4206–4214 *Nucleic Acids Research, 2001, Vol. 29, No. 20*

Arg660Ser mutation in *Thermus aquaticus* DNA polymerase I suppresses T→C transitions: implication of wobble base pair formation at the nucleotide incorporation step

Katsushi Yoshida, Aki Tosaka[1], Hiroyuki Kamiya[2], Takashi Murate[3], Hiroshi Kasai[4], Yuji Nimura, Masanori Ogawa[1], Shonen Yoshida[1] and Motoshi Suzuki[1,*]

Division of Surgical Oncology, Department of Surgery and [1]Laboratory of Cancer Cell Biology, Research Institute for Disease Mechanism and Control, Nagoya University Graduate School of Medicine, Tsurmai-cho, Showa-ku, Nagoya 466-8550, Japan, [2]Graduate School of Pharmaceutical Sciences, Hokkaido University, Sapporo 060-0812, Japan, [3]Nagoya University School of Health Science, Nagoya 461-8673, Japan and [4]Department of Environmental Oncology, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, Kitakyushu 807-8555, Japan

Received June 8, 2001; Revised and Accepted August 3, 2001

ABSTRACT

We examined the replication fidelity of an Arg660Ser (R660S) mutant of *Thermus aquaticus* DNA polymerase I (*Taq* pol I). In a forward mutation assay, R660S showed a marked reduction in T→C transitions, one of the most frequent errors made by the wild-type enzyme. Steady-state kinetics showed that R660S discriminates against dGTP incorporation at a template T 13-fold better than the wild-type. R660S was also 3.2-fold less efficient than the wild-type at extending a T:dG mismatch. These results indicate that R660S has enhanced fidelity during incorporation and extension, which reduces its T→C transition frequency. Interestingly, R660S also discriminated correct from incorrect nucleotides at the incorporation step of C:dATP, A:dATP, G:dATP and C:8-OH-dGTP mispairs 28-, 6.0-, 4.1- and 6.8-fold better, respectively, than the wild-type, although it may not always be as accurate as the wild-type at the extension step. A structural model suggests that Arg660 may participate in two interactions that influence fidelity; the guanidinium group of Arg660 might interact with the incoming guanine base at the major groove and it might compete for forming another interaction with the primer terminus. Substituting Arg with Ser may eliminate or alter these interactions and destabilize the closed complex with incorrect substrates. Our data also suggest that T:dGTP and C:dATP base pairs form 'wobble' structures at the incorporation step of *Taq* pol I.

INTRODUCTION

DNA polymerases have been characterized in many organisms and are a highly conserved group of proteins with regions of high homology in the aligned amino acid sequences (1). Crystal structures show that polymerases share a common structure with subdomains called 'fingers', 'palm' and 'thumb' and an active site located in a deep cleft that binds DNA (2). The palm subdomain forms the floor of the active site cleft and includes the highly conserved motif A and a catalytically essential aspartic acid (3,4). The fingers subdomain defines one wall of the cleft and includes conserved motif B [most of the long O-helix (5)].

Some of the catalytically essential polymerase residues involved in DNA synthesis have been well characterized (for details see 2,6). In addition, amino acid residues have been identified that play a role in DNA polymerase fidelity. Amino acid substitutions in these residues lower or increase fidelity and can alter the interaction between enzyme and substrate or template–primer (for details see 7). For example, a *Taq* pol I mutant, F667L, is a transversion antimutator in which the stacking force against bulky purine:purine base pairs may be reduced (8). Mutations affecting polymerase fidelity have also been identified in amino acids that do not contact the substrate or template–primer. Polymerase β mutant Y265H produces errors at a 40-fold higher frequency than the wild-type. In this mutant, Tyr265 does not make direct contact with the DNA substrate, however, the mutant has lowered dNTP discrimination at the level of k_{pol} (9). Another mutant in *Taq* pol I, T664P, produces base substitution and frameshift errors at a 5–150-fold higher frequency than the wild-type. In this mutant, the O-helix may bend and enlarge the catalytic pocket and increase the rate of nucleotide misincorporation (10).

In this manuscript, we describe a *Taq* pol I mutant with a Ser substitution at Arg660. Our data show that Arg660 is involved in T→C base substitution errors and 8-OH-dGTP incorporation.

*To whom correspondence should be addressed. Tel: +81 52 744 2456; Fax: +81 52 744 2457; Email: msuzuki@med.nagoya-u.ac.jp

```
T1    3′-GCGCGGCTTAAGGGCGATCGTTATAAGACGTCGGTTCTAAGGCGCG
P2   *5′-CGCGCCGAATTCCC

T3    3′-GCTTTCCCCCTACACGACGTTCCGCTAATTCAACCCATTG
T3G   3′-GCTTTCCCCCTACACGACGTTGCGCTAATTCAACCCATTG
P4   *5′-CGAAAGGGGGATGTGCTGCA
P5   *5′-CGAAAGGGGGATGTGCTGCAA
P6   *5′-CGAAAGGGGGATGTGCTGCAG

T7    3′-GCTTTCCCCCTACACGACGTTCCGCTAATTCAACCCATTG
P8              *5′-ATGTGCTGCAAGGCGATTAA
P9              *5′-ATGTGCTGCAAGGCGATTAAG
P10             *5′-ATGTGCTGCAAGGCGATTAAA

T11   3′-ACGGATTACTCACTCGATTGAGTGTAATTAACGCAACGCG
P12  *5′-TGCCTAATGAGTGAGCTAAC
P13  *5′-TGCCTAATGAGTGAGCTAACT
P14  *5′-TGCCTAATGAGTGAGCTAACA

T15   3′-CCCATTGCGGTCCCAAAAGGGTCAGTGCTGCAACATTTTG
P16  *5′-GGGTAACGCCAGGGTTTTCC
P17  *5′-GGGTAACGCCAGGGTTTTCCC
P18  *5′-GGGTAACGCCAGGGTTTTCCA
```

Figure 1. Oligonucleotides used in this study. Oligonucleotides T1 and P2 were annealed and used for primer extension and kinetic analyses of 8-OH-dGTP incorporation. For single nucleotide incorporation kinetics, oligonucleotides T3/P4, T3G/P4, T7/P8, T11/P12 and T15/P16 were used for T:dATP/dGTP, C:dGTP/dATP, A:dTTP/dATP and G:dCTP/dATP incorporation, respectively. For extension kinetics, oligonucleotides T3/P5 and T3/P6, T3G/P5 and T3G/P6, T7/P9 and T7/P10, T11/P13 and T11/P14 and T15/P17 and T15/P18 were used for incorporation of dGTP, dTTP, dCTP and dATP, respectively. Primer strands are indicated by an asterisk.

Crystal structures and a derived model suggest that this side chain may interact with the incoming guanine base and the primer terminus in the closed complex (8,11). We discuss possible mechanisms for the altered fidelity of Arg660Ser *Taq* pol I.

MATERIALS AND METHODS

Materials

Wild-type and R660S *Taq* pol I were purified and activity was measured as described previously (12). The specific activities of the wild-type and R660S polymerases were 37 000 and 30 000 U/mg, respectively. Enzyme specific activity was lower with the newly prepared activated DNA substrate than with a previous one (8). Gel-purified oligonucleotides (Fig. 1) and Ultra-pure dNTPs were purchased from Amersham Pharmacia Biotech (Little Chalfont, UK). Preparation of 8-OH-dGTP was as described (13).

Primer extension assay

The DNA primer (Fig. 1, P2) was ^{32}P-labeled at the 5′-end by incubation with [γ-^{32}P]ATP and T4 polynucleotide kinase and annealed in a 2-fold molar excess to the DNA template 46mer (Fig. 1, T1). Primer elongation was catalyzed by purified wild-type or R660S *Taq* pol I. An aliquot of 0.06–0.6 U pol I was incubated at 45°C for 30 min in reaction mixtures containing 50 mM Tris–HCl, pH 8.0, 2 mM $MgCl_2$, 50 mM KCl, 100 µM dNTPs as specified and 12.5 µM labeled template–primer in a final volume of 10 µl. The reaction was terminated by addition of an equal amount of 2× loading buffer containing 90% formamide, 20 mM EDTA, 0.05% xylene cyanol and 0.05% bromophenol blue. Reaction products were separated using a 14% polyacrylamide gel containing 8 M urea and exposed to X-ray film.

Forward mutation analysis

The gapped M13mp2 dsDNA substrate (200 ng) was incubated with 10 U wild-type or R660S *Taq* pol I in 20 µl with 200 µM each dNTP, 50 mM Tris–HCl, pH 8.0, 7 mM $MgCl_2$, 50 mM KCl. After incubation at 72°C for 5 min, the DNA was transfected into *Escherichia coli* and wild-type (dark blue) and mutant (pale blue or white) plaques were scored. In the present analysis, the forward mutation assay was performed with 0.14 ng/µl enzyme; this concentration is similar to the 'high polymerase concentration' (0.19 ng/µl) referred to in a previous report (8). Mutant frequency of wild-type *Taq* pol I was thus similar to that determined previously [compare Table 1 in this manuscript with table I in Suzuki *et al.* (8)].

Single nucleotide incorporation kinetics

Incorporation efficiency for correct and incorrect nucleotides was measured essentially as described (14). The ^{32}P-labeled template–primer (T1/P2, T3/P4, T3G/P4, T7/P8, T11/P12 and T15/P16 in Fig. 1) was used at a final concentration of 12.5 nM in a 10 µl reaction containing 50 mM Tris–HCl, pH 8.0, 2 mM $MgCl_2$, 50 mM KCl with various concentrations of enzyme and dNTP at 45°C for 5 min. For incorporation of 8-OH-dGTP, reactions were carried out for 120 min. The reaction was terminated with <20% template–primer utilization. Apparent values for K_m and V_{max} were calculated from Hanes–Woolf plots as the averages of three determinations.

Single nucleotide extension kinetics

Primer extension efficiency for matched or mismatched terminal base pairs was measured essentially as described (14,15). The ^{32}P-labeled template–primer (T3/P5, T3/P6, T3G/P5, T3G/P6, T7/P9, T7/P10, T11/P13, T11/P14, T15/P17 and T15/P18 in Fig. 1) was used at a final concentration of 12.5 nM in a 10 µl reaction containing 50 mM Tris–HCl, pH 8.0, 2 mM $MgCl_2$, 50 mM KCl and various concentrations of enzyme and dNTP at 45°C for 5 min. The reaction was terminated with <20% template–primer utilization. Apparent values for K_m and V_{max} were calculated from Hanes–Woolf plots as the averages of three determinations.

RESULTS

Generation of the R660S mutant

We previously created a library of *Taq* DNA pol I mutants (5). We prepared extracts of *E.coli* expressing mutant *Taq* pol I enzymes and screened 67 such mutant extracts in a primer elongation assay (12). Host DNA polymerases and nucleases were inactivated by heating at 72°C for 20 min and the ability of heat-treated extracts to elongate primers in the absence of a complete complement of four dNTPs was then determined. Relative to wild-type *Taq* pol I, mutants with increased fidelity would extend a smaller proportion of primers up to and beyond the first template position for which the complementary dNTP is absent. Our examination of 67 extracts revealed 15 mutants which we judged to be possible high fidelity variants. Notably, it is more difficult to identify potential high fidelity mutants than low fidelity mutants (12). A major factor is that the longer

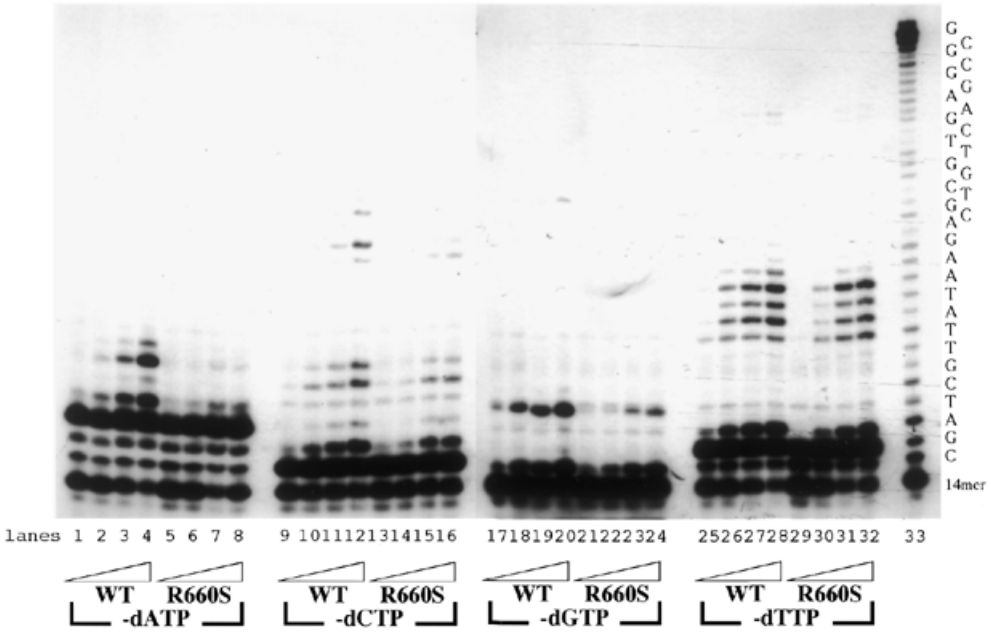

Figure 2. Primer extension by wild-type and R660S *Taq* pol I. Wild-type or R660S *Taq* pol I was incubated with ^{32}P-labeled primer–template at 45°C for 30 min. Enzyme was added at the following amounts: 0.06 (lanes 1, 5, 9, 13, 17, 21, 25 and 29), 0.2 (lanes 2, 6, 10, 14, 18, 22, 26 and 30), 0.4 (lanes 3, 7, 11, 15, 19, 23, 27 and 31) or 0.6 U (lanes 4, 8, 12, 16, 20, 24, 28 and 32). Reactions were carried out in the absence of dATP (lanes 1–8), dCTP (lanes 9–16), dGTP (lanes 17–24) or dTTP (lanes 25–32). In lane 33, wild-type *Taq* pol I was incubated with the four dNTPs. The unextended primer is a 14mer. The template sequence (T1 in Fig. 1) is given on the right.

products synthesized by unfaithful polymerases tend to be conspicuous in gel electrophoretic patterns and may have no counterpart in the wild-type pattern. For more faithful polymerases, however, quantitative differences in short products with similar mobilities that appear in both the mutant and wild-type gel patterns can be crucial. Another factor is that crude extracts may contain considerable chromosomal DNA that could bind to the mutant polymerases. Sequestration of a mutant polymerase might limit extension of the competing template–primer, thus mimicking high fidelity.

We purified the 15 respective polymerases and evaluated their PCR activity and fidelity. We identified two mutant *Taq* pol I enzymes among the 67 mutants screened that possess nearly wild-type specific activity, are sufficiently processive to amplify a 0.5 kb DNA sequence in a PCR assay and exhibit less elongation than the wild-type enzyme in the primer extension assay. One of these candidate anti-mutator pol I enzymes was the triple substituted variant A661E:I665T:F667L (8). In this manuscript, we describe characterization of another mutant, R660S. The specific activity of purified R660S is 80% of the wild-type and the enzyme amplifies a 2 kb PCR fragment as efficiently as the wild-type (data not shown).

Primer extension assay

We performed primer extension DNA synthesis in reactions that lack one of the four dNTPs (Fig. 2). Under these conditions the polymerase pauses one nucleotide upstream of template positions where the complementary dNTP is missing. DNA synthesis continues only when DNA polymerase incorporates an incorrect dNTP. Pausing was examined at a template T in reactions lacking dATP and at a high enzyme concentration. Wild-type pol I extended 19% of the primer beyond the first template T, while R660S extended 4% of the primer to a similar extent (lanes 4 and 8). Similar results were observed when other dNTPs were missing from the reaction. These results suggest that R660S incorporates incorrect nucleotides less efficiently than wild-type *Taq* pol I.

Forward mutation assay

The fidelities of wild-type and R660S *Taq* pol I were also measured quantitatively using the M13mp2 forward mutation assay (16). This assay detects polymerase misincorporations during DNA synthesis of 214 nt of the *lacZ*α gene in a ssDNA region of gapped circular M13mp2 dsDNA. To our surprise, the mutant frequency with R660S, measured twice, showed values similar to wild-type *Taq* pol I (Table 1). Forty-three mutant plaques generated by R660S in the forward mutation assay were isolated and the DNA sequence of the *lacZ*α gene was determined. The mutation spectrum of R660S included 36 base substitutions and seven frameshifts (six deletions and one addition; Table 2). T→C substitutions are one of the most frequent errors made by wild-type *Taq* pol I (8), and the frequency of T→C was much lower for R660S than for the wild-type. The base substitutions made by R660S included 16 G→T, seven A→T, six C→T, two T→A and one each of A→G, A→C, G→A and G→C. The percentages of G→T and C→T substitutions made by R660S are higher than for the wild-type, which may primarily reflect the low number of T→C substitutions.

Single nucleotide incorporation efficiency

The primer extension assay primarily measures misincorporation at the first template base for which no complementary

Table 1. Mutant frequencies in DNA copied by wild-type and R660S *Taq* pol I in the *lacZ*α forward mutation assay

	total plaque	mutant plaque	mutant frequency [a] ($x10^{-3}$)
experiment 1			
wild type	7968	160	19.2
R660S	1318	20	14.3
experiment 2			
wild type	3285	76	22.2
R660S	5097	101	19.8

[a]Background frequency (0.87×10^{-3}) was subtracted.

Table 2. Types of errors observed for wild-type and mutant *Taq* pol I in the forward mutation assay

	R660S		wild type [a]
error type	N	%	%
base substitutions	36	100	100
T->C	1	-	23
T->A	2	-	-
T->G	0	-	-
A->T	7	19	34
A->G	1	-	-
A->C	1	-	-
G->A	1	-	-
G->T	16	44	21
G->C	1	-	-
C->T	6	17	10
C->A	0	-	-
C->G	0	-	-
frameshifts [b]	7	100	100
deletions	6	86	100
addition	1	-	-

[a]Data from Suzuki *et al.* (8).
[b]Frameshifts of ≥2 nt were not observed.

dNTP is available, and DNA synthesis after the misincorporation event may not contribute significantly to the total amount of reaction product. In contrast, mutations detected by the forward mutation assay require *in vitro* incorporation of an incorrect nucleotide and extension of the mismatched primer terminus, because an unextended mismatched terminus is repaired in *E.coli* and does not generate a mutant plaque. The above results suggest that wild-type and R660S *Taq* pol I have different abilities to discriminate between correct and incorrect nucleotides at the incorporation and primer extension steps. Therefore, the kinetic constants were determined for T→C, C→T, A→T and G→T substitutions, which are frequent errors for wild-type and R660S *Taq* pol I. For correct nucleotides, R660S showed an incorporation efficiency (V_{max}/K_m) comparable to the wild-type (Tables 3–5). In contrast, T:dGTP incorporation efficiency for R660S was 6.3-fold lower than for the wild-type. Discrimination factors, the ratios of incorporation efficiency for incorrect and correct nucleotides, were calculated and compared for wild-type and R660S *Taq* pol I. For T:dGTP, C:dATP, A:dATP and G:dATP misincorporation, the discrimination factor for R660S was 13-, 28-, 6.0- and 4.1-fold lower, respectively, than for the wild-type. These data indicate that R660S is more accurate than the wild-type and discriminates against specific misincorporations more strongly than wild-type *Taq* pol I.

Single nucleotide extension efficiency

Extension efficiencies from mismatched primer termini T:dG and C:dA were also measured for wild-type and R660S *Taq* pol I by measuring incorporation of the downstream nucleotide. For the T:dG mismatch, the discrimination factor of R660S was 3.2-fold lower than that of the wild-type. Therefore, R660S is an antimutator mutant for T→C transition at both the incorporation and extension steps. These results are consistent with the spectrum generated by R660S mutant polymerase. In the forward mutation assay, T→C transitions occur at a frequency of 1 in 36 base substitutions; in contrast, wild-type *Taq* pol I generates T→C transitions at a frequency of 23% of all base substitutions (Table 2). In contrast, for a C:dA mismatch, the discrimination factor of R660S is 2.8-fold higher than that of the wild-type (Table 4). Therefore, dATP is less efficiently incorporated at template C (Table 3) but a C:dA mismatch primer–terminus is more efficiently extended by R660S than the wild-type. This may also be consistent with the observation that 17% of the base substitutions made by R660S are C→T transitions, which is similar to the 10% frequency of C→T transitions for wild-type *Taq* pol I (Table 2). For A:dA and G:dA mismatches, extension was very inefficient and the efficiency could not be determined. Poor extension from these

Table 3. Steady-state kinetic parameters for incorporation accuracy by wild-type and R660S *Taq* pol I

base pair [a]	polymerase	V_{max} min^{-1}	K_m μM	f [b] $\mu M^{-1}min^{-1}$	DF [c]	Ratio of DF [d]
T:dATP	Wild type	4.0	1.5	2.7	-	-
	R660S	6.7	1.3	5.3	-	-
T:dGTP	Wild type	$9.6x10^{-2}$	80	$1.2x10^{-3}$	$4.5x10^{-4}$	1
	R660S	$1.7x10^{-2}$	90	$0.19x10^{-3}$	$0.36x10^{-4}$	1/13
C:dGTP	Wild type	1.7	0.25	7.0	-	-
	R660S	8.1	0.42	19	-	-
C:dATP	Wild type	$3.2x10^{-2}$	221	$1.4x10^{-4}$	$2.0x10^{-5}$	1
	R660S	$0.40x10^{-2}$	282	$0.14x10^{-4}$	$0.073x10^{-5}$	1/28
A:dTTP	Wild type	0.31	0.49	0.63	-	-
	R660S	0.89	1.9	0.48	-	-
A:dATP	Wild type	$2.2x10^{-2}$	53	$4.1x10^{-4}$	$6.4x10^{-4}$	1
	R660S	$0.37x10^{-2}$	72	$0.52x10^{-4}$	$1.1x10^{-4}$	1/6.0
G:dCTP	Wild type	0.79	1.2	0.65	-	-
	R660S	1.6	2.5	0.66	-	-
G:dATP	Wild type	$1.5x10^{-2}$	223	$6.9x10^{-5}$	$1.1x10^{-4}$	1
	R660S	$0.27x10^{-2}$	162	$1.7x10^{-5}$	$0.26x10^{-4}$	1/4.1

The rate constants were derived from Hanes–Woolf plots. V_{max} and K_m were obtained for incorporation of correct versus incorrect nucleotides. The values are the averages of triplicate determinations.
[a]Template base, incoming dNTP base.
[b]f, enzyme efficiency = V_{max}/K_m.
[c]DF, discrimination factor = $f_{incorrect\ dNTP}/f_{correct\ dNTP}$.
[d]DF ratio = $DF_{R660S}/DF_{wild\text{-}type}$.

bulky mispairs has also been reported by other groups using *Taq* pol I and the exo⁻ Klenow fragment of *E.coli* pol I (15,17).

Incorporation efficiency of 8-OH-dGTP

Kinetic studies summarized in Table 3 indicate that R660S discriminates T:dGTP, C:dATP, A:dATP and G:dATP mispairs better than the wild-type at the incorporation step. The ability of R660S to incorporate the modified nucleotide 8-OH-dGTP, an oxidized form of dGTP, was also examined. 8-OH-dGTP is a mutagen *in vitro* and *in vivo* (13,18,19). Because 8-OH-dGTP can be incorporated opposite C and forms a Watson–Crick-type base pair (20), the geometry of this mispair may be more similar to the correct C:dGTP than to other incorrect base pairs that were examined in this study. Nevertheless, R660S discriminated 8-OH-dGTP better than the wild-type; the discrimination factor for 8-OH-dGTP over dGTP was 6.8-fold lower than that for wild-type *Taq* pol I. These results suggest that R660S suppresses misincorporation involving radical changes in the geometry of a nascent base pair and also discriminates against subtle modifications in the catalytic pocket.

Effect of 5′-sequence context in T→C transitions

In T→C transitions made by wild-type *Taq* pol I, 64% occurred at the sequence 5′-C*T* (8). We used one of those hot-spot sequences for kinetic studies (Tables 3 and 4). We further asked whether the specific reduction in T→C transitions is unique to this sequence context or whether it is also found in others. Kinetic parameters were measured for T:dGTP incorporation and extension from a T:dG primer terminus using a new template that has the same sequence except that the upstream C was replaced by G.

As shown in Table 6, at a new sequence context in the template, 5′-G*T*, R660S discriminated substrate dGTP in the incorporation assay and primer terminus dG in the extension assay to approximately the same fold better than the wild-type. These data indicate that R660S suppresses T→C transitions not only at the 5′-C*T* site, but also at other sites with different upstream sequences.

Table 4. Steady-state kinetic parameters for extension accuracy by wild-type and R660S *Taq* pol I

3'-termini	incoming base [a]	polymerase	V_{max} min^{-1}	K_m μM	f [b] $\mu M^{-1}min^{-1}$	DF [c]	Ratio of DF [d]
T:dA	dGTP	Wild type	2.5	0.51	4.8	-	-
	dGTP	R660S	5.8	1.1	5.2	-	-
T:dG	dGTP	Wild type	0.18	27	$6.8x10^{-3}$	$1.4x10^{-3}$	1
	dGTP	R660S	0.075	33	$2.3x10^{-3}$	$0.44x10^{-3}$	1/3.2
C:dG	dTTP	Wild type	1.2	1.7	0.72	-	-
	dTTP	R660S	3.0	4.9	0.63	-	-
C:dA	dTTP	Wild type	$2.6x10^{-3}$	106	$2.5x10^{-5}$	$3.5x10^{-5}$	1
	dTTP	R660S	$6.2x10^{-3}$	102	$6.1x10^{-5}$	$9.8x10^{-5}$	2.8
A:dT	dCTP	Wild type	0.81	0.21	3.8	-	-
	dCTP	R660S	1.7	0.56	3.0	-	-
A:dA	dCTP	Wild type	-	-	-	-	-
	dCTP	R660S	-	-	-	-	-
G:dC	dATP	Wild type	0.74	0.63	1.2	-	-
	dATP	R660S	2.4	1.4	1.7	-	-
G:dA	dATP	Wild type	-	-	-	-	-
	dATP	R660S	-	-	-	-	-

The rate constants were derived from Hanes–Woolf plots. V_{max} and K_m were obtained for extension of matched versus mismatched primer termini. The values are the averages of triplicate determinations.
[a]Incoming correct nucleotide.
[b]f, enzyme efficiency = V_{max}/K_m.
[c]DF, discrimination factor = $f_{incorrect\ dNTP}/f_{correct\ dNTP}$.
[d]DF ratio = $DF_{R660S}/DF_{wild\text{-}type}$.

DISCUSSION

Role of a major groove interaction in *Taq* pol I

Numerous residues are involved in the discrimination of incorrect base pairs by DNA polymerases. Some residues are essential for both catalysis and discrimination; incorrect base pairs may disturb the proper geometry and alter the efficiency of the phosphoryl transfer reaction. Using a *Taq* pol I mutant library we have isolated fidelity mutants, but we did not find any that changed the hydrophobic pocket structure (residues 665–671), with the exception of F667L, which was associated with two other compensatory substitutions (8). Our screening strategy might have excluded radical pocket mutations because they are likely to be associated with a deleterious reduction in catalytic activity (5).

Interestingly, Arg660 is a non-conserved, dispensable residue for catalysis and is not in the hydrophobic pocket wall but hangs over the major groove of a nascent base pair. Despite this structural evidence, substitution with Ser suppresses T→C transitions and a variety of misincorporations (see below). Our data show that polymerase fidelity is controlled not only by tight interactions in the pocket, but also by loose interactions at the major groove.

Structural implications at the incorporation step

In the closed, ternary complex of wild-type *Taq* pol I, Li *et al.* (11) indicated that the guanidinium group of Arg660 may interact with the O^6 and N^7 atoms of the incoming guanine (Fig. 3). They also suggested that these interactions may not form when Arg660 is mutated to Ser (or to another amino acid), because the mutants show a marked and selective reduction in ddGTP incorporation rate. Our data show that R660S generates T→C base substitutions less frequently in the forward mutation assay (Table 2) and that the mutant is less efficient in T:dGTP incorporation and extension from a T:dG primer terminus (Tables 3 and 4). We assume that interactions between Arg660 and the guanine base may play a role during formation of a T→C base substitution by wild-type *Taq* pol I, but similar interactions may not form with R660S.

In addition to the strong selective reduction in T→C base substitutions, R660S misincorporates deoxyribonucleotides in C:dATP, A:dATP, G:dATP and C:8-OH-dGTP base pairs less efficiently than the wild-type (Tables 3 and 5). These data

Table 5. Steady-state kinetic parameters for incorporation accuracy by wild-type and R660S *Taq* pol I

base pair [a]	polymerase	V_{max} min^{-1}	K_m μM	f [b]	DF [c]	Ratio of DF [d]
C:dGTP	Wild type	1.2	1.9	0.63	-	-
	R660S	2.7	8.8	0.3	-	-
C:8-OH-dGTP	Wild type	$1.5x10^{-4}$	56	$2.7x10^{-6}$	$4.3x10^{-6}$	1
	R660S	$0.075x10^{-4}$	39	$0.19x10^{-6}$	$0.63x10^{-6}$	1/6.8

The rate constants were derived from Hanes–Woolf plots. V_{max} and K_m were obtained for incorporation of dGTP versus 8-OH-dGTP. The values are the averages of triplicate determinations.
[a]Template base, incoming dNTP base.
[b]f, enzyme efficiency = V_{max}/K_m.
[c]DF, discrimination factor = $f_{incorrect\ dNTP}/f_{correct\ dNTP}$.
[d]DF ratio = $DF_{R660S}/DF_{wild\text{-}type}$.

Table 6. Effect of local 5′-nucleotide sequence on T→C transitions

mispair[a]	polymerase	template 5'-C*T*[b] f $\mu M^{-1}min^{-1}$	DF	Ratio of DF[c]	template 5'-G*T*[b] f $\mu M^{-1}min^{-1}$	DF	Ratio of DF[c]
T:dATP	Wild type	2.7	-	-	5.0	-	-
	R660S	5.3	-	-	5.4	-	-
T:dGTP	Wild type	$1.2x10^{-3}$	$4.5x10^{-4}$	1	$2.5x10^{-3}$	$5.1x10^{-4}$	1.1
	R660S	$0.19x10^{-3}$	$0.36x10^{-4}$	1/13	$0.32x10^{-3}$	$0.60x10^{-4}$	1/7.5
T:dA	Wild type	4.8	-	-	7.2	-	-
	R660S	5.2	-	-	7.0	-	-
T:dG	Wild type	$6.8x10^{-3}$	$1.4x10^{-3}$	1	$1.3x10^{-2}$	$1.8x10^{-3}$	1.2
	R660S	$2.3x10^{-3}$	$0.44x10^{-3}$	1/3.2	$0.29x10^{-2}$	$0.42x10^{-3}$	1/3.3

The rate constants were derived from Hanes–Woolf plots. The values are the averages of triplicate determinations.
[a]Top: template base, incoming dNTP base; bottom: template base, 3′-terminus base pair.
[b]The target sequence T is shown in italic. Enzyme efficiency (*f*), the discrimination factor (DF) and the DF ratio were calculated as in Tables 3 and 4.
[c]The wild-type DF for T:dGTP (misincorporation) and for T:dG (misextension) were regarded as 1.

suggest that Arg660 may interact with the ternary complex in several ways that affect misincorporation fidelity, not only with an incoming guanine base. This idea is supported by the fact that the Arg660 guanidinium group competes with the 3′-terminal phosphate group of the primer DNA (Fig. 3). We have proposed that interactions between the O-helix and the primer terminus may stabilize the closed complex, allowing time for the reaction to take place, and increasing the probability of a phosphoryl transfer reaction with incorrectly paired dNTPs (8). Our present results are consistent with this hypothesis because substitution of Arg by Ser may remove, or at least alter, the electrostatic interaction and reduce the stabilization effect.

Base pair structures at the incorporation step

In the closed complex with a T:dGTP base pair, the geometry of the O-helix with the incoming wrong nucleotide dGTP and the primer terminus is likely to be similar to that with the

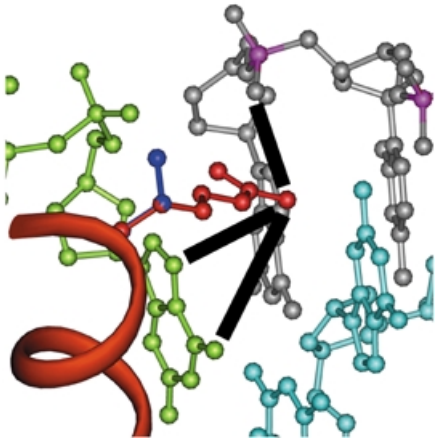

Figure 3. Structural model of *Taq* pol I. The N-terminal portion of the O-helix is shown using a ball and stick representation; a red ribbon indicates the backbone and side chain of Arg660 (wild-type), while blue indicates Ser660 (mutant). Substrate ddGTP, primer and template are in light green, gray and light blue, respectively. Phosphorus atoms in the primer are in purple. The guanidinium group of Arg660 interacts with the O^6 and N^7 atoms of the incoming guanine, as well as with the primer phosphate group (solid lines). The interaction between Arg660 and the O^6 and N^7 atoms of the incoming guanine (shown as ddGTP) is specific: in other respects, the structure of the O-helix also applies to incoming ddCTP, ddTTP and dATP (11). The coordinate sets were obtained from the Protein Data Bank (24). The drawing was made using the program package InsightII (Molecular Simulations, San Diego, CA).

correct C:dGTP base pair; if the wrong base pair T:dGTP made a big difference, the interaction between Arg660 and dGTP/primer phosphate would be lost and we would not have seen a difference in misincorporation between the wild-type and R660S. For example, if the T:dGTP base pair is much larger than C:dGTP, Arg660 may not come into close proximity to the primer phosphate.

So far the structure of the T:dGTP base pair is not known at the nucleotide incorporation step, although the crystal and MMR structures of the T:G base pair in a double-stranded oligomer have been solved (21,22). The T:dG base pair, sitting in the middle of the oligomer, is as compact as the T:dA base pairs and can adopt a 'wobble' structure with the thymine projecting into the major groove and the guanine into the minor groove. In the wobble form both T:dGTP and C:dATP base pairs do not greatly alter the C1′–C1′ distance. If these forms occur in the closed complex, Arg660 might come into close proximity to the primer phosphate and function as a stabilizing hook. Substitution of Arg by Ser destabilizes the closed complex and that may lead to 13- and 28-fold better discrimination of T:dGTP and C:dATP base pairs than for the wild-type (Table 3). Thus our data are consistent with both T:dGTP and C:dATP forming wobble base pairs in the catalytic pocket of *Taq* pol I.

On the other hand, A:dATP and G:dATP base pairs may form both bulky *anti–anti* and *anti–syn* base pairs. With these structures, the stabilizing effect of Arg660 could be weaker than wobble forms of T:dGTP and C:dATP for the wild-type. This structural difference may lead to less of an effect of Ser substitution on the misincorporation efficiency for A:dATP and G:dATP base pairs.

Evolution of Arg660

Arg660 is conserved among some other *E.coli* pol I class DNA polymerases (i.e. *E.coli*, *Thermus flavae* and *Streptococcus pneumoniae*; 1). In these organisms the corresponding Arg may contribute to certain types of spontaneous mutations by allowing misincorporation of dG opposite T and 8-OH-dGTP. *Thermus aquaticus* is an organism that is adapted to live at high temperatures and *Taq* pol I is more error-prone because it lacks proofreading exonuclease activity. Thus Arg660 might have been advantageous in the evolution and adaptation of *Taq*.

Among the 13 O-helix residues, we have found three residues that are important for polymerase accuracy (Arg660, Thr664 and Ala661; 8,12) but not for polymerase activity. For fidelity alteration, however, the target residues are not limited to the O-helix but may exist throughout the whole sequence (see for example 23). With robust polymerase activity, some mutations could survive and further mutations accumulate during evolution.

In DNA polymerases, it is important to note that mutations are interactive; a second or third mutation could cancel the strong effect of a first one and may sometimes convert an essential residue to a non-essential one (8). We believe that polymerases have diverged into classes through mutations of these primary and secondary non-essential residues. Thus we suggest a specific relationship between polymerase structure and fidelity, the structures of base pairs and the evolution of DNA polymerases.

ACKNOWLEDGEMENTS

We thank Drs Mariko Tada (Shukutoku University) and Brad Preston (University of Utah) for helpful discussions and Yasutomo Ito (Nagoya University) for constructing a model structure. We are also grateful to Ms Tazuko Tomita for technical assistance. This work was supported by the Yokoyama and Nitto Foundations and Grants-in-Aid from the Ministry of Education, Science, Sports and Culture of Japan.

REFERENCES

1. Braithwaite,D.K. and Ito,J. (1993) Compilation, alignment and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res.*, **21**, 787–802.
2. Steitz,T.A. (1999) DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.*, **274**, 17395–17398.
3. Patel,P.H. and Loeb,L.A. (2000) DNA polymerase active site is highly mutable: evolutionary consequences. *Proc. Natl Acad. Sci. USA*, **97**, 5095–5100.
4. Shinkai,A., Patel,P.H. and Loeb,L.A. (2001) *Escherichia coli* DNA polymerase I: the conserved active site motif A is highly mutable. *J. Biol. Chem.*, **12**, 12.
5. Suzuki,M., Baskin,D., Hood,L. and Loeb,L.A. (1996) Random mutagenesis of *Thermus aquaticus* DNA polymerase I: concordance of immutable sites *in vivo* with the crystal structure. *Proc. Natl Acad. Sci. USA*, **93**, 9670–9675.
6. Joyce,C.M. and Steitz,T.A. (1994) Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.*, **63**, 777–822.
7. Kunkel,T.A. and Bebenek,K. (2000) DNA replication fidelity. *Annu. Rev. Biochem.*, **69**, 497–529.
8. Suzuki,M., Yoshida,S., Adman,E.T., Blank,A. and Loeb,L.A. (2000) *Thermus aquaticus* DNA polymerase I mutants with altered fidelity. Interacting mutations in the O-helix. *J. Biol. Chem.*, **275**, 32728–32735.

9. Shah,A.M., Li,S.X., Anderson,K.S. and Sweasy,J.B. (2001) Y265H mutator mutant of DNA polymerase beta. Proper geometric alignment is critical for fidelity. *J. Biol. Chem.*, **276**, 10824–10831.
10. Tosaka,A., Ogawa,M., Yoshida,S. and Suzuki,M. (2001) O-helix mutant T664P of *Thermus aquaticus* DNA polymerase I. Altered catalytic properties for incorporation of incorrect nucleotides but not correct nucleotides. *J. Biol. Chem.*, **276**, 27562–27567.
11. Li,Y., Mitaxov,V. and Waksman,G. (1999) Structure-based design of *Taq* DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc. Natl Acad. Sci. USA*, **96**, 9491–9496.
12. Suzuki,M., Avicola,A.K., Hood,L. and Loeb,L.A. (1997) Low fidelity mutants in the O-helix of *Thermus aquaticus* DNA polymerase I. *J. Biol. Chem.*, **272**, 11228–11235.
13. 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.
14. 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.
15. Huang,M.M., Arnheim,N. and Goodman,M.F. (1992) Extension of base mispairs by *Taq* DNA polymerase: implications for single nucleotide discrimination in PCR. *Nucleic Acids Res.*, **20**, 4567–4573.
16. Bebenek,K. and Kunkel,T.A. (1995) Analyzing fidelity of DNA polymerases. *Methods Enzymol.*, **262**, 217–232.
17. Joyce,C.M., Sun,X.C. and Grindley,N.D. (1992) Reactions at the polymerase active site that contribute to the fidelity of *Escherichia coli* DNA polymerase I (Klenow fragment). *J. Biol. Chem.*, **267**, 24485–24500.
18. Maki,H. and Sekiguchi,M. (1992) MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature*, **355**, 273–275.
19. 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.
20. Oda,Y., Uesugi,S., Ikehara,M., Nishimura,S., Kawase,Y., Ishikawa,H., Inoue,H. and Ohtsuka,E. (1991) NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Res.*, **19**, 1407–1412.
21. Allawi,H.T. and SantaLucia,J.,Jr (1998) NMR solution structure of a DNA dodecamer containing single G:T mismatches. *Nucleic Acids Res.*, **26**, 4925–4934.
22. Hunter,W.N., Brown,T., Kneale,G., Anand,N.N., Rabinovich,D. and Kennard,O. (1987) The structure of guanosine-thymidine mismatches in B-DNA at 2.5-Å resolution. *J. Biol. Chem.*, **262**, 9962–9970.
23. Patel,P.H., Kawate,H., Adman,E., Ashbach,M. and Loeb,L.A. (2001) A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. *J. Biol. Chem.*, **276**, 5044–5051.
24. Bernstein,F.C., Koetzle,T.F., Williams,G.J., Meyer,E.,Jr, Brice,M.D., Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur. J. Biochem.*, **80**, 319–324.