and Bernhard, K. (1932) Recherches sur la graine de croton. I. Sur le crotonoside (2-oxy-6-amino-purine-d-riboside). *Helv. Chim. Acta* **15**, 464–471
- Bergmann, W. and Burke, D. C. (1956) Contributions to the study of marine products. XL. The nucleosides of sponges. IV. Spongiosine. *J. Org. Chem.* **21**, 226–228
- Fuhrman, F. A., Fuhrman, G. J., Nachman, R. J. and Mosher, H. S. (1981) Isoguanosine: Isolation from an animal. *Science* **212**, 557–558
- Kim, Y. H., Nachman, R. J., Pavelka, L., Mosher, H. S., Fuhrman, F. A. and Fuhrman, G. J. (1981) Doridosine, 1-methylisoguanosine, from *Anisodorus nobilis*; structure, pharmacological properties and synthesis. *J. Nat. Prod.* **44**, 206–214
- Dizdaroglu, M. (1992) Oxidative damage to DNA in mammalian chromatin. *Mutat. Res.* **275**, 331–342
- Kamiya, H. and Kasai, H. (1995) Formation of 2-hydroxydeoxyadenosine triphosphate, an oxidatively damaged nucleotide and its incorporation by DNA polymerases. Steady-state kinetics of the incorporation. *J. Biol. Chem.* **270**, 19446–19450
- Kamiya, H., Ueda, T., Ohgi, T., Matsukage, A. and Kasai, H. (1995) Misincorporation of dAMP opposite 2-hydroxyadenine, an oxidative form of adenine. *Nucleic Acids Res.* **23**, 761–766
- Bukowska, A. M. and Kuśmerek, J. T. (1996) Miscoding properties of isoguanine (2-oxoadenine) studied in an AMV reverse transcriptase *in vitro* system. *Acta Biochim. Polon.* **43**, 247–254

- 9 Kamiya, H. and Kasai, H. (1997) Mutations induced by 2-hydroxyadenine on a shuttle vector during leading and lagging strand syntheses in mammalian cells. *Biochemistry* **36**, 11125–11130
- 10 Inoue, M., Kamiya, H., Fujikawa, K., Ootsuyama, Y., Murata-Kamiya, N., Osaki, T., Yasumoto, K. and Kasai, H. (1998) Induction of chromosomal gene mutations in *Escherichia coli* by direct incorporation of oxidatively damaged nucleotides. New evaluation method for mutagenesis by damaged DNA precursors *in vivo*. *J. Biol. Chem.* **273**, 11069–11074
- 11 Fujikawa, K., Kamiya, H., Yakushiji, H., Fujii, Y., Nakabeppu, Y. and Kasai, H. (1999) The oxidized forms of dATP are substrates for the human MutT homologue, the hMTH1 protein. *J. Biol. Chem.* **274**, 18201–18205
- 12 Ohtsubo, T., Nishioka, K., Imaiso, Y., Iwai, S., Shimokawa, H., Oda, H., Fujiwara, T. and Nakabeppu, Y. (2000) Identification of human MutY homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. *Nucleic Acids Res.* **28**, 1355–1364
- 13 Switzer, C., Moroney, S. E. and Benner, S. A. (1989) Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **111**, 8322–8323
- 14 Piccirilli, J. A., Krauch, T., Moroney, S. E. and Benner, S. A. (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature (London)* **343**, 33–37
- 15 Switzer, C. Y., Moroney, S. E. and Benner, S. A. (1993) Enzymatic recognition of the base pair between isocytidine and isoguanosine. *Biochemistry* **32**, 10489–10496
- 16 Lutz, M. J., Horlacher, J. and Benner, S. A. (1998) Recognition of 2'-deoxysoguanosine triphosphate by HIV-1 reverse transcriptase and mammalian cellular DNA polymerases. *Bioorg. Med. Chem. Lett.* **8**, 499–504
- 17 Sepioł, J., Kazimierczuk, Z. and Shugar, D. (1976) Tautomerism of isoguanine and solvent-induced keto-enol equilibrium. *Z. Naturforsch.* **31C**, 361–370
- 18 Kazimierczuk, Z., Mertens, R., Kawczyński, W. and Seela, F. (1991) 2'-Deoxysoguanosine and base-modified analogues: chemical and photochemical synthesis. *Helv. Chim. Acta* **74**, 1742–1748
- 19 Seela, F., Chen, Y., Melenewski, A., Rosemeyer, H. and Wei, C. (1996) Synthesis and application of novel nucleoside phosphonates and phosphoramidites modified at the base moiety. *Acta Biochim. Polon.* **43**, 45–52
- 20 Seela, F. and Wei, C. (1997) Oligonucleotides containing consecutive 2'-deoxysoguanosine residues: synthesis, duplexes with parallel chain orientation and aggregation. *Helv. Chim. Acta* **80**, 73–86
- 21 Chenna, A. and Singer, B. (1995) Large scale synthesis of *p*-benzoquinone-2'-deoxycytidine and *p*-benzoquinone-2'-deoxyadenosine adducts and their site-specific incorporation into DNA oligonucleotides. *Chem. Res. Toxicol.* **8**, 865–874
- 22 Boosalis, M. S., Petruska, J. and Goodman, M. F. (1987) DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. *J. Biol. Chem.* **262**, 14689–14696
- 23 Goodman, M. F., Creighton, S., Bloom, L. B. and Petruska, J. (1993) Biochemical basis of DNA replication fidelity. *Crit. Rev. Biochem. Mol. Biol.* **28**, 83–126
- 24 Cornish-Bowden, A. and Eisenthal, R. (1978) Estimation of Michaelis constant and maximum velocity from the direct linear plot. *Biochim. Biophys. Acta* **523**, 268–272
- 25 Kamiński, Z. W. and Domino, E. F. (1987) Computer program for calculation of kinetic and pharmacologic parameters using a 'direct linear plot' derived algorithm. *Comput. Meth. Prog. Biomed.* **24**, 41–45
- 26 Horn, T., Chang, C.-A. and Collins, M. L. (1995) Hybridization properties of the 5-methyl-isocytidine/isoguanosine base pair in synthetic oligodeoxynucleotides. *Tetrahedron Lett.* **36**, 2033–2036
- 27 Seela, F. and Wei, C. (1999) The base-pairing properties of 7-deaza-2'-deoxysoguanosine and 2'-deoxysoguanosine in oligonucleotide duplexes with parallel and antiparallel chain orientation. *Helv. Chim. Acta* **82**, 726–745
- 28 Kunkel, T. A. and Bebenek, K. (2000) DNA replication fidelity. *Annu. Rev. Biochem.* **69**, 97–529
- 29 Seela, F., Wei, C. and Kazimierczuk, Z. (1995) Substituent reactivity and tautomerism of isoguanosine and related nucleosides. *Helv. Chim. Acta* **78**, 1843–1854
- 30 Cysewski, P., Madjar, C. V., Noinville, V. and Oliński, R. (1995) Structure and properties of 2-hydroxy-adenine and 2-hydroxy-2'-deoxy-adenosine. *Bull. Soc. Chim. Fr.* **132**, 453–460
- 31 Roberts, C., Bandaru, R. and Switzer, C. (1997) Theoretical and experimental study of isoguanine and isocytosine. Base pairing and expanded genetic system. *J. Am. Chem. Soc.* **119**, 4640–4649
- 32 Krishnanrathy, R., Pitsch, S., Minton, M., Miculka, C., Windhab, N. and Eschenmoscher, A. (1996) Pyranosyl-RNA: Base pairing between homochiral oligonucleotide strand of opposite sense of chirality. *Angew. Chem. Int. Ed. Engl.* **35**, 1537–1541
- 33 Groebke, K., Hunziker, J., Fraser, W., Peng, L., Diederichsen, U., Zimmermann, K., Holzner, A., Leumann, C. and Eschenmoscher, A. (1998) Warum Pentose- und nicht Hexose-Nucleinsäuren? Teil V. (Purin-Purin)-Basenpaarung in der homo-DNS-Reihe: Guanin, Isoguanin, 2,6-Diaminopurin und Xantin. *Helv. Chim. Acta* **81**, 375–473
- 34 Seela, F., He, Y., Reuter, H. and Heithoff, E. M. (2000) 2'-Deoxy-5-methylisocytidine. *Acta Crystallogr. C* **56**, 989–991
- 35 Leach, A. R. and Kollman, P. A. (1992) Theoretical investigations of novel nucleic acid bases. *J. Am. Chem. Soc.* **114**, 3675–3683
- 36 Robinson, H., Gao, Y. G., Bauer, C., Roberts, C., Switzer, C. and Wang, A. (1998) 2'-Deoxysoguanosine adopts more than one tautomer to form base pairs with thymidine observed by high-resolution crystal structure analysis. *Biochemistry* **37**, 10897–10905
- 37 Li, Y. and Waksman, G. (2001) Crystal structures of a ddATP-, ddTTP-, ddCTP- and ddGTP-trapped ternary complex of KlenTaq1: insights into nucleotide incorporation and selectivity. *Protein Sci.* **10**, 1225–1233
- 38 Yuh, S. H., Tibudan, M. and Hentosh, P. (1998) Analysis of 2-chloro-2'-deoxyadenosine incorporation into cellular DNA by quantitative polymerase chain reaction. *Mol. Pharmacol.* **45**, 955–961
- 39 Bommarito, S., Peyret, N. and SantaLucia, Jr, J. (2000) Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.* **28**, 1929–1934
- 40 Mroczkowska, M. M. and Kuśmierek, J. T. (1993) The effect of neighboring bases on miscoding properties of N²,3-ethenoguanine. *Z. Naturforsch.* **48C**, 63–67
- 41 Reeves, S. T. and Beattie, K. L. (1985) Base-pairing properties of N⁴-methoxydeoxycytidine 5'-triphosphate during DNA synthesis on natural templates, catalyzed by DNA polymerase I of *Escherichia coli*. *Biochemistry* **24**, 2262–2268
- 42 Hatahet, Z., Zhou, M., Reha-Krantz, L. J., Morrical, S. W. and Wallace, S. S. (1998) In search of a mutational hotspot. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8556–8561
- 43 Aida, M. and Nishimura, S. (1987) An *ab initio* molecular orbital study on the characteristics of 8-hydroxyguanine. *Mutat. Res.* **192**, 83–89
- 44 Culp, S. J., Cho, B. P., Kadlubar, F. F. and Evans, F. E. (1989) Structural and conformational analyses of 8-hydroxy-2'-deoxyguanosine. *Chem. Res. Toxicol.* **2**, 416–422
- 45 Cho, B. P., Kadlubar, F. F., Culp, S. J. and Evans, F. E. (1990) ¹⁵N nuclear magnetic resonance studies on the tautomerism of 8-hydroxy-2'-deoxyguanosine, 8-hydroxyguanosine, and other C8-substituted guanine nucleosides. *Chem. Res. Toxicol.* **3**, 445–452
- 46 Oda, Y., Uesugi, S., Ikehara, M., Kawase, Y. and Ohtsuka, E. (1991) NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Res.* **19**, 1407–1412
- 47 Kouchakdjian, M., Bodepudi, V., Shibutani, S., Eisenberg, M., Johnson, F., Grollman, A. P. and Patel, D. J. (1991) NMR structural studies of the ionizing radiation adduct 7-hydro-8-oxodeoxyguanosine (8-oxo-7H-dG) opposite deoxyadenosine in a DNA duplex. 8-Oxo-7H-dG(*syn*) . dA(*anti*) alignment at lesion site. *Biochemistry* **30**, 1403–1412
- 48 Kamiya, H. and Kasai, H. (1996) Effect of sequence contexts on misincorporation of nucleotides opposite 2-hydroxyadenine. *FEBS Lett.* **391**, 113–116
- 49 Kawakami, J., Kamiya, H., Yasuda, K., Fujiki, H., Kasai, H. and Sugimoto, N. (2001) Thermodynamic stability of base pairs between 2-hydroxyadenine and incoming nucleotides as a determinant of nucleotide incorporation specificity during replication. *Nucleic Acids Res.* **15**, 3289–3296

Received 14 June 2002/18 September 2002; accepted 21 October 2002

Published as BJ Immediate Publication 21 October 2002, DOI 10.1042/BJ20020922