Nucleic Acids Research, 2001, Vol. 29, No. 16 **3289–3296**

Thermodynamic stability of base pairs between 2-hydroxyadenine and incoming nucleotides as a determinant of nucleotide incorporation specificity during replication

Junji Kawakami[1,2], Hiroyuki Kamiya[3], Kyohko Yasuda[2], Hiroyoshi Fujiki[1], Hiroshi Kasai[3] and Naoki Sugimoto[1,2,*]

[1]Department of Chemistry, Faculty of Science and Engineering, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan, [2]High Technology Research Center, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan and [3]Department of Environmental Oncology, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, Kitakyushu 807-8555, Japan

Received as resubmission June 26, 2001; Accepted July 5, 2001

ABSTRACT

We investigated the thermodynamic stability of double-stranded DNAs with an oxidative DNA lesion, 2-hydroxyadenine (2-OH-Ade), in two different sequence contexts (5′-GA*C-3′ and 5′-TA*A-3′, A* represents 2-OH-Ade). When an A*–N pair (N, any nucleotide base) was located in the center of a duplex, the thermodynamic stabilities of the duplexes were similar for all the natural bases except A (N = T, C and G). On the other hand, for the duplexes with the A*–N pair at the end, which mimic the nucleotide incorporation step, the stabilities of the duplexes were dependent on their sequence. The order of stability is T > G > C >> A in the 5′-GA*C-3′ sequences and T > A > C > G in the 5′-TA*A-3′ sequences. Because T/G/C and T/A are nucleotides incorporated opposite to 2-OH-Ade in the 5′-GA*C-3′ and 5′-TA*A-3′ sequences, respectively, these results agree with the tendency of mutagenic misincorporation of the nucleotides opposite to 2-OH-Ade *in vitro*. Thus, the thermodynamic stability of the A*–N base pair may be an important factor for the mutation spectra of 2-OH-Ade.

INTRODUCTION

Many oxidative lesions are produced when reactive oxygen species attack the bases or sugars of DNA (1). One of the lesions, 8-hydroxyguanine (7,8-dihydro-8-oxoguanine; 8-OH-Gua), a guanine in which the C8 position is oxidized, is known as a highly mutagenic base (2). Many studies have revealed that G→T transversion would occur when 8-OH-Gua is generated in cells (3–7). Another important and well-known lesion is 2-hydroxyadenine (1,2-dihydro-2-oxoadenine or isoguanine; 2-OH-Ade). 2-OH-Ade, wherein the C2 position is oxidized, was found to be produced by oxygen radicals *in vitro* (8,9). 2-OH-Ade is formed in DNA by various treatments of animals and living cells and in certain human cancerous tissues (10). This oxidized adenine is as mutagenic as 8-OH-Gua in *Escherichia coli* and mammalian cells (11,12). Recently, we found that the human MutT homolog (hMTH1) 8-hydroxy-dGTPase very efficiently hydrolyzes 2-hydroxy-dATP (13). In addition, the human MutY homolog (hMYH), a repair enzyme for 8-OH-Gua, was reported to be a repair enzyme for 2-OH-Ade (14). These results strongly suggest the importance of 2-OH-Ade as an endogenous mutagen in cells.

2-OH-Ade and 8-OH-Gua are expected to cause the misincorporation of nucleotides during the replication of DNA by non-canonical base pairing. 8-OH-Gua in DNA induces G→T transversions by the formation of a mispair between 8-OH-Gua and adenine (15–17). On the contrary, 2-OH-Ade elicits various types of mutations in sequence and host cell-dependent manners, suggesting its ability to pair with all four natural bases (11,12). *In vitro*, 2-OH-Ade is incorporated as a complementary base to T/C by a mammalian DNA polymerase α and to T/G by the exonuclease-deficient Klenow fragment of *E.coli* DNA polymerase I (KF^{exo-}), respectively (8,18). In the opposite direction, as expected, dTMP/dCMP and dTMP/dGMP are incorporated opposite 2-OH-Ade by mammalian DNA polymerases (α and β) and KF^{exo-}, respectively (19). However, unexpectedly, dAMP along with dTMP are incorporated opposite to 2-OH-Ade in the 5′-TA*A-3′ (A* denotes 2-OH-Ade) sequences by all three DNA polymerases, α, β and KF^{exo-} (19,20).

The fact that nucleotide incorporation depended on the sequence context of the template, and not on the simple pairing of facing bases, strongly suggests the importance of the local DNA structure at the template–primer terminus for the polymerase reaction. The local structure of DNA and its determinants, the

*To whom correspondence should be addressed at: Department of Chemistry, Faculty of Science and Engineering, Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501, Japan. Tel: +81 78 435 2497; Fax: +81 78 435 2539; Email: sugimoto@konan-u.ac.jp
Present address:
Hiroyuki Kamiya, Graduate School of Pharmaceutical Sciences, Hokkaido University, Sapporo 060-0812, Japan

stacking of incoming nucleotide bases with the adjacent base, hydrogen bonding to the opposite nucleotide and so on, directly correlate with the thermodynamic stability of the base pairing (21). Furthermore, the stability of the base pair at the 3′ end of the primer strand is important for further elongation of the strand (22). Consequently, the thermodynamic stability of the base pair at the termini is expected to correlate with the efficiency of nucleotide incorporation. The mechanism of nucleotide incorporation by several DNA polymerases is under discussion and studies by Kool and co-workers (22–25) have revealed the importance of the base pairing geometry for a polymerase reaction. However, the relationship between the thermodynamic stabilities of the base pairs and their recognition by DNA polymerases remains unclear.

Almost all the thermodynamic studies have been carried out with a duplex containing the base pair of interest in its center, and no obvious relationship between the tendency of nucleotide incorporation and base pairing stability has yet been found. Because the situation of a base pair in the center of DNA may be quite different from the environment of the base pair at the end of the strand as previously suggested (26), the thermodynamic stabilities of duplexes that contain the base pair of interest at the end, mimicking the nucleotide incorporation step (27), should be used for evaluation of the relationship between the thermodynamic stability and misincorporation.

Here, we are focusing on the correlation between the stability of a base pair involving 2-OH-Ade (A*–N) and the incorporation pattern opposite 2-OH-Ade obtained by our *in vitro* experiments (19,20). In this study, we prepared a total of 20 DNA duplexes that could be categorized into four groups, with the 5′-TA*A-3′ or 5′-GA*C-3′ sequence in 'template' strands (Fig. 1). The thermodynamic parameters of the 5′-TA*A-3′ sequence that exhibits misincorporation of adenine (AT closing sequence) and those of the control 5′-GA*C-3′ sequence (GC closing sequence) were compared. As shown in Figure 1, 2-OH-Ade was introduced into the center or terminus of the oligodeoxyribonucleotides with the AT, GC closing sequences. An A–T base pair was used as the control for stability evaluation of the A*–N (N = A, G, C or T) base pair. The free energy change at 37°C for duplex formation (ΔG°_{37}) was calculated in order to compare the thermodynamic parameters and the results of the misincorporation experiments using DNA polymerases.

The stabilities of the duplexes containing 2-OH-Ade were similar when they paired with T, C and G but not with A when the base pair was located in the center. Similar results were obtained with a series [4] that contains the 5′-GA*C-3′ sequence at the end. However, the A*–A pair was more stable than the pair with C or G in the series [3] which has the 5′-TA*A-3′ sequence at the end. These results were in agreement with those of the nucleotide misincorporation by DNA polymerases *in vitro*, suggesting a relationship between the thermodynamic stability of the terminal base pair and the mutation spectra of 2-OH-Ade.

MATERIALS AND METHODS

Synthesis of oligodeoxyribonucleotides

All oligodeoxyribonucleotides were synthesized on a solid support using phosphoramidite chemistry and purified as previously described (20). β-cyanoethyl phosphoramidite of 2-OH-Ade, 9-[2′-deoxy-5′-*O* (4,4′-dimethoxytrityl)-β-D-*erythro*-pentofuranosyl]-6-[(dimethylamino)methylidene]-9*H*-isoguanine 3′-[(2-cyanoethyl)*N*,*N*-diisopropylphosphoramidite], was prepared as described (28). The synthesized oligonucleotides are listed in Figure 1 and their sequences are the same as those of the primed template used in *in vitro* nucleotide incorporation studies (19,20). A single strand concentration of purified oligodeoxynucleotides was determined by measuring the 260 nm absorbance at high temperature with a Hitachi U-3210 spectrophotometer connected to a Hitachi SPR-10 thermoprogrammer as previously described (29). Extinction coefficients were calculated from the mononucleotide and dinucleotide data using a nearest-neighbor approximation (30,31).

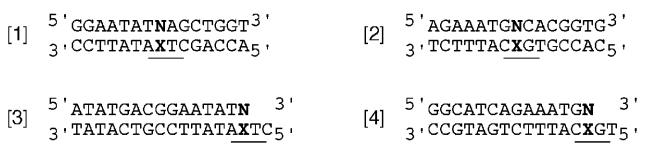

Figure 1. Oligodeoxyribonucleotides with a 5′-TA*A-3′ or 5′-GA*C-3′ sequence (underlined) used in this study. X denotes A or 2-OH-Ade; N represents A, G, C or T. A particular duplex is indicated as sequence series name followed by X–N base pair as subscript, e.g. $[1]_{A*T}$ indicates 5′-GGAATATTAGCTGGT-3′/5′-ACCAGCTA*ATATTCC-3′.

UV melting analysis

Measurements of the UV absorption were carried out on Hitachi U-3200, U-3210 and Beckman DU640 spectrophotometers. All experiments were carried out in a buffer containing 1 M NaCl, 10 mM Na_2HPO_4, 1 mM Na_2EDTA pH 7.0. A DNA strand and its complementary strand were mixed in a 1:1 concentration ratio and annealed to obtain the duplex. Melting curves (absorbance versus temperature curves) were measured at 260 nm using the spectrophotometers connected to a Hitachi SPR-10 thermoprogrammer at several concentration points for a duplex between 1 and 100 μM. The heating rate was 0.5 or 1.0°C/min. The water condensation on the cuvette exterior in the low temperature range was avoided by flushing with a constant stream of dry N_2 gas. The melting temperatures (T_m) were measured at various total concentration (C_t) points of oligonucleotides. The thermodynamic parameters (enthalpy change, ΔH°; entropy change, ΔS°) for each duplex formation were determined as average values obtained by the following two methods: (i) curve fitting of individual melting curves to two-state model with sloping baselines was averaged, and (ii) linear van't Hoff plot, T_m^{-1} versus $\ln(C_t/4)$ as shown in the supplemental Figure S1 (32–34). A free energy change associated with the duplex formation at 37°C, ΔG°_{37}, was calculated from the obtained ΔH° and ΔS° values using following the equation: $\Delta G^\circ_{37} = \Delta H^\circ - (310.15 \times \Delta S^\circ)$.

Molecular modeling

Molecular models of the duplex were built by molecular mechanic calculations performed on a QUANTA 97.1/CHARMm23.2 (Molecular Simulations Inc.). Parameters of the residue topology file (RTF) were modified based on that of the B-form DNA. The energy minimization of a structure was carried out with the adopted basis set Newton–Raphson

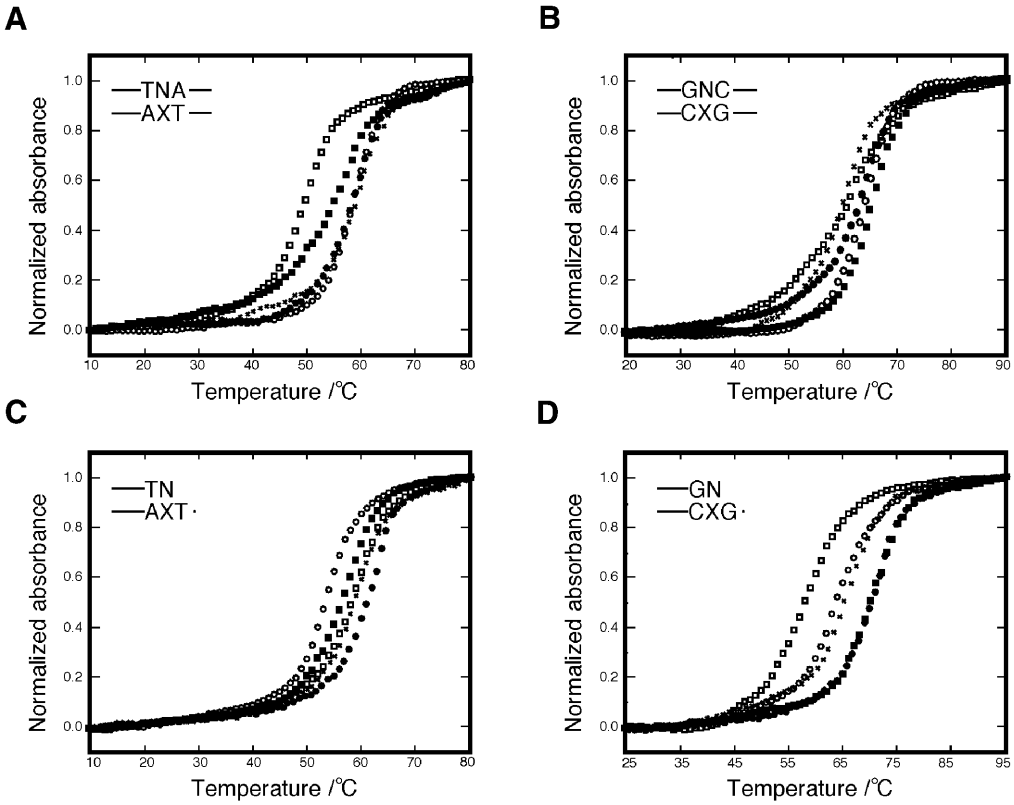

Figure 2. Normalized UV melting curves. (**A**–**D**) Results for duplexes [1], [2], [3] and [4] in Figure 1, respectively. Melting behavior of control duplexes having A–T base pair as X–N are shown as curves drawn by crosses. Melting curves of duplexes with A*–A, A*–G, A*–C and A*–T are drawn by open squares, filled squares, open circles and filled circles, respectively. Each oligomer concentration was 5 μM, and measurements were done in 1 M NaCl-phosphate buffer pH 7.0.

(ABNR) method. All the calculations were performed on a Silicon Graphics Indigo2 workstation running IRIX 5.3.

RESULTS

Determination of thermodynamic parameters from UV melting curves

All of the UV melting curves of the oligodeoxynucleotides shown in Figure 1 were measured in the range from 0 to 90°C. As shown in Figure 2, almost all the oligonucleotides with the A*–N base pair exhibit stability that is similar to those with an A–T base pair. Two isoforms, namely the keto and enol forms of 2-OH-Ade, were previously found in aqueous solution (30,35). Because the extent of each tautomer changes along with the temperature, the two (or more) tautomers of 2-OH-Ade could not be distinguished in the UV melting analysis. However, all the melting curves of the duplexes having 2-OH-Ade showed normal melting behaviors with linear lower and upper baselines as shown in Figure 2. All of these behaviors were regarded as the two-state transition. This phenomenon indicates that the tautomerization would be normalized during melting to produce a linear slope and be negligible in the discussion about the energetical differences between double-stranded and single-stranded DNAs. The fact that the parameters obtained from two independent methods were quite similar also justifies the parameters. The thermodynamic parameters for each duplex formation are listed in Tables 1 and 2.

Stability of the duplexes with internal 2-OH-Ade

As listed in Table 1, the $\Delta H°$ value of duplexes $[1]_{AT}$ and $[2]_{AT}$ were –123 and –125 kcal mol^{-1}, respectively. All the duplexes with an internal 2-OH-Ade residue had a smaller enthalpic gain, suggesting an unfavorable enthalpic effect for the duplex formation by the 2-OH-Ade introduction. On the other hand, the $\Delta S°$ value of duplexes $[1]_{AT}$ and $[2]_{AT}$ were the same value, –344 cal mol^{-1} K^{-1}, and those of the duplexes with the 2-OH-Ade had smaller entropic losses. The order of $\Delta H°$ for the base pairs was the same as that of the free energy change ($\Delta G°_{37}$). Thus, these data reveal that the introduction of the 2-OH-Ade destabilizes the duplex due to the unfavorable enthalpic effect.

For the AT closing [1] series, the order of $-\Delta G°_{37}$ was A–T > A*–C > A*–G ~ A*–T >> A*–A. The differences in the free energy change ($\Delta\Delta G°_{37}$) of A*–C, A*–G, A*–T and A*–A referenced to the stability of the A–T base pairs were 0.7, 2.1, 2.1 and 4.5 kcal mol^{-1}, respectively. The last value might suggest no hydrogen bonding at the A*–A site. Regarding the GC closing [2] series, the order of $-\Delta G°_{37}$ was A–T > A*–T > A*–G ~ A*–C >> A*–A and the $\Delta\Delta G°_{37}$ of these were 0.2, 2.2, 2.2 and 7.9 kcal mol^{-1}, respectively. Also, in this case, the A*–A base pairs should not be formed as in series [1]. The A*–A pair was the weakest base pair in all cases, in spite of the closing

Table 1. Thermodynamic parameters for DNA sequences containing a 2-OH-Ade[a]

	sequence	X[b]	N	symbol[c]	$-\Delta H^\circ$ /kcal mol^{-1}	$-\Delta S^\circ$ /cal mol^{-1} K^{-1}	$-\Delta G^\circ_{37}$ /kcal mol^{-1}	$\Delta\Delta G^\circ_{37}$[d] /kcal mol^{-1}	T_m[e] /°C
[1]	5'·····TNA·····3'	A	T	×	123 ± 3	344 ± 8	16.3 ± 0.5	—	63.8
	3'·····AXT·····5'	A*	A	□	96.5± 2.9	273 ± 9	11.8 ± 0.5	4.5	54.9
		A*	G	■	102 ± 4	283 ± 14	14.2 ± 0.8	2.1	62.2
		A*	C	○	114 ± 5	317 ± 15	15.6 ± 0.9	0.7	63.8
		A*	T	●	109 ± 4	306 ± 11	14.2 ± 0.6	2.1	60.5
[2]	5'·····GNC·····3'	A	T	×	125 ± 2	344 ± 6	18.4 ± 0.3	—	69.5
	3'·····CXG·····5'	A*	A	□	54.2± 4.2	141 ± 11	10.5 ± 0.8	7.9	61.5
		A*	G	■	99.4± 5.2	268 ± 14	16.2 ± 0.9	2.2	70.4
		A*	C	○	106 ± 4	290 ± 12	16.2 ± 0.8	2.2	68.1
		A*	T	●	116 ± 4.6	315 ± 13	18.2 ± 0.8	0.2	71.7

[a]All experiments were carried out in a buffer (pH 7.0) containing 1 M NaCl, 10 mM Na_2PO_4, 1 mM Na_2EDTA.
[b]A* denotes 2-OH-Ade.
[c]Symbols used in Figures 2, 3 and S1.
[d]Increments are defined as the following example: $\Delta\Delta G^\circ_{37}[1]_{A^*N} = \Delta G^\circ_{37}[1]_{A^*N} - \Delta G^\circ_{37}[1]_{AN}$.
[e]Melting temperatures were calculated at the total strand concentration of 100 μM.

Table 2. Thermodynamic parameters for DNA sequences containing a 2-OH-Ade[a]

	sequence	X[b]	N	symbol[c]	$-\Delta H^\circ$ /kcal mol^{-1}	$-\Delta S^\circ$ /cal mol^{-1} K^{-1}	$-\Delta G^\circ_{37}$ /kcal mol^{-1}	$\Delta\Delta G^\circ_{37}$[d] /kcal mol^{-1}	T_m[e] /°C
[3]	5'·····TN 3'	A	T	×	112 ± 3	311 ± 9	15.5 ± 0.5	—	64.0
	3'·····AXT· 5'	A*	A	□	106 ± 3	293 ± 9	15.0 ± 0.5	0.5	63.9
		A*	G	■	63.7± 4.3	167 ± 11	12.0 ± 0.8	3.5	66.1
		A*	C	○	96.4± 5.0	266 ± 14	13.8 ± 0.8	1.7	62.3
		A*	T	●	116 ± 3	324 ± 9	15.5 ± 0.5	0.0	63.0
[4]	5'·····GN 3'	A	T	×	142 ± 6	398 ± 17	18.7 ± 0.9	—	66.1
	3'·····CXG· 5'	A*	A	□	65.4± 1.6	171 ± 4	12.3 ± 0.3	6.4	67.0
		A*	G	■	121 ± 4	335 ± 13	17.2 ± 0.8	1.5	67.0
		A*	C	○	105 ± 4	288 ± 10	15.8 ± 0.6	2.9	67.0
		A*	T	●	127 ± 4	351 ± 11	18.0 ± 0.7	0.7	67.8

[a]All experiments were carried out in a buffer (pH 7.0) containing 1 M NaCl, 10 mM Na_2PO_4, 1 mM Na_2EDTA.
[b]A* denotes 2-OH-Ade.
[c]Symbols used in Figures 2, 3 and S1.
[d]Increments are defined as the following example: $\Delta\Delta G^\circ_{37}[1]_{A^*N} = \Delta G^\circ_{37}[1]_{A^*N} - \Delta G^\circ_{37}[1]_{AN}$.
[e]Melting temperatures were calculated at the total strand concentration of 100 μM.

base pair, as previously reported (36). These results are summarized in Figure 3A and B.

Stability of the duplexes with a terminal 2-OH-Ade pair at the dangling end

To mimic the nucleotide incorporation step during DNA replication, the thermodynamic parameters of the duplexes with 2-OH-Ade in the terminal 'pair' (corresponding to the opposite site of the primer end) were measured. In addition, we added a dangling end in order to consider the stacking interaction with a 5′-flanking base of 2-OH-Ade as previously reported (27). As listed in Table 2, the ΔH° value of duplexes $[3]_{AT}$ and $[4]_{AT}$ were –112 and –142 kcal mol^{-1}, respectively. As for the fully-paired duplexes, almost all the duplexes with the 2-OH-Ade lost the enthalpic gain and had unfavorable free energy changes with the exception, A*–T with AT closing ($[3]_{A^*T}$). The ΔH° and ΔG°_{37} of $[3]_{A^*T}$ were –116 and –15.5 kcal mol^{-1} while for $[3]_{AT}$, they were –112 and –15.5 kcal mol^{-1}, respectively. This A*–T pair was as stable as the A–T pair and more stable than the other pairs involving 2-OH-Ade in series [3]. The A*–T pair

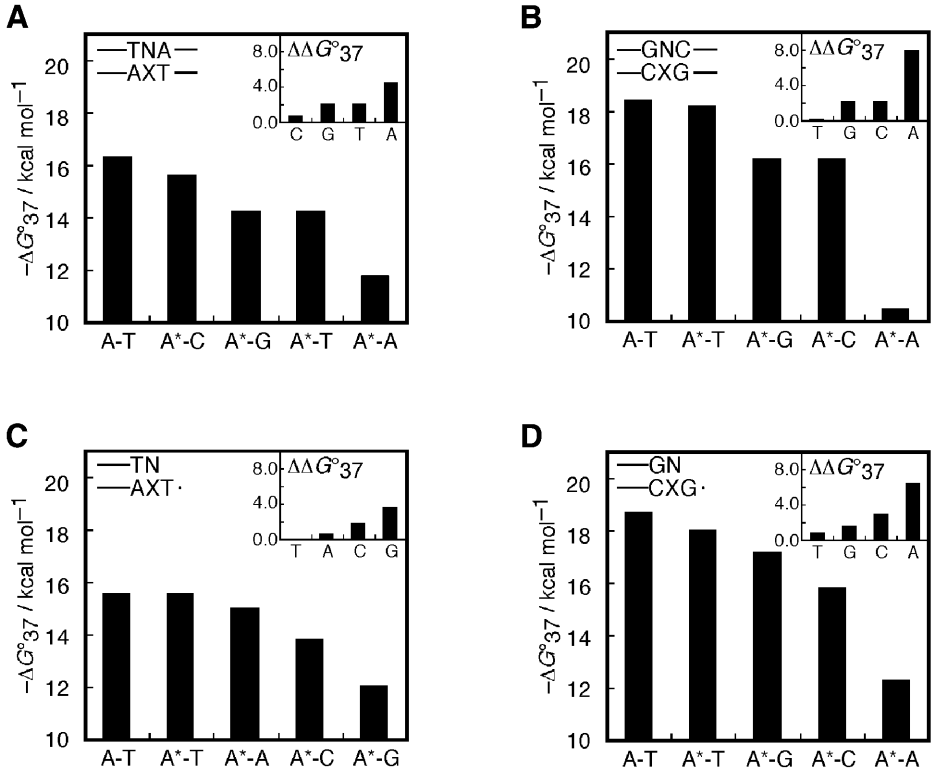

Figure 3. Summary of stability as ΔG°_{37} for all 20 model duplexes. (**A**–**D**) Results for duplexes [1], [2], [3] and [4], respectively. The data are aligned by the order of $-\Delta G^{\circ}_{37}$ value. Inset shows $\Delta\Delta G^{\circ}_{37}$ values against the ΔG°_{37} of control duplex with A–T base pair for evaluation of the effect of trinucleotides with 2-OH-Ade.

was also the most stable A*–N pair for sequence [4]. Another notable finding is the stability of AT closed $[3]_{A*A}$. The free energy change of the $[3]_{A*A}$ duplex formation was nearly the same as that of the original $[3]_{AT}$ ($\Delta\Delta G^{\circ}_{37}$ = 0.5 kcal mol^{-1}) although the A*–A base pair in the sequence series [4] was quite unstable as seen in the sequences [1] and [2]. These results are summarized in Figure 3C and D.

DISCUSSION

Natural nucleic acids recognize their complementary sequences specifically by Watson–Crick base pairing and their structures have been defined. However, most of the local structures of the double-stranded DNAs with a modified base, such as 2-OH-Ade, are still unclear.

The existence of many tautomers of 2-OH-Ade has been reported. The presence of the N1-H keto and O2-H enol forms was found in the 9-methyl derivative and in the deoxyriboside (30,35). Also, it was reported that the equilibrium between the two tautomers shifts depending on the hydrophobicity of the environment. In a duplex DNA, the N1-H keto form is predominant in aqueous solution (~95% at 0°C) and the population of another form (probably the O2-H enol form) gradually increased, reaching 40% at 40°C (37). In addition, the N3-H keto, imino oxo and imino enol with the N1-H or N3-H isoforms were also suggested based on a theoretical calculation (38). This conformational diversity of 2-OH-Ade appears to result in many structures: oligodeoxyribonucleotides containing 2-OH-Ade are able to form normal antiparallel duplexes and also parallel duplexes, quadruplexes and pentaplexes (39–43). The base pairing pattern concerned with 2-OH-Ade is expected to be vary because 2-OH-Ade could make hydrogen bonds with the all four natural bases through the many tautomeric forms. This ambiguous base recognition resembles the 'wobble' base pairing in the codon–anticodon recognition system, and actually 2-OH-Ade was applied for use in the construction of an expanded genetic code (38,44).

Thus, once 2-OH-Ade is produced by the oxidation of adenine, many types of mutations would be induced during the DNA replication. Indeed, various mispairing patterns of 2-OH-Ade have been found. It has also been clarified that the mispairing property of 2-OH-Ade may depend on the sequence context around the lesion (8,18–20,45), suggesting the relationship between the thermodynamic stability of the terminal base pair involving 2-OH-Ade and this sequence context-dependent nucleotide misincorporation.

Thermodynamic effect of 2-OH-Ade in the center of the duplexes

Only a few thermodynamic studies of a double-stranded oligonucleotide with the A*–N base pair have been carried out, including studies with RNA–DNA heteroduplexes and RNA–RNA duplexes (38,46). The results reported by Seela and Wei (36) are comparable with our results, although the $\Delta\Delta G^{\circ}$ values could not be directly compared because they used A*–isocytosine and G–C base pairs as the control. They introduced 2-OH-Ade into the center of the DNA duplexes and reported that this introduction led the enthalpic loss of the duplexes and that the order of ΔH° was same as the order of ΔG°_{37}. These data have the same tendency as our data in Tables 1 and 2. In their study, the order of $-\Delta G^{\circ}_{37}$ for the duplex formation with the A*–N base pair in 5′-TA*A-3′ that corresponds to our series [1] was

C ~ G > T >> A. The differences in the free energy change ($\Delta\Delta G^\circ_{37}$) of the unusual base pairs against the stability of the most stable A*–C or A*–G base pair were 0, 0, 0.7 and 2.6 kcal mol^{-1}, respectively. The order of $-\Delta G^\circ_{37}$ for duplexes with A*–N in 5′-GA*T-3′ that resembles our series [2] was T > G ~ C >> A ($\Delta\Delta G^\circ_{37}$ against A*–T were 0, 0.5, 0.5 and 2.5 kcal mol^{-1}, respectively). These results are very close to ours. In our present study, the order of $-\Delta G^\circ_{37}$ was C > G ~ T >> A in series [1] ($\Delta\Delta G^\circ_{37}$ against A*–C were 0, 1.4, 1.4 and 3.8 kcal mol^{-1}, respectively) and T > G ~ C >> A in series [2] ($\Delta\Delta G^\circ_{37}$ against A*–T were 0, 2.0, 2.0 and 7.7 kcal mol^{-1}, respectively). The discrepancy should indicate the necessity of considering the same theoretical aspect as normal DNA; the stability of a base pair depends on the combination of both nearest neighbor base pairs. Although it is necessary to compare the stability of all 64 trinucleotides with the A*–N base pair for a formal discussion, the following points could be generally stated for the A*–N pairs at the center of the duplexes in general. (i) A→A* conversion destabilizes a duplex due to an unfavorable enthalpic effect. (ii) Stability of the A*–N base pair depends on the surrounding sequence context. Stacking of the A*–N plane with nearest neighbor base pair on both sides should significantly contribute to the determination of the A*–N base pair stability. (iii) The A*–A base pair is the most unstable. (iv) The A*–T base pair might be stably formed in any sequence context.

Thermodynamic effect of 2-OH-Ade at the end of the strand

This is the only report that evaluates the thermodynamic stability of a duplex with a terminal A*–N pair and a dangling end. For the sequence series [3] and [4], the absolute value of the thermodynamic parameters of a trinucleotide with the A*–N base pair at its center could not be determined because the duplexes had a dangling end. However, the dangling effect should be excluded by subtraction of the duplex parameters having the same dangling end. Thus we compared the stability of the A*–N base pairs based on their $\Delta\Delta G^\circ_{37}$ values.

In the GC closing series [4], the order of stability of the A*–N base pair was T > G > C >> A ($\Delta\Delta G^\circ_{37}$ against A*–T were 0, 0.8, 2.2 and 5.7 kcal mol^{-1}, respectively). This tendency resembles that in the series [2] duplexes (Fig. 3B and D) and that reported for the 5′-GA*T-3′ sequences by Seela and Wei (36). Actually, $\Delta\Delta\Delta G^\circ_{37}$, that is the difference of $\Delta\Delta G^\circ_{37}$ for the A*–N base pair in the sequence series [2] and that in [4], were within the range of 0.7 kcal mol^{-1} for the three stable base pairs (N = T, G and C). Thus, the stability of a GC closed A*–N base pair at the end of the strand is analogous to that in the center of the strand.

On the other hand, in the AT closing series [3], the order of stability of the A*–N base pairs was T > A > C > G ($\Delta\Delta G^\circ_{37}$ against A*–T were 0, 0.5, 1.7 and 3.5 kcal mol^{-1}, respectively), and was very different from that in series [1] (Fig. 3A and C). The $\Delta\Delta\Delta G^\circ_{37}$ values (differences in $\Delta\Delta G^\circ_{37}$ for A*–N in [1] and those in [3]) were quite scattered. In particular, $\Delta\Delta\Delta G^\circ_{37}$ of the A*–A base pair was 4.0 kcal mol^{-1}; the value could be interpreted such that the terminal A*–A base pair in the 5′-TA*A-3′ sequence is nearly 10^3 times more stable than that within the duplex. This result strongly suggests that the 5′-TA*A-3′ sequence is quite unique. The structure of series [3] is a model of the situation during the DNA replication of a template containing 2-OH-Ade. Depending upon the thermodynamic stabilities of the series [3] duplexes, the incorporation of T or A opposite to 2-OH-Ade is expected. Indeed, dTMP and dAMP are incorporated opposite to 2-OH-Ade in the 5′-TA*A-3′ sequences by *E.coli* and mammalian DNA polymerases *in vitro* (19,20).

It is hard to compare the mutation spectra of 2-OH-Ade in living cells directly to those of the *in vitro* results because of the effects of repair enzymes and the 3′–5′ exonuclease activity of DNA polymerases. However, the results observed in *E.coli* and mammalian COS-7 cells agreed with the results of the *in vitro* experiments (11,12). Thus, the thermodynamic stability of the A*–A pair may be an important factor for the incorporation of A opposite to 2-OH-Ade. In addition, the results of the nucleotide insertion opposite to 2-OH-Ade in the 5′-GA*C-3′ sequence is in agreement with the thermodynamic stabilities of the series [4]. Therefore, the stability of a base pair of 2-OH-Ade with an incoming nucleotide may determine the specificity of the nucleotide incorporation.

Correlation of stability of the A*–N base pair and nucleotide incorporation by DNA polymerases

The nucleotide incorporated opposite to 2-OH-Ade depends on the kind of DNA polymerase except for the 5′-TA*A-3′ sequence. In general, the mammalian DNA polymerases α and β insert T and C, and KF^{exo-} incorporates T and G (19). Thus, factors other than the thermodynamic stability would be included in this nucleotide selection. Recently, Kool and coworkers (23–25) carried out incorporation experiments with several nucleotide analogs by DNA polymerases. Taking their studies together with the biochemical and structural data reported by others (47,48), it is suggested that geometric selection of an incoming nucleotide and a minor groove interaction between the base(s) and the polymerase are important in the polymerase reactions. Moreover, these studies indicate that the recognition mode of each DNA polymerase differs to some extent.

It appears that a stable base pair is formed more easily than an unstable pair even in the active site of a DNA polymerase because our three experiments with 2-OH-Ade, *in vitro*, *in vivo*, and the thermodynamic analyses correlate with each other. However, some DNA polymerase dependence would be involved in the selective incorporation of a nucleotide. Thus, the polymerase-dependent specificity of the nucleotide incorporation opposite 2-OH-Ade seems to also be influenced by different interaction patterns of newly formed base pairs with different polymerases. On the other hand, 2-OH-Ade in the 5′-TA*A-3′ sequence formed a stable pair with A at the terminus of an elongating DNA strand (Table 2 and Fig. 3C). The experiments *in vivo* also strongly suggest the presence of a stable A*–A pair because any mispair at the end of a strand should be removed by exonucleases and the mispair should inhibit further elongation of the strand (22,49). For the geometric selection, the fact that dAMP is incorporated opposite to 2-OH-Ade in the 5′-TA*A-3′ sequence indicates that the A*–A pair in the active site is not ‘rejected’ by all the DNA polymerases tested. We propose that the thermodynamic stability of a base pair at the end of a strand is the primary factor in the determination of nucleotide selection, at least for 2-OH-Ade with some ambiguity that depends on the nature of the polymerases.

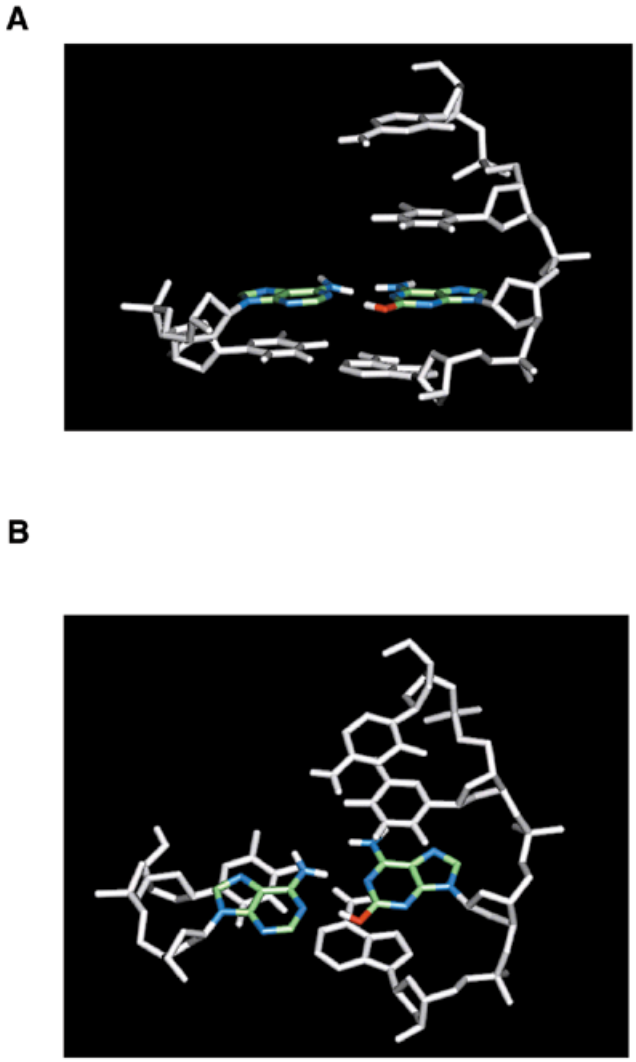

Figure 4. Molecular modeling of a base pair of 2-OH-Ade and adenine in a tetranucleotide model of AT closing 5′-TA*A-3′ sequence, 5′-CTA*A-3′/5′TA-3′. (**A**) Top view of the base pair. (**B**) Side view of the base pair with adjacent bases.

Why is the A*–A base pair stable in the AT closing series [3]?

The A*–A base pair in the AT closing [3] sequence is quite stable with a $\Delta\Delta G°_{37}$ of 0.5 kcal mol^{-1} against the natural A–T base pair. This extraordinary stability should be a reason for the A incorporation opposite to 2–OH–Ade regardless of the polymerase. Why does the A*–A base pair in the sequence [3] exhibit such a high stability? To evaluate this phenomenon, molecular modeling was carried out for the tetranucleotide models with the A*–A base pair, 5′-CTA*A-3′/5′-TAAG-3′ ($[1]_{A*A}$ model), 5′-TGA*C-3′/5′-GACA-3′ ($[2]_{A*A}$ model), 5′-CTA*A-3′/5′-TA-3’ ($[3]_{A*A}$ model), and 5′-TGA*C-3′/5′-GA-3′ ($[4]_{A*A}$ model). During this modeling, the stable A*–A pair could be formed only in the $[3]_{A*A}$ model for the O2-H enol form of 2-OH-Ade (Fig. 4). In this model, 2-OH-Ade and A make two hydrogen bonds in *anti–anti* orientation and the base pair plane was stacked in parallel to the adjacent A–T base pair plane. On the other hand, none of the A*–A pairs in the other models formed hydrogen bonds at the successively stacked position. This is the tetranucleotide model and detailed structural analyses have to be carried out to determine the structural features of this unusual base pairing. However, this result strongly supports our thermodynamic data.

CONCLUSION

We have investigated the thermodynamic stability of four distinct series of DNAs with 2-OH-Ade in the center or at the end. The thermodynamic stability of the A*–N base pair at the end correlates well with the misincorporation pattern opposite 2-OH-Ade. Incorporation of dAMP opposite 2-OH-Ade in a 5′-TA*A-3′ sequence can be explained by the unusual stability of the A*–A base pair in the 5′-TA*A-3′ sequence. The incorporation of dCMP and dGMP can also be explained by the same mechanism. Thus, thermodynamic stability of 2-OH-Ade and an incoming nucleotide may determine the nucleotide insertion specificity.

The apparent correlation may indicate the ability of rationalizing the wide mutation spectra of damaged or modified bases by the thermodynamic stability of a terminal base pair with a dangling end. Also, it may be possible to predict another ‘hot spot’ sequence at which specific modifications would occur like 5′-TA*A-3′ by thermodynamic analysis of several model DNAs.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Prof Ronald R. Breaker, MCDB Yale University, for critical reading of this manuscript. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports and Culture, Japan, and Grants from ‘Research for the Future’ Program of the Japan Society for the Promotion of Science.

REFERENCES

1. Wallace,S.S. (1994) DNA damages processed by base excision repair: biological consequences. *Int. J. Radiat. Biol.*, **66**, 579–589.
2. Kasai,H. and Nishimura,S. (1991) Formation of 8-hydroxydeoxyguanosine in DNA by oxygen radicals and its biological significance. In Sies,H. (ed.), *Oxidative Stress: Oxidants and Antioxidants*. Academic Press, San Diego, CA, pp. 99–116.
3. Kamiya,H., Miura,K., Ishikawa,H., Inoue,H., Nishimura,S. and Ohtsuka,E. (1992) c-Ha-ras containing 8-hydroxyguanine at codon 12 induces point mutations at the modified and adjacent positions. *Cancer Res.*, **52**, 3483–3485.
4. Kamiya,H., Murata-Kamiya,N., Koizume,S., Inoue,H., Nishimura,S. and Ohtsuka,E. (1995) 8-Hydroxyguanine (7,8-dihydro-8-oxoguanine) in hot spots of the c-Ha-ras gene: effects of sequence contexts on mutation spectra. *Carcinogenesis*, **16**, 883–889.
5. Wagner,J., Kamiya,H. and Fuchs,R.P. (1997) Leading versus lagging strand mutagenesis induced by 7,8-dihydro-8-oxo-2′-deoxyguanosine in *Escherichia coli*. *J. Mol. Biol.*, **265**, 302–309.
6. Wood,M.L., Dizdaroglu,M., Gajewski,E. and Essigmann,J.M. (1990) Mechanistic studies of ionizing radiation and oxidative mutagenesis: genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome. *Biochemistry*, **29**, 7024–7032.

7. Cheng,K.C., Cahill,D.S., Kasai,H., Nishimura,S. and Loeb,L.A. (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G→T and A→C substitutions. *J. Biol. Chem.*, **267**, 166–172.
8. 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.
9. Murata-Kamiya,N., Kamiya,H., Morioka,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.
10. Kamiya,H. and Kasai,H. (2000) 2-Hydroxyadenine formation by reactive oxygen species and mutagenic effects. *Recent Res. Devel. Biochem.*, **2**, 41–50.
11. Kamiya,H. and Kasai,H. (1997) Substitution and deletion mutations induced by 2-hydroxyadenine in *Escherichia coli*: effects of sequence contexts in leading and lagging strands. *Nucleic Acids Res.*, **25**, 304–311.
12. 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.
13. 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.
14. 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.
15. Shibutani,S., Takeshita,M. and Grollman,A.P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, **349**, 431–434.
16. Kamiya,H., Sakaguchi,T., Murata,N., Fujimuro,M., Miura,H., Ishikawa,H., Shimizu,M., Inoue,H., Nishimura,S., Matsukage,A. *et al.* (1992) *In vitro* replication study of modified bases in ras sequences. *Chem. Pharm. Bull.*, **40**, 2792–2795.
17. Kamiya,H., Murata-Kamiya,N., Fujimuro,M., Kido,K., Inoue,H., Nishimura,S., Masutani,C., Hanaoka,F. and Ohtsuka,E. (1995) Comparison of incorporation and extension of nucleotides *in vitro* opposite 8-hydroxyguanine (7,8-dihydro-8-oxoguanine) in hot spots of the c-Ha-ras gene. *Jpn. J. Cancer Res.*, **86**, 270–276.
18. Kamiya,H., Maki,H. and Kasai,H. (2000) Two DNA polymerases of *Escherichia coli* display distinct misinsertion specificities for 2-hydroxy-dATP during DNA synthesis. *Biochemistry*, **39**, 9508–9513.
19. Kamiya,H. and Kasai,H. (1996) Effect of sequence contexts on misincorporation of nucleotides opposite 2-hydroxyadenine. *FEBS Lett.*, **391**, 113–116.
20. 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.
21. Tinoco,I.,Jr, Uhlenbeck,O.C. and Levine,M.D. (1971) Estimation of secondary structure in ribonucleic acids. *Nature*, **230**, 362–367.
22. Morales,J.C. and Kool,E.T. (1999) Minor groove interactions between polymerase and DNA: more essential to replication than Watson–Crick hydrogen bonds? *J. Am. Chem. Soc.*, **121**, 2323–2324.
23. Moran,S., Ren,R.X. and Kool,E.T. (1997) A thymidine triphosphate shape analog lacking Watson–Crick pairing ability is replicated with high sequence selectivity. *Proc. Natl Acad. Sci. USA*, **94**, 10506–10511.
24. Morales,J.C. and Kool,E.T. (1998) Efficient replication between non-hydrogen-bonded nucleoside shape analogs. *Nature Struct. Biol.*, **5**, 950–954.
25. 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.
26. Kool,E.T. (1998) Replication of non-hydrogen bonded bases by DNA polymerases: a mechanism for steric matching. *Biopolymers*, **48**, 3–17.
27. Koizume,S., Kamiya,H., Inoue,H. and Ohtsuka,E. (1994) Synthesis and thermodynamic stabilities of damaged DNA involving 8-hydroxyguanine (7,8-dihydro-8-oxoguanine) in a ras gene fragment. *Nucl. Nucl.*, **13**, 1517–1534.
28. Seela,F., Mertens,R. and Kazimierczuk,Z. (1992) Synthesis of phosphonates and oligodeoxyribonucleotides derived from 2′-deoxyisoguanosine and 2′-deoxy-2-haloadenosines. *Helv. Chim. Acta*, **75**, 2298–2306.
29. Sugimoto,N., Nakano,S., Katoh,M., Matsumura,A., Nakamuta,H., Ohmichi,T., Yoneyama,M. and Sasaki,M. (1995) Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry*, **35**, 11211–11216.
30. Seela,F., Wei,C. and Kazimierczuk,Z. (1995) Substituent reactivity and tautomerism of isoguanosine and related nucleotides. *Helv. Chim. Acta*, **78**, 1843–1854.
31. Richerds,E.G. (1975) Use of tables in calculation of absorption, optical rotatory dispersion, and circular dichroism of polyribonucleotides. In Fasman,G.D. (ed.), *Handbook of Biochemistry and Molecular Biology*, 3rd edn. Nucleic Acids Vol. 1. CRC Press, Cleveland, OH, pp. 596–603.
32. Sugimoto,N., Kierzek,R., Freier,S.M. and Turner,D.H. (1986) Energetics of internal GU mismatches in ribooligonucleotide helixes. *Biochemistry*, **25**, 5755–5759.
33. Petersheim,M. and Turner,D.H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry*, **22**, 256–263.
34. Longfellow,C.E., Kierzek,R. and Turner,D.H. (1990) Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry*, **29**, 278–285.
35. Sepiol,J., Kazimierczuk,Z. and Shugar,D. (1976) Tautomerism of isoguanosine and solvent-induced keto-enol equilibrium. *Z. Naturforsch C*, **31**, 361–370.
36. Seela,F. and Wei,C. (1999) The base-pairing properties of 7-deaza-2′-deoxyisoguanosine and 2′-deoxyisoguanosine in oligonucleotide duplexes with parallel and antiparallel chain orientation. *Helv. Chim. Acta*, **82**, 726–745.
37. Robinson,H., Gao,Y.-G., Bauer,C., Roberts,C., Switzer,C. and Wang,A.H.-J. (1998) 2′-Deoxyisoguanosine adopts more than one tautomer to form base pairs with thymidine observed by high-resolution crystal structure analysis. *Biochemistry*, **37**, 10897–10905.
38. Roberts,C., Bandaru,R. and Switzer,C. (1997) Varied molecular interactions at the active sites of several DNA polymerases: nonpolar nucleoside isosteres as probes *J. Am. Chem. Soc.*, **119**, 4640–4649.
39. Seela,F. and Wei,C. (1997) Oligonucleotide containing consecutive 2′-deoxyisoguanosine residues: synthesis, duplexes with parallel chain orientation, and aggregation. *Helv. Chim. Acta*, **80**, 73–85.
40. Seela,F., Wei,C., Melenewski,A. and Feiling,E. (1998) Parallel-stranded duplex DNA and self-assembled quartet structures formed by isoguanine and related bases. *Nucl. Nucl.*, **17**, 2045–2052.
41. Seela,F., Wei,C. and Melenewski,A. (1996) Isoguanine quartets formed by d(T4isoG4T4): tetraplex identification and stability. *Nucleic Acids Res.*, **24**, 4940–4945.
42. Chaput,J.C. and Switzer,C. (1999) A DNA pentaplex incorporating nucleobase quintets. *Proc. Natl Acad. Sci. USA*, **96**, 10614–10619.
43. Yang,X.-L., Sugiyama,H., Ikeda,S., Saito,I. and Wang,A.H.-J. (1998) Structural studies of a stable parallel-stranded DNA duplex incorporating isoguanine:cytosine and isocytosine:guanine basepairs by nuclear magnetic resonance spectroscopy. *Biophys. J.*, **75**, 1163–1171.
44. Bain,J.D., Chamberlin,A.R., Switzer,C.Y. and Benner,S.A. (1992) Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code. *Nature*, **356**, 537–539.
45. Kamiya,H. and Kasai,H. (2000) 2-Hydroxy-dATP is incorporated opposite G by *Escherichia coli* DNA polymerase III resulting in high mutagenicity. *Nucleic Acids Res.*, **28**, 1640–1646.
46. Chen,X., McDowell,J.A., Kierzek,R., Krugh,T.R. and Turner,D.H. (2000) Nuclear magnetic resonance spectroscopy and molecular modeling reveal that different hydrogen bonding patterns are possible for G.U pairs: one hydrogen bond for each G.U pair in r(GGCGUGCC)$_2$ and two for each G.U pair in r(GAGUGCUC)$_2$. *Biochemistry*, **39**, 8970–8982.
47. Horlacher,J., Hottiger,M., Podust,V.N., Hübscher,U. and Benner,S.A. (1995) Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Natl Acad. Sci. USA*, **92**, 6329–6333.
48. Pelletier,H., Sawaya,M.R., Kumar,A., Wilson,S.H. and Kraut,J. (1994) Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science*, **264**, 1891–1903.
49. Morales,J.C. and Kool,E.T. (2000) Importance of terminal base pair hydrogen-bonding in 3′-end proofreading by the Klenow fragment of DNA polymerase I. *Biochemistry*, **39**, 2626–2632.