THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 44, Issue of October 31, pp. 27931–27935, 1997

*Printed in U.S.A.*

# Appendix

## GEL KINETIC ANALYSIS OF POLYMERASE FIDELITY IN THE PRESENCE OF MULTIPLE ENZYME DNA ENCOUNTERS*

(Received for publication, March 19, 1997, and in revised form, July 28, 1997)

**D. Kuchnir Fygenson‡ and Myron F. Goodman§**

*From the Department of Biological Sciences, Hedco Molecular Biology Laboratories, University of Southern California, Los Angeles, California 90089-1340*

The kinetic analysis of the gel fidelity assay for polymerases with exonucleolytic activity was derived assuming SCH[1] (1). Occasionally, however, polymerase fidelity is too high or processivity is too low to detect misincorporation under standard SCH conditions (defined by Poisson statistics as <20% of p/t DNA extended). Greater quantities of p/t DNA can be exposed to SCH conditions by using trap DNA, which sequesters a polymerase after it dissociates from a p/t. Alternatively, the reaction can be run into MCH conditions. In this case, more p/t DNA is processed and the incorporation signal (gel bands at or beyond the target site) may even be enhanced at the expense of the unincorporated signal (gel band before the target site), but the interpretation of the band intensity ratios must be re-evaluated.

This appendix is concerned with extending the SCH kinetic analysis to MCH conditions. Consequences of MCH are very different for polymerases with and without exonuclease activity and for incorporation of right and wrong nucleotides. Polymerases that lack exonuclease activity eventually incorporate a nucleotide (right or wrong) opposite a target site on every available p/t DNA. When incorporation is rare, as for the wrong nucleotide, the ratio of gel band intensities increases in proportion to the number of hits. When incorporation is frequent, as for the right nucleotide, the ratio of band intensities increases exponentially with the number of hits. Polymerases with exonuclease activity establish a steady state distribution between incorporation and nonincorporation after a second hit. At steady state the gel band intensity ratio is always greater than for SCH, but only to the extent that the polymerase refuses to excise the nucleotide incorporated at the target site. For the wrong nucleotide the enhancement is generally less than a factor of two; for the right one it can be an order of magnitude. The conclusion is that MCH conditions can tractably and usefully enhance the incorporation of wrong nucleotides, but are best avoided (and generally unnecessary) for measuring incorporation of a right nucleotide.

The presentation follows. The first section reviews the original kinetic model and SCH analysis. The second section derives expressions for interpreting gel band intensity ratios caused by multiple completed hits. This section has three parts: one on the case without exonuclease activity, one on the case with exonuclease activity, and one on the effect of suppressing exonuclease activity with increased concentrations of next correct (rescue) nucleotide. To apply the various MCH expressions in the interpretation of experimental results, it is essential to account for the distribution in the number of hits across the population of p/t DNA in a reaction. This is the topic of the third section. The net result is a set of guidelines and correction factors for the practical application and interpretation of MCH conditions in the measurement of polymerase fidelity. The fourth, and final, section summarizes the conclusions and emphasizes the underlying assumptions.

*Kinetic Analysis of Polymerase Activity Assuming Single Completed Hits*—A kinetic analysis based on a Markov model describing polymerase action, including exonucleolytic proofreading, was introduced earlier (1) and is summarized in Fig. 1. Three absorbing states (1, 2, and 3) represent polymerase dissociation at or near the target site, with corresponding gel band intensities $I_{\mathbf{T-1}}$, $I_{\mathbf{T}}$, and $I_{\mathbf{T+1}}$. Two transient states (4 and 5) represent addition or excision of a nucleotide at the target site. Transitions between the states are described by their associated probabilities in terms of rate constants and can be expressed in matrix form as follows.

$$P(\mathrm{d}t) = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline k_{\mathrm{off}}\mathrm{d}t & 0 & 0 & 1-k_{\mathrm{off}}\mathrm{d}t-k_{\mathrm{pol}}\mathrm{d}t & k_{\mathrm{pol}}\mathrm{d}t \\ 0 & k'_{\mathrm{off}}\mathrm{d}t & k_{\mathrm{res}}\mathrm{d}t & k_{\mathrm{exo}}\mathrm{d}t & 1-k'_{\mathrm{off}}\mathrm{d}t-k_{\mathrm{res}}\mathrm{d}t-k_{\mathrm{exo}}\mathrm{d}t \end{array}\right] \quad \text{(Eq. 1)}$$

This model traces the movement of a polymerase during a single encounter with an unextended p/t DNA. Upon binding, a polymerase rapidly adds several correct (running start) nucleotides to a p/t until it enters transient state 4 at the **T−1** site. From state 4 it either moves to absorbing state 1 by dissociating, with rate $k_{\mathrm{off}}$, or to transient state 5 by adding a target nucleotide, with rate $k_{\mathrm{pol}}$. From state 5, it can either return to transient state 4 by excising the target nucleotide, with rate $k_{\mathrm{exo}}$, move to absorbing state 2 by dissociating, with rate $k'_{\mathrm{off}}$, or move to absorbing state 3 by adding the next, correct (rescue) nucleotide, with rate $k_{\mathrm{res}}$. It is assumed that polymerase does not excise the running start or rescue nucleotides. (We have verified that this assumption is satisfied using our standard assay conditions (2).

The relative probabilities for a polymerase to dissociate into states 1, 2, or 3 (*i.e.* the relative intensities of gel bands $I_{\mathbf{T-1}}$, $I_{\mathbf{T}}$,

* This work was supported by National Institutes of Health Grants GM21422 and GM42554 and by a grant from The Jane Coffin Childs Memorial Fund for Medical Research. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Fellow of the Jane Coffin Childs Memorial Fund for Medical Research.

§ To whom correspondence should be addressed: Dept. of Biological Sciences, University of Southern California, SHS Rm. 172, University Park, Los Angeles, CA 90089-1340. Tel.: 213-740-5190; Fax: 213-740-8631; E-mail: mgoodman@mizar.usc.edu.

[1] The abbreviations used are: SCH, single completed hit conditions referring to extension of a primer via interaction with a DNA polymerase at most once, followed by polymerase dissociation; MCH, multiple completed hit conditions referring to extension of a primer via multiple interactions with DNA polymerase; p/t, primer/template; **T** (in boldface), refers to the template target site at which fidelity is measured and should not be confused with T which refers to a template thymine.

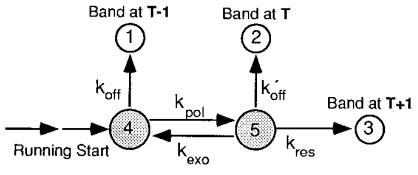


FIG. 1. **Minimal Markov model for a DNA polymerase with an associated exonuclease activity.** Transient states (*4* and *5*) represent insertion or excision of a nucleotide at the target site and are indicated by *large gray circles*. Absorbing states (*1, 2,* and *3*) represent polymerase dissociation at or near the target site and are indicated by *small white circles*. Transitions between states are indicated by *single arrows* labeled with the corresponding transition rate. Upon binding, a polymerase rapidly adds several correct (running start) nucleotides to a p/t until it enters transient state 4 at the **T–1** site. From state 4 it either moves to absorbing state 1 by dissociating with rate $k_{off}$, or to transient state 5 by inserting a target nucleotide, with rate $k_{pol}$. From state 5, it can either return to transient state 4 by excising the target nucleotide, with rate $k_{exo}$, move to absorbing state 2 by dissociating, with rate $k_{off}$, or move to absorbing state 3 by inserting the next, correct (rescue) nucleotide, with rate $k_{res}$. It is assumed that the polymerase does not excise the running start or rescue nucleotides.

or $I_{\mathbf{T+1}}$) can be calculated by infinite application of the transition matrix (3) (Equation 1). The meaningful element of the resulting "hit" matrix is the lower left quadrant (LLQ), which lists the contribution of each transient state to each absorbing state after a completed hit.

LLQ =

$$\begin{array}{c} \\ 4 \rightarrow \\ 5 \rightarrow \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \\ \uparrow & \uparrow & \uparrow \end{array} \\ \begin{bmatrix} k_{off}(k_{res} + k'_{off} + k_{exo})dt^2 & k_{pol}k'_{off}dt^2 & k_{pol}k_{res}dt^2 \\ k_{exo}k_{off}dt^2 & k'_{off}(k_{pol} + k_{off})dt^2 & k_{res}(k_{pol} + k_{off})dt^2 \end{bmatrix} \end{array} \quad \text{(Eq. 2)}$$

The matrix entries are actually probabilities, with the normalization factor,

$$(k_{off}k_{res}dt^2 + k_{off}k'_{off}dt^2 + k_{off}k_{exo}dt^2 + k_{pol}k'_{off}dt^2 + k_{pol}k_{res}dt^2)^{-1} \quad \text{(Eq. 3)}$$

removed for clarity. Entries in the top row of this matrix are the probabilities that a polymerase, beginning in transient state 4, ends up in absorbing states 1, 2, or 3. Similarly, entries in the bottom row characterize the redistribution of polymerase initially in transient state 5 into the three absorbing states.

Under SCH conditions, every polymerase approaches the target site from state 4, and relative gel band intensities reflect rate constants as listed in the top row of the matrix. In general,

$$R_{SCH} \equiv \frac{I^{\Sigma}_{\mathbf{T}}}{I_{\mathbf{T-1}}} = \frac{k_{pol}(k_{res} + k'_{off})}{k_{off}(k_{res} + k'_{off} + k_{exo})} \quad \text{(Eq. 4a)}$$

where $I^{\Sigma}_{\mathbf{T}} = I_{\mathbf{T}} + I_{\mathbf{T+1}} \ldots$ is the sum of gel band intensities at and beyond the target site. The sum $(k_{res} + k'_{off})$ represents the net rate of moving "beyond" the target site. Experimental conditions can be adjusted to maximize this rate by introducing high concentrations of rescue dNTP (*i.e.* $k_{res} >> k'_{off}$), and then

$$R_{SCH} \equiv \frac{I^{\Sigma}_{\mathbf{T}}}{I_{\mathbf{T-1}}} = \frac{k_{pol}k_{res}}{k_{off}(k_{res} + k_{exo})} \quad \text{(Eq. 4b)}$$

or to minimize it, by the absence of rescue dNTP (*i.e.* $k'_{off} >> k_{res} = 0$), in which case,

$$R_{SCH} \equiv \frac{I^{\Sigma}_{\mathbf{T}}}{I_{\mathbf{T-1}}} = \frac{k_{pol}k'_{off}}{k_{off}(k'_{off} + k_{exo})} \quad \text{(Eq. 4c)}$$

For polymerases lacking exonuclease activity, $k_{exo} = 0$, and the equations above reduce to the following.

$$R_{SCH} \equiv \frac{I^{\Sigma}_{\mathbf{T}}}{I_{\mathbf{T-1}}} = \frac{k_{pol}}{k_{off}}. \quad \text{(Eq. 4d)}$$

It has been shown that the ratio of gel band intensities calculated above corresponds to the $V_{max}/K_m$ of the nucleotide incorporation reaction (1). A general expression for the fidelity in terms of these microscopic rate constants is given by the following.

Fidelity ≡

$$\frac{(V_{max}/K_m)_R}{(V_{max}/K_m)_W} = \frac{R^R_{SCH}}{R^W_{SCH}} = \frac{k^R_{pol}(k^R_{res} + k'^R_{off})}{(k^R_{res} + k'^R_{off} + k^R_{exo})} \Big/ \frac{k^W_{pol}(k^W_{res} + k'^W_{off})}{(k^W_{res} + k'^W_{off} + k^W_{exo})} \quad \text{(Eq. 5)}$$

Since $k^R_{res} > k^W_{res}$ and $k^R_{exo} > k^W_{exo}$, it can be seen from Equation 5 that the fidelity increases monotonically with decreasing values of $k'_{off}$. Thus, the effect of increasing the lifetime of the polymerase on the p/t DNA, *e.g.* the presence of processivity proteins, can only act to increase polymerase fidelity.

*Kinetic Analysis of Polymerase Activity Assuming Multiple Completed Hits*—The model outlined above for a single hit can be applied multiple times to trace the movement of a polymerase through multiple encounters with a p/t DNA. p/t DNA that was left in state 1 after a hit re-enters state 4 on a subsequent hit and is redistributed between states 1, 2, and 3 in the same manner (Equation 2). p/t DNA left in state 2 re-enters state 5 on a subsequent hit and is redistributed between states 1, 2, and 3 according to the lower row of the lower left quadrant (Equation 2). p/t DNA left in state 3 presumably remains in state 3 on all subsequent hits. Outcomes differ qualitatively when the polymerase lacks or possesses exonuclease activity, so the two cases are best dealt with separately.

*MCH in the Absence of Exonuclease Activity*—In the absence of exonuclease activity, dissociation into state 2 or state 3 means the same thing: permanent incorporation opposite the target site. Without loss of generality, then, we can lump states 2 and 3 together and consider only four transition probabilities: $p_{4\rightarrow1}$, $p_{4\rightarrow\Sigma} \equiv p_{4\rightarrow2} + p_{4\rightarrow3}$, $p_{5\rightarrow1}$, and $p_{5\rightarrow\Sigma} \equiv p_{5\rightarrow2} + p_{5\rightarrow3}$, with values given by the entries in the hit matrix (Equation 2) divided by the normalization factor (Equation 3). Note that $p_{4\rightarrow1} + p_{4\rightarrow\Sigma} = p_{5\rightarrow1} + p_{5\rightarrow\Sigma} = 1$. The consequences of a hit are calculated by multiplying the transpose of the hit matrix by the vector of relative band intensities,

$$\begin{bmatrix} p_{4\rightarrow1} & p_{5\rightarrow1} \\ p_{4\rightarrow\Sigma} & p_{5\rightarrow\Sigma} \end{bmatrix} \begin{bmatrix} I_{\mathbf{T-1}}(n) \\ I^{\Sigma}_{\mathbf{T}}(n) \end{bmatrix} = \begin{bmatrix} I_{\mathbf{T-1}}(n+1) \\ I^{\Sigma}_{\mathbf{T}}(n+1) \end{bmatrix}, \quad \text{(Eq. 6a)}$$

$$\begin{bmatrix} I_{\mathbf{T-1}}(0) \\ I^{\Sigma}_{\mathbf{T}}(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \text{(Eq. 6b)}$$

where $n$ is the number of previous hits.

Lack of exonuclease activity, $k_{exo} = 0$, implies $p_{5\rightarrow1} = 0$, and $p_{5\rightarrow\Sigma} = 1$. In this case, as noted (Equation 4d), the ratio of incorporated $I^{\Sigma}_{\mathbf{T}}$ to unincorporated $I_{\mathbf{T-1}}$ after a first hit is

$$R_1 = R_{SCH} = \frac{p_{4\rightarrow\Sigma}}{p_{4\rightarrow1}} = \frac{I^{\Sigma}_{\mathbf{T}}}{I_{\mathbf{T-1}}} = \frac{k_{pol}}{k_{off}}. \quad \text{(Eq. 7)}$$

After a second hit, it becomes

$$R_2 = \frac{p_{4\rightarrow\Sigma} + (p_{4\rightarrow1})p_{4\rightarrow\Sigma}}{{p_{4\rightarrow1}}^2} = R_1\left(1 + \frac{1}{p_{4\rightarrow1}}\right), \quad \text{(Eq. 8)}$$

and, after a third hit, it becomes

$$R_3 = \frac{p_{4\rightarrow\Sigma} + (p_{4\rightarrow1})p_{4\rightarrow\Sigma} + (p_{4\rightarrow1})^2p_{4\rightarrow\Sigma}}{{p_{4\rightarrow1}}^3} = R_1\left(1 + \frac{1}{p_{4\rightarrow1}} + \frac{1}{{p_{4\rightarrow1}}^2}\right). \quad \text{(Eq. 9)}$$

In the limit of $n$ hits, the ratio is

$$R_n = R_1 \sum_{i=0}^{n-1} p^{-i}_{4\rightarrow1}. \quad \text{(Eq. 10)}$$

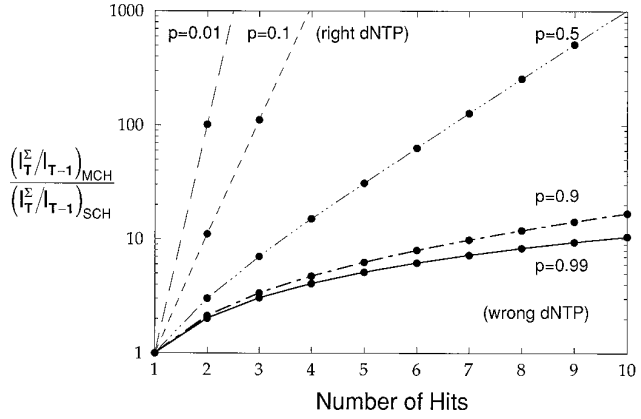


FIG. 2. **Correction factor for gel band intensity ratio as a function of the number of completed hits.** The correction factor relating $R_n$ to $R_{SCH}$ is calculated based on Equation 10, for different values of the probability of failing to insert opposite the target site: ~1 for the wrong dNTP; ~0 for the right dNTP. The correction factor is discrete like its argument, the number of hits. *Lines* are drawn connecting the points to guide the eye. Note that for the wrong dNTP (~1), the dependence is essentially linear.

When the probability, $p_{4\to1}$, of dissociating into state 1 (*i.e.* not incorporating a base opposite the target site) is small (*e.g.* when the right nucleotide is available for incorporation opposite the target site), its inverse is large and the MCH correction factor, dominated by the term with $i = n - 1$, grows exponentially with the number of hits. By contrast, when $p_{4\to1}$ is close to 1 (*e.g.* when only wrong nucleotide is available for incorporation opposite the target site), every term in the sum is similarly of order 1, and the correction factor is approximately $n$, the number of hits.

The obvious and important lesson from this analysis is that, when both right and wrong nucleotides are run under identical MCH conditions, the correction factors will not cancel in the ratio of their apparent $V_{max}/K_m$. In fact, a serious *over*-estimate of polymerase fidelity may result for, while misincorporation is enhanced by multiple hits, proper incorporation is enhanced much more.[2] MCH correction factors applicable in the absence of exonuclease activity are plotted as a function of $n$, the number of hits, for a range of probabilities $p_{4\to1}$ in Fig. 2.

*MCH in the Presence of Exonuclease Activity without Rescue Nucleotide*—In the presence of exonuclease activity, multiple hits can redistribute p/t DNA from state 5 among all three absorbing states. However, when the rescue nucleotide is not available, state 3 is not populated: $k_{res} = 0$, which implies $p_{4\to3} = p_{5\to3} = 0$. Referring back to the simplified hit matrix (Equation 6), $p_{5\to\Sigma} = p_{5\to2}$, $p_{4\to\Sigma} = p_{4\to2}$.

As mentioned before (Equation 4c), the gel band intensity ratio after the first hit is

$$R_1 = R_{SCH} = \frac{p_{4\to2}}{p_{4\to1}} = \frac{k_{pol}k'_{off}}{k_{off}(k'_{off} + k_{exo})}. \qquad \text{(Eq. 11)}$$

After a second hit, it becomes

$$R_2 = \frac{(p_{4\to2})p_{4\to1} + (p_{5\to2})p_{4\to2}}{p_{4\to1}^2 + p_{5\to1}p_{4\to2}}. \qquad \text{(Eq. 12)}$$

Eventually it reaches a steady state, when $p_{4\to2} \times I_{\mathbf{T-1}}(n) = p_{5\to1} \times I_{\mathbf{T}}^{\Sigma}(n)$, which implies a band intensity ratio of

[2] As was noted in an earlier work (1) the exception to this rule arises when, the concentration of right nucleotides is suppressed to the point that their insertion rate is as low as that of a wrong nucleotide. In this special case, and when all other reaction conditions are identical, the time evolution of the reactions for right and wrong nucleotides are identical, and the ratio of their apparent $V_{max}/K_m$ values is independent of time (*i.e.* number of hits).

TABLE I
*Proportion of primer/template DNA encountering DNA polymerase multiple times*

Values in the table were calculated based on a Poisson process for the amount of primer/template used listed in the first column. Weighting of MCH correction factors is based on the relative proportions of p/t that experience various numbers of hits. Relative proportions are calculated by dividing the absolute proportions listed below by the proportion of p/t used, listed in the first column.

| p/t used | Hit once | Hit twice | Hit three times | Hit four times |
|---|---|---|---|---|
| % | % | % | % | % |
| 10 | 9.5 | 0.5 | 0.0 | 0.0 |
| 20 | 17.9 | 2.0 | 0.1 | 0.0 |
| 30 | 25.0 | 4.5 | 0.5 | 0.0 |
| 40 | 30.7 | 7.8 | 1.3 | 0.2 |
| 50 | 34.7 | 12.0 | 2.8 | 0.5 |
| 60 | 36.7 | 16.8 | 5.1 | 1.2 |
| 70 | 36.1 | 21.7 | 8.7 | 2.6 |
| 80 | 32.2 | 25.9 | 13.9 | 5.6 |
| 90 | 23.0 | 26.5 | 20.3 | 11.7 |

$$R_{steady\ state} = \frac{p_{4\to2}}{p_{5\to1}} = \frac{k_{pol}k'_{off}}{k_{off}k_{exo}} = R_{SCH}\left(1 + \frac{k'_{off}}{k_{exo}}\right). \qquad \text{(Eq. 13)}$$

Experimental error in the measurement of gel band intensity ratios is typically ± 30%. It is an algebraic exercise to show that $R_2 \geq 0.75 \times R_{steady\ state}$ when either $p_{5\to1} > 0.5$ (typical for the wrong nucleotide) or $k'_{off} > k_{exo}$ (typical for the right nucleotide). Thus, after two hits in the presence of exonuclease activity the gel band intensity ratio is effectively independent of the number of hits.

*MCH in the Presence of Exonuclease Activity with Rescue dNTP*—The availability of rescue nucleotide introduces another path for dissociation from the target site, namely through sites beyond the target. Whether polymerase will return to the target site from state 3 by excising correct nucleotides incorporated downstream is not certain.

When the right nucleotide is incorporated at the target site, it is reasonable to assume that $k_{res} > k'_{off} >> k_{exo}$. Then state 3 is well approximated as an absorbing state and the effect of multiple hits is analogous to that which applies in the absence of exonuclease activity (Equation 10).

When the wrong nucleotide is incorporated at the target site, the absence of a gel band at the target site justifies the assumption that state 3 is an absorbing state under SCH conditions. However, MCH introduce new opportunities for excision of rescue nucleotides, since a polymerase may bind a p/t DNA with its exo site on a second hit more easily than switch from polymerase to exo site (without dissociating) on a first hit. This would be consistent with the finding that excision of a mismatched base-analog 2-aminopurine is favored from as far as 5 bases upstream (4).

Based on this reasoning, it is most appropriate treat state 3 in our model as an absorbing state for the purposes of an individual hit, but lump states 2 and 3 together in the hit matrix, since p/t DNA from either one re-enters state 5 at the beginning of a new hit. The effect of rescue nucleotide is then simply to increase $p_{4\to\Sigma} \equiv p_{4\to2} + p_{4\to3}$ and $p_{5\to\Sigma} \equiv p_{5\to2} + p_{5\to3}$ by increasing $k_{res}$. The SCH and steady state band intensity ratios (Equations 11 and 13) become

$$R_{SCH} = \frac{p_{4\to\Sigma}}{p_{4\to1}} = \frac{k_{pol}(k'_{off} + k_{res})}{k_{off}(k'_{off} + k_{res} + k_{exo})} \qquad \text{(Eq. 14)}$$

and

$$R_{steady\ state} = \frac{p_{4\to\Sigma}}{p_{5\to1}} = \frac{k_{pol}(k'_{off} + k_{res})}{k_{off}k_{exo}} = R_{SCH}\left(1 + \frac{(k'_{off} + k_{res})}{k_{exo}}\right). \qquad \text{(Eq. 15)}$$

Without or with rescue nucleotide, the steady state MCH correction factor (Equation 13 or 15) is never less than 1.

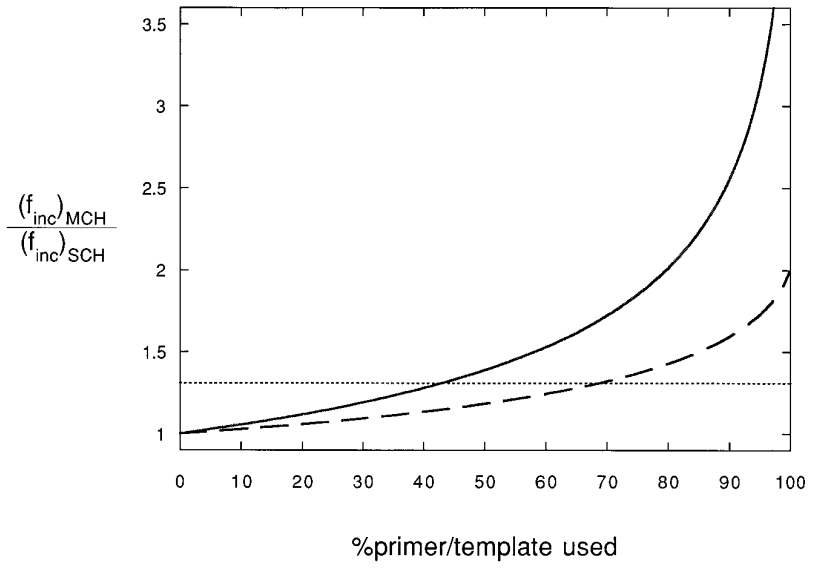


FIG. 3. **Composite MCH correction factor for the misincorporation efficiency as a function of the percentage of primer/template used in the reaction.** Assuming MCH conditions are used when measuring incorporation of the wrong dNTP and SCH conditions are used when measuring incorporation of the right dNTP, the correction factor relating $R_{\mathrm{MCH}}$ to $R_{\mathrm{SCH}}$ is calculated. The *solid line* is the correction factor for polymerases which lack exonuclease activity (Equation 20). The *broken line* is an upper bound on the correction factor for polymerases with exonuclease activity (Equation 22). The *dotted line* is at a value of 1.3, reprresenting the sensitivity limit of the assay.

However, since the rate of excising a wrong nucleotide is generally faster than the rate of dissociation from a mismatch, the correction factor for the wrong nucleotide is generally less than 2 and potentially even negligible (*e.g.* in the presence of processivity proteins). For the right nucleotide, as in the absence of exonuclease activity, the correction factor can be large.

Again, the obvious and important lesson is that, when both right and wrong nucleotides are run under identical MCH conditions, the correction factors will not cancel in the ratio of their apparent $V_{\max}/K_m$. Since incorporation of the right nucleotide is easily measured under SCH conditions, the recommendation is to use MCH conditions only when necessary for measuring incorporation of a wrong nucleotide. Proper application of MCH correction factors depends on how many times the p/t DNA in a reaction were hit by polymerase, which is the subject of the next section.

*Distribution of the Number of Hits*—By adjusting the relative concentration of polymerase (low) and p/t DNA (high), it is possible to make encounters between polymerase and p/t rare, so that the number of times a DNA is hit by polymerase is appropriately modeled as a Poisson process. In this case, the distribution of the number of hits is well known,

$$f(n,t) = e^{-kt}\frac{(kt)^n}{n!} \qquad \text{(Eq. 16)}$$

where $n$ is the number of hits, $t$ is the reaction time, and $k$ is the frequency of hits. In a running start experiment, the reaction can be clocked without explicitly measuring $k$ and $t$, according to the decay in the number of unextended DNA molecules, $f(0,t) = e^{-kt}$. For example, if 50% of the primers remain unextended, $kt = -\ln(0.5) = 0.69$; $f(1) = 0.347$; $f(2) = 0.120$; $f(3) = 0.028$; $f(4) = 0.005$ and, thus, 0.347/0.5 = 69% of the extended p/t DNA experienced only one hit, 0.12/0.5 = 24% experienced two, 0.028/0.5 = 6% experienced three, and only 0.005/0.5 = 1% experienced four. The absolute percentages of p/t DNA hit different numbers of times are listed in Table I for different amounts of p/t DNA used.

The observed band intensities and their ratio must be interpreted in terms of the distribution of the number of hits in the population of p/t DNA and their respective contributions to the bands. The intensity observed in bands at and beyond the target site, $I_{\mathbf{T}}^{\Sigma}$, is the weighted sum of the numerators, $I_{\mathbf{T}}^{\Sigma}(n)$, of the band ratios from different numbers of hits, $n$. Similarly, the observed intensity in the band before the target band, $I_{\mathbf{T-1}}$, is the weighted sum of the denominators, $I_{\mathbf{T-1}}(n)$, of the band ratios from different numbers of hits,

$$R_{\mathrm{observed}} = \frac{\sum_n f(n) I_{\mathbf{T}}^{\Sigma}(n)}{\sum_n f(n) I_{\mathbf{T-1}}(n)} = \frac{f(1)I_{\mathbf{T}}^{\Sigma}(1) + f(2)I_{\mathbf{T}}^{\Sigma}(2) + f(3)I_{\mathbf{T}}^{\Sigma}(3) + \cdots}{f(1)I_{\mathbf{T-1}}(1) + f(2)I_{\mathbf{T-1}}(2) + f(3)I_{\mathbf{T-1}}(3) + \cdots} \qquad \text{(Eq. 17a)}$$

where the weighting factor $f(n)$ is the fraction of the population that experienced $n$ hits and

$$\sum_{n=1}^{\infty} f(n) = 1 - f(0). \qquad \text{(Eq. 17b)}$$

Combining these results with those of the previous two sections, we can calculate the composite correction factor for MCH conditions that should be used when incorporating the wrong nucleotide opposite a target site.

In the absence of exonuclease activity, the denominators $I_{\mathbf{T-1}}$ are all roughly equal, so the observed MCH ratio is approximately the weighted sum of the individual ratios.

$$R_{\mathrm{MCH,\,no\,exo}}^{\mathrm{wrong}} \approx \frac{f(1)I_{\mathbf{T}}^{\Sigma}(1) + f(2)I_{\mathbf{T}}^{\Sigma}(2) + f(3)I_{\mathbf{T}}^{\Sigma}(3) + \cdots}{[1 - f(0)]I_{\mathbf{T-1}}(n)} = \frac{f(1)R_1 + f(2)R_2 + f(3)R_3 + \cdots}{1 - f(0)} \qquad \text{(Eq. 18)}$$

And, since the ratios are approximately the number of hits times the ratio for a single hit (*cf.* discussion following Equation 10),

$$R_{\mathrm{MCH,\,no\,exo}}^{\mathrm{wrong}} \approx \frac{f(1)R_1 + 2f(2)R_1 + 3f(3)R_1 + \cdots}{1 - f(0)} \qquad \text{(Eq. 19)}$$

the observed ratio is approximately the SCH ratio times the average number of hits in the population,

$$R_{\mathrm{MCH,\,no\,exo}}^{\mathrm{wrong}} = R_{\mathrm{SCH}}\frac{\sum_n nf(n)}{1 - f(0)} = R_{\mathrm{SCH}}\frac{-\ln f(0)}{1 - f(0)} \qquad \text{(Eq. 20)}$$

where

$$\sum_n nf(n) = \langle n \rangle = kt = -\ln f(0). \qquad \text{(Eq. 20a)}$$

This MCH correction factor is plotted in Fig. 3 (*solid line*) as a

function of $1 - f(0)$, the proportion of p/t DNA used. It becomes significant (>1.3) when more than 40% of the primer/template has been used.

In the presence of exonuclease activity, the numerators $I_{\mathbf{T}}^{\Sigma}(n)$ are all equal and the denominators for 2 or more hits are at steady state (ss).

$$R_{\text{observed exo}} \approx \frac{[1 - f(0)]I_{\mathbf{T}}^{\Sigma}(1)}{f(1)I_{\mathbf{T-1}}(1) + [1 - f(1) - f(0)]I_{\mathbf{T-1}}(\text{ss})} \quad \text{(Eq. 21)}$$

The exact correction factor depends on the relation between $I_{\mathbf{T-1}}(1) = k_{\text{off}}(k'_{\text{off}} + k_{\text{res}} + k_{\text{exo}})$ and $I_{\mathbf{T-1}}(\text{ss}) = k_{\text{off}}k_{\text{exo}}$ (*cf.* Equations 14 and 15). A worst case scenario for incorporation of the wrong nucleotide places $k_{\text{exo}} = k'_{\text{off}} + k_{\text{res}}$ for which $I_{\mathbf{T-1}}(\text{ss}) = I_{\mathbf{T}-1}(1)/2$. Therefore, the ratio is at most,

$$R_{\text{MCH, exo}}^{\text{wrong}} \leq \frac{[1 - f(0)]I_{\mathbf{T}}^{\Sigma}(1)}{(\tfrac{1}{2})[f(1) + 1 - f(0)]I_{\mathbf{T-1}}(1)} = R_{\text{SCH}} \times 2\left(1 + \frac{f(1)}{1 - f(0)}\right)^{-1} \quad \text{(Eq. 22)}$$

where $f(1) - f(0)\ln f(0)$. This worst case MCH correction factor is plotted in Fig. 3 (*broken line*) as a function of $1 - f(0)$, the proportion of p/t DNA used. It is only significant when more than 70% p/t has been used. Thus, to a good approximation, there is no need for a MCH correction factor when incorporating a wrong nucleotide in the presence of exonuclease activity.

*Conclusions*—In closing, two assumptions implicit in the preceding analysis merit emphasis. The working model was based on a primer template design that presents several running-start bases before the target site. It is assumed that the polymerase inserts opposite these sites without dissociation until it reaches the **T−1** site immediately before the target. This assumption is important for properly clocking the reaction in terms of the amount of primer used. On subsequent hits, the polymerase interacts with the p/t from a standing start. It is assumed that the rate constants for the reactions involved in DNA synthesis are the same for running and standing starts. This assumption is essential for rewriting the MCH band intensity ratio in terms of the SCH ratio.

The lesson of the MCH analysis is that correction factors for the right nucleotide are always much greater than those for the wrong nucleotide. Thus, the most accurate, sensitive, and convenient way to measure the fidelity using the gel kinetic assay is to use SCH conditions for incorporation of the right nucleotide and, if necessary, MCH conditions for the wrong nucleotide. Fidelity (or its inverse, the misincorporation efficiency, $f_{\text{inc}}$) is measured from the slope, $V_{\text{max}}/K_m$, of the line graph of the gel band intensity ratio as a function of (right or wrong) nucleotide concentration. MCH correction factors for the gel band intensity ratios computed in this appendix simply factor out of the slope, yielding a single overall correction factor for the misincorporation efficiency (Equations 20 and 22 and Fig. 3). These guidelines and correction factors enable the practical application and interpretation of MCH conditions in the measurement of polymerase fidelity.

REFERENCES

1. Creighton, S., and Goodman, M. F. (1995) *J. Biol. Chem.* **270,** 4759–4774
2. Bloom, L. B., Chen, X., Fygenson, D. K., Turner, J., O'Donnell, M., and Goodman, M. F. (1997) *J. Biol. Chem.* **272,** 27919–27930
3. Goodman, R. (1988) *Introduction to Stochastic Models*, pp. 157–164, The Benjamin/Cummings Publishing Co., Inc., Menlo Park, CA
4. Petruska, J., and Goodman, M. F. (1985) *J. Biol. Chem.* **260,** 7533–7539