

# DNA Polymerase Chain Reaction: A Model of Error Frequencies and Extension Rates

**Mark Griep**

Dept. of Chemistry, University of Nebraska, Lincoln, NE 68588

**Scott Whitney and Michael Nelson**

Megabase Research Products, Lincoln, NE 68504

**Hendrik Viljoen**

Dept. of Chemical Engineering, University of Nebraska, Lincoln, NE 68588

DOI 10.1002/aic.10604

Published online September 13, 2005 in Wiley InterScience (www.interscience.wiley.com).

*The polymerase chain reaction (PCR) is one of the most important reactions in molecular biology. The detailed mechanistic studies of the polymerase chain reaction have revealed a complex sequence of reversible reactions that involve intermediaries and activated complexes. The DNA polymerase does not merely facilitate the insertion of dNMP (deoxynucleotide monophosphates), but it also performs rapid screening of substrates to ensure a high degree of fidelity. The main result of this study is an expression for the average extension rate of the enzyme. The model is versatile and additional complexities, such as the type of nucleotide to be inserted, the GC content of the sequence in the vicinity of the insertion site, and the topology of the template, such as kinks and hairpins, are easy to incorporate. The insertion of incorrect nucleotides into the sequence is also addressed. Expressions to predict error frequencies are presented. It is shown that a relation exists between error frequency and extension rate: the error frequency is a minimum when the extension rate is optimal. © 2005 American Institute of Chemical Engineers AICHE J, 52: 384–392, 2006*

*Keywords: molecular biology, polymerase chain reaction, mathematical model*

## Introduction

The polymerase chain reaction (PCR) is arguably one of the most important reactions in biochemistry. Multiple copies of a DNA template can be made with high accuracy. Polymerases are the family of enzymes that facilitate the template-copying process. When a double-stranded DNA template is heated to its melting point, the two strands come apart and the template is in a single-stranded form. In the presence of a buffer solution that also contains deoxynucleotide triphosphates (dNTPs) and re-

verse and forward primers, the polymerase catalyzes the addition of deoxynucleotide monophosphates (and release of pyrophosphate) to the single-stranded template to form a new double-stranded DNA molecule. The copying process is better known as the elongation or extension step in PCR. The enzyme attaches to the single-stranded template, deftly manipulates incoming dNTPs to determine whether it complements the sequence (A–T, C–G pairs), and inserts the nucleotide if it is correct and translocates to the next position or it rejects the nucleotide if it is incorrect. On occasion an erroneous nucleotide is inserted and certain polymerases have the ability to detect and correct the mistakes (proofreading, exonuclease). If the temperature increases, the polymerase has the tendency to detach from the template, at which point the elongation is

Correspondence concerning this article should be addressed to H. Viljoen at hviljoen1@unl.edu.

stalled until the polymerase has reattached to the template. Both the processing of incorrect nucleotides and the detachment of the polymerase from the template contribute to slowing down the elongation rate. In this study a model that describes the function of the polymerase (with the exception of the exonuclease step) is presented. A competition exists between the delay processes and successful insertions.

Template-directed nucleic acid synthesis is one of the greatest biochemical discoveries. Since the activity of the first RNA polymerase and DNA polymerase were studied in 1955,<sup>1,2</sup> template-directed polymerases have contributed to our understanding of many subcellular processes, such as DNA replication and repair, transcription, and telomere homeostasis. The accuracy, relative ability to bypass lesions, and ability to elongate polymers processively makes these enzymes crucial to understanding such biological processes as infection, cancer, and aging and also identifies them as important drug targets.<sup>3,4</sup> Finally, template-directed nucleic acid polymerases are used in many procedures basic to modern biotechnology including PCR and DNA sequencing.<sup>5,6</sup>

The microscopic kinetic mechanisms and structure/function relationships of *Escherichia coli* DNA polymerase I, T7 DNA polymerase, KlenTaq DNA polymerase, and a *Bacillus* DNA polymerase have been worked out in some detail.<sup>7-9</sup> Comparative analysis suggests that all DNA polymerases will have the same broadly conserved molecular properties and mechanisms, but that each polymerase will have its unique set of microscopic rates and equilibria.<sup>10-12</sup> The mechanisms of nucleotide insertion and incorporation are probably highly conserved, although the details of the nucleotide selection process are just beginning to be understood. Although the studies have significantly increased our understanding of the detailed catalytic pathway, there is a need for a macroscopic kinetic model that captures the essential features of the reaction.

Despite the advances on the molecular level, we still lack an understanding of the relationship of these parameters to macroscopic properties such as chromosomal elongation rates. This lack of theoretical understanding has practical consequences. For instance, it is one of the obstacles that prevent users from selecting the best DNA polymerase for a given PCR application. Currently researchers choose a polymerase, work out the PCR conditions semiempirically, and then adjust if necessary. Our poor theoretical understanding is also an obstacle for those researchers trying to solve the many PCR problem areas such as the amplification of GC-rich sequences, amplification of long sequences, and maintaining low error frequencies. There is a commercial impact in that many thermophilic DNA polymerases are available<sup>13</sup> with various known biochemical properties (elongation rate, processivity, nucleotide selectivity, thermal half-life at 95°C, and so on) but it is not clear which properties are best for which PCR protocol. To bridge the gap between the microscopic and macroscopic in the realm of DNA polymerases, we have developed a model for DNA synthesis that accounts for the template nucleotide sequence and dNTP pool to predict DNA synthesis rates and error frequencies. The derivation of the macroscopic model is based on a stochastic approach.<sup>14</sup> The behavior of a single enzyme/template complex is tracked over time to determine its average behavior. Based on an ergodic argument, there is equivalence between the average behavior of a single complex over a long period of time and the average behavior of many complexes at any

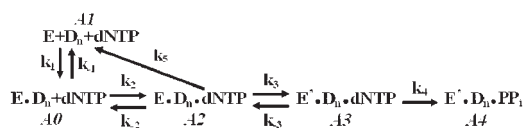
specific moment in time. The macroscopic model enables us to compare the performance of different polymerases, investigate the roles of temperature and pool compositions, and optimize PCR experiments.

An important result of the analysis is the role that the dNTP pool plays. Imbalances of the intracellular dNTP pool are linked with enhanced rates of biological mutagenesis, cancer,<sup>15</sup> and genetic diseases. To begin with, deoxyguanosine 5'-triphosphate (dGTP) is naturally underrepresented in eukaryotic cells and its molar fraction typically accounts for 5–10% of the total pool (that is, the molar fraction of dGTP,  $X_G = 0.05$  to 0.10).<sup>16</sup> Biochemical and cellular experiments show that eukaryotic polymerases are slightly but significantly more mutagenic in these naturally imbalanced dNTP pools than in equimolar pools.<sup>17</sup> This is exactly the modest outcome predicted by our model. When the pool is biased beyond the natural, it should cause substantial increases in the mutagenic rates. In fact, a wide variety of external agents do cause severe changes to cytosolic, nuclear, and mitochondrial dNTP pool compositions and increases to cellular mutation rates.

Most changes that affect the nucleotide biosynthetic or salvaging pathways lead to dNTP pool imbalances such as by heat-induced cytosine deamination,<sup>18</sup> by ultraviolet light and cytosine arabinoside,<sup>19,20</sup> by fluorouracil,<sup>21</sup> by chlorodeoxyadenosine,<sup>22</sup> by changes to the single-carbon transfer pathway through dietary folate or methotrexate treatment,<sup>23-25</sup> by the antineoplastic tropoline,<sup>26</sup> by difluorodeoxycytidine,<sup>27</sup> by hypoxia,<sup>28</sup> and by ribonuclease reductase inhibitors such as hydroxyurea. The exceptions seem to be the nucleoside analog drugs azidothymidine and azidocytidine, which do not cause changes to the dNTP pool, even though they inhibit intracellular DNA synthesis.<sup>29,30</sup> Finally, mitochondrial DNA deletion syndrome (MDS) is a rare disease caused by mutations to a variety of genes. Specifically, mutations in thymidine phosphorylase<sup>31</sup> or thymidine kinase 2 (TK2) have been associated with different forms of MDS.<sup>32,33</sup> Results demonstrate that dNTP pool imbalances have to increase or decrease substantially before any effects will be observed on DNA polymerase elongation and error rates. The cited biological results indicate that such drastic changes to the dNTP pools do occur when cells are treated with many nucleoside analogs or when the nucleoside biosynthetic or salvaging pathways are mutated or altered in some way.

## Mathematical Model

The formulation of a macroscopic model that describes the average dynamics of a polymerase/template complex is based on the details of a kinetic pathway. In Figure 1 a kinetic scheme is presented that is based on the mechanism that Johnson<sup>7</sup> proposed for DNA propagation by T7 polymerase. There is a small modification to the original scheme of Johnson, notably the detachment of the polymerase from the state marked **A2**. The template DNA is designated as  $D_n$  to indicate that the extension has progressed to position  $n$  and the insertion of the  $n + 1$  nucleotide is considered. The polymerase is  $E$  and the deoxynucleotide triphosphate is  $dNTP$ . The incorporation of the deoxynucleotide monophosphate into the sequence leads to the release of the pyrophosphate ( $PP_i$ ). For the purpose of the mathematical derivation, we assign a symbol to each state of the enzyme. The state before the enzyme binds to the primer/



**Figure 1. Polymerase-catalyzed DNA elongation.**

**A1** presents the primer/template complex, **A0** is the ternary complex that includes the polymerase, and **A2** is the quaternary complex that forms after addition of a dNTP. The model accounts for polymerase detachment from states **A0** and **A2**. The quaternary complex adopts an activated state **A3** followed by phosphodiester bond synthesis, the release of pyrophosphate, and translocation of the enzyme. The schematic also identifies the notation to label the different states and the respective rate constants.

template complex is denoted as **A1**. The complex binds a polymerase molecule to form a ternary complex **A0**. The dNTP pool is composed of dATP, dCTP, dGTP, dTTP, and dUTP. Uracil is formed by thermal deamination of dCTP, and thus it must be included in the pool. After one of these components has bound to the ternary complex to form the quaternary complex **A2**, the polymerase may undergo a conformational change denoted as **A3**. The final step to **A4** involves the release of  $PP_i$ . Patel et al.<sup>34</sup> noted that the rate-limiting step is the formation of the activated complex **A3**. The state **A4** is a ternary complex similar to **A0**, but the number of base pairs has increased by one. A reaction has been added to the original scheme of Johnson to account for the detachment of the polymerase from the quaternary state, that is,  $A2 \rightarrow A1$ . (Because of the paucity of kinetic data, it is reasonable to assume that  $k_5$  and  $k_{-1}$  would be similar.)

### Reaction rates and error frequencies

Starting at the state **A1** of the reaction scheme in Figure 1, the binary complex (formed during the annealing step) binds the polymerase with rate  $k_1[E]$  to form a ternary complex, identified as state **A0**. The free polymerase concentration is denoted as  $[E]$ . The polymerase may detach from the ternary complex at a rate  $k_{-1}$  and return to state **A1**, or it may bind a nucleotide with rate  $k_2[X]$  and proceed to state **A2**. The term  $[X]$  presents the sum of all species that may attach to the ternary complex. No selection of dNTP precedes the step. Therefore the random nature of thermal motion in the solution ensures that the probability of the nucleotide being of type  $i$  (where  $i = A, C, T, G$ ) is determined by the molar fraction of each of those species. For example, the molar fraction of dATP is  $X_A = [dATP]/[X]$ . Suppose A is the correct nucleotide; then the probability that a dATP binds to the ternary complex is  $X_A$  and the probability that it is not dATP is  $1 - X_A$ . If the correct nucleotide is bound to the ternary complex, the kinetics favors insertion. If the incorrect nucleotide has been bound, insertion may still occur, but it is not kinetically favored.

At this point we make an important distinction. The forward rate constants  $k_3$  and  $k_4$  change to lower values  $k_{3w}$  and  $k_{4w}$  if a wrong nucleotide is selected.<sup>35,36</sup> Note, however, that the nucleotides still bind to the ternary complex with the same rate constant  $k_2$ .

**Remark.** Apart from the nucleotides, dUTP (which forms by thermal deamination of dCTP) and  $PP_i$  (a product of extension) may compete for attachment. Competitive binding by dUTP or  $PP_i$  is reduced by an excess dNTP, as well as the use

of enzymes such as dUTPase and pyrophosphatase. In the absence of competitive binding by dUTP or  $PP_i$ ,  $[X] = [dNTP]$ .

To progress from state **A0** to state **A4** means that one nucleotide has been added, although several events may delay the progress:

(1) The polymerase may detach when the complex is in state **A0** or in state **A2** (reactions  $A0 \rightarrow A1$  and  $A2 \rightarrow A1$  in Figure 1).

(2) The polymerase may reject the nucleotide in state **A2** (even if it is correct) and the reverse reaction  $A2 \rightarrow A0$  occurs or it may proceed to the activated complex **A3**.

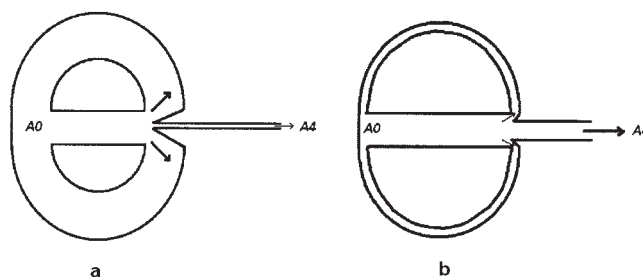
(3) The nucleotide fails to insert and the activated complex may return to the quaternary complex  $A3 \rightarrow A2$ .

Thus the advancement of a single enzyme/primer/template complex must be viewed as a probabilistic affair that consists of the series of events listed above. Figure 2 offers a schematic equivalent of the mathematical model. The initial state **A0** is marked on the left of the figures. The time-wasting loops are presented as two lanes that return to state **A0** and the center lane on the right connects to state **A4**. If an incorrect nucleotide binds to the insertion site, there is a high probability that it may return to state **A0** by one of the time-wasting loops and in Figure 2a the broad return lanes indicate the likelihood of this occurrence; moreover, the probability of an incorrect insertion is slim, as indicated by the narrow center lane in Figure 2a. In the case of a correct nucleotide, there is a much higher probability that state **A4** is reached; note that the center lane on the right of Figure 2b is much broader, but the time-wasting loops are narrow, to mark the small probability that the correct nucleotide may follow one of them. Although the effect of temperature is not shown in Figure 2, the return lanes would become broader, to present the higher probability of polymerase detachment.

A successful event **S** starts at **A0** and ends at **A4**, but it may idle between states **A2** and **A3**. Therefore **S** can be presented as follows:  $A0 \rightarrow (A2 \leftrightarrow A3)^n \rightarrow A4$ , where  $n$  = number of cycles between **A2** and **A3**.

The probabilities to proceed from **A0** to **A2**, **A2** to **A3**, **A3** to **A2**, and **A3** to **A4** are

$$P_{02} = \frac{k_2[X]}{k_2[X] + k_{-1}} \quad (1a)$$



**Figure 2. Passage behavior for: (a) an incorrect nucleotide (the pathways that return the enzyme to state **A0** are more favored than the pathway to insertion); (b) a correct nucleotide has a high probability to reach **A4** and it seldom participates in one of the return loops to state **A0**.**

$$P_{23} = \frac{k_3}{k_3 + k_{-2} + k_{-1}} \quad (1b)$$

$$P_{32} = \frac{k_{-3}}{k_4 + k_{-3}} \quad (1c)$$

$$P_{34} = 1 - P_{32} = \frac{k_4}{k_4 + k_{-3}} \quad (1d)$$

Note that in Eqs. 1a and 1b we have used  $k_{-1}$  instead of  $k_5$ , because the two rates are expected to be close. If data become available,  $k_5$  should be used instead.

If an incorrect nucleotide has attached to the ternary complex, the probabilities are designated by the subscript  $w$  (for wrong) as follows

$$P_{23w} = \frac{k_{3w}}{k_{3w} + k_{-2} + k_{-1}} \quad (2a)$$

$$P_{32w} = \frac{k_{-3}}{k_{4w} + k_{-3}} \quad (2b)$$

$$P_{34w} = 1 - P_{32w} = \frac{k_{4w}}{k_{4w} + k_{-3}} \quad (2c)$$

The probability to insert a correct nucleotide includes the idling

$$P_S = P_{02}P_{23}P_{34} + P_{02}(P_{23}P_{32})P_{23}P_{34} + \dots + P_{02}(P_{23}P_{32})^n P_{23}P_{34} \quad n \rightarrow \infty \quad (3)$$

$$P_S = P_{02}P_{23}P_{34}/[1 - P_{23}P_{32}] \quad (4)$$

If  $P_{23}$ ,  $P_{32}$ , and  $P_{34}$  in Eq. 4 are replaced by  $P_{23w}$ ,  $P_{32w}$ , and  $P_{34w}$  the probability to insert an incorrect nucleotide is

$$P_{Sw} = P_{02}P_{23w}P_{34w}/[1 - P_{23w}P_{32w}] \quad (5)$$

The second event is detachment (denoted as **D1**) from state **A2**. The same idling may occur as before:  $\mathbf{A0} \rightarrow (\mathbf{A2} \Leftrightarrow \mathbf{A3})^n \rightarrow \mathbf{A4} \rightarrow \mathbf{A0}$ .

The probability for event **D1** is

$$P_{D1} = P_{02}P_{21}/[1 - P_{23}P_{32}] \quad (6)$$

*Remark.* The probability to go from **A1** to **A0** is one and is therefore not explicitly included in Eq. 6. If the nucleotide is incorrect, the probability is

$$P_{D1w} = P_{02}P_{21w}/[1 - P_{23w}P_{32w}] \quad (7)$$

and

$$P_{21} = \frac{k_{-1}}{k_3 + k_{-2} + k_{-1}} \quad P_{21w} = \frac{k_{-1}}{k_{3w} + k_{-2} + k_{-1}}$$

The third event (**F**) is rejection of the nucleotide in state **A2** and return to state **A0**. Idling must be considered and the probabilities for correct and incorrect nucleotides become

$$P_F = P_{02}P_{20}/[1 - P_{23}P_{32}] \quad (8)$$

$$P_{Fw} = P_{02}P_{20w}/[1 - P_{23w}P_{32w}] \quad (9)$$

The last event (**D2**) is detachment from state **A0** and the loop is  $\mathbf{A0} \rightarrow \mathbf{A1} \rightarrow \mathbf{A0}$ . The probability for this event to occur is

$$P_{D2} = P_{01} = \frac{k_{-1}}{k_2[X] + k_{-1}} \quad (10)$$

Associated with each probability is its characteristic time. Again we need to distinguish between correct and incorrect nucleotides. The times associated with  $P_{01}$ ,  $P_{02}$ ,  $P_{20}$ ,  $P_{21}$ ,  $P_{23}$ ,  $P_{32}$ , and  $P_{34}$  are

$$t_{01} = t_{02} = \frac{1}{k_2[X] + k_{-1}} \quad (11a)$$

$$t_{20} = t_{21} = t_{23} = \frac{1}{k_3 + k_{-2} + k_{-1}} \quad (11b)$$

$$t_{32} = t_{34} = \frac{1}{k_4 + k_{-3}} \quad (11c)$$

Similarly we define  $t_{20w}$ ,  $t_{21w}$ ,  $t_{23w}$ ,  $t_{32w}$ , and  $t_{34w}$  by replacing  $k_{3,4}$  with  $k_{3w,4w}$  in the above equations.

The time for a successful event is

$$\begin{aligned} \tau_S = & P_{02}P_{23}P_{34}[t_{02} + t_{23} + t_{34}] + P_{02}(P_{23}P_{32})P_{23}P_{34}[t_{02} \\ & + (t_{23} + t_{32}) + t_{23} + t_{34}] + \dots + P_{02}(P_{23}P_{32})^n P_{23}P_{34} \\ & \times [t_{02} + n(t_{23} + t_{32}) + t_{23} + t_{34}] = P_S[t_{02} + t_{23} + t_{34}] \\ & + P_S P_{23}P_{32}[t_{23} + t_{32}]/(1 - P_{23}P_{32}) \end{aligned} \quad (12)$$

The times for the other events are

$$\tau_{D1} = P_{D1}[t_{02} + t_{21} + t_{10}] + P_{D1}P_{23}P_{32}[t_{23} + t_{32}]/(1 - P_{23}P_{32}) \quad (13a)$$

$$\tau_F = P_F[t_{02} + t_{20}] + P_F P_{23}P_{32}[t_{23} + t_{32}]/(1 - P_{23}P_{32}) \quad (13b)$$

$$\tau_{D2} = P_{01}[t_{01} + t_{10}] \quad (13c)$$

In Eqs. 13a and 13c,  $t_{10} = 1/k_1[E]$  is the average time it takes the binary complex (template/primer) to bind a polymerase. The average passage time  $\tau(\mathbf{A0} \rightarrow \mathbf{A0}, \mathbf{A4})$  is the expectancy value of all possible events:

$$\begin{aligned}\tau &= \tau_S + \tau_{DI} + \tau_F + \tau_{D2} \\ &= t_{02} + P_{01}t_{10} + \frac{P_{02}[t_{20} + P_{23}t_{34} + P_{21}t_{10}]}{1 - P_{23}P_{32}}\end{aligned}\quad (14)$$

The average passage time of an incorrect nucleotide is  $\tau_w$ . The template sequence determines the correct nucleotide and it is realistic to expect that the rate constants of different nucleotide types differ. Theoretically speaking we should distinguish between passage times for different nucleotide types:  $\tau_A$ ,  $\tau_C$ ,  $\tau_T$ ,  $\tau_G$  and  $\tau_{wA}$ ,  $\tau_{wC}$ ,  $\tau_{wT}$ ,  $\tau_{wG}$ , but in practice the paucity of kinetic data rules out such a distinction.

If we track a primer/template complex over a large number of passage times and observe that the template extends with  $N_A + N_T + N_C + N_G$ , each nucleotide type requires on the average

$$\frac{N_i}{X_i P_S + (1 - X_i) P_{Sw}}$$

passage times ( $i = A, C, T, G$ ). The number of passage times associated with correct/incorrect nucleotide is, respectively

$$\frac{X_i N_i}{X_i P_S + (1 - X_i) P_{Sw}} \quad \text{and} \quad \frac{[1 - X_i] N_i}{X_i P_S + (1 - X_i) P_{Sw}}$$

Therefore the time to insert  $N_i$  nucleotides is

$$\frac{X_i N_i \tau + [1 - X_i] N_i \tau_w}{X_i P_S + (1 - X_i) P_{Sw}}$$

Thus the extension of the template at site  $i$  occurs at the rate

$$\nu_i = \frac{X_i P_S + (1 - X_i) P_{Sw}}{X_i \tau + (1 - X_i) \tau_w} \quad (15)$$

This is a local rate at a point in the sequence and  $1/\nu_i$  is the average time to insert nucleotide  $i$ . The average rate for a template with composition  $N = N_A + N_T + N_C + N_G$  is

$$\nu_{ave} = \frac{N}{\sum_{i=A,C,T,G} \frac{N_i [X_i \tau + (1 - X_i) \tau_w]}{X_i P_S + (1 - X_i) P_{Sw}}} \quad (16)$$

An explicit expression of the error frequency, expressed as errors/ $10^6$  base pairs, is

$$F = \left\{ \sum_{i=A,C,T,G} \frac{N_i (1 - X_i) P_{Sw}}{X_i P_S + (1 - X_i) P_{Sw}} \right\} / (10^{-6} \times N) \quad (17)$$

Values of  $F$  vary from 8 for *Taq pol.*,<sup>37</sup> 3 for KOD,<sup>38</sup> and 1 for Pfu.<sup>37</sup> The four parameters  $\tau$ ,  $\tau_i$ ,  $P_S$ , and  $P_{Sw}$  in Eqs. 16 and 17 are functions of the rate constants in Figure 1. The physical interpretation of the four parameters is as follows. The parameters  $\tau$  and  $\tau_i$  are the average passage times for a correct ( $\tau$ ) or an incorrect nucleotide ( $\tau_i$ ). An enzyme spends its time in

different states. The passage time is the average time from the moment an enzyme leaves state **A0**, until it returns to either **A0** or **A4**. In the case of a correct nucleotide, it is most likely that the enzyme ends up in state **A4** because  $P_S$  is close to unity. If an incorrect nucleotide is selected, the most likely outcome is a return of the enzyme to state **A0**. The probability to insert an incorrect nucleotide is  $P_{Sw} \ll 1$ . Thus the incorrect nucleotides participate with higher probability in the time-wasting loops **A0-A1-A0**, **A0-A2-(A3-A2)<sup>n</sup>-A0** and **A0-A2-(A3-A2)<sup>n</sup>-A1-A0** ( $n = 0, 1, 2, \dots$ ).

### Refinements of the model

It was pointed out in the discussion following Eq. 14 that the local extension rate (alternatively the local insertion time) depends on the type of nucleotide. It has been experimentally observed<sup>39</sup> that GC-rich sequences copy slower, in other words, we expect that  $\tau_G > \tau_A, \tau_T, \tau_C$ . A refinement on that idea is to include the nearest neighbors, next-nearest neighbors, and subsequently the sequence in the vicinity of the insertion site into the kinetic model. Kinetic rates are affected by the local sequence. The problem with the implementation of more sophisticated models is paucity of data. However, the ideas may be explored on a phenomenological level. Therefore we propose that the rate constant  $k_3$  is nucleotide dependent, thus  $k_3(i)$ . The model accounts for the dependency of rate on the four base pairs that immediately precede the insertion site. At position  $j$  in the sequence nucleotide type  $i$  must be inserted. Check the four base pairs before position  $j$  and set their number of G-C pairs equal to  $N_G$ . The value of  $N_G$  can vary between 0 and 4. The phenomenological model prescribes that the rate constant is lowered by 10% if a  $G$  must be inserted and by a further 10% for each additional GC pair in the previous four base pairs. When a GC-rich region is encountered, the rate could be reduced by as much as 50%

$$k_3(i, j) = \left\{ \begin{array}{l} k_3, i \neq G \\ 0.8k_3, i = G \end{array} \right\} - 0.1k_3 \times N_G \quad (18)$$

where  $i$  is nucleotide type and  $j$  is sequence position.

The model is tested for two computer-generated sequences of a template of 1020 base pairs. A random generator assigns G, C, T, or A values to the template based on probabilities. The first sequence (S1) has equal probability for any nucleotide type. A second sequence (S2) is generated, but the GC content is 70%. The average time to insert a nucleotide of type  $i$  is calculated from Eq. 15 and Eq. 18 has been used to evaluate  $k_3$ . The parameter values are listed in Table 1. It is assumed that the primers are 20 base pairs long, and thus copying starts at site 21. In Figure 3 the extension rate for template S1 is shown for the region  $301 \leq j \leq 400$ . The extension rate is stochastic because of the local interrogation of the sequence by the enzyme. The average rate is 90 nucleotides/s.

When the calculation is repeated for template S2, the extension is slower (Figure 4). In this case the average rate drops to 86 nucleotides/s. The increase in GC content has led to only a small decrease in the overall elongation rate. Keep in mind that this is a phenomenological study and the GC effect was included only in  $k_3$ . A comparison of Figure 3 and Figure 4 shows that the local dynamics have changed. The polymerase rates drop to 51 nucleotides/s in sequences that are GC rich.



Table 1. Parameter Values

	Parameter											
	$k_4$	$k_{4w}$	$k_3$	$k_{3w}$	$k_{-3}$	$k_1$	$k_{-1}$	$k_2$	$k_{-2}$	$k_5$	[dNTP]	[E]
Value	3000	0.1	1000	0.01	1	40	0.2	4	$k_2[X]/1.6$	0.2	800 $\mu\text{M}$	$\mu\text{M}$

Besides  $k_3$ , there may be other parameters that are also affected by sequence composition, but at this stage the relations are not known.

It is further known that polymerases of the *Pyrococcus* family have the ability to assess the sequence ahead of the insertion site and they detach if a U is detected.<sup>40-43</sup> To describe this property, the sequence is not only interrogated at a few positions before  $j$ , but also a few positions ahead of the insertion site. If a U is detected,  $k_3 = 0$  and the polymerase effectively detaches from the bound state. This property can be easily added to the model.

A further sophistication of the model is to consider the topology of the template; ssDNA may fold up on itself and kinks and hairpins may form that influence the extension rate. Again it is lack of data that prevents a more quantitative description. The only challenge is to relate the template sequence to the topology complexity and express the topology complexity in terms of modified kinetic parameters.

### Estimates of Elongation Rates and Error Frequencies with Varying dNTP Pools

The composition of the dNTP pool plays an important role in the overall kinetics. Not only does it influence the extension rate, but it also affects the error frequencies. Changes in  $[X]$  will lead to changes in the probabilities and passage times. Template S1 is composed of equal amounts of all four nucleotides and it is expected that an equimolar dNTP pool will perform best. To investigate the effect of varying dNTP pools, the total dNTP concentration is fixed at 800  $\mu\text{M}$  and [dGTP] is varied from 20 to 740  $\mu\text{M}$  with the balance equally distributed among the other three dNTPs. The elongation rate is calculated as a function of  $X_G$ , the molar fraction of dGTP in the pool. As expected, the elongation rate is a maximum when  $X_G$  is 0.25 (Figure 5). The rate decreases sharply when the molar fraction is lowered, but the decrease in rate is less pronounced when  $X_G$  is increased. This finding is logical. If  $X_G$  is reduced to 0.01

(that is, 0.25 – 0.24), there is a 1% chance that dGTP presents at the binding site. If  $X_G$  is (0.25 + 0.24), the molar fraction of each one of the rest is  $(1 - 0.49)/3 = 0.17$ , which means a reduction from 25 to 17% to be present at the binding site. Clearly an increase in  $X_G$  does not exhibit the same sensitivity as a reduction in its value. Of course, the reaction is inhibited when  $X_G$  approaches one.

In Figure 5 the error frequency is shown as a function of  $X_G$ . The molar fraction  $X_G$  is varied between 0.025 and 0.925. The parameters are listed in Table 1. The parameter  $k_3$  has been kept constant in this case. The lowest error frequency of  $4.09/10^6$  corresponds to equimolar conditions,  $X_G = 0.25$ . The extreme values are  $15.70/10^6$  at  $X_G = 0.025$  and  $39.60/10^6$  at  $X_G = 0.925$ . At equimolar conditions, the rate is a maximum and the complete extension of the template requires the least number of passage times. The results show that the error frequency is a minimum when the number of passage times is a minimum. Error frequencies are relatively insensitive to changes in the molar fraction over a wide range from about 0.1 to 0.7 mole fraction of dGTP.

When the GC-rich template S2 is used (35% G) the maximum elongation rate occurs at  $X_G = 0.300$  and this molar fraction coincides with the minimum error frequency of  $4.05/10^6$  (Figure 6). However, the optimum value  $X_G = 0.300$  does not correspond to the nucleotide composition of the template, which is 35%. Although the optimum molar fraction is more than the equimolar value of 25%, an increase to  $X_G = 0.35$  impedes the extension rate of the other three nucleotides. Interestingly, the values of the error frequencies of S1 and S2 at their optimum compositions are the same, but at  $X_G = 0.925$  the error rate of S1 ( $39.60/10^6$ ) exceeds the error rate of S2 ( $35.20/10^6$ ).

### Implementation and Practical Applications

The implementation of Eq. 16 is straightforward, especially if we do not include secondary factors that influence the kinet-

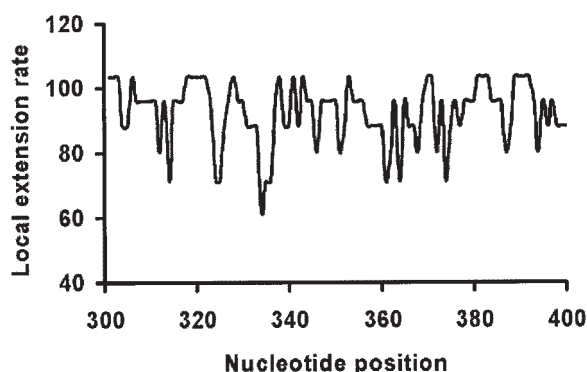


Figure 3. Local extension rate vs. sequence S1 (50% G + C).

The part of S1 between 300 and 400 is shown.

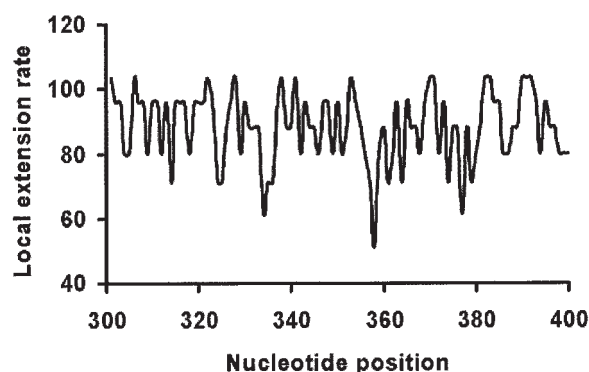


Figure 4. Local extension rate vs. sequence S2 (70% G + C).

The part of S2 between 300 and 400 is shown.

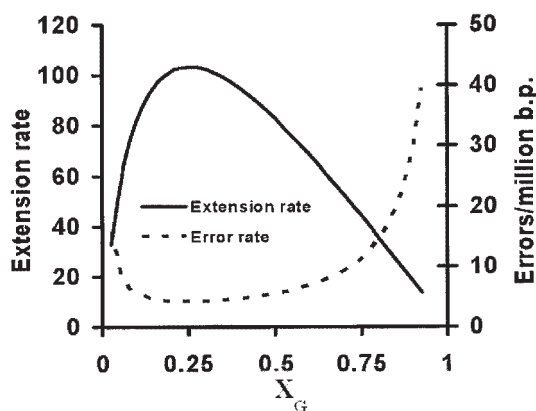


Figure 5. Elongation rate and error frequency vs.  $X_G$  for sequence S1.

ics as discussed in earlier. The inverse of Eq. 16 may be written as follows

$$\tau_{ave} = \frac{1}{\nu_{ave}} = \frac{1}{N} \sum_{i=A,C,T,G} \frac{N_i [x_i \tau/P_S + (1 - x_i) \tau_w/P_S]}{x_i + (1 - x_i) P_{Sw}/P_S} \quad (19)$$

The denominator contains the term  $P_{Sw}/P_S$  and it is expected that  $P_{Sw}/P_S \ll 1$  because polymerases are efficient. The argument can be made more quantitatively if one solves Eq. 17 for  $P_{Sw}/P_S$ . For example, in the case of an equimolar composition of dNTPs and an equal composition of A, C, T, and G in the sequence, the value for KOD Pol is

$$\frac{P_{Sw}}{P_S} = \frac{F}{3 \times 10^6 - 3F} = \frac{3}{3 \times 10^6 - 9} \approx 10^{-6}$$

Under most circumstances the denominator in Eq. 19 is approximately  $x_i$ —an exception occurs if an extreme dNTP pool composition  $x_i \ll 1$  is used. Therefore it is argued that Eq. 19 may be simplified as follows

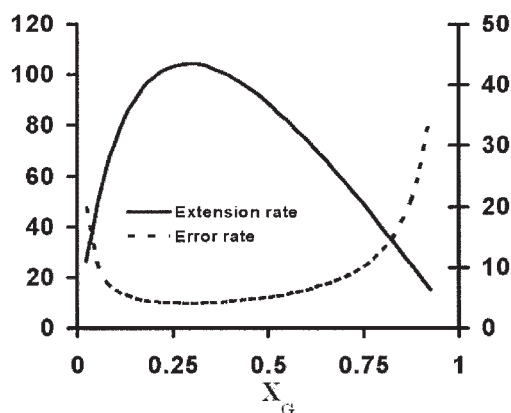


Figure 6. Elongation rate and error frequency vs.  $X_G$  for sequence S2.

Table 2. Extension Times for Balanced and Skew dNTP Pools at 68 and 78°C

Experiment	Elongation Temperature (°C)	Molar Fraction $G$	Elongation Time
1	68	0.25	2.5
2	78	0.25	3.0
3	68	0.82	7.5
4	78	0.82	11.0

$$\tau_{ave} \approx \frac{\tau}{P_S} + \left( \sum_{i=A,C,T,G} \frac{N_i(1 - x_i)}{Nx_i} \right) \frac{\tau_w}{P_S} \quad (20)$$

Equation 20 contains two parameters:  $\Gamma = \tau/P_S$  and  $\Gamma_w = \tau_w/P_S$ . The parameters are unique to the type of polymerase that is used for the enzymatic extension of a template. The parameters are also influenced by template-specific factors, such as structural complexities and sequence composition.

The experiments to determine  $\Gamma$  and  $\Gamma_w$  are designed in the following manner. A template with a known sequence is PCR amplified at an elongation temperature  $T_1$  with an equimolar dNTP pool. The experiment is repeated with progressively shorter elongation times until the cutoff is reached—the cutoff value is the minimum elongation time at temperature  $T_1$ . The experiments are repeated with a skewed dNTP pool, but the same elongation temperature is used. The experiments are repeated at a different elongation temperature  $T_2$  to obtain four extension times.

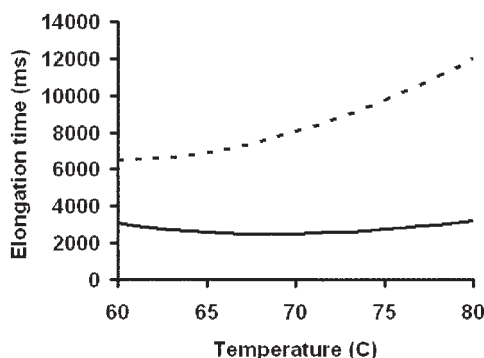
To illustrate the application of the measurements, let us assign numbers to the extension times, listed in Table 2. It must be pointed out that the data of Table 2 have been obtained from preliminary experiments and serve the purpose only to demonstrate how the key kinetic parameters can be obtained from such measurements. In a future contribution validated experimental data will be reported. The template is *B. cereus* and the length is 600 base pairs. Primer lengths of 12 base pairs are used and therefore the effective extension length is 588 base pairs, and thus  $N = 588$  and  $N_{A,C,T,G} = 147$ . The polymerase KOD Pol has been used.

First the extension times of experiments 1 and 3 from Table 2 are substituted into Eq. 20. For  $X_G = X_A = X_T = X_C = 0.25$ , Eq. 19 becomes  $2500 \text{ ms}/588 = \Gamma(68^\circ\text{C}) + 4 \times 147 \times (1 - 0.25)/(588 \times 0.25)\Gamma_w(68^\circ\text{C})$ .

The skewed dNTP composition is  $x_G = 0.82$ ,  $x_A = x_T = x_C = 0.06$ . If the data of experiment 3 are substituted into Eq. 20 one obtains  $7500 \text{ ms}/588 = \Gamma(68^\circ\text{C}) + [147 \times (1 - 0.82)/(588 \times 0.82) + 3 \times 147 \times (1 - 0.06)/(588 \times 0.06)]\Gamma_w(68^\circ\text{C})$ .

The solution at 68°C is  $(\Gamma_w, \Gamma) = (0.97 \text{ ms}, 1.35 \text{ ms})$ . Using the data of experiments 2 and 4, the solution at 78°C is  $(\Gamma_w, \Gamma) = (1.55 \text{ ms}, 0.47 \text{ ms})$ .

If the solutions at 68 and 78°C are compared, one notes that  $\Gamma$  decreases with temperature, but  $\Gamma_w$  increases with temperature. The solutions are used to find the parameters  $(E, E_w)$ . Once the parameters are resolved in the form  $\Gamma = \Gamma_0 e^{-E/RT}$ ;  $\Gamma_w = \Gamma_{w0} e^{-E_w/RT}$ , the effect of temperature on the extension time is determined. In Figure 7 the extension time for *B. cereus* is shown as a function of temperature. The dashed curve corresponds to the skewed dNTP pool as listed in Table 2 and the solid line corresponds to the balanced pool. The interesting result is that the minimum extension time is a function of pool



**Figure 7. Elongation times (in ms) for a balanced dNTP pool (solid line) and a skewed dNTP pool (dash line).**

composition. In retrospect it is an expected result, because the pool composition affects the role of the time-delaying and successful-insertion processes, all of which have different temperature dependencies. The optimum temperature for a balanced pool lies at 68°C. The imbalanced pool has a minimum extension time at 60°C. Quite clearly the delay processes (which have an opposite Arrhenius effect—that is, time increases with temperature) play a dominant role in the case of the skewed pool.

## Conclusions

The rate expression Eq. 15 has been derived for a multisubstrate enzyme-catalyzed DNA extension reaction. A reaction scheme, which is assumed to be generic to most polymerases, has been used for the derivation. If we track the behavior of a single enzyme/template complex over a period of time (let us call this period a cycle), we have complete knowledge about this complex over the span of the cycle. If the “tracking experiment” is repeated over a large number of cycles, we learn more about the average behavior of a single complex. The derivation is based on the equivalence of the average behavior of a single complex over many cycles and the average behavior of many complexes over one cycle.

The main results of the study are:

(1) The average reaction rate and the average time to insert a base pair have been derived. The insertion time consists of the time to insert a correct nucleotide plus delays to edit for incorrect nucleotides and the assembly of enzyme/template complexes after polymerase detachment from the complex.

(2) The insertion times depend on the type of nucleotide to be inserted and the composition of the template in the vicinity of the insertion site. The roles of sequence composition and topology can be included in the model. The effect of sequence composition has been demonstrated in a phenomenological model, although more experimental data are needed for a quantitative analysis.

(3) An expression for error frequencies has been derived. From this expression the error frequency can be easily calculated and once again the limitation is availability of experimental kinetic data.

(4) The optimum extension rate coincides with the minimum error frequency. It can be concluded that once the conditions for maximum extension have been determined, they

also correspond to the conditions where the minimum number of errors will be made.

(5) The composition of the dNTP pool plays a very important role in extension rates and error frequencies. At the optimum pool composition extension occurs with the minimum number of rejections. Normally the optimum composition is close to equimolar. The effect has been demonstrated by varying the [dGTP] concentration. When [dGTP] is lowered with respect to the other dNTP concentrations, the average time to insert dGMP increases. The enzyme spends more time to edit and reject other nucleotides at G-sites. When the [dGTP] concentration is increased with respect to the other dNTP concentrations, the insertion times for the other dNTPs increase because the enzyme spends more time editing and rejecting dGTP from non-G-insertion sites.

(6) Temperature plays an important role in extension rates. An illustrative example has been given to show the opposing effects of successful insertion processes vs. time-delaying processes. The optimum temperature shifts to lower values as the dNTP pool becomes more imbalanced.

(7) The effect of  $k_5$  has also been investigated; as the reader may recall it has been added to the reaction pathway as another way the polymerase may detach. For the conditions of Table 1 (and template S1), the value of  $k_5$  has been varied from 0 to  $10 \times k_{-1}$  and the maximum extension rate has changed accordingly from 103.68 to 100.80 nucleotides/s, which shows only slight variation from the value of 103.38 nucleotides/s if  $k_5 = k_{-1}$ .

(8) The macroscopic model provides the necessary tools for quantitative modeling of PCR.<sup>44</sup>

## Acknowledgments

H.J.V. acknowledges the financial support of the National Institutes of Health, Grant R21RR-20219.

## Notation

$[E]$	= concentration of polymerase, $\mu\text{M}$
$E$	= activation energy, $\text{J/mol}$
$F$	= error frequency, defined by Eq. 17
$k_j$	= rate constant, as indicated in Figure 1
$N$	= total length of DNA template
$n$	= $n$ th position along template
$N_j$	= number of deoxynucleotide monophosphates of type $J$ in DNA template
$P_{JK}$	= probability to proceed from state $AJ$ to state $AK$ ; the states are marked in Figure 1
$P_s$	= defined by Eq. 4
$R$	= universal gas constant, $\text{J mol}^{-1} \text{K}^{-1}$
$T$	= temperature, $\text{K}$
$t_{JK}$	= average time to go from state $AJ$ to state $AK$
$[X]$	= concentration of all deoxynucleotide triphosphates, dNTP, $\mu\text{M}$
$x_j$	= molar fraction of deoxynucleotide triphosphate, $J = \text{A, C, T, G}$

## Greek letters

$\nu$	= elongation rate, nucleotides/s
$\Gamma$	= defined as $\tau/P_s$
$\tau_K$	= average passage time for event $K$

## Literature Cited

- Grundberg-Manago M, Oritz PJ, Ochoa S. Enzymatic synthesis of nucleic acid-like polynucleotides. *Science*. 1955;122:907-910.



2. Lehman IR. Discovery of DNA polymerase. *J Biol Chem.* 2003;278:34733-34738.
3. Ahmed A, Tollefsbol TO. Telomerase, telomerase inhibition, and cancer. *J Anti Aging Med.* 2003;6:315-325.
4. Miura S, Izuta S. DNA polymerases as targets of anticancer nucleosides. *Curr Drug Targets.* 2004;5:191-195.
5. Smith AJ. DNA sequence analysis by primed synthesis. *Methods Enzymol.* 1980;65:560-580.
6. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Anheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science.* 1985;230.
7. Johnson KA. Conformational coupling in DNA polymerase fidelity. *Annu Rev Biochem.* 1993;62:685-713.
8. Kiefer JR, Mao C, Braman JC, Beese LS. Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal. *Nature.* 1998;391:304-307.
9. Li Y, Waksman G. Structural studies of the KlenTaq1 DNA polymerase. *Curr Org Chem.* 2001;5:871-883.
10. Jager J, Pata JD. Getting a grip: Polymerases and their substrate complexes. *Curr Opin Struct Biol.* 1999;9:21-28.
11. Kunkel TA. DNA replication fidelity. *J Biol Chem.* 2004;279:16895-16898.
12. Joyce CM, Benkovic SJ. DNA polymerase fidelity: Kinetics, structure, and checkpoints. *Biochemistry.* 2004;43:14317-14324.
13. Vieille C, Zeikus GJ. Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev.* 2001;65:1-43.
14. Ninio J. Alternative to the steady-state method: Derivation of reaction rates from first-passage times and pathway probabilities. *Proc Natl Acad Sci USA.* 1987;84:663-667.
15. Yamato M, Hirota Y, Yoshida S, Tanaka S, Morita T, Sakai J, Hashigaki K, Hayatsu H, Wataya Y. Imbalance of deoxyribonucleoside triphosphates and DNA double-strand breaks in mouse mammary tumor FM3A cells treated in vitro with an antineoplastic tropolone derivative. *Jpn J Cancer Res.* 1992;83:661-668.
16. Mathews CK, Ji J. DNA precursor asymmetries, replication fidelity, and variable genome evolution. *Bioessays.* 1992;14:295-301.
17. Martomo SA, Mathews CK. Effects of biological DNA precursor pool asymmetry upon accuracy of DNA replication in vitro. *Mutat Res.* 2002;499:197-211.
18. Lindahl T, Nyberg B. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry.* 1974;13:3405-3410.
19. Das SK, Benditt EP, Loeb LA. Rapid changes in deoxynucleoside triphosphate pools in mammalian cells treated with mutagens. *Biochem Biophys Res Commun.* 1983;114:458-464.
20. Suzuki K, Miyaki M, Ono T, Mori H, Moriya H, Kato T. UV-induced imbalance of the deoxyribonucleoside triphosphate pool in *E. coli*. *Mutat Res.* 1983;122:293-298.
21. Yoshioka A, Tanaka S, Hiraoka O, Koyama Y, Hirota Y, Ayusawa D, Seno T, Garrett C, Wataya Y. Deoxyribonucleoside triphosphate imbalance. 5-Fluorodeoxyuridine-induced DNA double strand breaks in mouse FM3A cells and the mechanism of cell death. *J Biol Chem.* 1987;262:8235-8241.
22. Wataya Y, Hazawa T, Watanabe K, Hirota Y, Yoshioka-Hiramoto A. Detection of deoxyribonucleoside-triphosphate imbalance death-induced DNA double strand breakage in FM3A cells by orthogonal-field-alternation gel electrophoresis (OFAGE). *Nucleic Acids Symp Ser.* 1988;53-55.
23. James SJ, Cross DR, Miller BJ. Alterations in nucleotide pools in rats fed diets deficient in choline, methionine and/or folic acid. *Carcinogenesis.* 1992;13:2471-2474.
24. Kasahara Y, Nakai Y, Miura D, Kanatani N, Yagi K, Hirabayashi K, Takahashi Y, Izawa Y. Decrease in deoxyribonucleotide triphosphate pools and induction of alkaline-labile sites in mouse bone marrow cells by multiple treatments with methotrexate. *Mutat Res.* 1993;319:143-149.
25. Melnyk S, Pogribwa M, Miller BJ, Basnakian AG, Pogribny IB, James SJ. Uracil misincorporation, DNA strand breaks, and gene amplification are associated with tumorigenic cell transformation in folate deficient/repleted Chinese hamster ovary cells. *Cancer Lett.* 1999;146:35-44.
26. Yamato M, Hirota Y, Yoshida S, Tanaka S, Morita T, Sakai J, Hashigaki K, Hayatsu H, Wataya Y. Imbalance of deoxyribonucleoside triphosphates and DNA double-strand breaks in mouse mammary tumor FM3A cells treated in vitro with an antineoplastic tropolone derivative. *Jpn J Cancer Res.* 1992;83:661-668.
27. Heinemann V, Schulz L, Issels RD, Plunkett W. Gemcitabine: A modulator of intracellular nucleotide and deoxynucleotide metabolism. *Semin Oncol.* 1995;22:11-18.
28. Chimpoy K, Tassotto ML, Mathews CK. Ribonucleotide reductase, a possible agent in deoxyribonucleotide pool asymmetries induced by hypoxia. *J Biol Chem.* 2000;275:39267-39271.
29. Akerblom L, Pontis E, Reichard P. Effects of azidocytidine on DNA synthesis and deoxynucleotide pools of mouse fibroblast cell lines. *J Biol Chem.* 1982;257:6776-6782.
30. Julius JG, Pathak VK. Deoxyribonucleoside triphosphate pool imbalances in vivo are associated with an increased retroviral mutation rate. *J Virol.* 1998;72:7941-7949.
31. Song S, Wheeler LJ, Mathews CK. Deoxyribonucleotide pool imbalance stimulates deletions in HeLa cell mitochondrial DNA. *J Biol Chem.* 2003;278:43893-43896.
32. Saada A, Shaag A, Mandel N, Nevo Y, Eriksson S, Elpeleg O. Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. *Nat Genet.* 2001;29:342-344.
33. Wang L, Saada A, Eriksson S. Kinetic properties of mutant human thymidine kinase 2 suggest a mechanism for mitochondrial DNA depletion myopathy. *J Biol Chem.* 2003;278:6963-6968.
34. Patel SS, Wong I, Johnson KA. Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry.* 1991;30:511-525.
35. Fersht AR, Shi JP, Tsui WC. Kinetics of base misinsertion by DNA polymerase I of *Escherichia coli*. *J Mol Biol.* 1983;165:655-667.
36. Florian J, Goodman MF, Warshel A. Theoretical investigation of the binding free energies and key substrate-recognition components of the replication fidelity of human DNA polymerase beta. *J Phys Chem B.* 2002;106:5739-5753.
37. Cline J, Braman JC, Hogrefe HH. PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* 1996;24:3546-3551.
38. Mizuguchi H, Nakatsuji M, Fujiwara S, Takagi M, Imanaka T. Characterization and application to hot start PCR of neutralizing monoclonal antibodies against KOD DNA polymerase. *J Biochem.* 1999;126:762-768.
39. Arezi B, Xing W, Sorge JA, Hogrefe HH. Amplification efficiency of thermostable DNA polymerases. *Anal Biochem.* 2003;321:226-235.
40. Greagg MA, Fogg MJ, Panayotou G, Evans SJ, Connolly BA, Pearl LH. A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. *Proc Natl Acad Sci USA.* 1999;96:9045-9050.
41. Fogg MJ, Pearl LH, Connolly BA. Structural basis for uracil recognition by archaeal family B DNA polymerases. *Nat Struct Biol.* 2002;9:922-927.
42. Connolly BA, Fogg MJ, Shuttleworth G, Wilson BT. Uracil recognition by archaeal family B DNA polymerases. *Biochem Soc Trans.* 2003;31:699-702.
43. Shuttleworth G, Fogg MJ, Kurpiewski MR, Jen-Jacobson L, Connolly BA. Recognition of the pro-mutagenic base uracil by family B DNA polymerases from archaea. *J Mol Biol.* 2004;337:621-634.
44. Whitney SE, Alugupally S, Nelson RM, Viljoen HJ. Principles of rapid polymerase chain reactions: Mathematical modeling and experimental verification. *Comp Biol Chem.* 2004;28:195-209.

Manuscript received Mar. 24, 2005, and revision received May 24, 2005.