Critical Reviews in Biochemistry and Molecular Biology, 28(2):83–126 (1993)

Biochemical Basis of DNA Replication Fidelity

Myron F. Goodman, Steven Creighton, Linda B. Bloom, and John Petruska

University of Southern California, Department of Biological Sciences, Los Angeles, California 90089-1340

Referee: Dr. Thomas A. Kunkel, Natl. Inst. of Environmental Health Sciences, Laboratory of Molecular Genetics, Research Triangle Park, NC

ABSTRACT: DNA polymerase is the critical enzyme maintaining genetic integrity during DNA replication. Individual steps in the replication process that contribute to DNA synthesis fidelity include nucleotide insertion, exonucleolytic proofreading, and binding to and elongation of matched and mismatched primer termini. Each process has been investigated using polyacrylamide gel electrophoresis (PAGE) to resolve ^{32}P-labeled primer molecules extended by polymerase. We describe how integrated gel band intensities can be used to obtain site-specific velocities for addition of correct and incorrect nucleotides, extending mismatched compared to correctly matched primer termini and measuring polymerase dissociation rates and equilibrium DNA binding constants. The analysis is based on steady-state "single completed hit conditions", where polymerases encounter many DNA molecules but where each DNA encounters an enzyme at most once. Specific topics addressed include nucleotide misinsertion, mismatch extension, exonucleolytic proofreading, single nucleotide discrimination using PCR, promiscuous mismatch extension by HIV-1 and AMV reverse transcriptases, sequence context effects on fidelity and polymerase dissociation, structural and kinetic properties of mispairs relating to fidelity, error avoidance mechanisms, kinetics of copying template lesions, the "A-rule" for insertion at abasic template lesions, an interesting exception to the "A-rule", thermodynamic and kinetic determinants of base pair discrimination by polymerases.

KEY WORDS: DNA polymerases, DNA synthesis fidelity, mutagenesis, base pair structures, discrimination free energy.

I. INTRODUCTION

Watson and Crick's discovery of DNA base pairing[1,2] followed by Kornberg's biochemical studies on *Escherichia coli* DNA polymerase I[3] showed that specificity of nucleotide incorporation in DNA replication is a template-directed process. Because of similarities in the structures

List of Abbreviations

- HIV-1 RT: human immunodeficiency virus reverse transcriptase
- AMV RT: avian myeloblastosis virus reverse transcriptase
- pol: DNA polymerase
- X: abasic (apurinic/apyrimidinic) DNA template site
- dNTP: deoxynucleoside 5′-triphosphate
- dWTP and dRTP: substrates for insertion by DNA polymerase that correspond to wrong and right base pairs respectively at a specific template position
- AP: 2-aminopurine
- B: bromodeoxyuridine
- F: fluorodeoxyuridine
- PCR: polymerase chain reaction

1080-9238/93/$.50

of base pairs and mispairs, template-directed base selection is far from error free, and DNA replication errors occur at frequencies sufficient to require the intervention of auxiliary error correcting systems. Examples in *E. coli* include polymerase-associated proofreading exonucleases,[4–7] postreplication methyl-directed mismatch repair enzymes,[8–10] and mismatch-specific DNA glycoases.[11–12a] These enzymes in conjunction with polymerases maintain mutation rates at levels suitable for both cell survival and evolution.

Recent technical advances have made it possible to address a range of fundamental questions germane to mechanisms of DNA synthesis fidelity and mutagenesis. In this article, we focus on topics relating to mechanistic aspects of fidelity. The topics include probing the structure of base mispairs using physical and enzymatic measurements; the effects of sequence context on the rates of nucleotide insertion and fidelity of DNA synthesis by various polymerases; behavior of polymerases in the vicinity of coding and noncoding DNA template lesions; and possible origins and magnitudes of free energy differences allowing polymerases to discriminate between correct and incorrect base pairs.

Much of our work has relied on a gel fidelity assay designed for rapid and convenient measurement of polymerase activity at arbitrary positions along a DNA template strand. In Section II, we show how gel band intensities can be analyzed to obtain polymerase kinetic parameters and fidelities. In Sections III and IV, we summarize experimental findings relating to the topics introduced earlier. Recently, in collaboration with J. Beechem, Vanderbilt University, we have begun to examine nucleotide insertion and removal kinetics in the transient time domain using time-resolved fluorescence spectroscopy.

II. DNA POLYMERASE FIDELITY

There are two steady-state methods for measuring fidelity of DNA polymerases, one called "competition" and the other "kinetic", reviewed in Reference 13. The first involves direct competition between right (R) and wrong (W) dNTP substrates for insertion into DNA by the enzyme. The second measures incorporation kinetics, velocity vs. [dNTP], separately for each substrate, to evaluate kinetic parameters used to deduce fidelity. Application of the kinetic method to a wide range of questions concerning fidelity is the main subject of this article.

We examine three contributions to fidelity: (1) misinsertion efficiency (f_{ins}), which measures the relative rates of inserting W compared to R nucleotides and is the reciprocal of insertion fidelity; (2) mismatch extension efficiency (f_{ext}), giving the relative rates of extending W compared to R primer 3′-termini; and (3) misincorporation efficiency, which takes into account the effects of exonucleolytic proofreading on misinsertion. Extension and proofreading can be regarded as competing reactions in which there is either elongation from W or R primer termini or excision of a W or R terminal nucleotide.

A. Misinsertion Efficiency

Labeled f_{ins}, misinsertion efficiency is defined as the ratio of the velocities for inserting W and R nucleotides opposite a given template base when dWTP and dRTP are present at equal concentration in solution. In a direct competition experiment, to detect W incorporation in the presence of R, a significant pool bias is needed, i.e., large excess of dWTP. Also, because the right nucleotide is inserted preferentially over the wrong, a condition must be selected to maintain [dWTP]/[dRTP] approximately constant. Although misincorporation can be measured directly in a double label experiment having W and R substrates present in the same reaction (see, e.g., Reference 14), f_{ins} can be determined more conveniently by measuring V_{max}/K_m for each substrate separately,[15]

$$f_{ins} = (V_{max,W}/K_{m,W})/(V_{max,R}/K_{m,R}) \quad (1)$$

The V_{max}/K_m value for each substrate corresponds to the linear slope of the Michaelis-Menten curve at low [dNTP].

The proof that relative reaction rates of two competing nucleotides can be deduced from V_{max}/K_m evaluated for each nucleotide separately was discussed by Fersht.[15] A general expression for nucleotide insertion velocity is $v = (k_{cat}/K_m)$ [polymerase-DNA] × [dNTP], where $k_{cat} = V_{max}/$[total polymerase-DNA]. For W and R sub-

strates competing at the *same* site, the polymerase-DNA concentrations for the two substrates are the same and hence cancel in the expression for the velocity ratio, v_W/v_R. It follows that v_W/v_R equals the ratio of V_{max}/K_m for W and R substrates multiplied by [dWTP]/[dRTP], the pool bias ratio. Because f_{ins} is defined as v_W/v_R when [dWTP] = [dRTP], its value is given by Equation 1.

1. Steady-State Polymerization Kinetics

Three kinetic parameters can be evaluated from steady-state measurements on polymerases in the absence of proofreading, V_{max} and K_m for nucleotide insertion, and K_D, the polymerase-DNA equilibrium binding constant.[16] A gel fidelity assay developed in our laboratory can be used to make such kinetic measurements rapidly, accurately, and at a large number of sites.[17–19]

The gel assay uses band intensities of 5′-^{32}P-labeled primers and extended products separated electrophoretically to measure nucleotide insertion velocities. In Section II.B, we describe how velocities are derived from band intensities measured by densitometry and integration. Models for polymerase action are discussed that enable one to evaluate k_{cat}/K_m at any template site by using ratios of band intensities. From band intensity ratios as a function of [dNTP] or [DNA] one can evaluate polymerase k_{cat} and K_m values for each dNTP[17–20] or polymerase-DNA dissociation rate constants, k_{off},[20] and polymerase-DNA equilibrium binding constants, K_D.[21]

B. Gel Fidelity Assay

The objective of the gel assay is to measure polymerase fidelity at any target site along a DNA or RNA template strand. Two types of reaction are possible, a "running start" reaction, in which template-bound primer is extended by several bases to reach a designated site where reaction velocity is determined (Figure 1a), and a "standing start" reaction where reaction velocity is determined at the first template site next to the original primer terminus (Figure 1b).

Polymerase-catalyzed reactions are performed with 5′-^{32}P-labeled primers annealed to unlabeled templates. The primers are extended at their 3′-ends by template-directed addition of one or more nucleotides. Reactions are quenched after a short reaction time, t, when less than 20% of the original primers are extended. Separate reactions are performed for correct and incorrect nucleotides, and unextended primers along with extended products are resolved as bands on gel by PAGE (Figure 2).

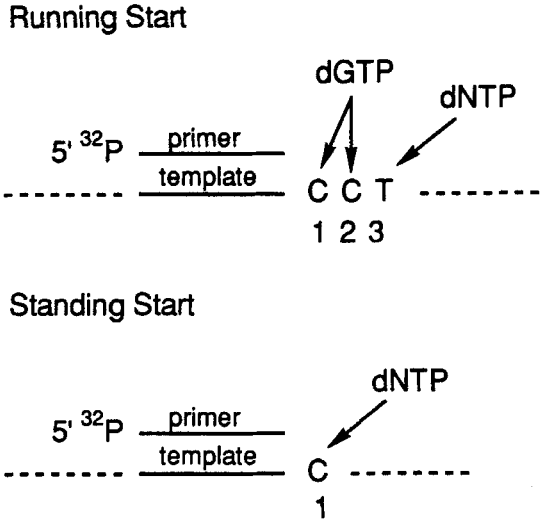


FIGURE 1. Examples of primer-template sites at which reaction kinetics are measured for right and wrong nucleotides. (a) Running start reaction in which reaction velocity vs. [dNTP] is determined for insertion opposite template base T. A "saturating" concentration of dGTP is used to extend the primer to reach the template T site. The substrate dNTP (right or wrong) is used at various concentrations for insertion opposite T. (b) Standing start reaction in which velocity vs. [dNTP] is measured for insertion opposite template base C, using either a right or wrong dNTP substrate.

The end product of a gel assay experiment is an autoradiogram of gel lanes, as shown in Figure 2. From a running start experiment, each lane has an intense primer band followed by one or more running start bands and a target (T) band. From a standing start, there are only the primer and target bands. The intensities of individual bands in each lane are integrated and normalized relative to the total integrated intensity of all relevant bands in the lane. The normalized (relative) band intensities in each lane are designated I_0, I_1, . . . , I_{T-1}, I_T with 0 indicating the original primer band and T the target band.

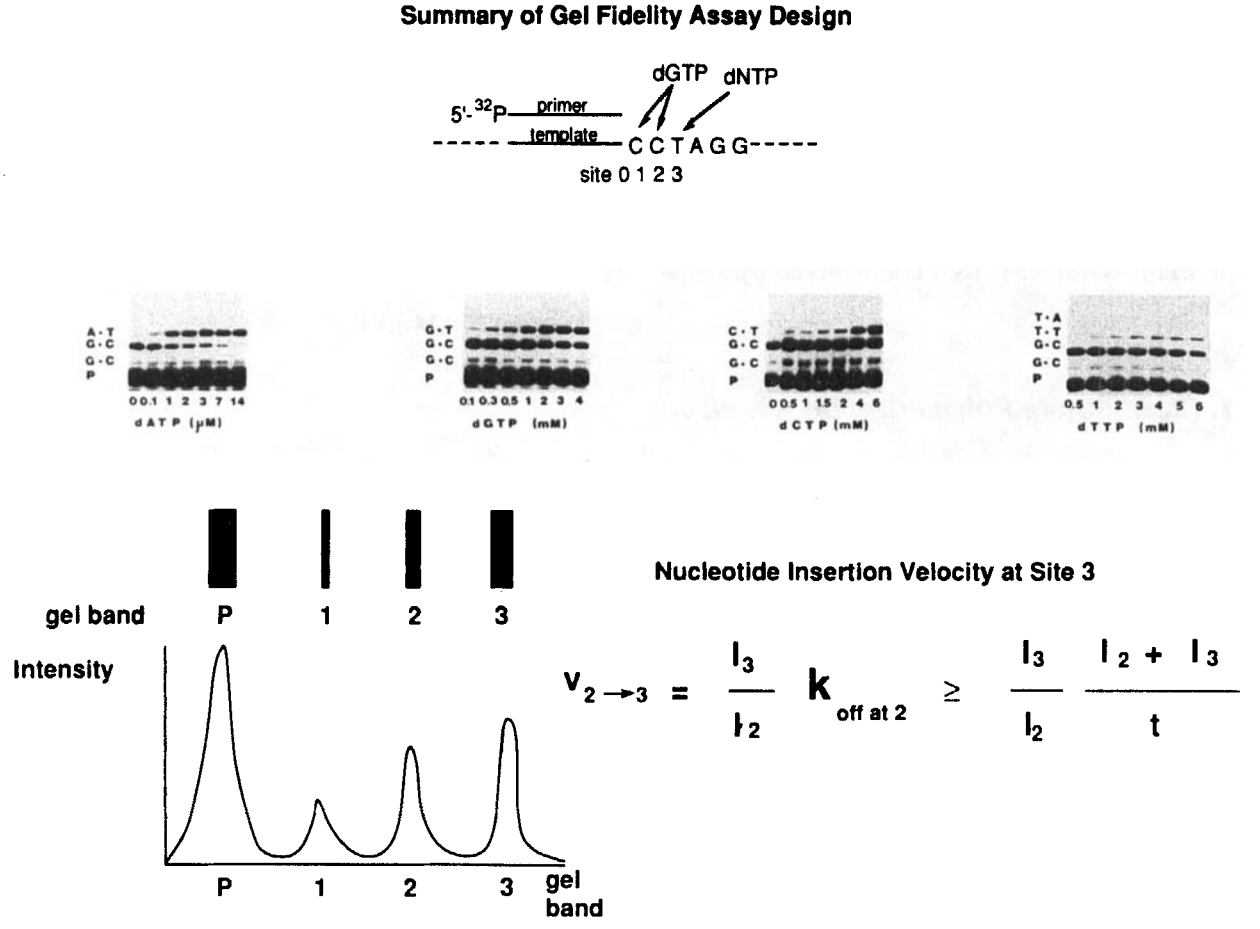


FIGURE 2. Summary of the gel fidelity assay. A running start reaction to extend a ^{32}P-labeled primer (Figure 1a) is used to measure the correct insertion opposite T of dAMP or incorrect insertions of dGMP, dCMP, or dTMP as a function of [dNTP]. Gel bands corresponding to the target template T, site 3, and previous template C, site 2, are integrated by densitometry or phosphorimaging, and the velocity to go from site 2 to 3 is calculated. A relative velocity, used for determining misinsertion efficiencies, f_{ins}, Equation 1, is obtained by taking the ratio of integrated band intensities at the target and previous template sites, I_3/I_2. In cases where there is extension beyond the target site, T, $I_{T+1} + I_{T+2} + \ldots$ should be added to I_T. The polymerase-DNA dissociation rate from the site prior to the target, required to obtain the "absolute" velocity for insertion,[20] is obtained experimentally as shown in Figure 5.

Previously, on the basis of a steady-state kinetics analysis,[17,18] we have used the following expression for determining nucleotide insertion velocity in a running start experiment,

$$V_{T-1\rightarrow T} = (I_T/I_{T-1})(I_T + I_{T-1})/t \qquad (2)$$

with t being the reaction time. This expression was derived by considering additivity of primer extension times (reciprocals of extension velocities) without considering microscopic steps in the enzymatic rate of converting primers to extended products. Under conditions where $(I_T + I_{T-1})/t$ is held constant, $V_{T-1\rightarrow T}$, the velocity of extending primer from T − 1 to T, is simply proportional to the band intensity ratio, I_T/I_{T-1}*. The band intensity ratio is then a convenient *relative* velocity measurement. For a standing start experiment, with T = 1 and $I_1 \ll I_0$, the equation for insertion velocity is simply

* In cases where there is extension beyond T, in the absence of proofreading exonuclease, $I_{T+1} + I_{T+2} + \ldots$ should be added to I_T; in the presence of exonuclease, the intensities $I_{T+2} + I_{T+3} + \ldots$ should be added to I_{T+1}.

$$V_{0\rightarrow 1} = I_1/t \quad (3)$$

Velocities measured by Equations 2 and 3 for running and standing starts, respectively, are found to vary with substrate concentration in accordance with Michaelis-Menten kinetics.[17–19] In both cases, V_{max} and K_m values can be obtained for W and R substrates and f_{ins} can be determined by Equation 1. While running starts tend to give higher V_{max} and K_m values than standing starts at the same target site, they yield similar f_{ins} values.[19]

From the V_{max} may be calculated a catalytic rate constant, $k_{cat} = V_{max}/[\text{total polymerase-DNA}]$. However, in order to estimate total polymerase at the target site, one needs models to interpret the velocity evaluated by Equation 2. This is a global velocity from site T − 1 to T, to which various microscopic steps (binding, enzyme conformational change, catalysis and dissociation) may contribute. Of particular concern is the possibility that the observed V_{max} may be limited by polymerase dissociation from primer template, and therefore the actual rate of catalysis may be underestimated.

The determination of k_{cat} is easiest to make under "single-hit" reaction conditions, to which models like those described in the following sections apply. With these models one can extract rate constants having a clearly defined microscopic meaning.

1. Analysis of Running Start Reactions

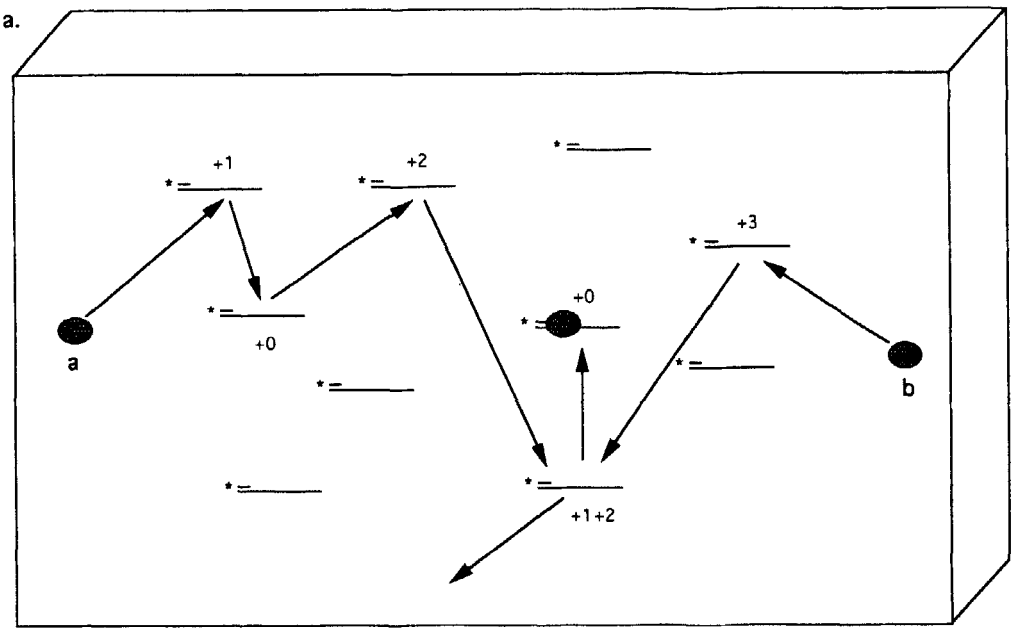
Figure 2 depicts a typical running start experiment, with [DNA] being the concentration of annealed primer template, [Pol] the concentration of DNA polymerase, and [dNTP] the substrate concentration for insertion at target site T in addition to whatever dNTP is needed for extending primer to the site prior to target, T − 1. The reaction is initiated at time 0 by mixing dNTP and Pol-DNA solutions and terminated at time t by adding a quenching agent (EDTA).

Using [Pol] much less than [DNA], we assume that each polymerase molecule collides with a randomly selected template-bound primer and remains associated for a period of time before dissociating and moving on to a new randomly selected template primer (Figure 3a). The reaction time, t, is chosen short enough so that only a small fraction of primers is extended, and the probability that a polymerase will reassociate with a primer already encountered is small. During its association with template primer, polymerase has an opportunity to add one or more nucleotides to the primer 3′-terminus. The probability of nucleotide addition depends on the identity of the template base next to the primer terminus and the concentration of dNTP in solution.

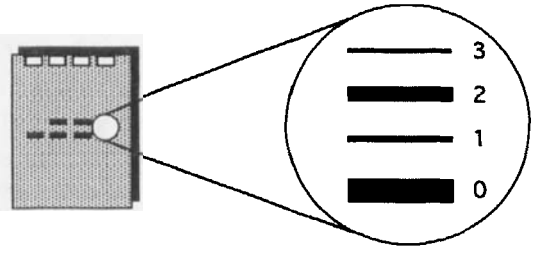
Under single encounter conditions, the probability that polymerase extends primer is closely related to the probability that polymerase dissociates from primer template. The ratio of the two probabilities determines the relative intensities of bands T − 1 and T on the gel. Band T − 1 becomes intense when the dissociation rate at site T − 1 is large compared with the extension rate for T − 1 → T. Band T − 1 becomes faint and T becomes intense for the opposite reason, i.e., rapid extension with respect to dissociation.

While it is impossible to force polymerase molecules to associate with a given primer template only once during the reaction, it is possible to engineer reaction conditions so that single hits predominate. Note that there are four classes of primer-template molecules present in the reaction at any given time: (a) unhit, (b) singly hit, (c) multiply hit, and (d) incompletely hit, where the reaction is quenched before the enzyme either adds a nucleotide or dissociates. We want to see bands arising mainly from classes a and b when products are run on the gel. Thus it is important to determine ways of reducing the proportions of c and d.

The mechanistic interpretation of band intensities requires microscopic models to describe the mechanics of band formation. Here we present a *Standard* model for polymerization-dissociation, a *Translocation* model allowing polymerase to translocate in a dNTP-independent manner, and a *Proofreading* model that includes 3′-exonuclease activity. In these models polymerases are assumed to perform under "single completed hit" conditions, i.e., a polymerase molecule encounters any given primer template only once during the time course of reaction and either adds one or more nucleotides or dissociates from primer template before the reaction is quenched.



b.



Gel Band Intensities

I_3 : hit + extended by 3

I_2 : hit + extended by 2

I_1 : hit + extended by 1

I_0 : unhit or hit + unextended

Single hits predominate when $I_0/(I_0+I_1+I_2+I_3) > 0.80$

Probability of base addition in a single encounter

P(Add base 1) = c

P(Add base 2) = c $(I_2+I_3)/(I_1+I_2+I_3)$

P(Add base 3) = c $(I_3)/(I_1+I_2+I_3)$

FIGURE 3. (a) Schematic representation of the course of a typical kinetics reaction in a small volume of the reaction mixture. Gray ovals labeled "a" and "b" represent polymerase molecules and the large arrows represent their movement during the time of the reaction. The pairs of small and large lines represent labeled (*) DNA primer/template molecules and the numbers, +0, +1, +2, +3 indicate primer templates "hit" by polymerase and the number of nucleotides added to the primer during the course of the encounter. Thus, "a" is shown adding a single nucleotide (indicated by +1) to the primer template at the top left, then associating and dissociating from another primer template without adding a nucleotide (+0), then adding two nucleotides to another primer template (+2), etc. Note that polymerase "b" ended the reaction still bound to a primer/template, ("incomplete hit") and a single primer/template suffered two encounters with a polymerase ("multiple hit"). (b) Gel bands observed when the reaction mixture is fractionated by PAGE. Regardless of the model used to explain the gel bands, the intensities of the bands reflect the population of extended primers. If most of the hits during the course of the reaction are single hits, then the intensities of the bands will reflect relative probabilities that the polymerase dissociated at the given location along the template by the equations shown. Note that "c" is the (unknown) *probability that the polymerase adds the first base in a given encounter.*

a. Single Completed Hit Conditions

Assuming that each primer template has the same probability of interacting with a polymerase, then the number of hits that a primer template experiences is a random variable with a Poisson distribution.[22] When the reaction is quenched, the fraction of primer templates with single and multiple hits can be calculated from a Poisson distribution, by measuring the fraction of unextended primers. The ratio of single to multiple hits declines as the fraction of unextended primers declines. A simple rule of thumb is that multiple hits can be neglected if less than 20% of the original primer templates are extended. Under these conditions less than 2% of primer templates experience multiple hits. The use of either excess unlabeled DNA or heparin to trap polymerases following dissociation from a labeled primer template can also ensure that single hit conditions are satisfied. However, excessively high trap concentrations may have the undesirable effect of stimulating dissociation.[23]

A predominance of complete over incomplete hits is also desired. An incomplete hit occurs when quenching interrupts the polymerase before nucleotide addition or dissociation occurs. The contribution of incomplete hits to gel band intensities can be neglected if [Pol] $\ll$ [DNA], e.g., less than 1 polymerase per 50 primer templates. Also, the effect of incomplete hits can be virtually eliminated by polymerase trapping.

For single completed hit conditions the pattern of extended bands can be easily analyzed. The probability of extending a primer beyond a given template site depends only on the probability that polymerase terminates processive polymerization at that site (Figure 3b). The ratio of relative probabilities for polymerase to extend primer from T − 1 to T or to dissociate at site T − 1 is given by the ratio of band intensities, I_T/I_{T-1}. Interpretation of this band intensity ratio as a function of [dNTP] depends on the model chosen to describe the probabilities.

b. Standard Model

This model for interpreting running start band intensities is perhaps the simplest one possible. A polymerase molecule first binds to an original primer template (site 0) and, with a probability dependent on [dNTP] for primer extension, adds one or more nucleotides to reach site T − 1. After this, the polymerase will either dissociate with some probability independent of [dNTP] or add the next nucleotide with a probability, k_{pol}, related to [dNTP] for insertion at site T by the Michaelis-Menten equation (Figure 4a).

When associated with primer template, the polymerase is continuously vulnerable to thermally driven dissociation either in the presence or absence of dNTP substrate molecules. For simplicity, we assume that the polymerase has fixed probabilities of dissociation and incorporation at all sites along the DNA template with respect to reaction time, but not position. In other words, at any template position there is a [dNTP]-independent dissociation probability (k_{off}) and a [dNTP]-dependent polymerization probability (k_{pol}).

At the target site, the enzyme either inserts a nucleotide opposite the target base on template and contributes to I_T (the intensity of gel band T) or dissociates from template and contributes to I_{T-1} (Figure 4a). Thus, I_T is proportional to the probability of insertion, $k_{pol}/(k_{pol} + k_{off})$, and I_{T-1} is proportional to the probability of dissociation, $k_{off}/(k_{pol} + k_{off})$. Hence, $(I_T/I_{T-1}) = k_{pol}/k_{off}$ so that the nucleotide insertion rate for T − 1 $\rightarrow$ T is given by

$$k_{pol} = k_{off}(I_T/I_{T-1}) \quad (4)$$

The intensity ratio I_T/I_{T-1}* is measured at different dNTP concentrations and fit to a rectangular hyperbola to obtain apparent values of K_m and relative $V_{max} = (I_T/I_{T-1})_{max}$ for each dNTP. The product of $(I_T/I_{T-1})_{max}$ and k_{off} is the maximum value of k_{pol}, according to Equation 4. From the k_{pol} maximum is obtained the catalytic rate constant,

$$k_{cat} = k_{off}(I_T/I_{T-1})_{max} \quad (5)$$

In Figure 5, we show how a k_{off} measurement can be made to obtain representative data for DNA polymerase off-rates required to calculate k_{pol}.

* See footnote under Section B.

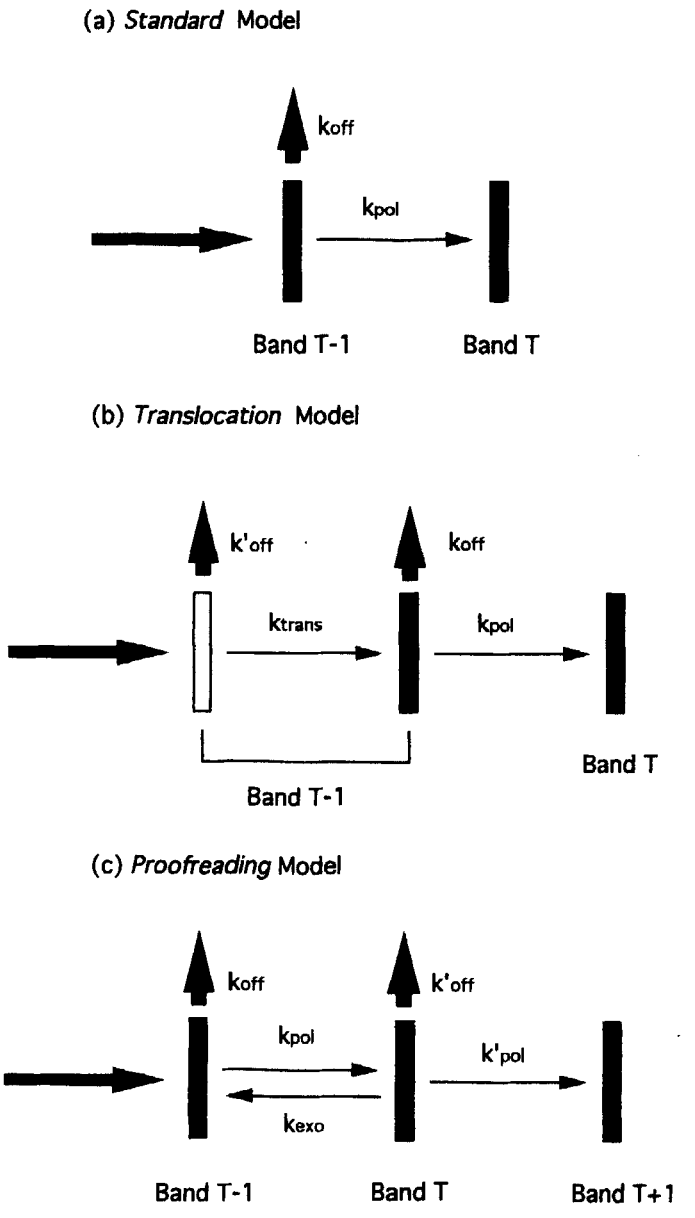


FIGURE 4. Three models of the elementary processes occurring during progressive DNA polymerization that give rise to gel bands. (a) *Standard* Model — after the addition of each base the polymerase can either dissociate with constant probability per unit time (k_{off}) or add the next base with probability per unit time depending on the [dNTP] through the Michaelis-Menten relation: $k_{pol} = k_{cat}[dNTP]/(K_m + [dNTP])$. (b) *Translocation* Model — after addition of each base the polymerase will first translocate to the next site at constant probability per unit time (k_{trans}) after which it is able to add the next nucleotide as in the Standard Model. (c) Proofreading Model — as in the Standard model except that after adding the first base and reaching position T there is a constant probability per unit time (k_{exo}) that the exonuclease will excise the base and return to the unextended state (position T − 1). This excision process is in competition with a constant rate of dissociation at position T (k'_{off}) and a Michaelis-Menten rate (k'_{pol}) of polymerization to position T + 1 — both of these processes will mitigate the effectiveness of the exonuclease.

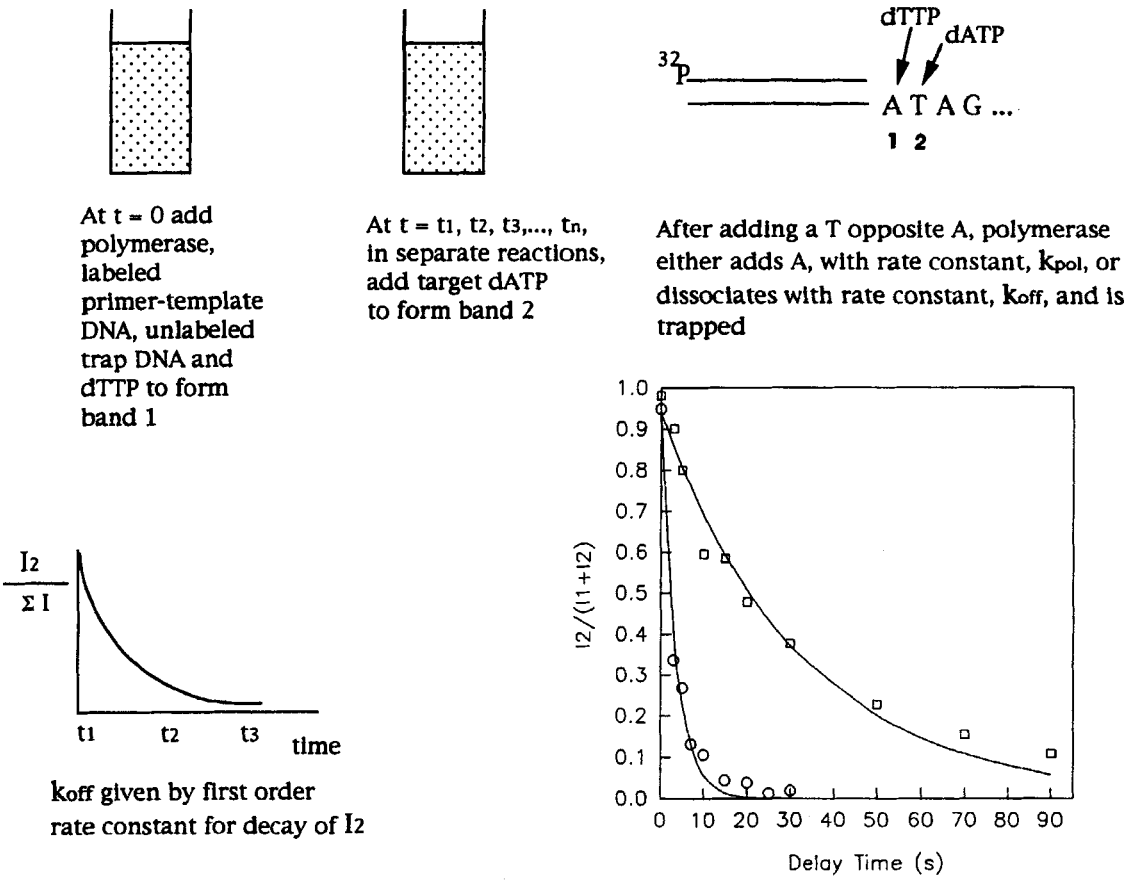


FIGURE 5. DNA trapping experiment to measure k_{off} for polymerase in running start. The data shown (bottom right) are for a one base running start experiment using *E. coli* pol I (Klenow fragment, exo$^-$), a ^{32}P-labeled primer bound to unlabeled template, and excess unlabeled trap DNA. The squares (□) represent data obtained with the template sequence shown (top right); dTTP being used for insertion at site 1 and after a delay time, t, dATP is added for insertion at site 2. The band intensity ratio, $I_2/(I_1 + I_2)$, is plotted against t to obtain an exponential decay curve, $e^{-k_{off}t}$, from which k_{off} is evaluated. The decay occurs because polymerase dissociates from primer template in the absence of dATP, and is captured by the DNA trap. For the template shown (□), 3′ . . . ATAG . . . , k_{off} = 0.03 s^{-1}. For the other template used (O), 3′ . . . CTAG . . . with dGTP instead of dTTP for insertion in site 1, and [dATP] kept the same for insertion at site 2, k_{off} = 0.28 s^{-1}.

c. Translocation Model

To the *Standard* model (Figure 4a) we now add a translocation step with a rate constant, k_{trans}, prior to dNTP binding and catalysis (Figure 4b). Contributions to I_{T-1} can now arise from enzyme dissociation events occurring either before or after translocation. Analysis of the band intensity ratio, I_T/I_{T-1}*, derived from the steps shown in Figure 4b gives

$$k_{pol} = k_{off}(I_T/I_{T-1})[(1 + \sigma)/\{1 - \sigma (I_T/I_{T-1})\}] \quad (6)$$

* See footnote under Section B.

where $\sigma = k'_{off}/k_{trans}$, with k'_{off} and k_{off} being the polymerase dissociation constants before and after the k_{trans} step, respectively. The dNTP-dependent insertion rate k_{pol} is now the product of the rate obtained by Equation 4 and the factor shown in square brackets that depends on k_{trans}.

For processive polymerases, translocation rates along DNA are typically much faster than dissociation rates from primer-template termini. Our measurements for a variety of exonuclease deficient polymerases, including *E. coli* pol I exo$^-$, bacteriophage T7 exo$^-$, AMV and HIV-1 reverse transcriptases, indicate that translocation rates must be at least several nucleotides per second, while dissociation rates, although strongly

dependent on template position,[20,24] are around 0.1/s. Thus, for polymerases showing moderate to large procession the square-bracketed factor in Equation 6 is close to unity, and to a good approximation, k_{pol} is again given by Equation 4 and k_{cat} by Equation 5.

For the case of distributive enzymes, e.g., pol β, knowledge of σ is required to determine k_{pol} by Equation 6. However, k_{pol} evaluation is not needed to obtain nucleotide misinsertion frequency, f_{ins}, by Equation 1. Only the ratios of V_{max}/K_m for wrong and right insertions are required to calculate f_{ins}. The V_{max}/K_m values can be determined with dNTP concentrations low enough so that k_{pol} is much less than k_{trans}. When [dNTP] is well below K_m, I_T/I_{T-1} increases linearly with [dNTP] and the slope equals V_{max}/K_m.

d. Proofreading Model

In the case of polymerases having 3′ to 5′ exonuclease activity it is necessary to modify the interpretation of the insertion probabilities. It is now possible for an enzyme at site T − 1 to add a nucleotide to get to site T, then remove the inserted nucleotide and dissociate from site T − 1. This means that the band intensity ratio I_T/I_{T-1} depends not only on [dNTP] for insertion opposite site T, but also depends on [dNTP] for insertion opposite T + 1. The reason is that exonuclease at site T competes not only with dissociation at that site, but also with polymerization as well. The effect of the exonuclease is reduced if the rate of polymerization to the next site is made large. Such a reduction is known as the *next correct nucleotide effect*.[25,26]

The *Proofreading* model is depicted in Figure 4c. As before, the target and preceding sites are labeled T and T − 1, respectively. The site T + 1, immediately following the target site, is now included in the analysis. We assume that reaction conditions are chosen so that production of band T + 1 is final and cannot be reversed by the exonuclease, while reversal is possible from T to T − 1. When the system is solved to obtain band intensity ratios there are now two expressions:

$$(I_T + I_{T+1})/I_{T-1} = \kappa = (k_{pol}/k_{off})[(k'_{pol} + k'_{off})/(k_{exo} + k'_{pol} + k'_{off})] \quad (7a)$$

$$I_{T+1}/I_T = \mu = (k'_{pol}/k'_{off}) \quad (7b)$$

To determine fidelity in the presence of exonuclease an apparent relative insertion rate, $\kappa_0 = (k_{pol}/k_{off})\,[k'_{pol}/(k_{exo} + k'_{pol})]$, needs to be determined. If k'_{pol} is much greater than k'_{off} then κ_0 is approximately the same as κ in Equation 7a. We use the ratio $I_{T+1}/I_T = \mu$ as the indicator. If μ is much greater than 1, then κ_0, the relative insertion velocity at the target site, can be measured as $(I_T + I_{T+1})/I_{T-1}$.

In the event that μ is small then additional information is required to estimate κ_0. Determination of k'_{off} at site T and measurement of the band intensities at two different concentrations of the next correct dNTP are needed. If these measurements are made then the exonuclease rate constant can be calculated,

$$k_{exo} = k'_{off}\,(\kappa_1 - \kappa_2)/[\kappa_2/(1 + \mu_1) - \kappa_1/(1 + \mu_2)] \quad (8)$$

where κ_1, μ_1 and κ_2, μ_2 are measurements of κ and μ at the two different concentrations. Knowing k'_{off}, one can determine k'_{pol} by Equation 7b and in combination with k_{exo} in Equation 7a find κ_0 at each [dNTP].

It is important to remember that κ_0, depends on the concentration of the next correct nucleotide, because k'_{pol} is dependent on this concentration. As the concentration is raised, the apparent insertion rate also increases. This increase in rate is more pronounced for misinsertions because correct insertions are less likely to be excised by exonuclease. Thus, high concentrations of the next correct nucleotide tend to be mutagenic.

2. Analysis of Standing Start Reactions

In the standing start reaction, a relative reaction velocity can be determined at a template

site immediately adjacent to the primer 3′-terminus (Figure 1b). As with running starts, the initiation of reaction occurs at time 0 by mixing dNTP and Pol-DNA solutions and termination occurs at time t, by adding quenching agent. The main difference between standing and running start experiments is that standing starts only yield gel bands for unextended primer and the product of extension by a single nucleotide. Because there are no bands present for extension to the site prior to the target site, we have no information about the probability that an enzyme-DNA complex added a nucleotide before it dissociated from the primer template.

Accordingly, the rate measured in a standing start experiment is the rate of the slowest step for the cycle of polymerase association with DNA, nucleotide insertion, polymerase dissociation and reassociation with a new primer template. The V_{max} for a standing start is not necessarily a measure of the rate of polymerization from one site to the next but may be the rate of enzyme-DNA association or dissociation. Even so, ratios of V_{max}/K_m for wrong and right nucleotides measured in the linear region at low [dNTP] still yield misinsertion efficiency in accordance with Equation 1.

C. Proofreading Contribution

Proofreading refers to the excision of polymerase errors at the replication fork by a 3′ to 5′ exonuclease activity. A variety of prokaryotic and eukaryotic DNA polymerases contain such activity, either within a single polypeptide as for *E. coli* pol I, pol II, bacteriophage T4, T5, T7, eukaryotic pols δ and ϵ, mitochondrial pol γ, or as a separate polypeptide subunit, e.g., ϵ in *E. coli* pol III holoenzyme (for a review see Reference 27). Several polymerases, e.g., eukaryotic pol β, *Thermus aquaticus* polymerase, AMV and HIV-1 reverse transcriptases, have no known associated exonuclease activities. Whether these enzymes can act in concert with external exonucleases remains an open question.

Proofreading appears to remove approximately 95 to about 99.5% of single base substitutions.[28–32] However, *DnaQ* (*MutD*) mutant alleles in the pol III ϵ-proofreading subunit[6,7] in *E. coli* exhibit an elevation in mutation rates of four to five orders of magnitude.[29,33–35] A failure in proofreading may increase mutations by 10^2- to 10^3-fold, and saturation in methylation-directed mismatch repair may also contribute a 10^2-fold increase.[29,36,37]

Similar estimates have been made for the contribution of T4 polymerase associated exonuclease to fidelity.[30–32,38,39] Several mutator and antimutator polymerase (*ts gene43*) alleles in bacteriophage T4[40–42] known to perturb nuclease/polymerase activity ratios,[5,14,43] exhibit up to two orders of magnitude alteration in mutation rates in base substitution pathways.[40–42,44–46] However, the properties of these mutants are dependent on the specific pathways involved. It has been shown by Ripley and Shoemaker[47] that a T4 *gene43* ''antimutator'' allele, highly antimutagenic in an A · T → G · C pathway, is mutagenic for simple frameshifts.

Polymerase errors destabilize primer 3′-termini. The effect is to delay continued polymerization from mismatched termini.[23,25,26,48–52] Reduced extension from mismatched compared with correctly matched termini (see Section II.D) provides an opportunity for selective proofreading of misinserted nucleotides.[25,26,52]

A simple schematic diagram (Figure 6) illustrates how polymerase may switch between incorporation and proofreading modes.[25] A nucleotide is added when the primer terminus is in the annealed state (A), and the primer 3′-terminus is excised when in the melted state (M). The 3′ to 5′ exonuclease activity defined by the rate constant, k_{exo}, is the probability per unit time to excise a nucleotide from state M at a given template location. The fraction of nucleotides removed from state M is proportional to the nucleotide excision rate constant, k_{exo}, multiplied by the relative concentration of melted primer termini. Even though k_{exo} is the same for matched and mismatched termini, selective removal of mismatched primer termini occurs, because the population of mismatched termini in state M is much greater than in state A (i.e., $k_M \gg k_A$), while the reverse is true for correctly matched termini.

A potential problem with this simple picture is that melting temperature differences between oligonucleotide duplexes containing matched and

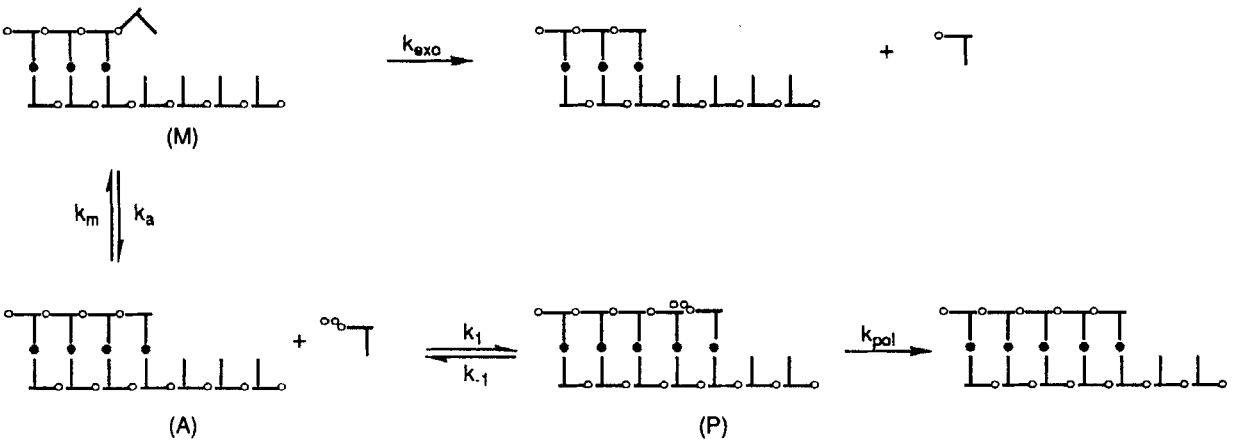


FIGURE 6. States and transitions in a simple kinetic model for nucleotide insertion and proofreading.[25,137] In state A, the dNTP binding site is empty and the terminal base is in its base-pairing position. A transition out of state A occurs by either the association of a dNTP (transition to state P, rate constant k_1), or dissociation of the terminal base from its base-pairing position (transition to state M, rate constant k_m). A transition from state M occurs by either the excision of the 3′-terminal base (rate constant k_{exo}) or reassociation of the terminal base to its pairing configuration (transition to state A, rate constant k_a). A transition from state P occurs by either the formation of a phosphodiester bond ([dNTP] dependent Michaelis rate k_{pol}) or the dissociation of the dNTP (transition to state A, rate constant k_{-1}). Following either excision or insertion, a shift occurs one base backward or forward to allow the cycle to repeat. When cycling occurs, the terminal base is assumed to be in an equilibrium distribution between states A and M.[25]

mismatched termini are much too small to account for high proofreading specificities.[48] Perhaps by exclusion of water in a polymerase active cleft, it may be possible to amplify free energy differences between matched and mismatched base pairs over those observed in aqueous solution.[53] However, a key experimental observation is that a significant fraction of correctly inserted nucleotides are excised as free dNMP *in vitro*, at rates approaching V_{max} rates for nucleotide incorporation.[5,14,25,28] In the case of wild-type T4 polymerase, about 20% of dAMP inserted opposite template T was removed.[25] For the case of *E. coli* pol III, correct A · T and G · C pairs were edited approximately 12 and 6%, respectively.[28] For the T4 *tsL141* antimutator polymerase, containing an anomalously high exonuclease/polymerase ratio,[5,43] about 40% dAMP inserted opposite T was turned over.[14,25] Even if proofreading exonucleases can impose additional active site specificity to excise mismatched termini, it is clear that a sizable fraction of correctly matched termini are also destroyed.

Thus, proofreading is an expensive means of reducing replication errors. This is especially true for T4 bacteriophage, which not only codes for most of its own deoxyribonucleotide metabolizing enzymes,[54,55] but appears to have a multienzyme complex dedicated to the *de novo* synthesis of dNTP substrates.[56–58] An important conclusion to be drawn is that exonucleolytic proofreading must be providing the cell with an essential component for maintaining genetic integrity, otherwise it would probably have been replaced much earlier by more accurate polymerases or by more efficient postreplication mismatch repair systems. From an evolutionary standpoint, it is interesting that T4 antimutators having high nuclease/polymerase ratios may be unable to compete effectively with wild-type (*43*$^+$) because of the high nucleotide turnover cost/accuracy ratio.[25,28]

Polymerase and exonuclease domains for *E. coli* pol I and T4 polymerase appear spatially distinct, based on X-ray crystallographic studies on Klenow fragment[59,60] and T4 *gene43* mapping and mutagenesis studies.[61,62] It has been proposed that pol 1 switches from synthetic to degradation modes by a mechanism involving an initial melting of DNA in the enzyme cleft and subsequent sliding to bring the melted primer terminus to the exonuclease active site.[60] The original Brutlag

and Kornberg proposal[4] that 3′ to 5′ exonuclease functions as a proofreader was based on the ability of pol I to excise single terminal mismatches more rapidly than correctly matched termini. It now appears that as many as four base pairs at the primer terminus (one mispair and three correct pairs) have to be melted to excise a single terminal mismatch.[63] When presented with a mismatched terminus, a significant fraction of pol I molecules appear to dissociate from primer template prior to excision.[64] In contrast, T4 proofreading may be occurring in a highly processive manner, in which excision is likely to occur prior to dissociation.[23] For T4 polymerase, there is also evidence suggesting that excision of a single terminal mismatch is optimized during forward DNA synthesis (at saturating dNTP concentrations) when melting of two to three base pairs is occurring.[23,63,65,66] Based on a comparison of the removal of matched and mismatched nucleotides from primer 3′-termini using the ϵ subunit of *E. coli* pol III holoenzyme, it was concluded that the source of proofreading specificity is the greater melting capacity of mispaired termini.[67]

At this stage of our knowledge about proofreading, it is useful to reemphasize some early ideas suggested by Bessman and colleagues.[5,14,68,69] An important correlation was noted between the exonuclease/polymerase ratios of certain T4 mutant DNA polymerases and the genetic properties of the T4 mutants; antimutators, notably *tsL141*, had higher exo/pol ratios than wild type, while mutators, e.g., *tsL56*, had lower ratios.[5,14,43] An extensive study, correlating the *in vivo* mutagenic behavior of a large number of T4 mutator and antimutator alleles with the properties of the polymerases purified from each,[68] strengthened the general notion that exo/pol ratio served as one important determinant of base substitution mutagenesis, keeping in mind the *caveat* mentioned earlier that the properties of the mutants are dependent on the specific pathways involved; for example, a T4 *gene43* antimutator allele, highly antimutagenic in an A · T → G · C pathway, was found to be mutagenic for simple frameshifts.[47] The work of Nossal and Hershfield[70–72] clearly demonstrated that other parameters were equally important. The T4 *gene43* mutator, *tsL88*, had an exo/pol ratio similar to wild type, but had a higher nucleotide misinsertion efficiency.[72] They also showed that the T4 *CB120* antimutator allele, now known to be identical to *tsL141*,[61,73] had difficulty in strand displacement.[74] Perhaps, this difficulty enabled the 3′-exonuclease to compete more effectively with the polymerase reaction to enhance proofreading. However, L141 antimutator polymerase was found to have a much higher nuclease activity than wild type on single stranded DNA in the absence of dNTP substrates.[5,43] Perhaps there is an additional contributing factor to the high replication fidelity associated with the mutant *tsL141* bacteriophage, an impediment to forward translocation that acts in concert with a highly active proofreading exonuclease.

Proofreading by exonuclease may become more proficient when there is an inhibition in polymerization or forward translocation, caused either by an intrinsic fault in the polypeptide, or by a diminished dNTP substrate pool.[25,26] Comparing polymerases from mutator L56, T4 wild type, and antimutator L141, Clayton et al.[25] showed that the large differences in incorporation fidelity and nucleotide turnover present at saturating concentrations of dNTP substrates, disappeared at low [dNTP]. The point is that while there may be many independent ways to alter the absolute synthetic rates of the excision and insertion steps, the exo/pol ratio proposed by Bessman remains an important parameter in fidelity.

In Section IV.D (Figure 17) we mention an interesting T4 mutant polymerase isolated by L. Reha-Krantz[61,75] that appears to have a normal exo/pol ratio, but has a moderately strong mutator phenotype. The purified polymerase appears to exhibit an aberrant translocation property that enables it to bypass DNA template lesions with abnormally high efficiency.[76] Clearly, there is still much to be learned concerning how proofreading and polymerization are balanced to optimize DNA synthesis fidelity, and, as suggested recently by von Hipple and co-workers,[23] a deeper understanding of proofreading mechanisms awaits future presteady state kinetic analyses.

D. Extension of Mismatched Base Pairs

In the previous section, we discussed how DNA synthesis fidelity may be modulated by competing reactions, nucleotide insertion, and exonucleolytic proofreading. An important factor

that couples the two reactions is the ability of polymerase to elongate primers containing mismatched 3′-termini.

1. General Properties of Primer Mismatch Extension

Compared with nucleotide insertion efficiency, f_{ins} given by Equation 1, the corresponding expression for mismatch extension efficiency, f^0_{ext}, is more complex.[50]

$$f^0_{ext} = \frac{V_{max,W}}{V_{max,R}} \times \frac{K_{D,R}}{K_{D,W}} \times \frac{(K_{m,R} + p_R[S])}{(K_{m,W} + p_W[S])} \quad (9)$$

By definition, f^0_{ext} is the relative rate of adding the next correct nucleotide onto a mismatched compared with a correctly matched primer terminus, where $V_{max,R}$ and $V_{max,W}$ are the maximum velocities for addition of the next correct dNMP to matched and mismatched primer termini at **equal** concentration. The parameters p_R and p_W are the catalytic rate constants for insertion divided by the polymerase dissociation rate constants for matched and mismatched primer termini, and $K_{D,R}$ and $K_{D,W}$ are the equilibrium constants for polymerase disassociation from matched and mismatched primer-template termini. [S] is the concentration of next correct dNTP.

Diagrams of the competition events for f_{ins} and f^0_{ext} are shown in Figures 7a and b. In the case of nucleotide insertion, right and wrong dNTP substrates are competing for the same primer terminus (Figure 7a), while for extension, right and wrong primer termini are competing for the same correct dNTP to be inserted at the next template site (Figure 7b). Unlike the expression for f_{ins}, which does not depend on [dNTP] or polymerase-DNA dissociation constants, K_D, the expression for f^0_{ext}, Equation 9, depends explicitly on substrate concentrations and K_D for R and W termini.[50]

a. Binding Affinities to Matched and Mismatched Primer Termini

We recently used AMV RT to measure values for f^0_{ext}[21,50] and binding constants[21] for all combinations of matched and mismatched primer-template termini. We and others have found that K_D values were roughly similar for both matched and mismatched termini,[21,77] usually within a factor of ten, while mispair extension efficiencies are generally less than correct pair extension efficiencies by 10^3 to 10^6. Thus, in most instances, inefficient elongation of mispairs is caused by a kinetic block[21,48–52,77] as opposed to a more rapid dissociation of the polymerase from a mispaired terminus.

b. Mispair Extension Efficiency Dependence on Next Nucleotide Concentration

Equation 9 can be used to make a plot of f^0_{ext} as a function of the concentration of next correct dNTP (Figure 8). The relative efficiency for extending mismatches decreases dramatically with decreasing [dNTP]. Maximum discrimination is achieved in the limit [dNTP] $\rightarrow$ 0, in which case $f^0_{ext} \rightarrow f^0_{min}$. The latter is the ''intrinsic'' mismatch extension efficiency,[50] obtained by extrapolation to [dNTP] = 0,

$$f^0_{min} = [(V_{max,W}/K_{m,W})/(V_{max,R}/K_{m,R})] \times (K_{D,R})/(K_{D,W}) \quad (10a)$$

In case $K_{D,W} \sim K_{D,R}$,

$$f^0_{min} \sim (V_{max,W}/K_{m,W})/(V_{max,R}/K_{m,R}) \quad (10b)$$

The intrinsic f^0_{ext} value, f^0_{min}, is analogous to the misinsertion efficiency, f_{ins}, Equation 1, in the sense that both are defined by V_{max}/K_m for reactions involving wrong compared with right substrates. In both cases, V_{max} and K_m represent the so-called true values, i.e., calculated from extrapolating to infinite primer-template DNA concentration. However, there is a subtle, but important, difference in the interpretation of V_{max}/K_m for the insertion and extension reactions.

Nucleotide insertion proceeds by an ordered bisubstrate reaction, where polymerase binds the primer template first, and dNTP substrates bind subsequently to the polymerase-primer-template complex.[78,79] The ''apparent'' V_{max}/K_m for the reaction $(V_{max}/K_m)_{app} = (V_{max}/K_m)_{true}/(1 + K_D/[D])$. The misinsertion efficiency, f_{ins}, can be

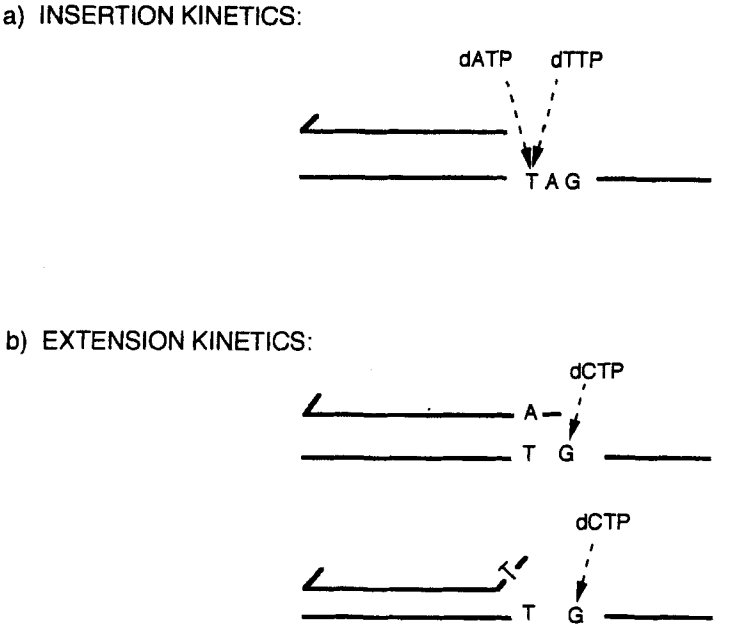


FIGURE 7. Competition events for nucleotide insertion and mismatch extension. (a) Insertion kinetics with right (dATP) and wrong (dTTP) substrates competing for addition onto the same 3′-primer terminus. (b) Extension kinetics for addition of the "next correct" (dCTP) nucleotide onto either matched or mismatched primer 3′-termini.

calculated using any convenient primer-template concentration (saturating [DNA] is generally used); the DNA dependent, $1 + K_D/[D]$ term, cancels when taking the ratio in Equation 1, because the polymerase is bound at the same primer-template site when right and wrong dNTP substrates are competing for insertion (see Figure 7a).

However, the DNA dependent terms do not cancel in the expression for f^0_{ext} unless the binding constants for matched and mismatched primer termini are about equal. Fortunately, this condition appears to be satisfied for all of the current K_D measurements. If future studies using different polymerases do reveal significant differences in K_D values, we have shown that intrinsic f^0_{ext} values, f^0_{min}, can be measured accurately by Equation 10a when $[D] < K_D$, i.e., when nucleotide insertion velocity varies linearly with DNA concentration.[21,50]

c. Application to Single Nucleotide Discrimination Using PCR

An important technique that exploits PCR methodology[80,81] is the selective amplification of a desired DNA molecule in the presence of another DNA molecule differing from the first by one or more single base changes.[82–88] In one version of this technique[83] amplification occurs from primer templates that are either perfectly matched or from ones that contain a single base mismatch at primer 3′-terminus corresponding to the single base change in the two templates; in other versions multiple mismatches are present at the 3′-terminus.[88]

The analysis for f^0_{ext} described in the previous section (Equation 9, Figure 8) can be used as a practical model to predict conditions to achieve optimized allele-selective amplification.[21,50,87] Selective amplification of an allele during the first PCR cycle should, in accordance with Equation 9, increase as the concentration of next correct dNTP decreases. An experiment by Ehlen and Dubeau[83] confirmed this point, because selectivity went from greater than 10^5 at low [dNTP] to no measurable amount at high [dNTP]. Because *Taq* polymerase has no associated proofreading activity, it has been used to carry out allele-selective amplification at high temperatures. An important feature of the *Taq* polymerase is that, in comparison with other polymerases

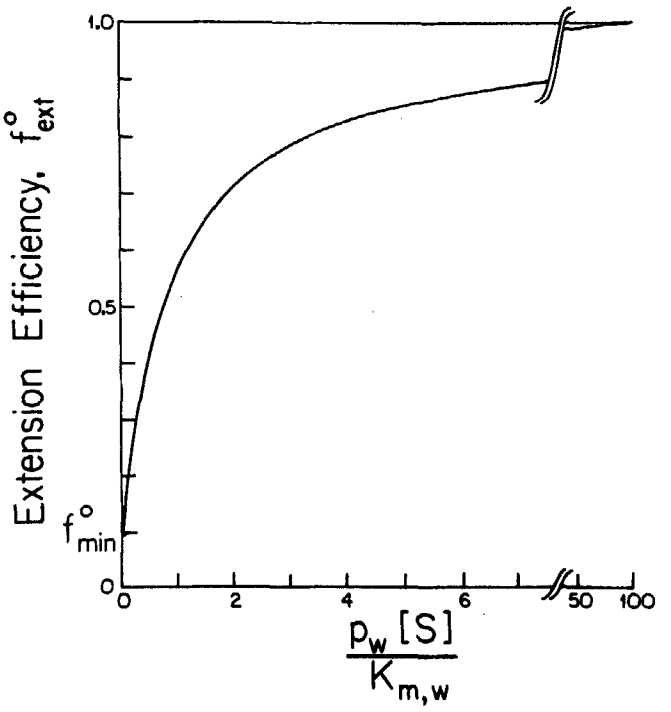


FIGURE 8. Predicted change in mismatch extension efficiency as a function of normalized substrate concentration.[60] In graphing Equation 7 (see text), we have assumed for simplicity that the concentrations of DNA with matched and mismatched primer termini are the same, and that their equilibrium dissociation constants are the same. f^{o}_{min} has been set equal to 0.1, corresponding roughly to extension of G · T mismatches by AMV RT. f^{o}_{min} is the minimum value of f^{o}_{ext}, when [S] → 0. p_w is the ratio of the polymerase rate constant, $k_{pol,w}$ to the rate constant for dissociation from a mismatched primer terminus. Note that as concentration of the next correct nucleotide, [S], increases, mismatched termini are extended with efficiencies that approach those for correctly paired termini, a prediction that was borne out in an experiment by Ehlen and Dubeau.[83]

(see Section II.D.2 following), it exhibits extremely low mismatch extension efficiencies for most of the common mismatches.[89,90] *Taq* extends most mismatches with efficiencies in the range of 10^{-3} to 10^{-5}, and values $<10^{-6}$ for extension of pur · pur and C · C mismatches. As a function of temperature, between 45°C and 70°C, the apparent V_{max}/K_m values measured for *Taq* polymerase, increased in parallel for extension of matched and mismatched primer termini.[89] Thus, the low values observed for f^{0}_{min} (Equation 10b) appear to be approximately temperature independent.

2. Promiscuous Mismatch Extension by HIV-1 and AMV Reverse Transcriptase

The virus believed to be responsible for AIDS, HIV-1, is believed to mutate rapidly. An important component in the mutagenic behavior is the viral encoded reverse transcriptase, HIV-1RT. The

enzyme contains an associated RNase H activity, but there is currently no evidence for the presence of 3′-exonuclease proofreading.

Nucleotide insertion kinetics at steady state[19,20,50,91–96a] and presteady state[97,98] have been investigated for HIV-1 and AMV RT. Relative efficiencies of nucleotide misinsertion and mismatch extension on DNA and RNA templates were investigated using the gel assay,[19,20,50,91,92,96a] and mutational spectra were determined on both templates by an *in vitro* transfection assay.[51,91,99–102]

Misinsertion efficiencies for these reverse transcriptases fall within a general range of f_{ins} $\sim 10^{-3}$ to 10^{-5}.[19,20,50,51,91,92,99–101] The values depend on the type of mispair formed, e.g., insertion of T opposite template G is usually considerably easier than G or A opposite G, and also on the surrounding sequence context.[19,20] These f_{ins} values do not appear to be systematically greater than for eukaryotic DNA pol α, having no associated exonuclease activity.[19,50,103] Indeed, AMV RT was generally observed to have lower f_{ins} than *Drosophila* pol α.[19,50] However, as observed by Kunkel and co-workers, the mutagenic spectrum of HIV-1 RT contained mutational hot spots attributed to primer-template misalignment, whereas AMV RT did not.[101]

Although f_{ins} values for reverse transcriptases appear "normal", f^0_{ext} values do not. Both HIV-1 RT and AMV RT show an ability to extend most mismatches with high efficiencies. For example, AMV RT extended a G(primer) · T(template) mismatch with about 18% efficiency compared with A · T and a T · G mismatch with 2.5% efficiency relative to C · G![50] The f^0_{ext} values for AMV RT[50] are one to three orders of magnitude larger than corresponding values for *Taq* polymerase.[89]

A plot comparing the relative efficiencies of mismatch extension, $f^0_{ext} = f^0_{min}$ evaluated by Equation 10b (*y-axis*), and f_{ins} by Equation 1 (*x-axis*) for AMV RT and pol α from *Drosophila* and calf thymus is shown in Figure 9.[50] Most of the data points for AMV RT fall above the diagonal signifying that mismatch extension efficiencies are larger than misinsertion efficiencies. In the cases of G · T and T · G, f^0_{ext} is between 3 and 4 logs greater than f_{ins}. In contrast, the two sets of pol α data are scattered on both sides of the diagonal.

A comparison of HIV-1 RT nucleotide misinsertion and mismatch extension efficiencies with those of AMV RT, using DNA and RNA templates, showed that the two reverse transcriptases have similar general properties in forming and extending single base mismatches.[20] As in the case of AMV RT, HIV-1 RT also extended mismatched termini with high efficiency. An example of the impressive mispair extension ability of HIV-1 RT is shown in Figure 10.[20]

In Figure 10a, addition of 13 nucleotides to a primer 3′-terminus are visible on the gel. These include a run of three consecutive mispairs, C · A, A · C, C · A. Typically, there are strong pause bands visible at primer-template sites containing mispairs. However, for HIV-1 RT, the low intensity of the bands representing the two mispairs following the initial C · A mispair illustrates the ease with which the enzyme can extend the A · C and C · A mispairs. This extension without pausing is unlikely to be the result of deamination of C to U, because the concentration of template or substrate deamination products are negligible in the assay. Also, AMV RT shows strong pause bands at the same mispaired primer-template sites bypassed by HIV-1 RT. A second illustration of the absence of pausing is shown in Figure 10b, where there are not visible pause bands at either T · T or T · C positions.

It may be biologically significant that extension efficiencies are largest for G · T and T · G, that is, for the same mispairs that are easiest to form. One might expect, therefore, to find these two mispairs overrepresented in base substitutional mutational spectra *in vivo*. Use of an *in vitro* transfection assay has shown that a significant class of HIV-1 RT mutational hot spots, both base substitutions and simple frameshifts, arise from transient misalignments.[101] However, it was also shown that AMV RT is unable to catalyze mispairs using a slippage or misalignment mechanism.[101] In Section IV.D. we discuss findings that HIV-1 RT, but not AMV RT, can efficiently bypass abasic template lesions using misaligned primer-template termini. A possible model to explain the exceptionally high rate of

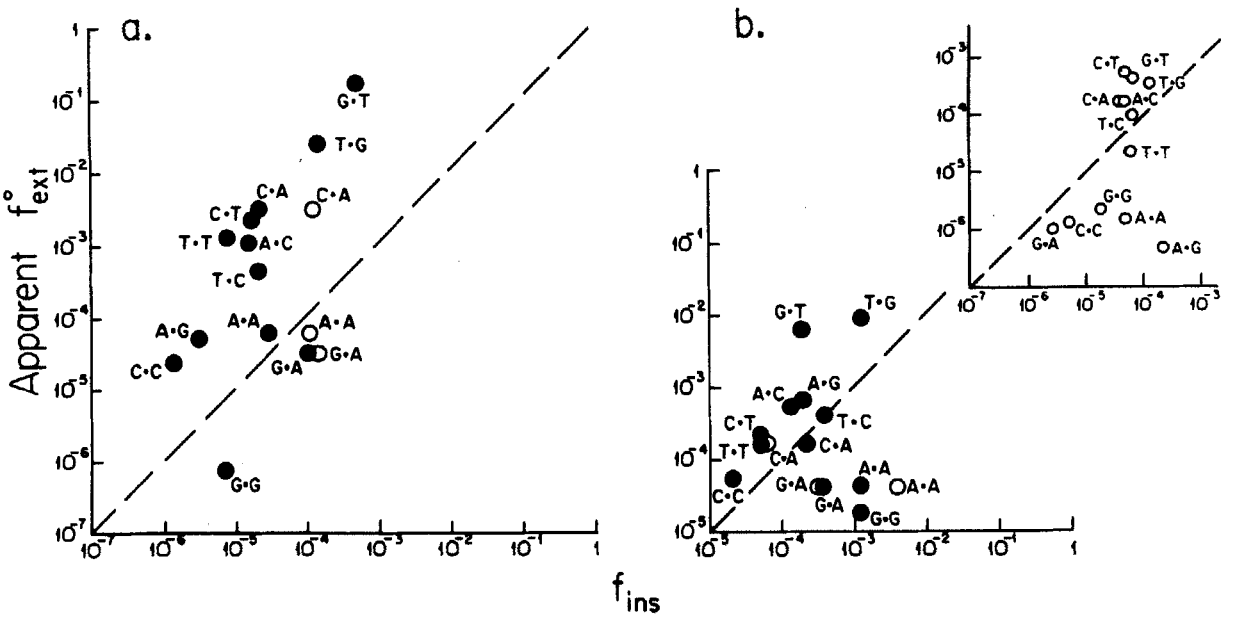


FIGURE 9. Comparison of extension efficiencies and insertion efficiencies for AMV reverse transcriptase and pol α from *Drosophila* and calf thymus.[50] The f^{o}_{ext} value determined for each mispair is plotted against the corresponding f_{ins} value for AMV RT (a) and pol α (b) in cases having the same primer 5′-nearest neighbor (A or G). The dashed line corresponds to $f^{o}_{ext} = f_{ins}$. Closed circles indicate f_{ins} values determined in "running start" reactions (Figure 1a) where polymerase inserts two correct nucleotides prior to the mispair. Open circles are the corresponding "standing start" reactions (Figure 1b) in which the mispair is made directly off the primer end. All f_{ins} values for calf thymus pol α (inset) are obtained in standing start reactions with a primer terminus having nearest neighbor A.[49]

HIV-1 mutagenesis is that HIV-1 RT is able to catalyze nucleotide addition at distorted termini that cannot be used by cellular polymerases or even by non HIV-1 reverse transcriptases.

III. SEQUENCE CONTEXT EFFECTS ON POLYMERASE DISCRIMINATION

As a replication fork proceeds along a DNA template strand, it encounters a variety of DNA sequences. There are different mechanisms to account for base sequence perturbations on fidelity. In a model proposed by Streisinger,[104] frameshift mutations can occur in homopolymer template regions by a strand slippage or displacement mechanism. Kunkel and collaborators[105–107] have shown how small displacements of one or a few nucleotides can stimulate both base substitution and frameshift mutational hot spots.[107a] Ripley and co-workers[108–110] have shown that larger displacements giving rise to hairpin loops can also generate base substitutions and frameshifts.

A. Effects of Sequence and Polymerase Characteristics on Insertion Fidelity

Mutational hot and cold spots in DNA are dependent on surrounding sequence. Therefore, it is necessary to study interactions between enzyme, primer-template, and dNTP substrates to gain an understanding of molecular events that determine polymerase error rates. There are several experiments suggesting that polymerases may be directly in contact with about five to eight bases in primer template.[111,112] It might be expected that local averaging over this many bases might act to dampen differences in nucleotide insertion fidelities. However, an example of nucleotide misinsertion rate dependence on neighboring sequence and enzyme identity is shown in Figure 11.[19] The enzymes used, *Drosophila* pol α and AMV RT, contain no detectible 3′-exonuclease activity.

At two separate template sites pol α was observed to incorporate dAMP opposite A (Figure 11a, left side), while AMV RT is able to form A • A mispairs at the first but not the second site

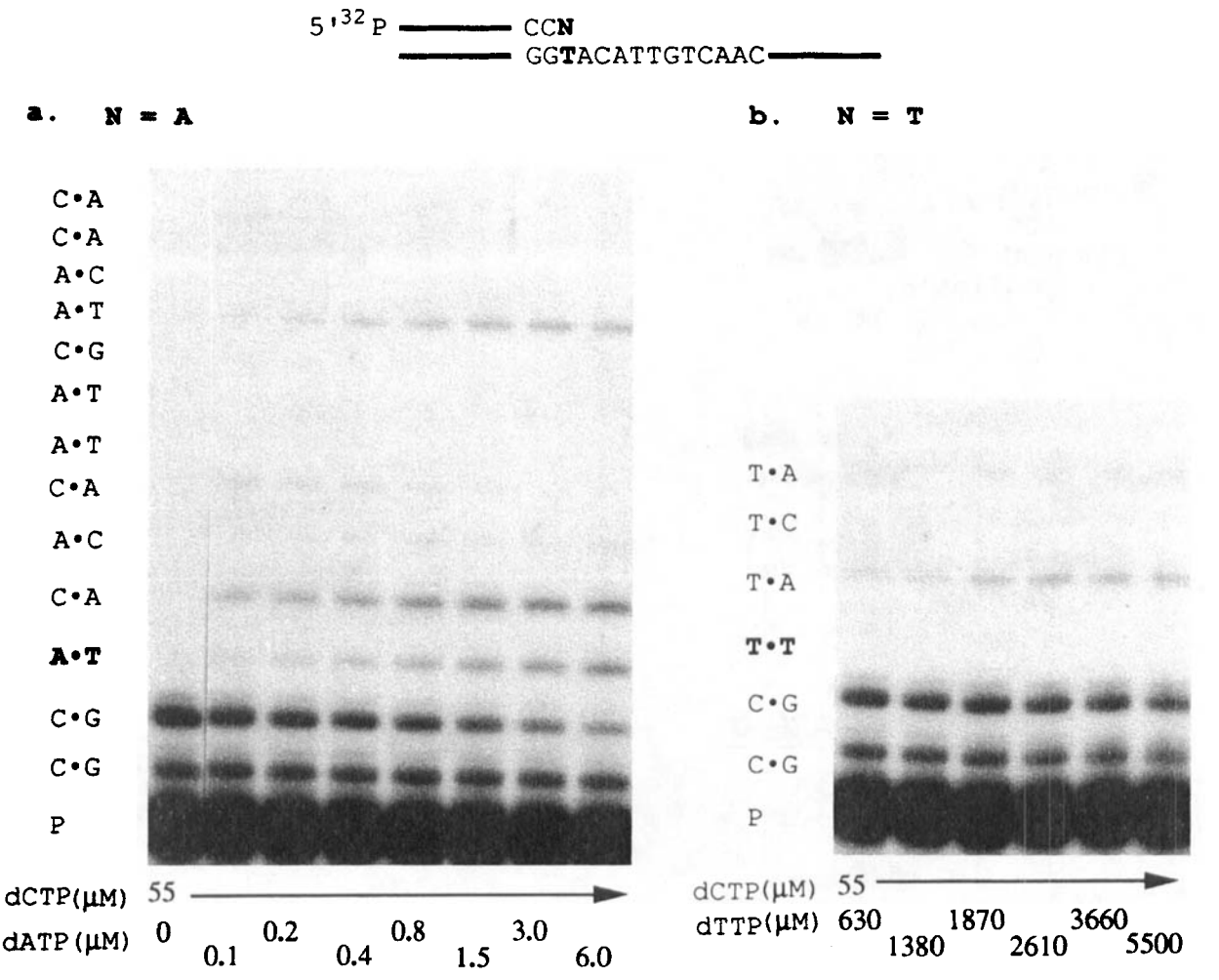


FIGURE 10. Gel autoradiograms showing HIV-1 RT catalyzed primer extensions forming correct A · T and incorrect T · T base pairs.[20] The variable dNTP substrate concentration used for insertion opposite the M13 DNA template target site is shown below the original primer band (p) in each lane. (a) the correct incorporation of A opposite T is followed by formation of three consecutive mispairs and four correct base pairs. (b) the incorrect incorporation of T opposite T is followed by a correct base pair, an incorrect base pair, and another correct base pair. A fixed, saturating, [dCTP] = 55 μM is used for elongation of the primer P to reach the template target site T. The 0 μM dTTP control is shown in the first gel lane.

(Figure 11a, right side). At a different set of template sites, pol α was able to incorporate dGMP opposite G (Figure 11b, left side), whereas AMV RT barely incorporated G opposite G at the first site and there was no detectible G misincorporation at the second site (Figure 11b, right side). These data illustrate a second important point, namely, the dependence of base mispair extension on surrounding sequence. The sites designated as p5 and p8 (Figure 11b) both contain G followed by a C so that in the presence of dGTP, the initial G · G mispair could have been extended by addition of the next correct base to form a G · C base pair. Pol α is able to catalyze mispair extension at site p8, but not at site p5, even though at p5 initial formation of the G · G mispair is catalyzed more efficiently (Figure 11b).

The data in Figure 11 provide a clear qualitative illustration of the fact that there can be significant effects of sequence context on nucleotide misinsertion and mismatch extension, and that these effects can differ depending on polymerase identity. The effect of sequence context on the efficiency of forming the four transition and eight transversion mispairs was investigated with pol α and AMV RT by measuring f_{ins}, Equa-

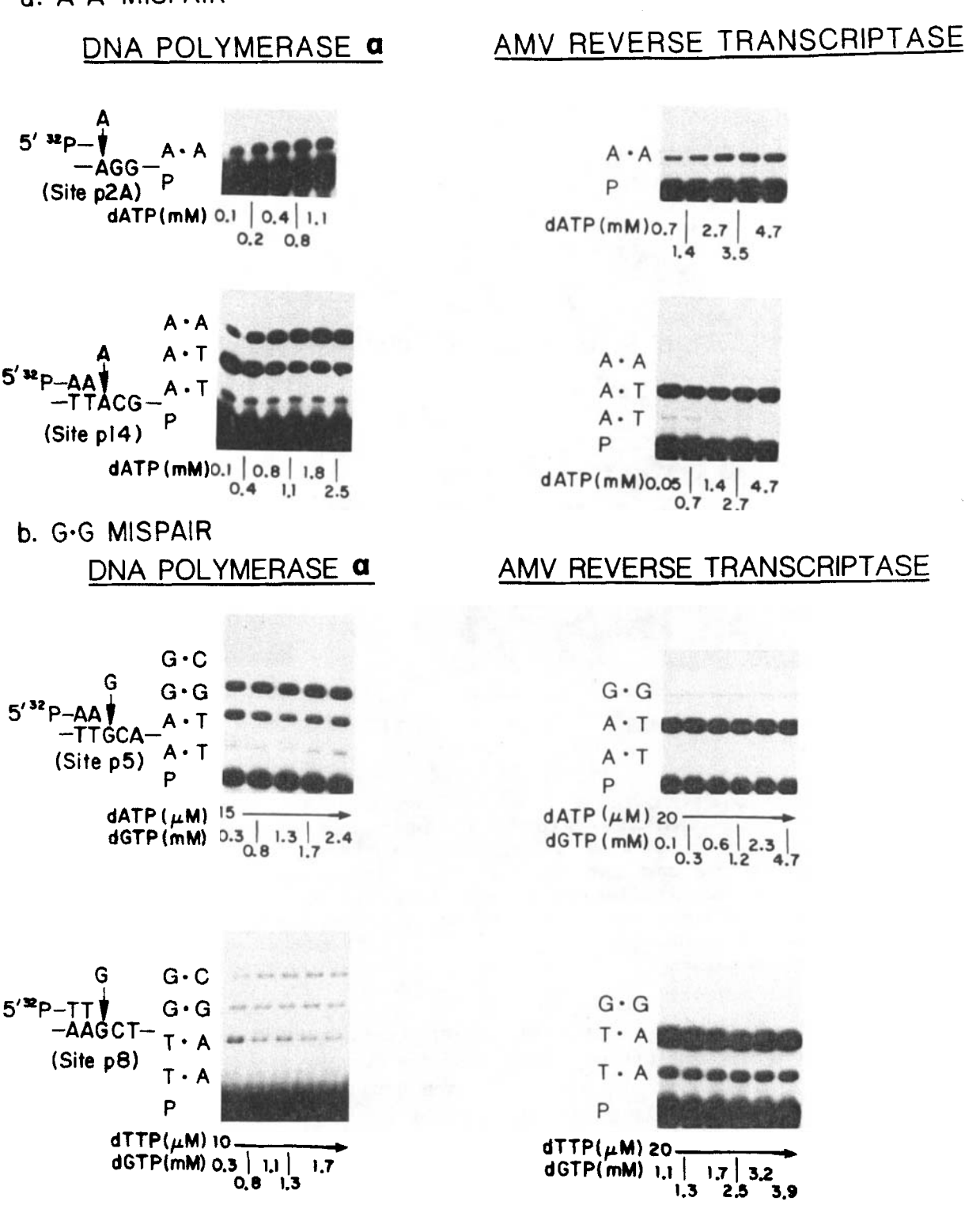


FIGURE 11. Gel autoradiograms showing primer extensions with mispairs of type A · A or G · G formed by increasing substrate concentrations.[19] In (a) the misinsertion of dAMP opposite template A (indicated by arrow) is observed for pol α at sites p2A and p14 and for AMV RT at site p2A. At site p14, the A · A misinsertion by AMV RT is much less frequent and can be seen only in autoradiograms with longer exposure times. In (b) the corresponding results for misinsertion of dGMP opposite template G at sites p5 and p8 are shown. Only slight misinsertion by AMV RT is found at site p5 and no detectable misinsertion at site p8.

tion 1, using at least three separate template target sites for each mispair. For every mispair, both enzymes showed at least a 5-fold effect and, more generally, a 10- to 50-fold range of misinsertion efficiencies.[19] Singer and collaborators[113,114] and Topal and co-workers[115] have made a similar series of observations of localized sequence effects on formation of base pairs using the base analogue O^6-methyl G.

B. Model Systems Using 2-Aminopurine to Investigate Sequence Context Effects on Fidelity

Bessman developed an exceptionally useful model system to study DNA synthesis fidelity using the mutagenic nucleotide analogue, 2-aminopurine deoxyribonucleoside triphosphate (dAPTP) as a substrate for bacteriophage T4 mutator, wild-type, and antimutator DNA polymerases.[14,65,116–118] AP can form normal basepairs with T and base mispairs with C to drive bidirectional A · T ↔ G · C transition mutations.[119–125] An important advantage of the system is the easily measurable insertion efficiencies of the analogue as a nucleotide substrate. In competition with dATP for insertion opposite template T, f_{ins} (AP · T) ~14%;[14,25,125,126] in competition with dGTP for insertion opposite C, f_{ins} (AP · C) ~1%;[127] and when AP is present on the template, f_{ins} (C · AP) ~5%,[122,125,126,128–130] representing insertion of dCMP vs. dTMP opposite AP. These high-insertion efficiencies are in contrast with values on the order of 10^{-3} to 10^{-6} for naturally occurring base mispairs.[13,19,52,77,131–133] Although AP · T base pairs are stabilized by two H-bonds in a normal Watson-Crick configuration,[134] they are considerably less stable than A · T base pairs based on differences in melting temperatures,[135] and are removed preferentially by proofreading exonucleases.[14,25] A description of base-pairing structures involving AP as well as other mispairing structures is presented in Section IV.A.

1. Incorporation of AP Deoxyribonucleotides at Many Template T Sites in vitro

Pless and Bessman performed an experiment in which equimolar concentrations of dAPTP and dATP competed for incorporation opposite T sites on ΦX174 DNA template, using T4 wild-type and *tsL141* antimutator T4 DNA polymerases. Measurements of AP and A incorporation were made at 57 different template T sites.[118] The most striking result was that the range of AP/A incorporation ratios varied, for the antimutator polymerase, from a high of 20 and 14% to a low of 0% for 5′-neighboring pyrimidines T and C, respectively, and from about 7 to 0% for neighboring purines.

These data using *tsL141* antimutator polymerase were extensive enough to investigate how perturbations in AP insertion and excision efficiencies correlate with sequence context.[136] For nucleotide insertion, it appeared that AP competed with A more effectively when pyrimidines were present as primer 3′-termini. Base-stacking effects appeared most pronounced for 3′-primer (i.e., 5′-neighbor) T; f_{ins} (AP) was predicted to be about 35% when stacking occurred next to T.[136] In contrast, insertion of the analogue next to G or A appeared to occur four-to fivefold less efficiently, f_{ins} ~7 to 8%. The highly active *tsL141* proofreading exonuclease[5,43] appeared to exert a dominant effect on the relative fraction of AP incorporation opposite T. Here the data showed no correlation between nearest-neighbor stacking partner and removal of AP. However, there was a marked correlation between editing and the average DNA stability in the ten-base region (one helical turn) surrounding AP.[136] Exonucleolytic editing appeared most effective in AT-rich and least effective in GC-rich regions, in accordance with an earlier report of Bessman and Reha-Krantz.[69]

Nucleotide insertion efficiencies are likely to be sensitive to base stacking forces between incoming dNTP substrates and primer-template termini at the replication fork. Base-pairing free energies (including base stacking and H-bonding) between incoming dNTP substrates and at the primer terminus should affect the orientation and residence time of right and wrong substrates in the polymerase active cleft. However, it is not the absolute magnitudes of the stacking interactions that should affect insertion fidelity. Rather, it is the *differences* in base-pairing free energies for matched vs. mismatched dNTP substrates that is expected to determine f_{ins}.[25,48,125,130,137,138] It is possible that the relative interactions between in-

coming dNTPs and the base at the primer 3′-terminus act as a major determinant of insertion specificity, and extension may also be influenced by 5′-neighbors.

Forces that tend to stabilize or destabilize newly formed matched or mismatched base pairs should also affect proofreading. However, because X-ray crystallographic data on Klenow fragment[59,60] and kinetic[23,64,139] and sequence comparison data[61,140–148] corresponding to a variety of other polymerases demonstrate that polymerase and exonuclease activities exist either as separate polypeptides or as structurally distinct domains on a single polypeptide, melting of two or more base pairs[23,63,65,66] may be required to transfer the DNA from polymerase to exonuclease sites. Therefore, effects of sequence context on the proofreading component of fidelity is likely to be affected to a considerably greater extent by changes in the average stability of groups of base pairs in the vicinity of the polymerase cleft rather than by the relative stability of a single stacked doublet containing either a correct or incorrect base pair at the primer terminus.

2. Fluorescence Detection of 2-Aminopurine Insertion

Free dAPTP exhibits strong fluorescence emission at 350 nm, which is quenched when incorporated into DNA[149,150] A reduction in dAPTP fluorescence can be measured during DNA synthesis to follow the kinetics of incorporating dAPMP into DNA. In a series of experiments that we have begun in collaboration with Beechem and co-workers at Vanderbilt University, AP insertion kinetics were measured using exo$^-$ Klenow fragment. The main experimental objective was to determine whether there are significant differences in the velocities of AP incorporation next to different nearest-neighbor base-stacking partners located at the primer 3′-terminus. Previously, an analysis of the AP insertion component[136] using experimental data from Pless and Bessman[118] indicated that AP insertion rates differ depending on whether insertion occurs next to purines or pyrimidines. These results were for T4 L141 polymerase; Pless and Bessman also used Klenow fragment, but did not see as much variation from site to site.

Insertion of dAPTP opposite T has been measured on synthetic primer templates of identical sequence except for the base at the primer terminus (Figure 12). AP insertion rates appear greatest next to G, C, and A, and smallest next to T. The largest difference in insertion rates is about a two- to threefold reduction when T is situated at the primer 3′-terminus compared with G. The rate of AP insertion by the Klenow fragment correlates with the thermodynamic stabilities of DNA doublets containing AP · T base pairs derived from DNA melting data.[135,136,151] However, speculations concerning the general significance of this observation should be postponed until future measurements are made using different sequence contexts and other polymerases.

The key point is that AP fluorescence measurements are extremely sensitive and provide direct measurements, in real time, on catalytic rates for AP incorporation into DNA and for exonucleolytic removal as dAPMP from 3′-termini. Using a stopped flow syringe system, fluorescence measurements can be made from a few milliseconds (presteady state) to several minutes (steady state). The fluorescence measurements in the presteady state kinetics region and the running start gel assay are directly comparable, because both assays measure the rate of extension of the primer from one site to the next.

3. Sequence Context Effects on Polymerase Dissociation Rates and Proofreading Fidelity

In order to calculate absolute (as opposed to relative) nucleotide insertion rates by the gel assay, it was necessary to measure the first-order dissociation rate constant of the enzyme from the template $T-1$ position, prior to insertion at the target site, (see Equation 4 and Figure 5). For AP insertion by exo$^-$ Klenow fragment, we found that k_{off} was about sevenfold faster when G was present in place of T at the primer terminus. Thus, the change of just a single base at the primer terminus can cause a significant alteration in the rate constant governing polymerase dis-

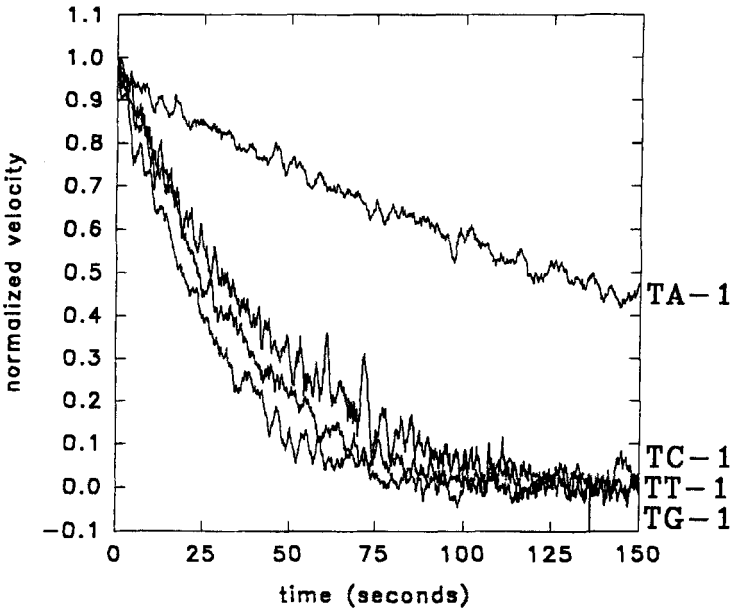


FIGURE 12. Reaction time courses showing decrease in fluorescence as dAPMP is incorporated into primer/templates having identical sequence except for the 5′-nearest neighbor base stacking partners. Reactions were performed at 20°C using 5 n*M* Klenow exo⁻ (D355A, E357A), 70 n*M* primer/template, and 0.2 μ*M* dAPTP. The primer/templates associated with each time course are indicated in the figure and have the following sequence: 5′-Y

3′-Z TAG

where Y · Z = T · A (TA – 1), G · C (TC – 1), A · T (TT – 1), and C · G (TG – 1).

sociation from primer-template DNA. Given a five- to eight-base region over which the polymerase cleft is in contact with DNA,[111,139] a single base change at a primer terminus can result in a binding free energy difference of about 1 kcal/mol.

If, at some sites, polymerase dissociation can compete with forward translocation, then an order of magnitude reduction in polymerase binding could affect fidelity by reducing the probability to excise a misinserted nucleotide during a processive synthesis mode. If there are differences in the probability to excise a terminal mismatch from running vs. standing start reactions, then sequence context effects on polymerase dissociation may be responsible for hot and cold base substitution via their effects on proofreading.

IV. STRUCTURAL FEATURES OF BASE MISPAIRS RELATING TO DNA SYNTHESIS FIDELITY

In the two 1954 articles by Watson and Crick describing the structure of DNA, the question of spontaneous errors in DNA was addressed.[1,2] It was proposed that base mispairs would occur during DNA synthesis when the common bases were present as disfavored tautomers. A disfavored imino tautomer of either adenine or cytosine could form a double H-bonded Watson-Crick structure with a favored amino tautomer of its partner resulting in an A · C mispair. Also, a disfavored enol tautomer of either guanine or thymine could form a triple H-bonded Watson-Crick structure with a favored keto tautomer of its partner resulting in a G · T mispair. Despite lack of any

supporting evidence, these structures are presented in genetics and molecular biology textbooks as the explanation for the occurrence of base substitution transition mutations. To account for transversion mutations, Topal and Fresco[152] proposed a model in which one of the bases in a mispair is a disfavored tautomer in a *syn* conformation.

A. Evidence for Ionized and Wobble Base Mispairs

We are unaware of any experiments demonstrating the involvement of disfavored tautomers in forming base mispairs. Two lines of evidence, NMR studies and X-ray crystallographic data, suggest that in duplex DNA, both base mispair partners are present as favored amino or keto tautomers. Examples include: (1) A · C mispairs[153] stabilized by two H-bonds in a protonated wobble conformation,[154–156] which undergoes a transition at pH above 7.5 to form a reverse wobble containing a single H-bond[157]; (2) wobble G · T mispairs[154,158,159]; (3) wobble G · A mispairs in which either both bases are *anti*[158,160] or are *G(anti)* · A*(syn)*[161] or *G(syn)* · A*(anti)*[162]; (4) 2-AP · C mispairs present either as a protonated Watson-Crick[134] or neutral wobble structure[163]; and (5) 5-B · G and 5-F · G that exist in a pH-dependent equilibrium between ionized Watson-Crick and neutral wobble structures.[164,165] Structures of the base mispairs mentioned above are shown in Figure 13.

B. Formation of Base Mispairs by Polymerase as a Function of pH

Based on thermodynamic considerations, Ramsay Shaw and colleagues have proposed that base mispairing may involve ionized base pairs.[166] Structural studies using physical measurements on duplex DNA containing internal base mispairs were in agreement with a model of ionized structures for base mispairs involving bromouracil as originally proposed by Lawley and Brooks.[167] However, it is still possible that such structures may not accurately reflect the state of purines and pyrimidines while incorporated as mispairs during DNA replication. It has been observed that when present in an *in vitro* DNA synthesis system either as a template base or as a deoxynucleoside triphosphate substrate, polymerases form B · G mispairs much more easily than T · G mispairs.[168–170] It was reported by Driggers and Beattie[171] that B · G mispairs were incorporated more readily by polymerase with increasing pH, suggesting that B · G mispairs can be formed with the anionic form of the analogue (Figure 13k) during DNA synthesis.

The equilibrium between keto and enol forms of bromouracil or fluorouracil is not dependent on pH. As pH is increased to produce an increase in the concentration of anionic B or F, a corresponding decrease occurs in the concentration of keto and enol forms. Accordingly, for B · G and F · G mispair frequencies, resulting either from insertion of 5-halodUMP opposite template G or dGMP opposite template 5-haloU, various predictions can be made: (1) if the ionized form of B or F can form mispairs with G either as a template or substrate nucleotide during DNA synthesis, then f_{ins} for haloU · G by Equation 1 should increase with increasing pH since the fraction of the ionized species increases with increasing pH; (2) if the enol form of haloU is required to form mispairs with G, then f_{ins} should decrease with increasing pH since the fraction of the neutral enol species decreases with increasing pH; and (3) the corresponding efficiencies of correct base pair formation, haloU · A should decrease with increasing pH since ionized B or F are not likely to stably base pair with A.

In recent experiments using AMV RT,[172] we have observed that the pH–rate (f_{ins} is a relative rate) profiles for B · G and F · G (Figure 14) appear to support a mechanism that involves incorporation of ionizated bases rather than disfavored enol tautomers since f_{ins} increases with increasing pH. Similar observations have also been made with B and F present on the template. However, the mispairing efficiencies, f_{ins} for B · G and F · G, are between one to two orders of magnitude higher when B and F are present on the template.[171,172] Although we cannot completely eliminate the possibility that ionization of an amino acid side chain in the enzyme active site contributes to the observed fidelity dependence on pH, the absence of a titration-like curve for

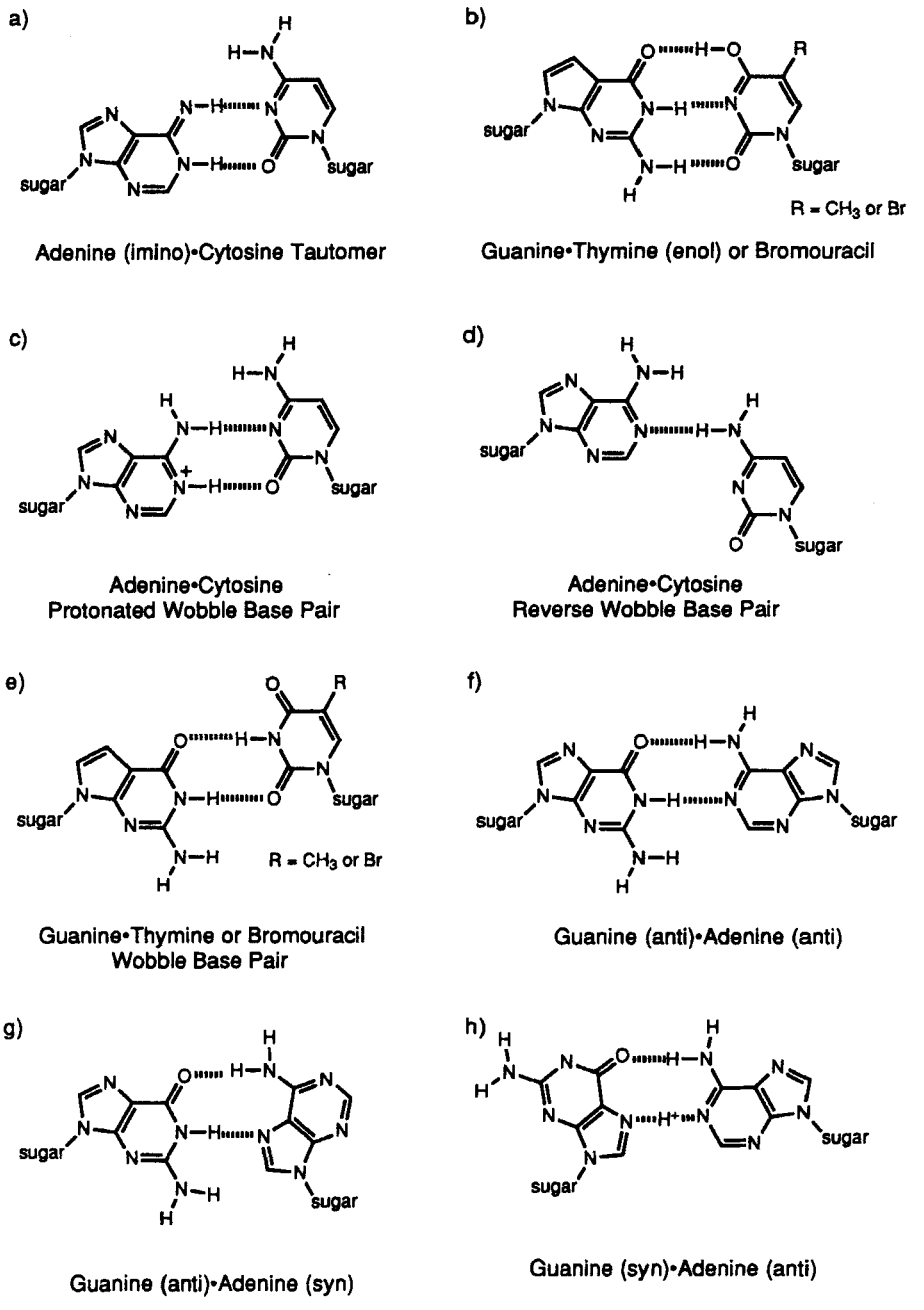
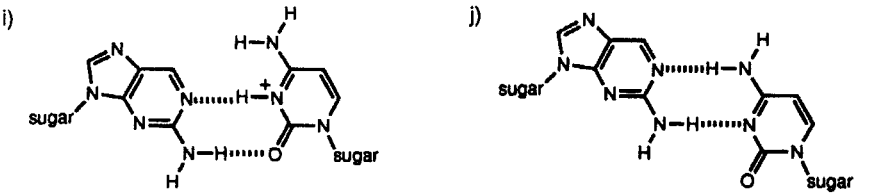


FIGURE 13. Proposed structures of base mispairs.

incorporating C opposite G supports the idea that DNA base ionization is the major cause of base mispairs during DNA synthesis. Thus, in light of the recent NMR, X-ray, and enzymatic studies on base mispair structures, it no longer seems reasonable to require involvement of imino and enol tautomers in spontaneous mutagenesis, although a role of disfavored tautomers cannot be formally eliminated.

C. An Error Avoidance Mechanism in *E. coli* Dependent on the Product of the *MutT* Gene

E. coli has a substantial number of genes devoted to cleansing the genome of errors made during DNA replication by pol III or during repair by pol I and possibly pol II. Mutations in these genes generally exhibit strong mutator pheno-



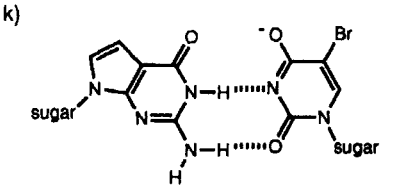


FIGURE 13 continued.

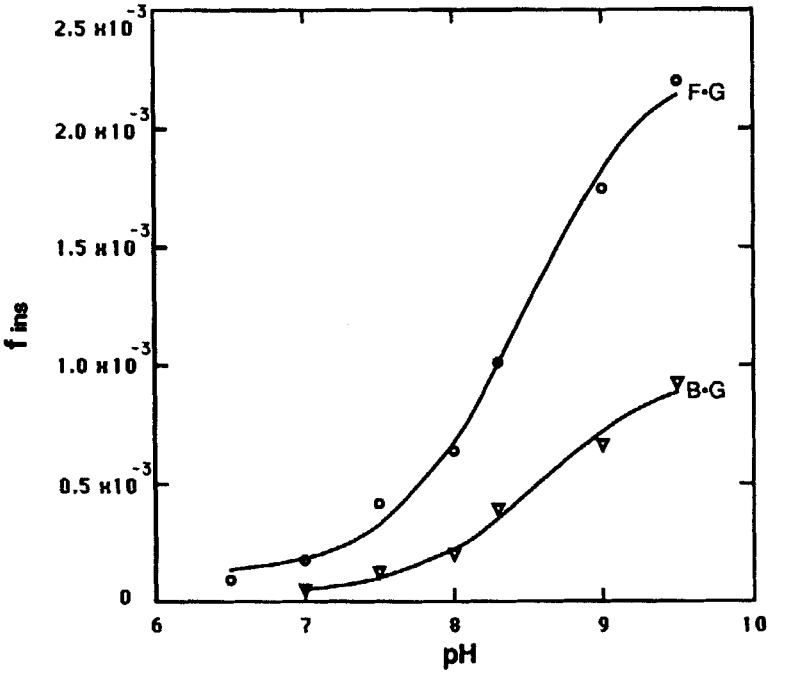


FIGURE 14. Misinsertion efficiency, f_{ins}, of FdUMP and BdUMP opposite template G as a function of pH using AMV RT.

types for specific mutational pathways. Examples of some well-characterized mutator gene products include: (1) *mutD*, the ϵ-subunit of the pol III holoenzyme complex responsible for proofreading at the replication fork;[6,7] (2) *mut H,L,S*, the system responsible for postreplicative methyl-directed mismatch repair[8-10] and the destruction of heteroduplex DNA resulting from interspecies recombination;[172a] (3) *mutY*, a glycosylase specific for removal of a misinserted A in A · G, A · C,[11,173] and A · 8-oxoG mismatches;[174] and (4) *mutM*, 8-oxyguanine DNA glycosylase,[175] that cuts duplexes containing 8-oxoG · C base pairs.[12] The adenine and 8-oxyG glycosylases function independently of the state of d(GATC) methylation and do not require *mutH*, *mutL*, *mutS*, or *mutU* gene products. All of these enzymes, methylation dependent and independent, share the common property of operating on DNA following an error.

Mutations at the *mutT* locus in *E. coli* result in 1000-fold increase specifically in the T → G transversion pathway.[33,176,177] The product of the *mutT* gene is a dGTPase.[178,179] In contrast to other mutator genes, by preventing formation of G · A mispairs,[180,181] the dGTPase apparently has an important role in error avoidance rather than in postreplication error correction. In a recent review, Reference 12a, it has been suggested that *mutT* is a part of a pathway, designated as the "GO system," which in *E. coli*, is composed of at least 3 proteins, *mutM*, *mutT*, and *mutY*.

Hydrolysis of 8-oxodGTP was observed to be 500-fold greater for the oxygen damaged substrate 8-oxodGTP compared with dGTP.[182] This interesting observation led Sekiguichi and coworkers to propose that the role of *mutT* gene product is to eradicate 8-oxodGTP from the nucleotide pool.[182] However, one might question the proposal on the grounds that it is unlikely that 8-oxodGTP pools are sufficiently large *in vivo* to mandate the presence of such a dedicated enzyme activity. There are many potential check points in both *de novo* and salvage pathways in deoxyribonucleotide biosynthesis that could serve, in principle, to eliminate oxidized nucleoside and nucleotide analogues before reaching the dNTP level. Bessman and co-workers[179] proposed a general role for the *mutT* dGTPase, in which this enzyme is required to hydrolyze *syn* isomers of dGTP and its analogues, because the *syn* isomer of G can form a relatively stable mispair with A.[162] Among the four common dNTPs, dGTP has the highest equilibrium concentration of *syn* isomers. Presumably, hydrolysis would have to be coupled to replication, because a rapid re-equilibration between *syn* and *anti* isomers would occur if the hydrolysis were occurring in the pool of free dGTP.

Because 8-oxodGTP is present predominately as a *syn* isomer, its elimination from deoxynucleotide pools may simply be a consequence of the more general requirement to exclude "normal" dGTP *syn* isomers. The observation by Sekiguichi and co-workers[182] that 8-oxodGTP is hydrolyzed much more efficiently than dGTP may reflect the fact that there is a much smaller *syn* component present in dGTP than in its oxidized analogue. A measurement of 8-oxodGTP pool size in *E. coli*, and a comparison of mutagenesis under aerobic and anaerobic conditions might help to clarify the role of the *mutT* dGTPase *in vivo*. However, it has recently been mentioned that "mutator strains grown aerobically or anaerobically have equal mutation rates,"[12a] suggesting, to us, that perhaps that *syn* G isomers other than 8-oxodGTP might be involved in the *mutT* mutagenesis pathway.

In contrast to the paucity of data on enzymes that might be present to protect dNTP substrate pools from contamination by exogenous damage, there are a wide variety of prokaryotic and eukaryotic enzymes available to excise DNA template lesions caused by radiation and chemical damage. One of these enzymes, the *mutM* gene product in *E. coli*,[12,175] identified originally as a formamidopyrimidine (FAPY) DNA glycosylase (*fpg* gene),[183] is able to remove 8-oxoG from DNA. When present on the template strand, 8-oxoG preferentially mispairs with A during replication causing G → T transversions.[184,185]

D. Copying DNA Containing Template Lesions

DNA lesions can arise spontaneously *in vivo* either from endogenous effects of temperature and pH, or from exogenous perturbations from many different sources. To cite a few familiar

examples, UV radiation causes cyclobutane dimers and 6-4 photoadducts; oxidative damage results in thymine glycol and 8-oxoguanine; alkylating compounds cause numerous base modifications including O^6-alkylguanine, O^2- and O^4-alkylthymine; AAF bulky adducts result from exposure of DNA to *N*-acetoxy-2-acetyl aminofluorine; abasic lesions occur as a result of glycosylases acting on either damaged or chemically modified bases or by spontaneous depurination and depyrimidination reactions. Eukaryotic and prokaryotic cells contain constitutive and induced enzymatic activities that repair specific types of DNA damage. For example, methyl transferases that repair O^6-MeG are expressed constitutively and can be induced in bacteria and animal cells.[186–188] SOS-induced DNA repair[189,190] is a negatively regulated pathway in *E. coli* consisting of more than 20 genes, under lex A repressor control.[190–193] SOS "error prone" repair is turned on in the presence of sufficient DNA damage to saturate the "error free" repair pathways,[194] permitting the cell to replicate past lesions, such as pyrimidine dimers and abasic sites, that would otherwise act to inhibit or block normal replication.[13,193,195–197a]

Lesions can be classified according to their coding or miscoding properties and according to the degree of DNA synthesis inhibition. Cyclobutane dimers, e.g., T^T, that continue to code for A insertion impose a strong[198,199] but not absolute barrier to bypass replication.[198,200,201] DNA containing thymine glycol, a nonplanar ring structure, appears to be nonmutagenic[202] but strongly inhibits replication.[203] Alkylated bases[204–210] and 8-OxoG[184,185] form mispairs *in vitro* and are mutagenic *in vivo* but are copied relatively easily.[113,114,211]

1. Insertion Opposite Abasic Lesions — Sequence Context and the "A-Rule"

Abasic (apurinic/apyrimidinic) lesions are important biologically and have served as an excellent model system to study the properties of noncoding lesions *in vivo* and *in vitro*.[212] *In vitro* measurements of nucleotide incorporation opposite abasic lesions (**X**) distributed randomly in DNA revealed a strong preference for incorporation of dAMP opposite **X**.[213–215] These results resulted in formulation of the A-rule,[216] which suggests that DNA polymerases tend to insert A, as a default mechanism, to continue synthesis in the absence of strong template coding signals. Because it is now straightforward to incorporate single abasic lesions into template DNA in a site-directed fashion,[17,135,217,218] it is possible to study abasic site mutagenesis and inhibition in defined sequence contexts.[17,197,218,219]

Drosophila pol α has been used in a running start gel assay to measure insertion of A, G, T, and C opposite a site-directed abasic lesion as a function of time (Figure 15).[17] The insertion efficiency, f_{ins}, for each of the four normal deoxynucleotide substrates opposite **X**, next to four different base stacking partners at the primer 3′-terminus, has been measured.[17] Insertion of A opposite **X** was found to be favored for each of the base stacking partners, by factors of about 6 to 10 over G and 20 to 40 over C and T. For inserting A opposite **X** where there was about a fivefold difference between the largest insertion efficiency, next to primer T, when compared with the smallest efficiency, next to primer G. Deoxynucleotide insertion data using HIV-1 RT were similar to that obtained with pol α. Sequence context effects on insertion opposite **X** was also observed in an *in vivo* transfection experiment in *E. coli*, where an abasic site was located in either of two locations.[197] At one location, the order was 80% A, 15% G, 4% T, and 1% C. At the other site, the order was 54% A, 25% T, 20% G, and 1% C.

2. Extension Past Abasic Lesions — Sequence Context and the "A-Rule"

In a recent set of experiments, HIV-1 RT and AMV RT were used in standing start gel assay to measure the effect of varying the template nucleotide downstream from the abasic site on extension of each of the four bases situated opposite **X**.[220] It was not possible to use *Drosophila* pol α to compare with the earlier insertion results, because pol α extension efficiencies were too small to quantitate. For HIV-1 RT, the extension efficiencies were highest when A was situated opposite **X** for each of the four downstream template bases. The extension efficiencies in descending order were A > G ⩾ C,T. Extension past the abasic lesion was virtually absent for

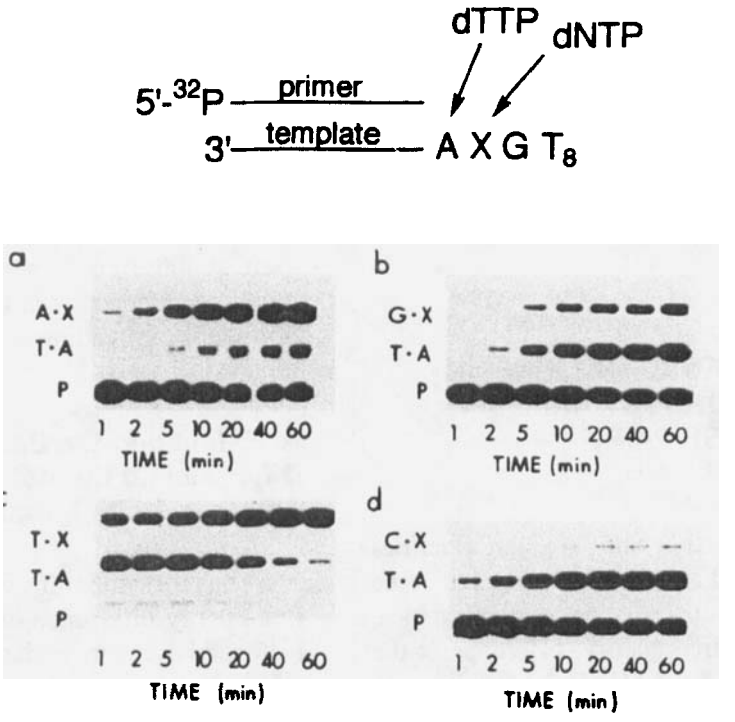


FIGURE 15. Gel autoradiogram showing the insertion of nucleotides opposite an abasic site X as a function of reaction time for a template oligomer 3'-AXGT8, using *Drosophila* pol α.[17] The dNTP substrates for insertion are a, dATP; b, dGTP; c, dTTP; d, dCTP. The lowest band represents unextended ^{32}P-labeled primer molecules; bands designated by T · A and N · X represent primers elongated by insertion of T opposite A (i.e., single base running start) and N = A, G, T, or C opposite X.

AMV RT. There are two additional observations centered on extension past abasic lesions that might have important biological ramifications.

a. Direct Compared with Misalignment Extension

Using HIV-1 RT, when C was located opposite **X**, we were unable to detect extension by **direct** addition of the next correct nucleotide. Also, beginning with a C · **X** terminus, we failed to detect incorporation of dCMP opposite a downstream template G (Figure 16 top right). However, beginning with precisely the same C · **X** terminus, we had no difficulty observing incorporation of dGMP. Apparently, G was incorporated opposite the template C located two nucleotides beyond the abasic site. This event could occur through a primer-template misalignment mechanism, described by Kunkel and co-workers,[105,106] by collapsing the abasic site and repositioning the primer terminal C opposite the downstream G (Figure 16 bottom right). Extension was abolished when the template C was replaced by another base. We also found that an A · **X** terminus, by far the easiest to extend by direct addition of a nucleotide, was inefficiently extended from a misalignment conformation.

These kinetic data appear to agree with recent NMR experiments showing that A stacks stably within the plane of the helix[221,222] when it is opposite an abasic site, while C and **X** are located extrahelically.[221] The stacking forces stabilizing A opposite **X** may act to preclude formation of a transiently misaligned primer template. The opposite may be true concerning the behavior of C, where its inherent instability when situated opposite **X** may favor transiently misaligned primer-template structures. If so, direct nucleotide addition onto the primer terminating in C may be strongly disfavored. Perhaps the A-rule in inser-

sociate in order to switch between polymerase and exonuclease active sites.[23] The conclusion is that while proofreading acts to inhibit bypass of abasic sites, there are additional important properties of the polymerase, and, perhaps, other proteins that can alter the extent of mismatch extension by limiting access of the exonuclease to repair the lesion.[223,224]

V. THERMODYNAMIC AND KINETIC ASPECTS OF CORRECT AND INCORRECT BASE PAIR DISCRIMINATION

Nucleotide insertion and excision fidelities can be characterized in terms of "effective" free energies between correct and incorrect base pairs.[25,125,137,138] A free energy difference at the polymerase active site, $\Delta\Delta G_p$, can be defined in terms of the ratio of wrong-to-right insertions given by Equation 1, as

$$f_{ins} = e^{-(\Delta\Delta G_p/RT)} \quad (11)$$

A rough estimate for the range of $\Delta\Delta G_p$ values is about 3 to 7 kcal/mol, covering a range of misinsertion efficiencies from 10^{-3} (e.g., G · T and T · G mispairs) to about 10^{-6} (e.g., pur · pur and pyr · pyr mispairs). For proofreading exonuclease, assuming that the turnover of newly inserted mismatched to correctly matched base pairs can occur with efficiencies that vary from about 90 to 99.5%, gives a range for $\Delta\Delta G_{ex} \sim 1$ to 3 kcal/mol.

Such $\Delta\Delta G_p$ values are phenomenological because they do not refer to an actual transition state in the catalysis pathway; however, they do provide an estimate of the energies involved in discrimination. If we make the assumption that the enzymatic pathway is the same for incorporation of right and wrong nucleotides, then the only place that base pairing free energies will show up will be as perturbations in the free energies of intermediate states. In this situation, the $\Delta\Delta G_p$ will always be less than or equal to the actual base paring free energy difference, because conservation of energy forbids it from being higher. Thus an observed fidelity of 10^{-6} must come from a base pairing free energy difference of at least 7 kcal/mol.

Is it reasonable to suppose that differences between correct base pair and mispair stabilities in the enzyme's active clefts are sufficient to account for insertion and excision fidelities? What active site constraints might the enzyme impose to optimize discrimination between matched and mismatched base pairs? One can pose questions such as these concerning possible sources that might account for or at least contribute to insertion and proofreading discrimination.

A. Estimating the Base Pairing Energy Difference in the Active Site Using Aqueous Melting Data

The most straightforward way to estimate the free energy difference of the matched and mismatched base pairs in the DNA is to measure the melting temperature and heat of melting for a series of DNA duplexes containing single defined base pair changes. When this is done, it appears that base pairing cannot provide any number near the requisite amount of free energy to account for the observed fidelities of DNA polymerases.[13,48,131] This makes it necessary to invoke models of the polymerase active site that modify the base pairing free energy differences observed in solution.

Thermal melting data have been obtained using duplex DNA oligonucleotides with either matched or mismatched base pairs at a defined location.[225,226] Comparing a matched and single mismatched base pair located internally, $\Delta\Delta G°$ values are typically between 1 to 3 kcal/mol.[225] Base pairing standard free energy differences are generally much lower when the mispair is placed at a primer terminus, for example, between 0.3 to 1 kcal/mol.[48] Thus, $\Delta\Delta G°$ values obtained by melting DNA in aqueous solutions are much smaller than those pertaining to nucleotide insertion fidelities. The $\Delta\Delta G°$ values, from the melting of terminal base pairs and mispairs, are closer to but still less than the effective free energy differences reflected by proofreading specificity.

Although values of $\Delta\Delta G°$ may be small, individual enthalpy ($\Delta\Delta H°$) and entropy ($\Delta\Delta S°$) terms in $\Delta\Delta G° = \Delta\Delta H° - T\,\Delta\Delta S°$, can be quite large. For example, comparing the melting of terminal mismatches G · T, C · T, and T · T with

A · T gives values of $\Delta\Delta G^\circ$ of only 0.25, 0.33, and 0.41 kcal/mol. However, the corresponding values for $\Delta\Delta H^\circ$ are much larger, i.e., 1.2, 2.2, and 4.5 kcal/mol.[48] The small $\Delta\Delta G^\circ$ values occur because large values of $\Delta\Delta H^\circ$ are compensated for by large values of $\Delta\Delta S^\circ$. The correlation between ΔH° and ΔS° for melting of DNA doublets containing correct and incorrect base pairs, and for the melting of proteins has been referred to as *enthalpy-entropy* compensation. Recently, we have shown that enthalpy-entropy compensation takes the form of a rectangular hyperbola, $\Delta S^\circ = a\Delta H^\circ/(b + \Delta H^\circ)$, where a and b are constants.[227] A significant number of correct and incorrect base pairs are located in the linear region of the curve so that a plot of $\Delta\Delta H^\circ$ vs. $\Delta\Delta S^\circ$ gives a straight line.

B. Discrimination between Base Pairs and Mispairs by DNA Polymerases

From a physical-chemical perspective, one would like to gain insight into how polymerases exploit free energy differences between right and wrong base pairs to achieve high levels of insertion and excision specificity. By what mechanisms and to what extent might polymerases exploit free energy differences between matched and mismatched base pairs? Are polymerases able to enhance specificities of insertion and excision by imposing additional base pairing constraints in the enzymes' active cleft?

We have suggested the possibility that the strong correlation observed between $\Delta\Delta H^\circ$ and $\Delta\Delta S^\circ$ comparing matched and mismatched base pairs in aqueous solution offers a plausible means for polymerases to exploit the enthalpic component as an energy source for discrimination during nucleotide insertion and excision.[48] In the relatively nonaqueous environment of an enzyme active cleft, an incoming dNTP substrate or a nucleotide present at the primer 3′-terminus can interact with a template base with reduced competition from water. Because water forms H-bonds with DNA bases, the partial exclusion of water might amplify differences in enthalpy between right and wrong base pairs.[53]

Alternatively, the active site of the polymerase may be designed in such a way that near perfect base pairing of the incoming base and the opposing base on the template is required for the proper alignment of the reactive groups of both substrates. Without proper alignment of the reactive groups, catalysis is impossible. The free energy required to put the two base pairs in proper alignment may be much higher than the free energy of the alignment adopted when the two bases are unconstrained in duplex DNA. This would explain why the solution melting data is inappropriate for estimating fidelities; mismatched bases in the duplex DNA simply do not adopt the required configuration needed for active site catalysis, and the free energy measured tells nothing about the energy of that conformation.

C. K_m and V_{max} Discrimination

The "kinetic" expression for the nucleotide misinsertion ratio, Equation 1, can be expressed as a product of K_m and V_{max} discrimination terms as follows:

$$\begin{aligned} f_{ins} &= (V_{max,W}/K_{m,W})/(V_{max,R}/K_{m,R}) \\ &= (K_{m,R}/K_{m,W}) \cdot (V_{max,W}/V_{max,R}) \\ &= f_K \cdot f_V \end{aligned} \quad (12)$$

The magnitudes of K_m are usually much smaller for insertion of right (R) compared with wrong (W) nucleotides, $f_K \ll 1$, and the magnitudes of V_{max} are usually larger for R compared to W insertions, $f_V < 1$.

Consider that the observed V_{max} and K_m values are for the reaction depicted in Figure 18a, the incorporation of a nucleoside monophosphate into the primer DNA with subsequent release of PP_i. Because the PP_i concentration is usually small in most reactions, the step where PP_i is released from the complex is assumed to be irreversible. Thus, the kinetic parameters measured represent the effect of all of the possible intermediate states of the complex between dNTP binding and PP_i release (Figure 18b). Rate constants from a variety of steps in the reaction pathway, including dNTP binding, enzyme conformational changes, phosphodiester bond catalysis and product release, contribute terms to the steady-state expressions for V_{max} and K_m. Presteady state measurements may give additional information about the system, because they measure the rate of incor-

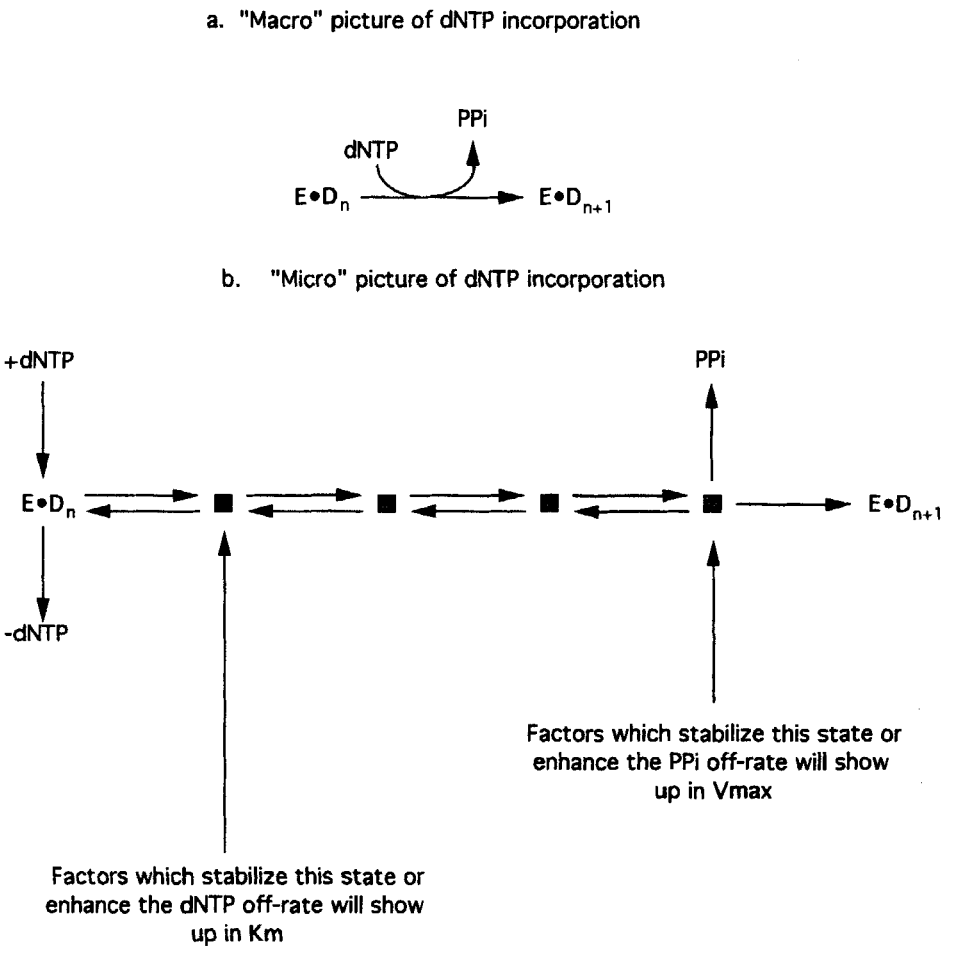


FIGURE 18. Two pictures of dNMP incorporation by a DNA polymerase. (a) The macroscopic picture – there is a single rate of incorporation dependent on [dNTP] and [PP_i]. Assuming that [PP_i] ~0, this gives a Michaelis-Menten relation between incorporation rate and [dNTP]. This is the picture that gives rise to gel bands in a steady-state assay. Note in particular that while the rate of PP_i dissociation may influence the observed incorporation rate, the rate of dissociation of the polymerase will not influence it. (b) The microscopic picture — after binding a dNTP molecule the polymerase-dNTP-primer/template complex can undergo a series of conformational and/or chemical changes culminating with the irreversible dissociation of PP_i (irreversible because we assume [PP_i] ~0). Some of the rate arrows in this diagram can be observed using presteady state kinetic gel assays because actual incorporation of dNMP into the primer may occur early in the pathway.

poration of dNMP and not the rate at which the enzyme dissociates from the product. Thus, the kinetics of fast steps before PP_i and product release can be measured by these techniques. Polymerases may have the same[77,228] or different[24,52] rate-limiting steps for insertion of right and wrong nucleotides.

While it is important to acknowledge this kinetic complexity in the polymerase fidelity pathway,[229] we believe that it is useful to investigate idealized models[25,125,130,137,138] to explore how K_m and V_{max} discrimination might affect insertion fidelity. That is, we attempt to give some meaning to the actual values of K_m and V_{max} in the system. First, note that the K_m term is indicative of the amount of time that the dNTP remains bound to the complex in a given encounter.[25,125,137,138] High K_m means that the dNTP either dissociates more rapidly when the complex is in its initial state in Figure 18b, or that it spends more time in that initial state. Conversely, high V_{max} means that the PP_i is released more rapidly

when the complex is in its final state, or that the system spends more time in the final state. Using these observations it is possible to make a few broad statements about the meaning of K_m and V_{max} discrimination for the insertion of the correct and incorrect nucleotide.

For the K_m discrimination component f_K, we assume that $f_K = K_{m,R}/K_{m,W}$ is an estimate of the relative residence times of wrong compared to right dNTP substrates bound in the polymerase active cleft.[25,125,137,138] The magnitude of this component will depend on how much of the total base pairing free energy is devoted to stabilization of the dNTP in the complex. Factors that perturb the differential stability of dNTP complexes will show up in f_K. A recent presteady state kinetic analysis[230] confirms our earlier suggestion[25,125,130,138] based on steady state analysis that T4 DNA polymerase exhibits significant discrimination against misinsertion during the nucleotide binding step.

The V_{max} discrimination component, $f_V = V_{max,W}/V_{max,R}$, is determined by the relative rates of creation of PP_i when the dNTP is engaged in the complex. The magnitude of this component will depend on the contribution of base pairing free energy to the rate of PP_i dissociation, or the stability of the state of the enzyme just before PP_i release. Because it is hard to see how base pairing free energies could affect the PP_i dissociation rate, we will assume that factors that perturb the differential rate of formation of bound PP_i for right and wrong dNTP complexes will show up in V_{max}.

Given these two different modes of discrimination available to the polymerase, it is possible that the effects of sequence context is different depending on whether the polymerase is in a K_m or a V_{max} mode of discrimination. That is, the polymerase may have different sensitivities to the small perturbing influence of nearby bases, depending on where the base pairing energy manifests itself in the catalytic cycle.

D. Correlations between K_m and V_{max} Discrimination Ratios and Base Misinsertion Hot Spots and Cold Spots

Measurements made with AMV RT showed an interesting correlation between misinsertion hot and cold spots and relative f_K and f_V values.[19] Misinsertion hot spots were found to occur predominantly next to primer-3′ pyrimidines, primarily T, when K_m discrimination predominates ($f_k \ll f_v$), while hot spots occurred mainly next to purines when V_{max} discrimination predominates ($f_v \ll f_k$). Thus, there appears to be a definite correlation between the sensitivity of the polymerase to different types of nearest neighbors and the mode of discrimination being used.

Base stacking can influence f_V and f_K. Strong base stacking from a primer terminating in G might reduce the degrees of freedom of bound right and wrong substrates immediately prior to phosphodiester bond catalysis. If most conformations for a wrong-bound dNTP do not result in catalysis but those for a right-bound dNTP do, then a differential reduction in the relative number of nonproductive-to-productive conformations will tend to increase $V_{max,W}$ relative to $V_{max,R}$. Insertional hot spots may occur more frequently next to primers terminating in G when $f_V \ll 1$, i.e., when V_{max} discrimination is large. Base stacking can also influence the relative residence times of wrong and right dNTPs during competition for binding to the enzyme-DNA complex. Because H-bonding and base stacking presumably act synergistically,[231] strong stacking may preferentially stabilize a complex involving a right dNTP. Insertional cold spots may occur more frequently next to primers terminating in G when $f_K \ll 1$, i.e., when K_m discrimination is large. For insertion of AP in competition with A opposite T by T4 L141 antimutator polymerase, K_m discrimination was found to be large,[25] and insertion cold spots appeared to occur mainly next to purines.[118,136] One should keep in mind that insertion hot and cold spots occur infrequently. We speculate that their occurrence might be related to specific base-stacking interactions that perturb E-DNA-dNTP complex stabilities and limit the number of degrees of freedom of bound dNTP substrates. At normal template sites, f/f_{avg} may be close to unity irrespective of sequence context.

In the future, we intend to expand our use of presteady state fluorescence techniques using 2-aminopurine (Section III.B.2) to measure insertion and proofreading of the analogue opposite T and C with a variety of polymerases in different sequence contexts. This approach may lead ultimately to a more complete understanding of the

mechanisms relating sequence dependent mutations to kinetic properties of DNA polymerases.

ACKNOWLEDGMENTS

The authors' work described in the review was supported by National Institutes of Health grants GM 21422 and GM 42554.

REFERENCES

1. **Watson, J. D. and Crick, F. H. C.,** Genetical implications of the structure of deoxyribonucleic acid, *Nature*, 171, 964, 1953.
2. **Watson, J. D. and Crick, F. H. C.,** The structure of DNA, *Cold Spring Harbor Symp. Quant. Biol.*, 18, 123, 1953.
3. **Kornberg, A. and Baker, T. A.,** *DNA Replication*, W. H. Freeman, New York, 1992, chap. 4.
4. **Brutlag, D. and Kornberg, A.,** Enzymatic synthesis of deoxyribonucleic acid. XXXVI. A proofreading function for the 3′ → 5′ exonuclease activity in deoxyribonucleic acid polymerases, *J. Biol. Chem.*, 247, 241, 1972.
5. **Muzyczka, N., Poland, R. L., and Bessman, M. J.,** Studies on the biochemical basis of spontaneous mutation. I. A. comparison of the deoxyribonucleic acid polymerase of mutator, antimutator, and wild-type strains of bacteriophage T4, *J. Biol. Chem.*, 247, 7116, 1972.
6. **Scheuermann, R., Tam, S., Burgers, P. M., Lu, C., and Echols, H.,** Identification of the epsilon-subunit of *Escherichia coli* DNA polymerase III holoenzyme as the *dnaQ* gene product: a fidelity subunit for DNA replication, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 7085, 1983.
7. **DiFrancesco, R., Bhatnagar, S. K., Brown, A., and Bessman, M. J.,** The interaction of DNA polymerase III and the product of the *Escherichia coli* mutator gene, *mutD*, *J. Biol. Chem.*, 259, 5567, 1984.
8. **Radman, M. and Wagner, R.,** Mismatch repair in *Escherichia coli*, *Annu. Rev. Genet.*, 20, 523, 1986.
9. **Modrich, P.,** DNA mismatch correction, *Annu. Rev. Biochem.*, 56, 435, 1987.
10. **Modrich, P.,** Methyl-directed DNA mismatch correction, *J. Biol. Chem.*, 264, 6597, 1989.
11. **Au, K. G., Clark, S., Miller, J. H., and Modrich, P.,** *Escherichia coli mutY* gene encodes an adenine glycosylase active in G · A mispairs, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 8877, 1989.
12. **Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P., and Nishimura, S.,** 8-Oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity, *Proc. Natl. Acad. Sci. U.S.A.*, 88, 4690, 1991.

12a. **Michaels, M. L. and Miller, J. H.,** The GO system protects organism from the mutagenic effect of spontaneous lesion 8-hydroxyguanine (7,8-Dihydro-8-Oxoguanine), *J. Bacteriol.*, 174, 6321, 1992.

13. **Echols, H. and Goodman, M. F.,** Fidelity mechanisms in DNA replication, *Annu. Rev. Biochem.*, 60, 477, 1991.
14. **Bessman, M. J., Muzyczka, N., Goodman, M. F., and Schnaar, R. L.,** Studies on the biochemical basis of spontaneous mutation. II. The incorporation of a base and its analogue into DNA by wild-type, mutator, and antimutator DNA polymerases, *J. Mol. Biol.*, 88, 409, 1974.
15. **Fersht, A. R.,** *Enzyme Structure and Mechanism*, W. H. Freeman, New York, 1985, 112.
16. **Segal, I. H.,** *Enzyme Kinetics*, John Wiley & Sons, New York, 1975, 564.
17. **Randall, S. K., Eritja, R., Kaplan, B. E., Petruska, J., and Goodman, M. F.,** Nucleotide insertion kinetics opposite abasic lesions in DNA, *J. Biol. Chem.*, 262, 6864, 1987.
18. **Boosalis, M. S., Petruska, J., and Goodman, M. F.,** DNA polymerase insertion fidelity: gel assay for site-specific kinetics, *J. Biol. Chem.*, 262, 14689, 1987.
19. **Mendelman, L. V., Boosalis, M. S., Petruska, J., and Goodman, M. F.,** Nearest neighbor influences on DNA polymerases insertion fidelity, *J. Biol. Chem.*, 264, 14415, 1989.
20. **Yu, H. and Goodman, M. F.,** Comparison of HIV-1 and avian myeloblastosis virus reverse transcriptase fidelity on RNA and DNA templates, *J. Biol. Chem.*, 267, 10888, 1992.
21. **Creighton, S., Huang, M.-M., Cai, H., Arnheim, N., and Goodman, M. F.,** Base mispair extension kinetics: binding of avian myeloblastosis reverse transcriptase to matched and mismatched base pair termini, *J. Biol. Chem.*, 267, 2633, 1992.
22. **Clarke, A. B. and Disney, R. L.,** *Probability And Random Processes: a First Course With Applications*, John Wiley & Sons, New York, 1985.
23. **Reddy, M. K., Weitzel, S. E., and von Hippel, P. H.,** Processive proofreading is intrinsic to T4 DNA polymerase, *J. Biol. Chem.*, 267, 14157, 1992.
24. **Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., and Benkovic, S. J.,** Kinetic mechanism of DNA polymerase I (Klenow), *Biochemistry*, 26, 8410, 1987.
25. **Clayton, L. K., Goodman, M. F., Branscomb, E. W., and Galas, D. J.,** Error induction and correction by mutant and wild-type T4 DNA polymerases: kinetic error discrimination mechanisms, *J. Biol. Chem.*, 254, 1902, 1979.
26. **Fersht, A. R.,** Fidelity of replication of phage ϕX174 DNA by DNA polymerase III holoenzyme; spontaneous mutation by misincorporation, *Proc. Natl. Acad. Sci. U.S.A.*, 76, 4946, 1979.

27. **Kornberg, A. and Baker, T. A.,** *DNA Replication*, W. H. Freeman, New York, 1992, chap. 5.
28. **Fersht, A. R., Knill-Jones, J. W., and Tsui, W. C.,** Kinetic basis of spontaneous mutation. Misinsertion frequencies, proofreading specificities and cost of proofreading by DNA polymerases of *Escherichia coli*, *J. Mol. Biol.*, 156, 37, 1982.
29. **Schaaper, R. M.,** Mechanisms of mutagenesis in the *Escherichia coli* mutator *mutD5*: role of DNA mismatch repair, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 8126, 1988.
30. **Sinha, N. K. and Haimes, M. D.,** Molecular mechanisms of substitution mutagenesis. An experimental test of the Watson-Crick and topal-fresco models of base mispairings, *J. Biol. Chem.*, 256, 10671, 1981.
31. **Sinha, N. K. and Goodman, M. F.,** Fidelity of DNA replication, in *Bacteriophage T4*, Mathews, C. K., Kutter, E. M., Mosig, G., and Berget, P. B., Eds., Am. Soc. Microbiol., Washington, D. C., 1983, 131.
32. **Kunkel, T. A., Loeb, L. A., and Goodman, M. F.,** On the fidelity of DNA replication. The accuracy of T4 DNA polymerases in copying ϕX174 DNA *in vitro*, *J. Biol. Chem.*, 259, 1539, 1984.
33. **Cox, E. C.,** Bacterial mutator genes and the control of spontaneous mutagenesis, *Annu. Rev. Genet.*, 10, 135, 1976.
34. **Cox, E. C. and Horner, D. L.,** Dominant mutators in *Escherichia coli*, *Genetics*, 100, 7, 1982.
35. **Degnan, G. E. and Cox, E. C.,** Conditional mutator gene in *Escherichia coli*. Isolation, mapping, and effector studies, *J. Bacteriol.*, 117, 477, 1974.
36. **Schaaper, R. M., Bond, B. I., and Fowler, R. G.,** A · T → C · G transversions and their prevention by the *Escherichia coli mutT* and *mut*HLS pathways, *Mol. Gen. Genet.*, 219, 1, 1989.
37. **Schaaper, R. M. and Radman, M.,** The extreme mutator effect of *Escherichia coli mut*D5 results from saturation of mismatch repair by excessive DNA replication errors, *EMBO J.*, 8, 3511, 1989.
38. **Reha-Krantz, L. J., Stocki, S., Nonay, R. L., Dimayuga, E., Goodrich, L. D., Konigsberg, W. H., and Spicer, E. K.,** DNA polymerization in the absence of exonucleolytic proofreading: *in vivo* and *in vitro* studies, *Proc. Natl. Acad. Sci. U.S.A.*, 88, 2417, 1991.
39. **Young, M. C., Reddy, M. K., and von Hippel, P. H.,** Structure and function of the bacteriophage T4 DNA polymerase holoenzyme, *Biochemistry*, 31, 8675, 1992.
40. **Speyer, J. F.,** Mutagenic DNA polymerase, *Biochem. Biophys. Res. Comm.*, 21, 6, 1965.
41. **Freese, E. B. and Freese, E. F.,** On the specificity of DNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.*, 57, 650, 1967.
42. **Drake, J. W. and Allen, E. F.,** Antimutagenic DNA polymerases of bacteriophage T4, *Cold Spring Harbor Symp. Quant. Biol.*, 33, 339, 1968.
43. **Lo, K.-Y. and Bessman, M. J.,** An antimutator deoxyribonucleic acid polymerase. I. Purification and properties of the enzyme, *J. Biol. Chem.*, 251, 2475, 1976.
44. **Speyer, J. F., Karam, J. D., and Lenny, A. B.,** On the role of DNA polymerase in base selection, *Cold Spring Harbor Symp. Quant. Biol.*, 31, 693, 1966.
45. **Drake, J. W.,** Comparative rates of spontaneous mutation, *Nature*, 221, 1132, 1969.
46. **Drake, J. W., Allen, E. F., Forsberg, S. A., Preparata, R., and Greening, E. O.,** Spontaneous mutation. Genetic control of mutation rates in bacteriophage T4, *Nature*, 221, 1128, 1969.
47. **Ripley, L. S. and Shoemaker, N. B.,** A major role for bacteriophage T4 DNA polymerase in frameshift mutagenesis, *Genetics*, 103, 353, 1983.
48. **Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I., Jr.,** Comparison between DNA melting thermodynamics and DNA polymerase fidelity, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 6252, 1988.
49. **Perrino, F. W. and Loeb, L. A.,** Differential extension of 3′ mispairs is a major contribution to the high fidelity of calf thymus DNA polymerase-alpha, *J. Biol. Chem.*, 264, 2898, 1989.
50. **Mendelman, L. V., Petruska, J., and Goodman, M. F.,** Base mispair extension kinetics: comparison of DNA polymerase α and reverse transcriptase, *J. Biol. Chem.*, 265, 2338, 1990.
51. **Perrino, F. W., Preston, B. D., Sandell, L. L., and Loeb, L. A.,** Extension of mismtached 3′ termini of DNA is a major determinant of the infidelity of human immunodeficiency virus type 1 reverse transcriptase, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 8343, 1989.
52. **Kuchta, R. D., Benkovic, P., and Benkovic, S. J.,** Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity, *Biochemistry*, 27, 6716, 1988.
53. **Petruska, J., Sowers, L. C., and Goodman, M. F.,** Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 1559, 1986.
54. **Cohen, S. S.,** *Virus-Induced Enzymes*, Columbia University Press, New York, 1968.
55. **Mathews, C. K.,** Bacteriophage biochemistry, *ACS Monograph No. 164*, Van Nostrand Reinhold, New York, 1971.
56. **Tomich, P. K., Chiu, C. S., Wovcha, M. G., and Greenberg, G. R.,** Evidence for a complex regulating the *in vivo* activities of early enzymes induced by bacteriophage T4, *J. Biol. Chem.*, 249, 7613, 1974.
57. **Reddy, G. P. V., Singh, A., Stafford, M. E., and Mathews, C. K.,** Enzyme associations in T4 phage DNA precursor synthesis, *Proc. Natl. Acad. Sci. U.S.A.*, 74, 3152, 1977.
58. **Thylen, C. and Mathews, C. K.,** Essential role of T4 phage deoxycytidylate hydroxymethylase in a multienzyme complex for deoxyribonucleotide synthesis, *J. Biol. Chem.*, 264, 15169, 1989.

59. **Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A.,** Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP, *Nature*, 313, 762, 1985.
60. **Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M., and Steitz, T. A.,** Genetic and crystallographic studies of the 3′,5′-exonucleolytic site of DNA polymerase I, *Science*, 240, 199, 1988.
61. **Reha-Krantz, L. J.,** Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants: relating structure to function, *J. Mol. Biol.*, 202, 711, 1988.
62. **Spicer, E. K., Rush, J., Fung, C., and Reha-Krantz, L. J.,** Primary structure of T4 DNA polymerase: evolutionary relatedness to eukaryotic and other prokaryotic DNA polymerases, *J. Biol. Chem.*, 263, 7478, 1988.
63. **Cowart, M. C., Gibson, K. J., Allen, D. J., and Benkovic, S. J.,** DNA substrate structural requirements for the exonuclease and polymerase activities of prokaryotic and phage DNA polymerases, *Biochemistry*, 28, 1975, 1989.
64. **Joyce, C. M.,** How DNA travels between the separate polymerase and 3′ → 5′ exonuclease sites of DNA polymerase I (Klenow fragment), *J. Biol. Chem.*, 264, 10858, 1989.
65. **Goodman, M. F., Gore, W. C., Muzyczka, N., and Bessman, M. J.,** Studies on the biochemical basis of spontaneous mutation. III. Rate model for DNA polymerase-effected nucleotide misincoporation, *J. Mol. Biol.*, 88, 423, 1974.
66. **Sinha, N. K.,** Specificity and efficiency of editing of mismatches involved in the formation of base substitution mutations by the 3′ → 5′ exonuclease activity of phage T4 DNA polymerase, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 915, 1987.
67. **Brenowitz, S., Kwack, S., Goodman, M. F., O'Donnell, M., and Echols, H.,** Specificity and enzymatic mechanism of the editing exonuclease of *Escherichia coli* DNA polymerase III, *J. Biol. Chem.*, 266, 7888, 1991.
68. **Reha-Krantz, L. J. and Bessman, M. J.,** Studies on the biochemical basis of spontaneous mutation. IV. Effect on amino acid substitution on the enzymatic and biological properties of the T4 DNA polymerase, *J. Mol. Biol.*, 116, 99, 1977.
69. **Bessman, M. J. and Reha-Krantz, L. J.,** Studies on the biochemical basis of spontaneous mutation. V. Effects of temperature on mutation frequency, *J. Mol. Biol.*, 116, 115, 1977.
70. **Hershfield, M. S. and Nossal, N. G.,** Hydrolysis of template and newly synthesized deoxyribonucleic acid by the 3′ to 5′ exonuclease activity of the T4 deoxyribonucleic acid polymerase, *J. Biol. Chem.*, 247, 3393, 1972.
71. **Nossal, N. G. and Hershfield, M. S.,** Exonuclease activity of wild-type and mutant T4 DNA polymerases: hydrolysis during DNA synthesis *in vitro*, in *DNA Synthesis In Vitro*, Wells, R. D., and Inman, R. B., Eds., University Park Press, Baltimore, 1973, 47.
72. **Hershfield, M. S.,** On the role of deoxyribonucleic acid polymerase in determining mutation rates: characterization of the defect in the T4 deoxyribonucleic acid polymerase caused by the *ts*L88 mutation, *J. Biol. Chem.*, 248, 1417, 1973.
73. **Reha-Krantz, L. J.,** Locations of amino acid substitutions in bacteriophage T4 *ts*L56 DNA polymerase predict an N-terminal exonuclease domain, *J. Virol.*, 63, 4762, 1989.
74. **Gillen, F. D. and Nossal, N. G.,** Control of mutation frequency by bacteriophage T4 DNA polymerase. I. The CB120 antimutator DNA polymerase is defective in strand displacement, *J. Biol. Chem.*, 251, 5219, 1976.
75. **Reha-Krantz, L. J.,** Genetic and biochemical studies of the bacteriophage T4 DNA polymerase, in *UCLA Symp. Mol. Cell. Biol. New Ser.*, 47, Kelly, T. J. and McMacken, R., Eds., Alan R. Liss, New York, 1987, 501.
76. **Cai, H., Bloom, L. B., Reha-Krantz, L. J., and Goodman, M. F.,** unpublished, 1992.
77. **Wong, I., Patel, S. S., and Johnson, K. A.,** An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics, *Biochemistry*, 30, 526, 1991.
78. **Detera, S. D., Becerra, S. P., Swack, J. A., and Wilson, S. H.,** Studies on the mechanism of DNA polymerase alpha. Nascent chain elongation, steady state kinetics, and the initiation phase of DNA synthesis, *J. Biol. Chem.*, 256, 6933, 1981.
79. **Fisher, P. A. and Korn, D.,** Ordered sequential mechanism of substrate recognition and binding by KB cell DNA polymerase alpha, *Biochemistry*, 20, 4560, 1981.
80. **Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N.,** Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, *Science*, 230, 1350, 1985.
81. **Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., and Horn, G. T.,** Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science*, 239, 487, 1988.
82. **Li, H., Cui, X., and Arnheim, N.,** Direct electrophoretic detection of the allelic state of single DNA molecules in human sperm by using the polymerase chain reaction, *Proc. Natl. Acad. Sci. U.S.A.*, 87, 4580, 1990.
83. **Ehlen, T. and Dubeau, L.,** Detection of *ras* point mutations by polymerase chain reaction using mutation-specific, inosine-containing oligonucleotide primers, *Biochem. Biophys. Res. Comm.*, 160, 441, 1989.
84. **Gibbs, R. A., Nguyen, P. N., and Caskey, C. T.,** Detection of single DNA base differences by competitive oligonucleotide priming, *Nucl. Acids Res.*, 17, 2437, 1989.

85. **Newton, C. R., Graham, A., Heptinstall, L. E., Powell, S. J., Summers, C., Kalsheker, N., Smith, J. C., and Markham, A. F.,** Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS), *Nucl. Acids Res.*, 17, 2503, 1989.
86. **Wu, D. Y., Ugozzoli, L., Pal, B. K., and Wallace, R. B.,** Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 2757, 1989.
87. **Ugozzoli, L. and Wallace, R. B.,** Allele-specific polymerase chain reaction, *METHODS:* A Companion to Methods in Enzyomology, 2, 42, 1991.
88. **Cha, R. S., Zarbl, H., Keohavong, P., and Thilly, W. G.,** Mismatch amplification mutation assay (MAMA): application to the c-H-*ras*-gene, PCR, *Methods and Applications*, 2, 14, 1992.
89. **Huang, M.-M., Arnheim, N., and Goodman, M. F.,** Extension of base mispairs by *Taq* DNA polymerase: implications for single nucleotide discrimination in PCR, *Nucl. Acids Res.*, 20, 4567, 1992.
90. **Kwok, S., Kellogg, D. E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J. J.,** Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies, *Nucl. Acids Res.*, 18, 999, 1990.
91. **Preston, B. D., Poiesz, B. J., and Loeb, L. A.,** Fidelity of HIV-1 reverse transcriptase, *Science*, 243, 1168, 1988.
92. **Ricchetti, M. and Buc, H.,** Reverse tanscriptases and genomic variability: the accuracy of DNA replication is enzyme specific and sequence dependent, *EMBO J.*, 9, 1583, 1990.
93. **Majumdar, C., Abbotts, J., Broder, S., and Wilson, S. H.,** Studies on the mechanism of human immunodeficiency virus reverse transcriptase: steady state kinetics, processivity, and polynucleotide inhibition, *J. Biol. Chem.*, 263, 15657, 1988.
94. **Majumdar, C., Stein, C. A., Cohen, J. S., Broder, S., and Wilson, S. H.,** Stepwise mechanism of HIV reverse transcriptase: primer function of phosphorothioate oligodeoxynucleotide, *Biochemistry*, 28, 1340, 1989.
95. **Huber, H. E., McCoy, J. M., Seehra, J. S., and Richardson, C. C.,** Human immunodeficiency virus 1 reverse transcriptase. Template binding, processivity, strand displacement synthesis, and template switching, *J. Biol. Chem.*, 264, 4669, 1989.
96. **Dudding, L. R., Nkabinde, N. C., and Mizrahi, V.,** Analysis of the RNA- and DNA-dependent DNA polymerase activities of point mutants of HIV-1 reverse transcriptase lacking ribonuclease H activity, *Biochemistry*, 30, 10498, 1991.
96a. **Bakhanashvili, M. and Hizi, A.,** Fidelity of the RNA-dependent DNA synthesis exhibited by the reverse transcriptases of Human Immunodeficiency Virus Types 1, and 2 and of Murine Leukemia Virus: Mispair extension frequencies, *Biochemistry*, 31, 9393, 1992.
97. **Reardon, J. E. and Miller, W. H.,** Human immunodeficiency virus reverse transcriptase. Substrate and inhibitor kinetics with thymidine 5′-triphosphate and 3′-azido-3′-deoxythymidine-5′-triphosphate, *J. Biol. Chem.*, 265, 20302, 1990.
98. **Krug, M. S. and Berger, S. L.,** Reverse transcriptase from human immunodeficiency virus: a single template-primer binding site serves two physically separable catalytic functions, *Biochemistry*, 30, 10614, 1991.
99. **Roberts, J. D., Bebenek, K., and Kunkel, T. A.,** The accuracy of reverse transcriptase from HIV-1, *Science*, 242, 1171, 1988.
100. **Weber, J. and Grosse, F.,** Fidelity of human immunodeficiency virus type I reverse transcriptase in copying natural DNA, *Nucl. Acids Res.*, 17, 1379, 1989.
101. **Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H., and Kunkel, T. A.,** Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase, *J. Biol. Chem.*, 264, 16948, 1989.
102. **Huebner, A., Kruhoffer, M., Grosse, F., and Krauss, G.,** Fidelity of human immunodeficiency virus type I reverse transcriptase in copying natural RNA, *J. Mol. Biol.*, 223, 595, 1992.
103. **Roberts, J. D., Preston, B. D., Johnston, L. A., Soni, A., Loeb, L. A., and Kunkel, T. A.,** Fidelity of two retroviral reverse transcriptases during DNA-dependent DNA synthesis *in vitro*, *Mol. Cell. Biol.*, 9, 469, 1989.
104. **Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., and Inouye, M.,** Frameshift mutations and the genetic code, *Cold Spring Harbor Symp. Quant. Biol.*, 31, 77, 1966.
105. **Kunkel, T. A.,** Frameshift mutagenesis by eukaryotic DNA polymerases *in vitro*, *J. Biol. Chem.*, 261, 13581, 1986.
106. **Kunkel, T. A. and Soni, A.,** Mutagenesis by transient misalignment, *J. Biol. Chem.*, 263, 14784, 1988.
107. **Boosalis, M. S., Mosbaugh, D. W., Hamatake, R., Sugino, A., Kunkel, T. A., and Goodman, M. F.,** Kinetic analysis of base substitution mutagenesis by transient misalignment of DNA and by miscoding, *J. Biol. Chem.*, 264, 11360, 1989.
107a. **Fowler, R. G., Degnan, G. E., and Cox, E. C.,** Mutational specificity of conditional *E. coli* mutator, *mutD5*, *Mol. Gen. Genet.*, 133, 179, 1974.
108. **de Boer, J. G. and Ripley, L. S.,** Demonstration of the production of frameshift and base-substitution mutations by quasipalindromic DNA sequences, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 5528, 1984.
109. **Papanicolaou, C. and Ripley, L. S.,** Polymerase-specific differences in the DNA intermediates of frameshift mutagenesis. *In vitro* synthesis errors of *Escherichia coli* DNA polymerase I and its large fragment derivative, *J. Mol. Biol.*, 207, 335, 1989.
110. **Ripley, L. S.,** Frameshift mutation: determinants of specificity, *Annu. Rev. Genet.*, 24, 189, 1990.

111. **Allen, D. J., Darke, P. L., and Benkovic, S. J.,** Fluorescent oligonucleotides and deoxynucleotide triphosphates: preparation and their interaction with the large (Klenow) fragment of *Escherichia coli* DNA polymerase I, *Biochemistry,* 28, 4601, 1989.
112. **Carroll, S. S. and Benkovic, S. J.,** Mechanistic aspects of DNA polymerases: *Escherichia coli* DNA polymerase I (Klenow fragment) as a paradigm, *Chem. Rev.,* 90, 1291, 1990.
113. **Singer, B., Chavez, F., Goodman, M. F., Essigmann, J. M., and Dosanjh, M. K.,** Effect of 3′ flanking neighbors on kinetics of pairing of dCTP or dTTP opposite O^6-methylguanine in a defined primed oligonucleotide when *Escherichia coli* DNA polymerase I is used, *Proc. Natl. Acad. Sci. U.S.A.,* 86, 8271, 1989.
114. **Dosanjh, M. K., Galeros, G., Goodman, M. F., and Singer, B.,** Kinetics of extension of O^6-methylguanine paired with cytosine or thymine in defined oligonucleotide sequences, *Biochemistry,* 30, 11595, 1991.
115. **Voigt, J. M. and Topal, M. D.,** O^6-methylguanine and A · C and G · T mismatches cause asymmetric structural defects in DNA that are affected by DNA sequence, *Biochemistry,* 29, 5012, 1990.
116. **Reha-Krantz, L. J. and Bessman, M. J.,** Studies on the biochemical basis of spontaneous mutation. VI. Selection and characterization of a new bacteriophage T4 mutator DNA polymerase, *J. Mol. Biol.,* 145, 677, 1981.
117. **Pless, R. C., Levitt, L. M., and Bessman, M. J.,** Nonrandom substitution of 2-aminopurine for adenine during deoxyribonucleic acid synthesis *in vitro, Biochemistry,* 20, 6235, 1981.
118. **Pless, R. C. and Bessman, M. J.,** Influence of local nucleotide sequence on substituion of 2-aminopurine for adenine during deoxyribonucleic acid synthesis *in vitro, Biochemistry,* 22, 4905, 1983.
119. **Freese, E.,** The specific mutagenic effect of base analogues on phage T4, *J. Mol. Biol.,* 1, 87, 1959.
120. **Rudner, R.,** Mutation as an error in base pairing. II. Kinetics of 5-bromodeoxyuridine and 2-aminopurine-induced mutagenesis, *Zeitschrift für Vererbungslehre,* 92, 361, 1961.
121. **Ronen, A.,** 2-aminopuririne, *Mutat. Res.,* 75, 1, 1979.
122. **Hopkins, R. and Goodman, M. F.,** Asymmetry in forming 2-aminopurine · hydroxymethylcytosine heteroduplexes; a model giving misincorporation frequencies and rounds of DNA replication from base-pair populations *in vivo, J. Mol. Biol.,* 135, 1, 1979.
123. **Hopkins, R. L. and Goodman, M. F.,** Deoxyribonucleotide pools, base pairing and sequence configuration affecting bromodeoxyuradine- and 2-aminopurine-induced mutagenesis, *Proc. Natl. Acad. Sci. U.S.A.,* 77, 1801, 1980.
124. **Persing, D. H., McGinty, L., Adams, C. W., and Fowler, R. G.,** Mutational specificity of the base analogue,2-aminopurine, in *Escherichia coli, Mutat. Res.,* 83, 25, 1981.
125. **Goodman, M. F. and Branscomb, E. W.,** DNA replication fidelity and base mispairing mutagenesis, in *Accuracy in Molecular Processes: its Control and Relevance to Living Systems,* Kirkwood, T. B. L., Rosenberger, R. F., and Galas, D. J., Eds., Chapman and Hall, New York, 1986, 191.
126. **Goodman, M. F., Hopkins, R., and Gore, W. C.,** 2-aminopurine-induced mutagenesis in T4 bacteriophage: a model relating mutation frequency to 2-aminopurine incorporation in DNA, *Proc. Natl. Acad. Sci. U.S.A.,* 74, 4806, 1977.
127. **Mhaskar, D. N. and Goodman, M. F.,** On the molecular basis of transition mutations. Frequency of forming 2-aminopurine · cytosine and adenine · cytosine base mispairs in the G · C → A · T mutational pathway by T4 DNA polymerase *in vitro, J. Biol. Chem.,* 259, 11713, 1984.
128. **Ripley, L. S.,** Influence of diverse gene *43* DNA polymerases on the incorporation and replication *in vivo* 2-aminopurine at A · T base pairs in bacteriophage T4, *J. Mol. Biol.,* 150, 197, 1981.
129. **Watanabe, S. M. and Goodman, M. F.,** On the molecular basis of transition mutations: frequencies of forming 2-aminopurine · cytosine and adenine · cytosine base mispairs *in vitro, Proc. Natl. Acad. Sci. U.S.A.,* 78, 2864, 1981.
130. **Watanabe, S. M. and Goodman, M. F.,** Kinetic measurement of 2-aminopurine · cytosine and 2-aminopurine · thymidine base pairs as a test of DNA polymerase fidelity mechanisms, *Proc. Natl. Acad. Sci. U.S.A.,* 79, 6429, 1982.
131. **Loeb, L. A. and Kunkel, T. A.,** Fidelity of DNA synthesis, *Annu. Rev. Biochem.,* 52, 429, 1982.
132. **Fersht, A. R. and Knill-Jones, J. W.,** Fidelity of replication of bacteriophage ϕX174DNA *in vitro* and *in vivo, J. Mol. Biol.,* 165, 633, 1983.
133. **Sloane, D. L., Goodman, M. F., and Echols, H.,** The fidelity of base selection by the polymerase subunit of DNA polymerase III holoenzyme, *Nucl. Acids Res.,* 16, 6465, 1988.
134. **Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., and Goodman, M. F.,** Base pairing and mutagenesis: observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR, *Proc. Natl. Acad. Sci. U.S.A.,* 83, 5434, 1986.
135. **Eritja, R. E., Kaplan, B. E., Mhaskar, D., Sowers, L. C., Petruska, J., and Goodman, M. F.,** Synthesis and properties of defined DNA oligomers containing base mispairs involving 2-aminopurine, *Nucl. Acids Res.,* 14, 5869, 1986.
136. **Petruska, J. and Goodman, M. F.,** Influence of neighboring bases on DNA polymerase insertion and proofreading fidelity, *J. Biol. Chem.,* 260, 7533, 1985.
137. **Galas, D. J. and Branscomb, E. W.,** Enzymatic determinants of DNA polymerase accuracy. Theory of coliphage T4 polymerase mechanisms, *J. Mol. Biol.,* 88, 653, 1978.

138. **Goodman, M. F.,** DNA replication fidelity: kinetics and thermodynamics, *Mutat. Res.*, 200, 11, 1988.
139. **Guest, C. R., Hochstrasser, R. A., Dupuy, C. G., Allen, D. J., Benkovic, S. J., and Miller, D. P.,** Interaction of DNA with the Klenow fragment of DNA polymerase I studied by time-resolved fluorescence spectroscopy, *Biochemistry,* 30, 8759, 1991.
140. **Jung, G. H., Leavitt, M. C., Hsieh, J. C., and Ito, J.,** Bacteriophage PRD1 DNA polymerase: evolution of DNA polymerases, *Proc. Natl. Acad. Sci. U.S.A.*, 84, 8287, 1987.
141. **Leavitt, M. C. and Ito, J.,** T5 DNA polymerase: structural—functional relationships to other DNA polymerases, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 4465, 1989.
142. **Bernad, A., Zaballos, A., Salas, M., and Blanco, L.,** Structural and functional relationships between prokaryotic and eukaryotic DNA polymerases, *EMBO J.*, 6, 4219, 1987.
143. **Bernad, A., Blanco, L., Lazaro, J. M., Martin, G., and Salas, M.,** A conserved 3′–5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases, *Cell,* 59, 219, 1989.
144. **Blanco, L., Bernad, A., Blasco, M. A., and Salas, M.,** A general structure for DNA-dependent DNA polymerases, *Gene,* 100, 1991.
145. **Blanco, L., Bernad, A., and Salas, M.,** Evidence favoring the hypothesis of a conserved 3′–5′ exonuclease active site in DNA-dependent DNA polymerases, *Gene,* 112, 139, 1992.
146. **Wong, S. W., Wahl, A. F., Yuan, P. M., Arai, N., Pearson, B. E., Arai, K., and Korn, D.,** Human DNA polymerase alpha gene expression is cell proliferation dependent, and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases, *EMBO J.*, 7, 37, 1988.
147. **Wang, T. S., Wong, S. W., and Korn, D.,** Human DNA polymerase alpha: predicted functional domains and relationships with viral DNA polymerases, *FASEB J.*, 3, 14, 1989.
148. **Wang, T. S.,** Eukaryotic DNA polymerases, *Annu. Rev. Biochem.*, 60, 513, 1991.
149. **Ward, D. C. and Reich, E.,** Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives, *J. Biol. Chem.*, 244, 1228, 1969.
150. **Scheit, K. H. and Rackwitz, H.-R.,** Synthesis and physicochemical properties of two analogs of poly(dA):poly(2-aminopurine-9-β-D-deoxyribonucleotide) and poly 2-amino-deoxyadenylic acid, *Nucl. Acids Res.*, 10, 4059, 1982.
151. **Goodman, M. F. and Ratliff, R. L.,** Evidence of 2-aminopurine-cytosine base mispairs involving two hydrogen bonds, *J. Biol. Chem.*, 258, 12842, 1983.
152. **Topal, M. D. and Fresco, J. R.,** Molecular basis for substitution mutations, *Nature,* 263, 285, 1976.
153. **Patel, D. J., Kozlowski, S. A., Ikuta, S., and Itakura, K.,** Deoxyadenosine-deoxycytidine pairing in the d(C-G-C-G-A-A-T-T-C-A-C-G) duplex: conformation and dynamics at and adjacent to the dA X dC mismatch site, *Biochemistry,* 23, 3218, 1984.
154. **Hunter, W. N., Brown, T., Anand, N. N., and Kennard, O.,** Structure of an adenine-cytosine base pair in DNA and its implications for mismatch repair, *Nature,* 320, 552, 1986.
155. **Sowers, L. C., Fazakerley, G. V., Kim, H., Dalton, L., and Goodman, M. F.,** Variation of nonexchangeable proton resonance chemical shifts as a probe of aberrant base pair formation in DNA, *Biochemistry,* 25, 3983, 1986.
156. **Wang, C., Gao, H., Gaffney, B. L., and Jones, R. A.,** Nitrogen-15-labeled oligodeoxynucleotides. III. Protonation of the adenine N1 in the A · C and A · G mispairs on the duplexes $\{d[CG(^{15}N^1)AGAATTCCCG]\}_2$ and $d[CGGGAATTC(^{15}N^1)ACG]\}_2$, *J. Am. Chem. Soc.*, 113, 5486, 1991.
157. **Boulard, Y., Cognet, J. A., Gabarro, A. J., Le, B. M., Sowers, L. C., and Fazakerley, G. V.,** The pH-dependent configurations of the C · A mispair in DNA, *Nucl. Acids Res.*, 20, 1933, 1992.
158. **Patel, D. J., Kozlowski, S. A., Ikuta, S., and Itakura, K.,** Dynamics of DNA duplexes containing internal G · T, G · A, A · C, and T · C pairs: hydrogen exchange at and adjacent to mismatch sites, *Fed. Proc., Fed. Am. Soc. Exp. Biol.*, 43, 2663, 1984.
159. **Hunter, W. N., Kneale, G., Brown, T., Rabinovich, D., and Kennard, O,** Refined crystal structure of an octanucleotide duplex with G · T mismatched base pairs, *J. Mol. Biol.*, 190, 605, 1986.
160. **Kan, L. S., Chandrasegaran, S., Pulford, S. M., and Miller, P. S.,** Detection of a guanine X adenine base pair in a decadeoxyribonucleotide by proton magnetic resonance spectroscopy, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 4263, 1983.
161. **Brown, T., Hunter, W. N., Kneale, G., and Kennard, O.,** Molecular structure of the G · A base pair in DNA and its implications for the mechanism of transversion mutations, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 2402, 1986.
162. **Brown, T., Leonard, G. A., Booth, E. D., and Chambers, J.,** Crystal structure and stability of a DNA duplex containing A(anti) · G(*syn*) base pairs, *J. Mol. Biol.*, 207, 455, 1989.
163. **Sowers, L. C., Eritja, R., Chen, F. M., Khwaja, T., Kaplan, B. E., Goodman, M. F., and Fazakerley, G. V.,** Characterization of the high pH wobble structure of the 2-aminopurine.cytosine mismatch by N-15 NMR spectroscopy, *Biochem. Biophys. Res. Comm.*, 165, 89, 1989.
164. **Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., and Fazakerly, G. V.,** Equilibrium between a wobble and ionized base pair formed between fluorouracil and guanine in DNA as studied by proton and fluorine NMR, *J. Biol. Chem.*, 263, 14794, 1988.
165. **Sowers, L. C., Goodman, M. F., Eritja, R., Kaplan, B., and Fazakerley, G. V.,** Ionized and wobble base-pairing for bromouracil-guanine in equilibrium under physiological conditions. A nuclear magnetic resonance study on an oligonucleotide con-

taining a bromouracil-guanine base pair as a function of pH, *J. Mol. Biol.*, 205, 437, 1989.
166. **Sowers, L. C., Ramsay Shaw, B., Veigl, M. L., and Sedwick, W. D.,** DNA base modification: ionized base pairs and mutagenesis, *Mutat. Res.*, 177, 201, 1987.
167. **Lawley, P. D. and Brooks, P.,** Ionization of DNA bases or base analogues as a possible explanation of mutagenesis, *J. Mol. Biol.*, 4, 216, 1962.
168. **Trautner, T. A., Swartz, M. N., and Kornberg, A.,** Enzymatic synthesis of deoxyribonucleic acid. X. Influence of bromouracil substitutions on replication, *Proc. Natl. Acad. Sci. U.S.A.*, 48, 449, 1962.
169. **Lasken, R. S. and Goodman, M. F.,** The biochemical basis of 5-bromouracil-induced mutagenesis. Heteroduplex base mispairs involving bromouracil in G · C → A · T and A · T → G · C mutational pathways, *J. Biol. Chem.*, 259, 11491, 1984.
170. **Lasken, R. S. and Goodman, M. F.,** A fidelity assay using "dideoxy" DNA sequencing: a measurement of sequence dependence and frequency of forming 5-bromouracil · guanine base mispairs, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 1301, 1985.
171. **Driggers, P. H. and Beattie, K. L.,** Effect of pH on base-mispairing properties of 5-bromouracil during DNA synthesis, *Biochemistry*, 27, 1729, 1988.
172. **Yu, H., Eritja, R., Bloom, L. B., and Goodman, M. F.,** Ionization of bromuracil and fluorourcial stimulates base mispairing frequencies with guanine, submitted for publication, 1992.
172a. **Rayssiguier, C., Thaler, D. S., and Radman, M.,** The barrier to recombination between *E. coli* and *Salomonella typhimurium* is disrupted in mismatch repair mutants, *Nature (London)*, 342, 396, 1989.
173. **Tsai-Wu, J.-J., Liu, H.-F., and Lu, A.-L.,** *Escherichia coli mutY* protein has both *N*-glycosylase and apurinic/apyrimidinic endonuclease activities on A · C and A · G mispairs, *Proc. Natl. Acad. Sci. U.S.A.*, 89, 8779, 1992.
174. **Michaels, M. L., Cruz, C., Grollman, A. P., and Miller, J. H.,** Evidence that *MutY* and *MutM* combine to prevent mutations by an oxidatively damaged form of guanine in DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 89, 7022, 1992.
175. **Michaels, M. L., Pham, L., Cruz, C., and Miller, J. H.,** *MutM*, a protein that prevents G · C → T · A transversions, is formamidopyrimidine-DNA glycosylase, *Nucl. Acids Res.*, 19, 3629, 1991.
176. **Treffers, H. P., Spinelli, C., and Belser, N. O.,** A factor (or mutator gene) influencing mutation rates in *Escherichia coli*, *Proc. Natl. Acad. Sci. U.S.A.*, 40, 1064, 1954.
177. **Cox, E. C.,** Mutator gene studies in *Escherichia coli*: the *mutT* gene, *Genetics*, Suppl. 73, 67, 1973.
178. **Bhatnagar, S. K. and Bessman, M. J.,** Studies on the mutator gene, *mutT* of *Escherichia coli*. Molecular cloning of the gene, purification of the gene product, and identification of a novel nucleoside triphosphatase, *J. Biol. Chem.*, 263, 8953, 1988.
179. **Bhatnagar, S. K., Bullions, L. C., and Bessman, M. J.,** Characterization of the *mutT* nucleoside triphosphatase of *Escherichia coli*, *J. Biol. Chem.*, 266, 9050, 1991.
180. **Schaaper, R. M. and Dunn, R. L.,** *Escherichia coli mutT* mutator effect during *in vitro* DNA synthesis. Enhanced A · G replicational errors, *J. Biol. Chem.*, 262, 16267, 1987.
181. **Akiyama, M., Maki, H., Sekiguchi, M., and Horiuchi, T.,** A specific role of *MutT* protein: to prevent dG · dA mispairing in DNA replication, *Proc. Natl. Acad. Sci. U.S.A.*, 86, 3949, 1989.
182. **Maki, H. and Sekiguchi, M.,** *MutT* protein specifically hydrolyzes a potent mutagenic substrate for DNA synthesis, *Nature*, 355, 273, 1992.
183. **Boiteux, S., O'Connor, T. R., Lederer, F., Gouyette, A., and Laval, J.,** Homogeneous *Escherichia coli* FPG protein. A DNA glycosylase which excised imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites, *J. Biol. Chem.*, 265, 3916, 1990.
184. **Shibutani, S., Takeshita, M., and Grollman, A. P.,** Insertion of specific bases during DNA synthesis past the hydroxyl-radical damaged base 8-oxodG, *Nature*, 349, 431, 1991.
185. **Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A.,** 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G → T and A → C substitutions, *J. Biol. Chem.*, 267, 166, 1992.
186. **Walker, G. C.,** Inducible DNA repair systems, *Annu. Rev. Biochem.*, 54, 425, 1985.
187. **Friedberg, E. C.,** *DNA Repair*, W. H. Freeman, New York, 1985.
188. **Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y.,** Regulation and expression of the adaptive response to alkylating agents, *Annu. Rev. Biochem.*, 57, 133, 1988.
189. **Radman, M.,** SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis, in *Molecular Mechanisms for the Repair of DNA, Part A*, Hanawalt, P. and Setlow, R. B., Eds., Plenum Press, New York, 1975, 355.
190. **Witkin, E. M.,** Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*, *Bacteriol. Rev.*, 40, 869, 1976.
191. **Little, J. W. and Mount, D. W.,** The SOS regulatory system of *Escherichia coli*, *Cell*, 29, 11, 1982.
192. **Walker, G. C.,** Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*, *Microbiol. Rev.*, 48, 60, 1984.
193. **Echols, H. and Goodman, M. F.,** Mutation induced by DNA damage: a many protein affair, *Mutat. Res.*, 236, 301, 1990.
194. **Sancar, A. and Sancar, G. B.,** DNA repair enzymes, *Annu. Rev. Biochem.*, 57, 29, 1988.
195. **Bridges, B. A. and Woodgate, R.,** Mutagenic repair in *Escherichia coli*: products of the *recA* gene and of the *umuD* and *umuC* genes act at different steps in

UV-induced mutagenesis, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 4193, 1985.

196. **Banerjee, S. K., Borden, A., Christensen, R. B., LeClerc, J. E., and Lawrence, C. W.,** SOS-dependent replication past a single *trans-syn* T-T cyclobutane dimer gives a different mutation spectrum and increased error rate compared with replication past this lesion in uninduced cells, *J. Bacteriol.*, 172, 2105, 1990.
197. **Lawrence, C. W., Borden, A., Banerjee, S. K., and LeClerc, J. E.,** Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector, *Nucl. Acids Res.*, 18, 2153, 1990.
197a. **Rajagopalan, M., Lu, C., Woodgate, R., O'Donnell, M., Goodman, M. F., and Echols, H.,** Activity of the purified mutagenesis proteins UmuC, UmuD', and RecA in replicative bypass of an abasic DNA lesion by DNA polymerase III, *Proc. Natl. Acad. Sci. U.S.A.*, 89, 10777, 1992.
198. **Moore, P. D., Bose, K. K., Rabkin, S. D., and Strauss, B. S.,** Sites of termination of *in vitro* DNA synthesis on ultraviolet- and *N*-acetylaminofluorene-treated ϕX174 templates by prokaryotic and eukaryotic DNA polymerases, *Proc. Natl. Acad. Sci. U.S.A.*, 78, 110, 1981.
199. **Banerjee, S. K., Christensen, R. B., Lawrence, C. W., and LeClerc, J. E.,** Frequency and spectrum of mutations produced by a single *cis-syn* thymine-thymine cyclobutane dimer in a single-stranded vector, *Proc. Natl. Acad. Sci. U.S.A.*, 85, 8141, 1988.
200. **Taylor, J.-S. and O'Day, C. L.,** *cis-syn* thymine dimers are not absolute blocks to replication by DNA polymerase I of *Escherichia coli in vitro, Biochemistry*, 29, 1624, 1990.
201. **Wang, C. I. and Taylor, J. S.,** *In vitro* evidence that UV-induced frameshift and substitution mutations at T tracts are the result of misalignment-mediated replication past a specific thymine dimer, *Biochemistry*, 31, 3671, 1992.
202. **Hayes, R. C., Petrullo, L. A., Huang, H. M., Wallace, S. S., and LeClerc, J. E.,** Oxidative damage in DNA. Lack of mutagenicity by thymine glycol lesions, *J. Mol. Biol.*, 201, 239, 1988.
203. **Ide, H., Kow, Y. W., and Wallace, S. S.,** Thymine glycols and urea residues in M13 DNA constitute replicative blocks *in vitro, Nucl. Acids Res.*, 13, 8035, 1985.
204. **Loechler, E. L., Green, C. L., and Essigmann, J. M.,** *In vivo* mutagenesis by O^6-methylguanine built into a unique site in a viral genome, *Proc. Natl. Acad. Sci. U.S.A.*, 81, 6271, 1984.
205. **Eadie, J. S., Conrad, M., Toorchen, D., and Topal, M. D.,** Mechanism of mutagenesis by O^6-methylguanine, *Nature*, 308, 8, 1984.
206. **Topal, M. D.,** Mutagenesis by incorporation of alkylated nucleotides, *Basic Life Sci.*, 31, 339, 1985.
207. **Chambers, R. W., Sledziewska, G. E., Hirani, H. S., and Borowy, B. H.,** *uvrA* and *recA* mutations inhibit a site-specific transition produced by a single O^6-methylguanine in gene *G* of bacteriophage ϕX174, *Proc. Natl. Acad. Sci. U.S.A.*, 82, 7173, 1985.
208. **Bhanot, O. P. and Ray, A.,** The *in vivo* mutagenic frequency and specificity of O^6-methylguanine in ϕX174 replicative form DNA, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 7348, 1986.
209. **Preston, B. D., Singer, B., and Loeb, L. A.,** Mutagenic potential of O^4-methylthymine *in vivo* determined by an enzymatic approach to site-specific mutagenesis, *Proc. Natl. Acad. Sci. U.S.A.*, 83, 8501, 1986.
210. **Rossi, S. C. and Topal, M. D.,** Mutagenic frequencies of site-specifically located O^6-methylguanine in wild-type *Escherichia coli* and in a strain deficient in ada-methylguanine, *J. Bacteriol.*, 173, 1201, 1991.
211. **Dosanjh, M. K., Essigmann, J. M., Goodman, M. F., and Singer, B.,** Comparative efficiency of forming m4T · G vs. m4T · A base pairs at a unique site by use of *Escherichia coli* DNA polymerase I (Klenow fragment) and *Drosophila melanogaster* polymerase alpha-primase complex, *Biochemistry*, 29, 4698, 1990.
212. **Loeb, L. A. and Preston, B. D.,** Mutagenesis by apurinic/apyrimidinic sites, *Annu. Rev. Genet.*, 20, 201, 1986.
213. **Boiteux, S. and Laval, J.,** Coding properties of poly(deoxycycytidylic acid) templates containing uracil or apyrimidinic sites: *in vitro* modulation of mutagenesis by DNA repair enzymes, *Biochemistry*, 21, 6746, 1982.
214. **Sagher, D. and Strauss, B. S.,** Insertion of nucleotides opposite apurinic/payrimidinic sites in deoxyribonucleic acid during *in vitro* synthesis: uniqueness of adenine nucleotides, *Biochemistry*, 22, 4518, 1983.
215. **Schaaper, R. M., Kunkel, T. A., and Loeb, L. A.,** Infidelity of DNA synthesis associated with bypass of apurinic sites, *Proc. Natl. Acad. Sci. U.S.A.*, 80, 487, 1983.
216. **Strauss, B. S.,** The 'A rule' of mutagen specificity: a consequence of DNA polymerase bypass of non-instructional lesions?, *Bioessays*, 13, 79, 1991.
217. **Millican, T. A., Mock, G. A., Chauncey, M. A., Patel, T. P., Eaton, M. A., Gunning, J., Cutbush, S. D., Neidle,S., and Mann, J.,** Synthesis and biophysical studies of short oligodeoxynucleotides with novel modifications: a possible approach to the problem of mixed base oligodeoxynucleotide synthesis, *Nucl. Acids Res.*, 12, 7435, 1984.
218. **Takeshita, M., Chang, C. N., Johnson, F., Will, S., and Grollman, A. P.,** Oligodeoxynucleotides containing synthetic abasic sites. Model substrates for DNA polymerases and apurnic/apyrimidinic endonucleases, *J. Biol. Chem.*, 262, 10171, 1987.
219. **Ng, L., Weiss, S. J., and Fisher, P. A.,** Recognition and binding of template primers containing defined abasic sites by *Drosophila* DNA polymerase alpha holoenzyme, *J. Biol. Chem.*, 264, 13018, 1989.

220. **Cai, H., Bloom, L. B., and Goodman, M. F.,** unpublished, 1992.
221. **Cuniasse, P., Fazakerley, G. V., Guschlbauer, W., Kaplan, B. E., and Sowers, L. C.,** The abasic site as a challenge to DNA polymerase. A nuclear magnetic resonance study of G, C and T opposite a model abasic site, *J. Mol. Biol.*, 213, 303, 1990.
222. **Kalnik, M. W., Chang, C. N., Grollman, A. P., and Patel, D. J.,** NMR studies of basic sites in DNA duplexes: deoxyadenosine stacks into the helix opposite the cyclic analogue of 2-deoxyribose, *Biochemistry*, 27, 924, 1988.
223. **Fersht, A. R. and Knill, J. J. W.,** Contribution of $3' \rightarrow 5'$ exonuclease activity of DNA polymerase III holoenzyme from *Escherichia coli* to specificity, *J. Mol. Biol.*, 165, 669, 1983.
224. **Lu, C., Scheuermann, R. H., and Echols, H.,** *RecA* protein and SOS: correlation of mutagenesis phenotype with binding of mutant *RecAs* to duplex DNA and *LexA* cleavage, *J. Mol. Biol.*, 196, 497, 1987.
225. **Aboul-ela, F., Koh, D., Tinoco, I. J., and Martin, F. H.,** Base-base mismatches. Thermodynamics of double helix formation for $dCA_3XA_3G + dCT_3YT_3G$ (X,Y = A,C,G,T), *Nucl. Acids Res.*, 13, 4811, 1985.
226. **Gaffney, B. L. and Jones, R. A.,** Thermodynamic comparison of the base pairs formed by the carcinogenic lesion O^6-methylguanine with reference both to Watson-Crick pairs and to mismatched pairs, *Biochemistry*, 28, 5881, 1989.
227. **Petruska, J. and Goodman, M. F.,** unpublished, 1992.
228. **Patel, S. S., Wong, I., and Johnson, K. A.,** Pre-steady state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant, *Biochemistry*, 30, 511, 1991.
229. **Eger, B. T. and Benkovic, S. J.,** Minimal kinetic mechanism for misincorporation by DNA polymerase I (Klenow fragment), *Biochemistry*, 31, 9227, 1992.
230. **Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., and Benkovic, S. J.,** Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4, *Biochemistry*, 31, 10984, 1992.
231. **Ornstein, R. L., Rein, R., Breen, D. L., and MacElroy, R. D.,** An optimized potential function for the calculation of nucleic acid interaction energies. I. Base stacking, *Biopolymers*, 17, 2341, 1978.