Kinetic assays of DNA polymerase fidelity: a new theoretical perspective beyond Michaelis-Menten kinetics

Qiu-Shi Li,[1] Yao-Gen Shu,[2] Zhong-Can Ou-Yang,[3] and Ming Li[1,*]

[1] *School of Physical Science, University of Chinese Academy of Sciences*
[2] *Wenzhou Institute, University of Chinese Academy of Sciences*
[3] *Institute of Theoretical Physics, Chinese Academy of Sciences*

(Dated: May 7, 2021)

The high fidelity of DNA polymerase (DNAP) is critical for the faithful replication of DNA. There are several quantitative approaches to measure DNAP fidelity. Directly counting the error frequency in the replication products gives the true fidelity but it turns out very hard to implement in practice. Two biochemical kinetic approaches, the steady-state assay and the transient-state assay, were then suggested and widely adopted. In these assays, the error frequency is indirectly estimated by using kinetic theories combined with the measured apparent kinetic rates. However, whether it is equivalent to the true fidelity has never been clarified theoretically, and in particular there are different strategies using these assays to quantify the proofreading efficiency of DNAP but often lead to inconsistent results. In this paper, we make a comprehensive examination on the theoretical foundation of the two kinetic assays, based on the theory of DNAP fidelity recently proposed by us. Our studies show that while the conventional kinetic assays are generally valid to quantify the discrimination efficiency of DNAP, they are valid to quantify the proofreading efficiency of DNAP only when the kinetic parameters satisfy some constraints which will be given explicitly in this paper.These results may inspire more carefully-designed experiments to quantify DNAP fidelity.

PACS numbers: 82.39.-k, 87.15.Rn, 87.16.A-

I. INTRODUCTION

The high fidelity of DNA polymerase (DNAP) is critical for faithful replication of genomic DNA. Quantitative studies on DNAP fidelity began in 1960s' and became an important issue in biochemistry and molecular biology. Intuitively, the DNAP fidelity can be roughly understood as the reciprocal of the overall mismatch (error) frequency when a given DNA template is replicated with both the matched dNTPs (denoted as dRTP or R) and the mismatched dNTPs (denoted as dWTP or W). For instance, by using artificial simple template and conducting the replication reaction with both dRTP and dWTP, one can measure the ratio of the incorporated dRTPs to dWTPs in the final products so as to quantify the overall error frequency[1, 2]. Beyond such overall fidelity, the site-specific fidelity was defined as the reciprocal of the error frequency at individual template sites. In principle, the error frequency at any template site can be directly counted if a sufficient amount of full-length replication products can be collected and sequenced(this will be denoted as true fidelity $\mathscr{F}$ in this paper), *e.g.* by using deep sequencing techniques [3, 4]. However, this type of sequencing-based approach always requires a huge workload and was rarely adopted in fidelity assay. It is also hard to specify the sequence-context influences on the fidelity. A much simpler strategy is to only investigate the error frequency at the assigned template site by single-nucleotide incorporation assays. Such assays are conducted for exo^--DNAP (exonuclease-deficient DNAP), in which dRTP and dWTP compete for the same assigned template site and the amount of the final products containing the incorporated dRTP or dWTP are then determined by gel analysis to give the error frequency, *e.g.* [5, 6]. By designing various template sequences, one can further dissect the sequence-context dependence of the site-specific error frequency. Although the direct measurements seem simple and intuitive, they are actually very challenging since mismatches occur with too low probability to be detected even when heavily-biased dNTP pools are used. Besides, the single-nucleotide incorporation assays do not apply to exo^+-DNAP (exonuclease-efficient DNAP) because the coexistence of the polymerase activity and the exonuclease activity makes the reaction products very complicated and hard to interpret. Hence two alternative kinetic approaches were proposed, inspired by the kinetic proofreading theory of biosynthetic processes proposed by Hopfield [7] and Ninio[8].

The steady-state assay was developed by A. Fersht for exo^--DNAP, which is based on the Michaelis-Menten kinetics of the incorporation of a single dRTP or dWTP at the same assigned template site[9, 10]. The two incorporation reactions are conducted separately under steady-state conditions to obtain the so-called specificity constant (the quasi-first order rate constant) $(k_{cat}/K_m)_R$ or $(k_{cat}/K_m)_W$ respectively, k_{cat} is the maximal steady-state turnover rate of dNTP incorporation and K_m is the Michaelis constant. The site-specific fidelity is then characterized as the ratio between the two incorporation velocities, *i.e.* $(k_{cat}/K_m)_R[\mathrm{dRTP}]/(k_{cat}/K_m)_W[\mathrm{dWTP}]$ (denoted as steady-state fidelity $f_{s\cdot s}$) . However, this is

*Electronic address: liming@ucas.ac.cn

only an operational definition of DNAP fidelity which is not necessarily equal to the true fidelity (see discussions in Sec.III B). So far as we know, there was only one experiment work which did the comparison between $f_{s\cdot s}$ and $\mathscr{F}$ and indicated their possible equivalence for exo^--Klenow fragment (KF^-)[6], but no theoretical works have ever been published to examine the equivalence in general.

Besides the steady-state method, the transient-state kinetic analysis was also proposed to obtain the specificity constant[11, 12]. Under the pre-steady-state condition or the single-turnover condition, one can obtain the parameter k_{pol}/K_d (a substitute for k_{cat}/K_m) for the single-nucleotide incorporation reactions with exo^--DNAP, and define the site-specific fidelity as $(k_{pol}/K_d)_R[\mathrm{dRTP}]/(k_{pol}/K_d)_W[\mathrm{dWTP}]$ (denoted as transient-state fidelity $f_{t\cdot s}$). But again the relation between $f_{t\cdot s}$ and $\mathscr{F}$ is not yet clarified. Although the experiment has indicated the possible equivalence of $f_{t\cdot s}$ to $\mathscr{F}$ for KF^-[6], a general theoretical examination is still needed.

Further, these methods fail to unambiguously measure the site-specific fidelity of exo^+-DNAP. For exo^+-DNAP, the total fidelity is assumed to consist of two multiplier factors. The first is the initial discrimination f_{ini} contributed solely by the polymerase domain, which is often characterized by $f_{s\cdot s}$ or $f_{t\cdot s}$. The second factor is the additional proofreading efficiency f_{pro} contributed by the exonuclease domain, which is defined by the ratio of the elongation probability of the terminal R ($P_{el,R}$) to that of the terminal W ($P_{el,W}$). Here the elongation probability is defined as $P_{el} = k_{el}/(k_{el} + k_{ex})$, k_{el} is the rate of elongation to the next site, and k_{ex} is the rate of excising the primer terminal nucleotide (*e.g.* Eq.(A1-A6) in Ref.[13]). $P_{el,R}$ is usually assumed close to 100%, so f_{pro} equals approximately to $1 + k_{ex,W}/k_{el,W}$. Although these expressions seem reasonable, there are some problems that were not clarified. First, the mathematical definition of f_{pro} is subjective though intuitive, so a rigorous theoretical foundation is needed. Second, the apparent rate parameters k_{el} and k_{ex} are not well defined since both the elongation and the excision are multi-step processes. k_{el} and k_{ex} are unknown functions of the involved rate constants but there is not a unique way to define them. They could be theoretically defined under steady-state assumptions (Eq.(6) in Ref.[14]) or operationally defined by experiment assays (*e.g.* steady-state assays[15, 16] or transient-state assays [17]), but different ways often lead to different estimates of f_{pro} (see Sec.III). Additionally, k_{el} should be properly understood as the ultimate elongation rate in the sense that the elongated terminal (the added nucleotide) is no longer excised. This condition is not met if the exo^+-DNAP can proofread the buried mismatches (*e.g.* the penultimate or antepenultimate mismatches, etc.) . In these cases, k_{el} is affected not only by the next template site but also by further sites. Such far-neighbor effects were not seriously considered in previous studies. For these reasons, the widely cited initial discrimination $f_{ini} \approx 10^{4\sim5}$ and the proofreading efficiency $f_{pro} \approx 10^{2\sim3}$ [18] are questionable.

Recently two equivalent rigorous theories were proposed to investigate the true fidelity of either exo^--DNAP or exo^+-DNAP, *i.e.* the iterated function systems by P.Gaspard [19] and the first-passage (FP) method by us [20]. In particular, we can numerically rigorously compute the true fidelity ($\mathscr{F}$) of exo^+-DNAP by the FP method and can also derive the approximate analytical expressions (F) of the true fidelity. With these firmly established results, we can address all the above questions systematically. In the following sections, we will first give a brief review of the FP method and the major conclusions already obtained for the minimal kinetic model of DNA replication. Then we will generalize these conclusions to more realistic kinetic models for exo^--DNAP and exo^+-DNAP, and carefully examine the relations between $f_{s\cdot s}$, $f_{t\cdot s}$ and $\mathscr{F}$.

II. METHODS

A. Basics of the FP method

The first-passage (FP) method was proposed to study the replication of the entire template by exo^+-DNAP [20].

$$\cdots \rightleftharpoons E\cdot D_{i-1} \underset{\overline{r}_i}{\overset{\overline{k}_i}{\rightleftharpoons}} E\cdot D_i \underset{\overline{r}_{i+1}}{\overset{\overline{k}_{i+1}}{\rightleftharpoons}} E\cdot D_{i+1} \rightleftharpoons \cdots$$

FIG. 1: The minimal reaction scheme of DNA replication. E: the enzyme DNAP. D_i: the primer-template duplex with primer terminal at the template site i.

Here the minimal reaction scheme Fig.1 is taken as an example to illustrate the basic logic of this method. $\overline{k}_i$ is the rate of incorporating dNTP to the primer at the template site $i-1$ (the dNTP-concentration dependence of $\overline{k}_i$ is not explicitly shown here), $\overline{r}_i$ is the rate of excising the primer terminal at the template site i. Intuitively, $\overline{k}_i$ and $\overline{r}_i$ depend on the identity (A, G, T or C) and the state (matched or mismatched) not only of the base pair at site i but also of the one or more preceding base pairs. If there is only the nearest-neighbor (first-order) effect, $\overline{k}_i$ and $\overline{r}_i$ can be written as $\overline{k}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}$ and $\overline{r}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}$, X_{i-1} (or X_i) represents the nucleotide at site $i-1$ (or i) on the template, α_{i-1} represents the nucleotide at site $i-1$ on the primer, α_i represents the next nucleotide to be incorporated to the primer terminal at site i (for $\overline{k}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}$) or the terminal nucleotide of the primer at site i to be excised (for $\overline{r}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}$). X and α can be any of the four types of nucleotides. Similarly, there are $\overline{k}^{X_{i-2}X_{i-1}X_i}_{\alpha_{i-2}\alpha_{i-1}\alpha_i}$ etc. for the second-order neighbor effects, and so on for farther-neighbor (higher-order) effects.

To calculate the true fidelity $\mathscr{F}$, we consider the replication of any given template $X_1X_2...X_L$ of length L. The

primer grows from the starting unit α_1 and reaches the last site L to generate the final product of the sequence $\alpha_1\alpha_2...\alpha_L$. Except the first unit α_1 and the last unit α_L, any incorporated dNTP can be excised during this replication. In other words, this is a first-passage process from the reflecting boundary at the first site to the absorbing boundary at the last site. Since there are four types of dNTP competing for the incorporation reaction at each site, 4^L kinds of final products with different sequences should be generated. By solving the corresponding equations numerically, one can obtain the probability distribution of the final product $P^{X_1...X_L}_{\alpha_1...\alpha_L}(t\to\infty)$ from which the true fidelity $\mathscr{F}$ can be precisely counted. One can also derive approximate analytical expressions for the true fidelity under some conditions of the kinetic parameters. A brief introduction of these calculations is given in Appendix A (see also Supplemental Material (SM) Sec.I A.) . It is worth noting that the FP method does not need any extra assumptions like the steady-state or the quasi-equilibrium assumptions. Below we list the major results in terms of $\overline{k}_i$ and $\overline{r}_i$.

B. The fidelity calculated by the FP method

For exo^+-DNAP which has first-order neighbor effects, we have derived the approximate analytical expression of the overall fidelity at site i [20],

$$\mathscr{F}_i^{ov} \approx \left[\sum_{\substack{W_{i,m}\neq R_i\\ m=1,2,3}} \frac{\overline{k}^{X_{i-1}X_i}_{R_{i-1}W_{i,m}}}{\overline{k}^{X_{i-1}X_i}_{R_{i-1}R_i}} \frac{\overline{k}^{X_iX_{i+1}}_{W_{i,m}R_{i+1}}}{\overline{k}^{X_iX_{i+1}}_{W_{i,m}R_{i+1}}+\overline{r}^{X_{i-1}X_i}_{R_{i-1}W_{i,m}}}\right]^{-1} \quad (1)$$

R represents the matched nucleotide, and $W_{i,m}$ represents one of the three types of mismatched nucleotides. This formula provides a quite good approximation of $\mathscr{F}$: the relative deviation from the precise numerical result is less than 10% in a wide range of the kinetic parameters (in this paper, the relative deviation between any two quantities a and b is defined as $|a-b|/\min(a,b)$, $\min(a,b)$ means the smaller one of a and b. Details can be found in SM Sec.I A). This formula shows that the true fidelity at site i is overwhelmingly determined by the nearest neighbor sites. For simplicity, we omit all the superscripts below unless it causes misunderstanding, and also use W_i to represent any type of mismatch. Each term in the sum represents the error frequency of a particular type of mismatch, whose reciprocal (denoted as F_i) corresponds to the mismatch-specific fidelity discussed in the conventional steady-state assay or transient-state assay.Here the mismatch-specific fidelity is denoted as $\mathscr{F}_i \equiv \mathscr{F}_{ini,i}\cdot\mathscr{F}_{pro,i}$, in which $\mathscr{F}_{ini,i}$ is the initial discrimination and $\mathscr{F}_{pro,i}$ is the proofreading efficiency. Similarly, we denote $F_i \equiv F_i^{pol}\cdot F_i^{exo}$, $F_i^{pol} \equiv \overline{k}_{R_{i-1}R_i}/\overline{k}_{R_{i-1}W_i}$, $F_i^{exo} \equiv 1+\overline{r}_{R_{i-1}W_i}/\overline{k}_{W_iR_{i+1}}$. According to Eq.(1), we have

$$\mathscr{F}_i \approx F_i \quad (2)$$

Since $\mathscr{F}_{ini,i}$ is actually the fidelity for exo^--DNAP, we can set $\overline{r}=0$ in Eq.(1) to obtain

$$\mathscr{F}_{ini,i} \approx F_i^{pol} \equiv \frac{\overline{k}_{R_{i-1}R_i}}{\overline{k}_{R_{i-1}W_i}} \quad (3)$$

Hence, $\mathscr{F}_{pro,i} \equiv \mathscr{F}_i/\mathscr{F}_{ini,i}$ can be estimated by $F_i^{exo} \equiv F_i/F_i^{pol}$, as below

$$\mathscr{F}_{pro,i} \approx F_i^{exo} = 1+\frac{\overline{r}_{R_{i-1}W_i}}{\overline{k}_{W_iR_{i+1}}} \quad (4)$$

F_i^{exo} is similar to f_{pro} defined in Sec.I, if $\overline{k}_{W_iR_{i+1}}, \overline{r}_{R_{i-1}W_i}$ are regarded as $k_{el,W}, k_{ex,W}$ respectively. In these equations, $\approx$ means that the relative deviations of $F_i, F_i^{pol}, F_i^{exo}$ from $\mathscr{F}_i, \mathscr{F}_{ini,i}, \mathscr{F}_{pro,i}$ are less than 10% (see details in SM Sec.I A).

For exo^+-DNAP which has second-order neighbor effects, we have also derived the approximate analytical expressions for the overall fidelity at site i [20],

$$\mathscr{F}_i^{ov} \approx \left[\sum_{W_i\neq R_i} \frac{\overline{k}_{R_{i-2}R_{i-1}W_i}}{\overline{k}_{R_{i-2}R_{i-1}R_i}} \frac{k^{el}_{R_{i-1}W_iR_{i+1}}}{k^{el}_{R_{i-1}W_iR_{i+1}}+\overline{r}_{R_{i-2}R_{i-1}W_i}}\right]^{-1}$$

$$k^{el}_{R_{i-1}W_iR_{i+1}} = \frac{\overline{k}_{R_{i-1}W_iR_{i+1}}\overline{k}_{W_iR_{i+1}R_{i+2}}}{\overline{k}_{W_iR_{i+1}R_{i+2}}+\overline{r}_{R_{i-1}W_iR_{i+1}}} \quad (5)$$

Each term in the sum represents the mismatch-specific error frequency at site i. Its reciprocal defines the mismatch-specific fidelity which again consists of the initial discrimination and the proofreading efficiency, but the latter differs significantly from f_{pro} defined in Sec.I, since the effective elongation rate is not $\overline{k}_{R_{i-1}W_iR_{i+1}}$ but instead $k^{el}_{R_{i-1}W_iR_{i+1}}$ which includes the next-nearest neighbor effects. The same logic can be readily generalized to higher-order neighbor effects where the proofreading efficiency will be more complicated [20, 21].

In real DNA replication, either the dNTP incorporation or the dNMP excision is a multi-step process. By using the FP method, the complex reaction scheme can be reduced (or mapped) to the minimal scheme Fig.1, and the fidelity can still be calculated by Eq.(2)or Eq.(5), with only one modification: $\overline{k}$ and $\overline{r}$ are now the effective incorporation rates and the effective excision rates respectively which are functions of the involved rate constants. In the following sections, we will derive these effective rates by FP method for different multi-step reaction models, and compare them with the effective rates given by the steady-state assays or the transient-state assays. For simplicity, we only discuss the first-order neighbor effects of DNAP in details, since almost all the existing literature focused on nearest-neighbor effects. Higher-order neighbor effects will also be mentioned in later sections.

Below shows a simple example of the effective rates calculated by the FP method for the single-nucleotide incorporation reaction in the direct competition assays.

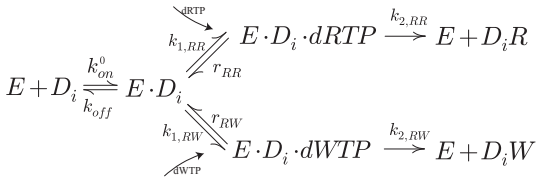

FIG. 2: The three-step reaction scheme of the competitive incorporation of dRTP and dWTP. E: exo^--DNAP. D_i: the primer-template duplex with the matched(R) terminal at site i. For brevity, the subscript i in each rate constant is omitted.

Fig.2 shows a three-step kinetic model of the competitive incorporation of a single dRTP or dWTP to site $i+1$, catalyzed by exo^--DNAP . The direct competition assay estimates the true fidelity by the ratio of the final product $[D_iR]$ to $[D_iW]$, which can be interpreted by the FP method as below

$$\begin{aligned} \mathscr{F} &\approx F = \overline{k}_{RR}\big/\overline{k}_{RW} \\ \overline{k}_{RR} &= \frac{k^0_{1,RR}k_{2,RR}}{k_{2,RR}+r_{RR}}[\text{dRTP}] \\ \overline{k}_{RW} &= \frac{k^0_{1,RW}k_{2,RW}}{k_{2,RW}+r_{RW}}[\text{dWTP}] \end{aligned} \tag{6}$$

F is exactly the initial discrimination defined by Eq.(3) with the two effective incorporation rates $\overline{k}_{RR}$ and $\overline{k}_{RW}$. These equations predict correctly the dNTP concentration-dependence of the fidelity which is consistent of the experimental observations in the direct competition assays [6]. Detailed discussion can be found in SM Sec.I B.

III. RESULTS AND DISCUSSION

As mentioned above, the total fidelity of exo^+-DNAP is assumed to consist of the initial discrimination f_{ini} and the proofreading efficiency f_{pro}. In this section we present a detailed analysis to show the availability and limitations of the conventional kinetic assays to characterize f_{ini} and f_{pro}. The reaction scheme under discussion is shown in Fig.3.

In this reaction scheme, the primer terminal can transfer between *Pol* and *Exo* in two different ways, *i.e.* the intramolecular transfer without DNAP dissociation (the transfer rates are denoted as k_{pe} and k_{ep}), and the intermolecular transfer in which DNAP can dissociate from and rebind to either *Pol* or *Exo* (the rates are denoted as k_{on} and k_{off}). These two modes have been revealed by single-turnover experiments[17] and directly observed by smFRET[22]. Here the quasi-first order rate k_{on} is proportional to the concentration of DNAP or DNA , *i.e.* $k_{on} = k^0_{on}[\text{E}]$ or $k_{on} = k^0_{on}[\text{DNA}]$ (see later sections). k^* is the effective incorporation rate, as explained below.

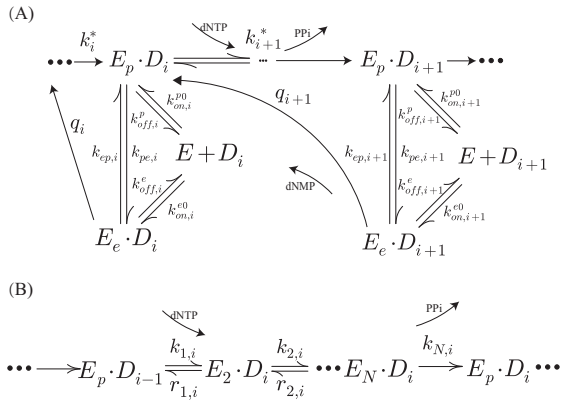

FIG. 3: (A): The multi-step reaction scheme of exo^+-DNAP including the multi-step incorporation process (indicated by k^* and details are shown in (B)) , the intramolecular/intermolecular transfer (*i.e.* without/with DNAP dissociation) of the primer terminal between *Pol* and *Exo* and the excision of the primer terminal nucleotide. (B): The multi-step incorporation scheme. The enzyme-substrate complex $(E{\cdot}D)$ goes through N states (indicated by subscripts $1, ..., N$) to successfully incorporate a single dNTP (indicated by subscript i). To simplify the notation, the superscripts indicating the template nucleotide X_i and the subscripts indicating the primer nucleotide α_i are omitted. This rule also applies to other figures in this paper, unless otherwise specified.

A. The FP method

Applying the FP method to complex reaction schemes like Fig.3(A), one can always reduce them to the simplified version Fig.1 with uniquely determined effective rates. For instance, for the multi-step incorporation schemes in Fig.3(B) , the effective incorporation rate k^* is given by (see details in Appendix B or SM Sec.I C)

$$\begin{aligned} k^* &= k'_{N-1}k_N\big/(r'_{N-1}+k_N) \\ k'_j &= k'_{j-1}k_j\big/(r'_{j-1}+k_j) \\ r'_j &= r'_{j-1}r_j\big/(r'_{j-1}+k_j) \\ &(j = N-1, ..., 3) \\ k'_2 &= k_1k_2\big/(r_1+k_2) \\ r'_2 &= r_1r_2\big/(r_1+k_2) \end{aligned} \tag{7}$$

Here $k_1 = k^0_1[\text{dNTP}]$, so $k^* = k^{*0}[\text{dNTP}]$.

For the complete scheme Fig.3(A), the effective rates are

$$\begin{aligned} \overline{k} = k^* \ &, \ \overline{r} = \frac{\widetilde{k}_{pe}q}{\widetilde{k}_{ep}+q} \\ \widetilde{k}_{pe} = k_{pe}+k_{p\to e} \ &, \ \widetilde{k}_{ep} = k_{ep}+k_{e\to p} \\ k_{p\to e} = \frac{k^p_{off}k^e_{on}}{k^p_{on}+k^e_{on}} \ &, \ k_{e\to p} = \frac{k^e_{off}k^p_{on}}{k^p_{on}+k^e_{on}} \end{aligned} \tag{8}$$

Details can be found in SM Sec.I D. These rates can be written more explicitly such as $\overline{k}_{RR} = k^*_{RR}$, if the

states of the base pairs at site $i, i-1, etc.$ are explicitly indicated. All the rates in the same equation have the same state-subscript. $k_{p\to e}$ and $k_{e\to p}$ define the effective intermolecular transfer rates between *Pol* and *Exo*. So $\widetilde{k}_{pe}$ and $\widetilde{k}_{ep}$ represent the total transfer rates *via* both the intramolecular and the intermolecular ways. Here $k^p_{on} = k^{p0}_{on}[\mathrm{E}]$, $k^e_{on} = k^{e0}_{on}[\mathrm{E}]$, so $k_{p\to e}$, $k_{e\to p}$ do not depend on [E]. With these effective rates, the initial discrimination and the proofreading efficiency can be estimated by $\mathscr{F}_{ini} \approx F_{ini} = k^*_{RR}/k^*_{RW}$ and $\mathscr{F}_{pro} \approx F_{pro} = 1 + \overline{r}_{RW}/k^*_{WR}$ respectively. Here we use the symbols F_{ini} and F_{pro} to replace F^{pol} and F^{exo} (see Sec.II) which are defined only for the minimal scheme. We keep these notations for any multi-step scheme in the following sections.

B. The steady-state assay

1. *The initial discrimination*

The steady-state assay is a standard method to analyze the catalytical capability of enzymes in biochemistry. The steady-state condition in experiment is usually established by two requirements, *i.e.* the substrate is in large excess to the enzyme, and the enzyme can dissociate quickly from the product once a single turnover is finished to resume its catalysis. The last dissociation step is reasonably assumed irreversible, since the enzyme will much unlikely rebind to the same substrate molecule after dissociation because the substrate is in large excess to the enzyme. Under such conditions, the amount of any intermediate product can be regarded approximately as constant (*i.e.* in steady state) in the initial stage of the reaction and thus the amount of the final product increases linearly with time. This initial growth velocity is defined as the steady-state turnover rate which is often used to characterize the catalytical capability of the enzyme.

It should be noted that, however, the steady-state assay is only valid for enzymes with single catalytical activity such as exo^--DNAP. So one should properly modify the exo^+-DNAP to obtain mutant exo^--DNAP (*e.g.* by deactivating the exonuclease activity or even eliminate the exonuclease domain by genetic mutations, without changing the polymerase activity of the DNAP), and then employ the steady-state assay to measure the initial discrimination of the DNAP. By measuring the initial velocity of the final product generation (*i.e.* single dNTP incorporation) under the steady-state condition, one can calculate the normalized velocity per enzyme which is in general given by the Michaelis-Menten equation

$$v^{pol}_{s\cdot s} = \frac{k_{cat}[\mathrm{dNTP}]}{[\mathrm{dNTP}] + K_m} \tag{9}$$

Here k_{cat} is the maximal steady-state turnover rate of dNTP incorporation and K_m is the Michaelis constant. The superscript *pol* indicates the polymerase activity, the subscript $s\cdot s$ indicates the steady state. Fitting the experimental data by this equation, one can get the specificity constant k_{cat}/K_m either for dRTP incorporation or dWTP incorporation, from which the initial discrimination was defined as $f_{s\cdot s,ini} = (k_{cat}/K_m)_R[\mathrm{dRTP}]/(k_{cat}/K_m)_W[\mathrm{dWTP}]$.

However, there is an apparent difference between the above-defined fidelity and the true fidelity: the former is measured under steady-state conditions, whereas the latter is defined without pre-assumptions like the steady-state approximation. So what is the relation between them?

To understand the exact meaning of k_{cat}/K_m, we have to consider the multi-step reaction scheme of dNTP incorporation (Fig.4) which explicitly includes the DNAP binding to DNA and dissociation from DNA. In general, we consider cases where the substrate DNA can bind either to the polymerase domain or to the deficient exonuclease domain (if the domain is only mutated but not eliminated). For the mutant exo^--DNAP, the transfer rate k'_{pe} and k'_{ep} are different from k_{pe} and k_{ep} of the wild type exo^+-DNAP. Under the steady-state condition, it can be easily shown

$$\left(\frac{k_{cat}[\mathrm{dNTP}]}{K_m}\right)_{\alpha_{i-1}\alpha_i} = \frac{k^*_{\alpha_{i-1}\alpha_i}}{(K_{s\cdot s})_{\alpha_{i-2}\alpha_{i-1}}} \tag{10}$$

Here k^* is defined in Eq.(7), $\alpha = R, W$. $K_{s\cdot s} = 1 + k'_{pe}/k'_{ep} + k^p_{off}/k^p_{on}$, $k^p_{on} = k^{p0}_{on}[\mathrm{DNA}]$.

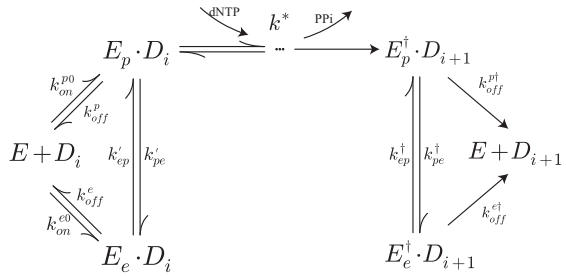

FIG. 4: The reaction scheme for the steady-state assay to measure the specificity constant of the nucleotide incorporation reaction of DNAP with deficient exonuclease domain. The omitted multi-step incorporation process is shown in Fig.3(B).

Theoretically, the real reaction scheme may be more complicated than Fig.4, *i.e.* there could be more pathways or more steps in DNAP binding/dissociation or DNA transfer. No matter how complicated the scheme is, it can be proven that the form of Eq.(10) is universal: k^* is exactly the effective incorporation rate defined by Eq.(7), and $K_{s\cdot s}$ is a simple function of the equilibrium constants of the steps just before dNTP binding(see details in SM Sec.II A,B).

Since $K_{s\cdot s}$ is independent on the incoming dNTP(α_i), the fidelity can be given generally by

$$f_{s\cdot s,ini} \equiv \frac{(k_{cat}[\mathrm{dNTP}]/K_m)_{R_{i-1}R_i}}{(k_{cat}[\mathrm{dNTP}]/K_m)_{R_{i-1}W_i}} = \frac{k^*_{R_{i-1}R_i}}{k^*_{R_{i-1}W_i}} \tag{11}$$

which is equal to F_{ini}. Since F_{ini} is a good approximation to $\mathscr{F}_{ini}$ with less than 10% deviation, the steady-state assay provides a quite good measure of the initial discrimination.

Below we give an example of using the measured specificity constants to estimate the site-specificity fidelity. In the steady-state assay of KF$^-$ [6], k_{cat} and K_m of the incorporation of 4 types of dNTP onto the template $\underline{\mathrm{T}}$ (the underline indicates the target site under investigation) has been measured separately, as listed in Table I. The initial discrimination $f_{s\cdot s,ini}$ of each mismatch can then be calculated by Eq.(11). The result is shown in the last column, indicating the significant mismatch-specific variations of the $f_{s\cdot s,ini}$. One can also calculate the site-specific initial discrimination, $\approx 7.6\times 10^3$, simply by using Eq.(1).

TABLE I: The steady-state assay of KF$^-$[6]

Template/Primer	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m $(\mu M^{-1}\cdot s^{-1})$	$f_{s\cdot s,ini}$
T$\underline{\mathrm{T}}^a$/AA	0.18	0.04	4.5	1
T$\underline{\mathrm{T}}$/AC	0.0093	390	2.4×10^{-5}	1.8×10^5
T$\underline{\mathrm{T}}$/AG	0.043	79	5.4×10^{-4}	8.4×10^3
T$\underline{\mathrm{T}}$/AT	0.0019	62	3.1×10^{-5}	1.5×10^5

^{a}: The template sequence is presented in the direction of $3' \to 5'$, and the primer is presented in $5' \to 3'$. The underlines indicate the target template site.

2. *The proofreading efficiency*

There were also some works using the steady-state assay to define the effective elongation rate $k_{el,W}$ and the effective excision rate $k_{ex,W}$ in order to characterize the proofreading efficiency $f_{s.s,pro} = 1+k_{ex,W}/k_{el,W}$ of exo^+-DNAP. If $k_{ex,W}$ equals to $\overline{r}_{RW}$ and $k_{el,W}$ equals to $\overline{k}_{WR}$, we then have $f_{pro} = F_{pro} \equiv 1+\overline{r}_{RW}/\overline{k}_{WR}$. Unfortunately, however, neither equation holds in general.

For instance, some works used the turnover velocity $v^{exo}_{s\cdot s}$ [17, 23], the specificity constant[24, 25] or the maximal turn-over rate k_{cat} [15] as $k_{el,W}$. As shown by Eq.(10), however, $\overline{k}_{WR}$ is not equal to any of the three quantities. So, the steady-state assay fails to measure $\overline{k}_{WR}$, unless $K_{s\cdot s} \approx 1$. For the reaction scheme Fig.4, this condition may be met, since the mutant exo^--DNAP has no exonuclease domain (k'_{pe}/k'_{ep} is absent) or it binds the DNA preferentially at the polymerase domain ($k'_{pe} \ll k'_{ep}$) and the DNA concentration in the experimental assay can be set large enough to ensure $k^p_{on} \gg k^p_{off}$. In such cases, the specificity constant, but not k_{cat} or $v^{exo}_{s\cdot s}$, can be regarded as $\overline{k}_{WR}$.

Of course $K_{s\cdot s} \approx 1$ may not hold if one considers reaction schemes more complicated than Fig.4. For instance, the primer terminal transfer may be a multi-step process rather than the one-step process shown in Fig.4. In this paper we do not discuss such schemes, since so far there are no experimental supporting evidences (except that there could be an additional DNAP translocation step before dNTP binding, as discussed in later sections). So we assume that $\overline{k}_{WR}$ can be reasonably measured by the specificity constant for cases discussed in this paper.

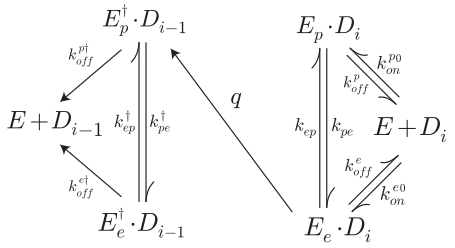

FIG. 5: The reaction scheme for the steady-state assay to measure the effective excision rate of exo^+-DNAP.

The steady-state assay was also employed to study the excision reaction of DNAP (Fig.5) . There was some studies which measured the initial velocity $(v^{exo}_{s\cdot s})_{RW}$ and interpreted it as $k_{ex,W}$ [15]. Whereas $(v^{exo}_{s\cdot s})_{RW}$ is determined by all the rate constants in Fig.5, some rate constants like $k^{p\dagger}_{off}$ and $k^\dagger_{pe}$ are absent from the effective excision rate $\overline{r}_{RW}$. So in principle, $(v^{exo}_{s\cdot s})_{RW}$ is not equal to $\overline{r}_{RW}$.

In summary, since $(k_{cat}[\mathrm{dNTP}]/K_m)_{WR} \approx \overline{k}_{WR}$ but $(v^{exo}_{s\cdot s})_{RW} \neq \overline{r}_{RW}$, $f_{s\cdot s,pro}$ may differ significantly from F_{pro}. A numerical example is shown in Fig.6 , where $F/f_{s\cdot s} = F_{pro}/f_{s\cdot s,pro} \approx \overline{r}_{RW}/(v^{exo}_{s\cdot s})_{RW}$ (here we have assumed $k_{ex,W} \gg k_{el,W}$ and $\overline{r}_{RW} \gg \overline{k}_{WR}$ for efficient proofreading).

Because the fidelity is a dimensionless number, the absolute values of the kinetic rates do not matter for our calculation. So we set $q = 1$ and other kinetic rates in unit of q with the magnitudes inspired by the measured values for T7 DNAP [17] (Table II), *e.g.* k_{ep} and k^e_{off} are set equal to q, and k_{pe} and k^p_{off} are smaller than q which are set to be 0.1. k^p_{on} and k^e_{on} are determined by DNA concentration which is often taken as $1\mu M$ in experiments(*e.g.* [15]). So these two values can be very large and we set them to be 10^3 in the calculations, which ensures $k^p_{on} > k^p_{off}$ and $k^e_{on} > k^e_{off}$. Larger values do not change the results. The dissociation rate from *Exo* are assumed irrelevant to the identity of the terminal, so $k^{e\dagger}_{off} = k^e_{off}$.

Fig.6 shows that the ratio $F/f_{s\cdot s}$ can even become larger than 10 when $k^\dagger_{pe}$, $k^{p\dagger}_{off}$ are much smaller than k_{pe}, k^p_{off}. This may unfortunately be true, since $k^\dagger_{pe}$ is the transfer rate of the newly formed matched terminal (after excision), which should be much smaller than the corresponding rate k_{pe} of the original mismatched terminal. $k^{p\dagger}_{off}$ is the dissociation rate of DNAP from the matched terminal, which should be much smaller than that from

the mismatched terminal k^p_{off}. So the steady-state assay are invalid to measure the proofreading efficiency of exo^+-DNAP. In the experiment to estimate $\mathscr{F}_{pro}$ for ap-polymerse[15], the authors wrongly interpreted k_{cat} and $v^{exo}_{s\cdot s}$ as $\overline{k}_{WR}$ and $\overline{r}_{RW}$ respectively , and gave that $f_{pro} \approx (v^{exo}_{s\cdot s})_{R_{i-1}W_i}/(k_{cat})_{W_iR_{i+1}}$. This measure is even totally irrelevant to $\mathscr{F}_{pro}$.

TABLE II: Kinetic rates of the excision of a mismatched T of A<u>G</u>/TT of T7 DNAP[17]

kinetic parameter	rate
q	$896s^{-1}$
k_{ep}	$714s^{-1}$
k_{pe}	$2.3s^{-1}$
k^p_{off}	$0.4s^{-1}$
k^e_{off}	$1022s^{-1}$
k^{p0}_{on}	$4 \times 10^7 M^{-1}s^{-1}$
k^{e0}_{on}	$(4-6) \times 10^8 M^{-1}s^{-1}$

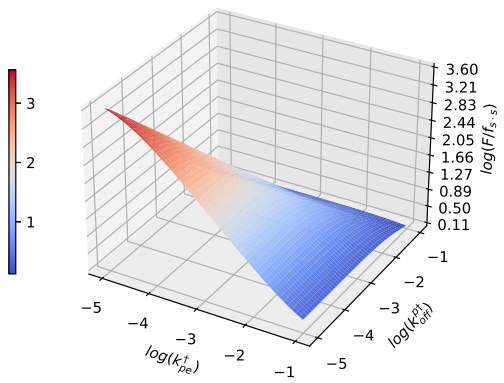

FIG. 6: The ratio $F/f_{s\cdot s}$ may become larger than 10 when $k^\dagger_{pe} \ll k_{pe}$ and $k^{p\dagger}_{off} \ll k^p_{off}$. The kinetic rates are set as $q = 1$, $k_{ep} = k^\dagger_{ep} = 1$, $k_{pe} = 0.1$, $k^e_{off} = k^{e\dagger}_{off} = 1$, $k^p_{off} = 0.1$, $k^p_{on} = 10^3$. k^e_{on} is calculated according to the thermodynamic constraint ($k^e_{on} = k^p_{on}k_{pe}k^e_{off}/k^p_{off}k_{ep} = 10^3$). The binding rates are large enough, *i.e.* $k^e_{on} > k^e_{off}$ and $k^p_{on} > k^p_{off}, k_{pe}$.

One can also change the concentration of the substrate DNA to obtain the specificity constant of the excision reaction in experiments[16], as can be shown theoretically

$$\left(\frac{k_{cat}[\mathrm{DNA}]}{K_m}\right)^{exo} = \frac{q(k_{pe}k^p_{on}/(k_{pe}+k^p_{off})+k^e_{on})}{q+k_{ep}k^p_{off}/(k_{pe}+k^p_{off})+k^e_{off}} \quad (12)$$

This expression includes explicitly the DNA concentration, whereas the effective rate $\overline{r}$ is independent of DNA concentration. So, it can not be used to estimate the effective excision rate. Details can be found in SM Sec.II C.

C. The transient-state assay

1. *The initial discrimination*

Similar to the steady-state assay, the transient-state assay can also be employed to study the polymerase and exonuclease of DNAP separately. The transient-state assay often refers to two different methods, the pre-steady-state assay or the single-turnover assay. Since the theoretical foundations of these two methods are the same, we only discuss the latter below for simplicity.

In single-turnover assays, the enzyme is in large excess to the substrate, and so the dissociation of the enzyme from the product is neglected. The time course of the product accumulation or the substrate consumption is monitored. The data is then fitted by exponential functions (single-exponential or multi-exponential) to give one or more exponents (*i.e.* the characteristic rates). In the initial discrimination assay, these rates are complex functions of all the involved rate constants and d-NTP concentration, which in principle can be analytically derived for any given kinetic model. For instance, for the commonly-used two-step model including only substrate binding and the subsequent irreversible chemical step, one can directly solve the kinetic equations to get two rates. On the other hand, the time course of the product accumulation can often be well fitted by single-exponential function at any given dNTP concentration, which implies that the two rates differ by more than one order of magnitude and only the smaller one is recorded in experiments . It was proved by K. Johnson that the smaller one obeys approximately the Michaelis-Menten-like equations [26](see also SM Sec.III A).

$$v^{pol}_{t\cdot s} \approx \frac{k_{pol}[\mathrm{dNTP}]}{[\mathrm{dNTP}]+K_d} \quad (13)$$

The subscript $t \cdot s$ indicates the transient state. Similar to the steady-state assays, k_{pol}/K_d is regarded as the specificity constant and thus the initial discrimination is defined as $f_{t\cdot s,ini} = (k_{pol}/K_d)_R[\mathrm{dRTP}]/(k_{pol}/K_d)_W[\mathrm{dWTP}]$.

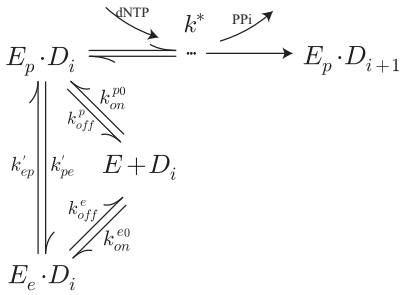

FIG. 7: The reaction scheme for the transient-state assay to measure the specificity constant of the nucleotide incorporation reaction of DNAP with deficient exonuclease domain. The omitted multi-step incorporation process is shown in Fig.3.(B).

To get a better understanding of k_{pol}/K_d, we consider the much more realistic multi-step scheme in Fig.7, in which the DNA can bind to either the polymerase domain or the deficient exonuclease domain and the dNTP incorporation is a multi-step process. One can prove that Eq.(13) still holds for such schemes and the specificity constant can be approximated as

$$\left(\frac{k_{pol}[\mathrm{dNTP}]}{K_d}\right)_{\alpha_{i-1}\alpha_i} \approx \frac{k^*_{\alpha_{i-1}\alpha_i}}{(K_{t\cdot s})_{\alpha_{i-2}\alpha_{i-1}}} \tag{14}$$

Eq.(14) provides a rough estimate of the real specificity constant with less than 200% relative deviation. The rigorous analysis is too lengthy to be presented here (for details see SM Sec.III B). Here $\alpha = R, W$. $K_{t\cdot s} = 1 + k_{pe}^{'}/k_{ep}^{'} + k_{off}^{p}/k_{on}^{p}$, $k_{on}^{p} = k_{on}^{p0}[\mathrm{E}]$. Hence, the relative deviation of $f_{t\cdot s,ini}$ from $F_{ini}(= k^*_{RR}/k^*_{RW})$ is less than 200%. Since $F_{ini} \approx \mathscr{F}_{ini}$ (with less than 10% deviation), the transient-state assay can give a rough estimate on $\mathscr{F}_{ini}$. In practice, both the steady-state assay and the transient-state assay are always used (say, Ref.[6]) to guarantee that the specificity constants obtained by both methods agree with one another (no order of magnitude difference) in order to make a good estimate on k^* and thus the true initial discrimination. Additionally, the specificity constant, but not k_{pol}, can be used to estimate $\overline{k}$ when $K_{t\cdot s} \approx 1$.

2. *The proofreading efficiency*

The transient-state assay of the exonuclease activity is often done under single-turnover conditions. The time course of product accumulation or substrate consumption is fitted by a single exponential or a double exponential to give one or two characteristic rates[17, 24, 25]. In the following, we show that the smallest one of the fitted exponents may probably be equal to $\overline{r}$ under some conditions.

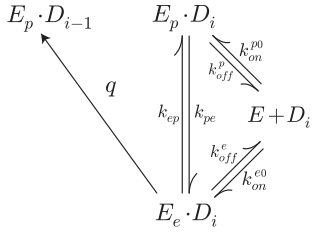

FIG. 8: The reaction scheme for the transient-state assay to measure the effective excision rate of exo^+-DNAP.

The simplest model for the transient-state assay of the excision reaction is depicted in Fig.8. By solving the corresponding kinetic equations, one can get three characteristic rates and the smallest one is given by

$$v_{t\cdot s}^{exo} \approx \frac{k_{pe}q(k_{on}^{p} + k_{on}^{e}) + qk_{off}^{p}k_{on}^{e}}{(k_{on}^{p} + k_{on}^{e})(q + k_{pe} + k_{ep}) + k_{off}^{e}k_{on}^{p} + k_{off}^{p}k_{on}^{e} + \epsilon} \tag{15}$$

Eq.(15) gives a rough estimate on the precise $v_{t\cdot s}^{exo}$ with less than than 200% deviation (details can be found in SM Sec.III C.). Here $\epsilon = k_{pe}q + k_{ep}k_{off}^{p} + k_{pe}k_{off}^{e} + qk_{off}^{p} + k_{off}^{p}k_{off}^{e}$, $k_{on} = k_{on}^{0}[\mathrm{E}]$. The DNAP concentration in the experiments can be set large enough, which ensures $k_{on}^{p} > k_{off}^{p}, k_{on}^{e} > k_{off}^{e}, k_{on}^{p} > k_{pe}$ and $\epsilon \approx 0$ (compared to other terms in the denominator), Eq.(15) can be simplified as

$$v_{t\cdot s}^{exo} \approx \frac{\widetilde{k}_{pe}q}{\widetilde{k}_{pe} + \widetilde{k}_{ep} + q} \tag{16}$$

If $\widetilde{k}_{ep} > \widetilde{k}_{pe}$ (when DNA binds preferentially to the polymerase domain) or $q > \widetilde{k}_{pe}$ (when the excision is a very fast process), one can get $v_{t\cdot s}^{exo} \approx \overline{r}$, with $\overline{r}$ defined by Eq.(8). So, if the real excision reaction follows the simplest scheme in Fig.8, $v_{t\cdot s}^{exo}$ may be interpreted as $\overline{r}$. However, if $q, \widetilde{k}_{ep} < \widetilde{k}_{pe}$, there could be large difference between $v_{t\cdot s}^{exo}$ and $\overline{r}$. Fig.9 gives an example showing the difference in a wide range of some key parameters.

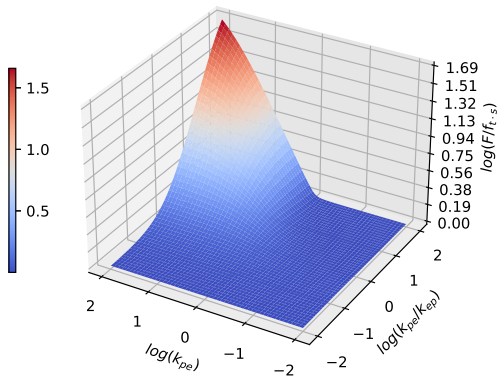

FIG. 9: The ratio $F/f_{t\cdot s}$ may become lager than 10 when $q, \widetilde{k}_{ep} < \widetilde{k}_{pe}$. The kinetic rates are set as $q = 1$, $k^e_{off} = 1$, $k^p_{off} = 0.1$, $k^p_{on} = 10^3$. k^e_{on} is calculated by the thermodynamic constraint ($k^e_{on} = k^p_{on} k_{pe} k^e_{off} / k^p_{off} k_{ep}$). The binding rates are always large enough when adjusting k_{pe} and k_{ep}, *i.e.* $k^e_{on} > k^e_{off}$ and $k^p_{on} > k^p_{off}, k_{pe}$.

In Fig.9 the proofreading efficiency is defined as $f_{t\cdot s} = 1 + k_{ex,W}/k_{el,W}$, $k_{ex,W}$ is taken as $(v^{exo}_{t\cdot s})_{RW}$ and $k_{el,W}$ is taken as $(k_{pol}[\text{dNTP}]/K_d)_{WR}$. Since $k^*_{WR} \approx (k_{pol}[\text{dNTP}]/K_d)_{WR}$, we have $F/f_{t\cdot s} = F_{pro}/f_{t\cdot s,pro} \approx \overline{r}_{RW}/(v^{exo}_{t\cdot s})_{RW}$. In this example, we show the precise value of $v^{exo}_{t\cdot s}$ by numerically solving the original kinetic equations, rather than using the approximate expression Eq.(15). The kinetic rates are set as explained in Sec.III B 2. The typical free energy difference between DNA binding to *Pol* and *Exo* may be only a few k_BT, so we set $10^{-2} < k_{pe}/k_{ep} < 10^2$. As shown, the ratio $F/f_{t\cdot s}$ may become even larger than 10 when $q, \widetilde{k}_{ep} \ll \widetilde{k}_{pe}$(the red area). So the transient-state assay can be used to roughly measure the proofreading efficiency only when $q > \widetilde{k}_{pe}$ or $\widetilde{k}_{ep} > \widetilde{k}_{pe}$ (the blue region in Fig.9) .

In Table II we lists all the kinetic rates of the excision reaction of T7 DNAP determined by the experiment[17], which shows $q, \widetilde{k}_{ep} > \widetilde{k}_{pe}$, so we can use $v^{exo}_{t\cdot s}$ to estimate $\overline{r}$ and calculate the F_{pro} by Eq.(4). In the original paper[17], however, the authors defined the proofreading efficiency as $f_{pro} = 1 + (k_{pe} + \theta k^p_{off})_W / k_{el,W}$, with the ambiguous quantity θ which was supposed between 0 and 1 (depending on the fate of the DNA after dissociation), and regarded $k_{el,W}$ as $v^{pol}_{s\cdot s}$ (Eq.(9)). This is essentially different from Eq.(4). Moreover, the authors ignored the template sequence-dependence of the proofreading efficiency in their calculations. In fact, Table II lists the rates of excising a mismatched T of A$\underline{\text{G}}$/TT, the corresponding elongation rate ($k_{el,W}$) should be the rate of incorporating the matched dNTP over the mismatched terminal T, *i.e.* $(k_{el,W})^{A\underline{G}X}_{TTR}$ (X can be one of A, T, G, C. R means the dNTP matched to X), according to Eq.(4). But the authors used $(k_{el,W})^{T\underline{A}G}_{AAC} = 0.012s^{-1}$ (calculated by Eq.(9) with the measured $k_{cat} = 0.025s^{-1}$, $K_m = 87\mu$M, [dNTP]=100μM in Ref.[23]) to estimate f_{pro} and finally gave $f_{pro} = 1 + (2.3 + 0.5 \times 0.4)/0.012 \approx 210$ ($\theta = 0.5$). So this estimate is highly questionable. Even if the template sequence-dependence of $k_{el,W}$ can be ignored, one should estimate F_{pro} by Eq.(4). The elongation rate $k_{el,W}$ (or k_{WR}) is approximately determined as $k_{cat}[\text{dNTP}]/K_m = 0.03s^{-1}$. The effective excision rate is determined by Eq.(8), $\overline{r}^{A\underline{G}}_{TT} \approx 1.23s^{-1}$. So we get a rough estimate of the proofreading efficiency for this template sequence, $\mathscr{F}_{pro,i} \approx F_{pro} \approx 1 + 1.23/0.03 \approx 40$ which is much smaller than that given by the authors.

There is another example possibly showing the second-order neighbor effects in proofreading of Human Mitochondrial DNAP. The excision velocities and the elongation rates were measured for a particular template sequence GT$\underline{\text{T}}$GG/CAT by the transient-state assays[24, 25] , as listed in Table III. Since there is no additional information on $q, \widetilde{k}_{pe}, \widetilde{k}_{ep}$, we can not use these values to estimate the proofreading efficiency. Here we just assume that the excision velocities can be taken as the effective excision rates to make a rough estimate. F_{pro} of a certain type of mismatch (Eq.(5)) can be rewritten as

$$\mathscr{F}_{pro,i} \approx F_{pro,i} = 1 + \frac{\overline{r}_{R_{i-2}R_{i-1}W_i}}{\overline{k}_{R_{i-1}W_iR_{i+1}}} \left(1 + \frac{\overline{r}_{R_{i-1}W_iR_{i+1}}}{\overline{k}_{W_iR_{i+1}R_{i+2}}}\right) \tag{17}$$

With the kinetic rates in Table III, we get

$$\begin{aligned} \mathscr{F}_{pro} \approx F_{pro} &\approx 1 + \frac{0.4}{0.1}\left(1 + \frac{3}{2.7}\right) \approx 9.4, (a) \\ \mathscr{F}_{pro} \approx F_{pro} &\approx 1 + \frac{0.4}{0.1}\left(1 + \frac{21.5}{2.7}\right) \approx 36.9, (b) \end{aligned} \tag{18}$$

These two values correspond to two different experimental conditions: (a) without dNTP in the solution; (b) in the presence of dNTP. In the latter case, the second-order proofreading (21.5/2.7) contributes to $\mathscr{F}_{pro}$ even more than the first-order proofreading (0.4/0.1).

TABLE III: Kinetic rates of human mitochondrial DNAP[24, 25]

the effective rate	the experiment measure	the measured rate
$\overline{r}^{GT\underline{T}}_{CA\underline{T}}$ [a]	$v^{exo}_{t\cdot s}$	$0.4s^{-1}$
$\overline{k}^{T\underline{T}G}_{A\underline{T}C}$	$k_{pol}[\mathrm{dNTP}]/K_d$ [b]	$0.1s^{-1}$
$\overline{r}^{T\underline{T}G}_{A\underline{T}C}$	$v^{exo}_{t\cdot s}$ [c]	$3.0s^{-1}$
$\overline{r}^{T\underline{T}G}_{A\underline{T}C}$	$\widehat{v}^{exo}_{t\cdot s}$ [d]	$21.5s^{-1}$
$\overline{k}^{\underline{T}GG}_{\underline{T}CC}$	$k_{pol}[\mathrm{dNTP}]/K_d$	$2.7s^{-1}$

[a]: The template sequence around the target site $\underline{\mathrm{T}}$ is presented in the direction of $3' \to 5'$: GT$\underline{\mathrm{T}}$GG
[b]: $[\mathrm{dNTP}] = 100\mu M$
[c]: without dNTP in the solution
[d]: in the presence of dNTP

It is worth emphasizing that the interpretation of $v^{exo}_{t\cdot s}$ is model-dependent. Theoretically, the reaction scheme could be more complicated than the simplest model in Fig.8, *e.g.* there may be multiple substeps in the intramolecular transfer process since the two domains are far apart (2-4 nm[18]). For any complex scheme, one can analytically calculate $v^{exo}_{t\cdot s}$ and $\overline{r}$. These two functions always differ greatly (examples can be found in SM Sec.III D). So the single-turnover assay *per se* is not a generally reliable method to measure the effective excision rate.

In total, neither the steady-state assay nor the transient-state assay can reliably measure the effective elongation rates and the effective excision rates, so it's invalid in principle to use these assays to estimate the proofreading efficiency of DNAP, unless one can do more detailed study to confirm that the required conditions on the key rate parameters are actually met.

D. More realistic models including DNAP translocation

So far we have not considered the important step, DNAP translocation, in the above kinetic models.DNAP should translocate forward along the template to the next site after dNTP incorporation, which empties the active pocket to accept the next dNTP for incorporation. Goodman *et al.* had discussed the possible effect of such a translocation on the transient-state gel assay very early[27], and recently DNAP translocation has been directly observed for phi29 DNAP by using nanopore techniques[28–31] or optical tweezers[32]. However, so far there is no any theory or experiment to seriously study the effect of translocation on the replication fidelity.

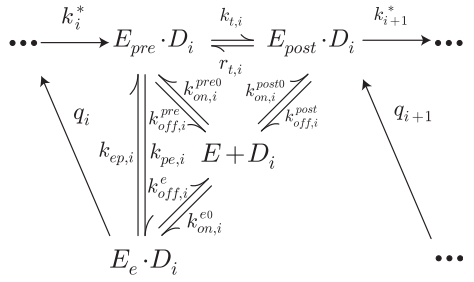

FIG. 10: The multi-step reaction scheme of exo^+-DNAP including the translocation step.

By using optical tweezers, Morin *et al.* had shown that DNAP translocation is not powered by PPi release or dNTP binding[32] and it's indeed a thermal ratchet process. So the simplest reaction scheme accounting for DNAP translocation can be depicted as Fig.10. k_t and r_t are the forward and the backward translocation rate respectively. $E_{pre}\cdot D_i$ and $E_{post}\cdot D_i$ indicate the pre-translocation and the post-translocation state of DNAP respectively. Here, the primer terminal can only switch intramolecularly between E_e and E_{pre} (but not E_{post}), according to the experimental observation [31]. We also assume DNAP can bind DNA either in state $E_{pre}\cdot D_i$ or in state $E_{post}\cdot D_i$ with possibly different binding rates and dissociation rates.

This complex scheme can be reduced to the minimal scheme Fig.1 by using the FP analysis. The obtained effective rates are given by (details see SM Sec.IV A)

$$\begin{aligned}
\overline{k} &= k^*\left(1 - q\widetilde{k}_{pe}/\xi\right) \\
\overline{r} &= q\eta/\xi \\
\eta &= k_{post\to e}(k_t + k_{pre\to post} + \widetilde{k}_{pe}) \\
&+ (r_t + k_{post\to pre})\widetilde{k}_{pe} \\
\xi &= (q + k_{e\to post})(k_t + \widetilde{k}_{pe} + k_{pre\to post}) \\
&+ \widetilde{k}_{ep}(k_t + k_{pre\to post}) \\
\widetilde{k}_{pe} &= k_{pe} + k_{pre\to e} \\
\widetilde{k}_{ep} &= k_{ep} + k_{e\to pre} \\
k_{a\to b} &= k^a_{off}k^b_{on}/(k^{pre}_{on} + k^{post}_{on} + k^e_{on}) \\
& a, b = pre, post, e
\end{aligned} \tag{19}$$

here k^* is defined by Eq.(7), $k^a_{on} = k^{a0}_{on}[\mathrm{E}]$. The total true fidelity $\mathscr{F}_i(\equiv \mathscr{F}_{ini,i}\cdot\mathscr{F}_{pro,i})$ can be approximated

by $F_i(\equiv F_{ini,i}\cdot F_{pro,i})$, as given below

$$\begin{aligned}
\mathscr{F}_i &\approx F_i \\
\mathscr{F}_{ini,i} &\approx F_{ini,i} = \frac{k^*_{R_{i-1}R_i}}{k^*_{R_{i-1}W_i}} \\
\mathscr{F}_{pro,i} &\approx F_{pro,i} \\
&= \frac{\left(1-q\widetilde{k}_{pe}/\xi\right)_{R_{i-1}R_i}}{\left(1-q\widetilde{k}_{pe}/\xi\right)_{R_{i-1}W_i}}\left(1+\frac{\overline{r}_{R_{i-1}W_i}}{\overline{k}_{W_iR_{i+1}}}\right) \\
&\approx \frac{q_{R_{i-1}W_i}}{k^*_{W_iR_{i+1}}} \\
&\times \frac{\left(1-q\widetilde{k}_{pe}/\xi\right)_{R_{i-1}R_i}}{\left(1-q\widetilde{k}_{pe}/\xi\right)_{W_iR_{i+1}}}\left(\frac{\eta}{\xi-q\widetilde{k}_{pe}}\right)_{R_{i-1}W_i} \qquad (20)
\end{aligned}$$

Here $\mathscr{F}_{ini,i}$ is the initial discrimination and $\mathscr{F}_{pro,i}$ is the proofreading efficiency, which are esitmated by $F_{ini,i}$ and $F_{pro,i}$ respectively. $F_{ini,i}$ is obtained by applying Eq.(3) with the effective rates $\overline{k}$ in Eq.(19) while setting $q=0$. The relative deviations of $F_i, F_{ini,i}, F_{pro,i}$ from $\mathscr{F}_i, \mathscr{F}_{ini,i}, \mathscr{F}_{pro,i}$ are less than 10% (for details see SM Sec.IV A).

On the other hand, we can calculate the fidelity defined by the kinetic assays. Following the same logic in Sec.III C 1, we obtain the specificity constant by the transient-state assay of the mutant exo^--DNAP (for details see SM Sec.IV B.)

$$\left(\frac{k_{pol}[\text{dNTP}]}{K_d}\right)_{\alpha_{i-1}\alpha_i} \approx \frac{k^*_{\alpha_{i-1}\alpha_i}}{(K_{t\cdot s})_{\alpha_{i-2}\alpha_{i-1}}} \qquad (21)$$

Here $\alpha = R, W$. $K_{t\cdot s} = 1 + r_t/k_t(1 + k'_{pe}/k'_{ep}) + k^{post}_{off}/k^{post}_{on}$, $k^{post}_{on} = k^{post,0}_{on}[\text{E}]$. Eq.(21) gives a rough estimate of the specificity constant with less than 200% relative deviation. Similarly, the steady-state assay also define another specificity constant, $(k_{cat}[\text{dNTP}]/K_m)_{\alpha_{i-1}\alpha_i} = k^*_{\alpha_{i-1}\alpha_i}/(K_{s\cdot s})_{\alpha_{i-2}\alpha_{i-1}}$, with $K_{s\cdot s} = 1 + r_t/k_t(1 + k'_{pe}/k'_{ep}) + k^{post}_{off}/k^{post}_{on}$, $k^{post}_{on} = k^{post,0}_{on}[\text{DNA}]$. Since $K_{t\cdot s}$ (or $K_{s\cdot s}$) is independent on the incoming dNTP α_i, the initial discrimination can be measured by the steady-state assay with high precision or roughly estimated by the transient-state assay, *i.e.* $F_{ini} = f_{s\cdot s,ini}$ and $F_{ini} \approx f_{t\cdot s,ini}$ (with no order of magnitude difference).

The transient-state excision velocity $v^{exo}_{t\cdot s}$ can be calculated as follows

$$\begin{aligned}
v^{exo}_{t\cdot s} &\approx q\eta/\kappa \\
\kappa &= (1+r_t/k_t)(\xi - q\widetilde{k}_{pe}) + \epsilon_1 + \epsilon_2 \qquad (22)
\end{aligned}$$

ϵ_1 and ϵ_2 are very complex functions of the kinetic rates which are too lengthy to give here. Eq.(22) offers a rough estimate of the real $v^{exo}_{t\cdot s}$ with less than 300% deviation (see details in SM Sec.IV B).

With Eq.(21)(22) , the proofreading efficiency f_{pro} is defined by the transient-state assay as

$$\begin{aligned}
f_{t\cdot s,pro,i} &\equiv 1 + \frac{(v^{exo}_{t\cdot s})_{R_{i-1}W_i}}{(k_{pol}[\text{dNTP}]/K_d)_{W_iR_{i+1}}} \\
&\approx 1 + \frac{q_{R_{i-1}W_i}}{k^*_{W_iR_{i+1}}}\left(\frac{\eta K_{t\cdot s}}{\kappa}\right)_{R_{i-1}W_i} \qquad (23)
\end{aligned}$$

here $\approx$ means less than 300% deviation. We compare this rough estimate on $f_{t\cdot s,pro}$ to F_{pro} defined by Eq.(20). They might be approximately equal only under some conditions. For instance, if the kinetic rates satisfy the following conditions

(1)$k_{t,RR} \gg k^{pre}_{off,RR}, k_{pe,RR}$

(2)$k_{t,WR} \gg k^{pre}_{off,WR}, k_{pe,WR}$

(3)$k_{t,RW} \gg k^{pre}_{off,RW}, k_{pe,RW}$ and $r_{t,RW} \gg k^{post}_{off,RW}$.

(4)$\widetilde{k}_{ep,RW} \gg \widetilde{k}_{pe,RW}$ and $q_{RW} \gg \widetilde{k}_{pe,RW}$.

Here $\gg$ means more than one order of magnitude higher. Under such conditions, one can estimate that $0.69 < F_{pro}/f_{t\cdot s,pro} < 1.20$, the relative deviation of $f_{t\cdot s,pro}$ from F_{pro} is less than 31% (details see SM Sec.IV B). Here we have assumed $K_{t\cdot s} \approx 1 + r_t/k_t$, since the DNAP concentration [E] can be set large enough in the transient-state assays to ensure $k^{post}_{on} \gg k^{post}_{off}$ and the transfer rates k'_{pe} and k'_{ep} of the mutant DNAP often satisfies $k'_{ep} \gg k'_{pe}$.

In general, $f_{t\cdot s,pro}$ can be much different from F_{pro} in a wide range of kinetic parameters, *e.g.* the translocation in the presence of a terminal mismatch may be very slow [33] so that the above condition (3) is violated. Fig.11 shows an example. Since some kinetic rates for terminal mismatch RW (for simplicity we omit the subscript in this example), *e.g.* k^{pre}_{off} and k^{post}_{off}, are unknown from the experiments, we just assume that $k^{pre}_{off} = k^{p}_{off} = 0.1$ (see k^{p}_{off} in Table II) and $k^{post}_{off} = 0.001$. Other rates is set as explained in Sec.III B 2. The ratio r_t/k_t ranges from 10^{-2} to 10^2, corresponding to a free energy difference (between the *Pre* and *Post* states) of a few k_BT. $q = 1, \widetilde{k}_{ep} > 1$ and $\widetilde{k}_{pe} < 0.2$, so the condition (4) still holds. As shown in the figure, the ratio $F/f_{t\cdot s} \approx F_{pro}/f_{t\cdot s,pro} \approx \left(q\eta/\left(\xi - q\widetilde{k}_{pe}\right)\right)_{RW} \Big/ (v^{exo}_{t\cdot s}K_{t\cdot s})_{RW}$ can become larger than 10 when k_t and r_t get small enough. Here $v^{exo}_{t\cdot s}$ is precisely computed by numerically solving the original kinetic equations, and $K_{t\cdot s} \approx 1 + r_t/k_t$ as discussed above.

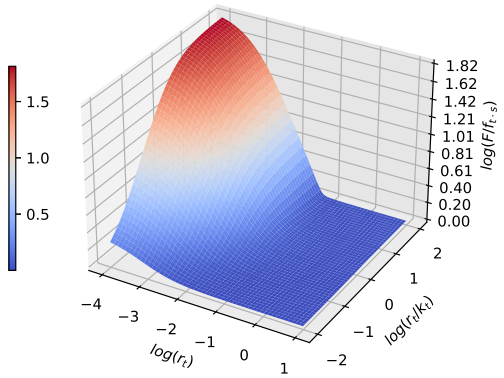

FIG. 11: The ratio $F/f_{t\cdot s}$ becomes larger than 10 when k_t and r_t get small enough. The kinetic rates (for terminal mismatch RW): $q = 1$, $k_{pe} = 0.1$, $k_{ep} = 1$, $k^e_{off} = 1$, $k^{pre}_{off} = 0.1$, $k^{post}_{off} = 0.001$, $k^{pre}_{on} = 10^3$. k^e_{on} and k^{post}_{on} are calculated by thermodynamic constraints ($k^e_{on} = k^p_{on} k_{pe} k^e_{off}/k^{pre}_{off} k_{ep} = 10^3$, $k^{post}_{on} = k^{pre}_{on} k_t k^{post}_{off}/k^{pre}_{off} r_t$). The binding rates are always large enough when adjusting k_t and r_t to ensure $k^a_{on} > k^a_{off}, a = pre, post, e$ and $k^{pre}_{on} > k_{pe}$.

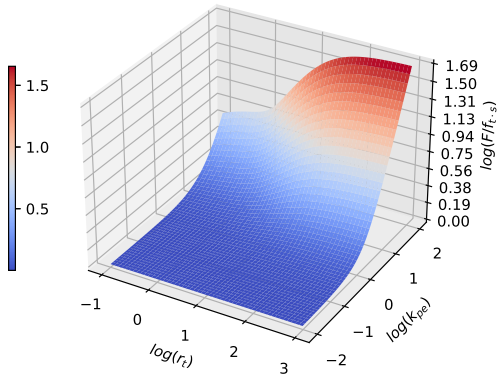

FIG. 12: The ratio $F/f_{t\cdot s}$ may become larger than 10 when k_{pe} gets large enough. The kinetic rates (for terminal mismatch RW): $q = 1$, $k_{ep} = 1$, $k^e_{off} = 1$, $k^{pre}_{off} = 0.1$, $k^{post}_{off} = 0.1$, $k_t = 10$, $k^{pre}_{on} = 10^3$. k^e_{on} and k^{post}_{on} are calculated by thermodynamic constraints($k^e_{on} = k^p_{on} k_{pe} k^e_{off}/k^{pre}_{off} k_{ep}$, $k^{post}_{on} = k^{pre}_{on} k_t k^{post}_{off}/k^{pre}_{off} r_t$). The binding rates are always large enough when adjusting k_{pe} and r_t to ensure $k^a_{on} > k^a_{off}, a = pre, post, e$ and $k^{pre}_{on} > k_{pe}$.

Fig.12 shows another example where the condition (1)(2)(3) hold but the condition (4) is violated. We assume for the terminal mismatch RW, $k^{pre}_{off} = k^{post}_{off} = k^p_{off} = 0.1$ (see k^p_{off} in Table II). $k_t = 10 > k^{pre}_{off}$ and $r_t > k^{post}_{off}$ which ensures the condition (1) . Other kinetic rates are set as explained in Sec.III B 2. As shown, the ratio $F/f_{t\cdot s}$ may become larger than 10 when k_{pe} gets large enough.

However, the condition (2) may also not hold for real DNAPs. Since the buried mismatch may slow down DNAP translocation, $\left(1 - q\widetilde{k}_{pe}/\xi\right)_{W_i R_{i+1}} \ll 1$ may hold if $\widetilde{k}_{pe,W_i R_{i+1}} \gg k_{t,W_i R_{i+1}}, k_{pre\to post,W_i R_{i+1}}$ and $q_{W_i R_{i+1}} \gg \widetilde{k}_{ep,W_i R_{i+1}}$. On the other hand, kinetic rates for terminal WR are present in F_{pro} (in the factor $1/\left(1 - q\widetilde{k}_{pe}/\xi\right)_{W_i R_{i+1}}$, see Eq.(20)) but totally absent from $f_{t\cdot s,pro}$ (see Eq.(23)). So, F_{pro} can become much larger if $\left(1 - q\widetilde{k}_{pe}/\xi\right)_{W_i R_{i+1}}$ gets much smaller than 1, while other kinetic rates for terminal RW are fixed. That is to say, $F/f_{t\cdot s}$ may become even more than one order of magnitude larger than that shown by the red region in Fig.11 and Fig.12.

In total, if the terminal mismatch or the buried mismatch severely slows down the DNAP translocation, the transient-state assay is likely to underestimate the fidelity even by one or more orders of magnitude. Unfortunately, so far as we know, there have been no experimental studies on the translocation kinetics of any DNAP in the presence of the terminal mismatch or the buried mismatch. So, in this paper we can not further evaluate the reliability of the reported transient-state assays on specific DNAPs in the literature. Future experimental studies should offer more information on this issue.

IV. SUMMARY

The conventional kinetic assays of DNAP fidelity, *i.e.* the steady-state assay or the transient-state assay, have indicated that the initial discrimination f_{ini} is about $10^{4\sim5}$ and the proofreading efficiency f_{pro} is about $10^{2\sim3}$ [18]. Although these assays have been widely used for decades and these estimates of f_{ini} and f_{pro} have been widely cited in the literature, they are not unquestionable since the logic underlying these methods are not well founded. No rigorous theories about the true fidelity $\mathscr{F}$ have ever been proposed, and its relation to the operationally defined $f_{s\cdot s}$ or $f_{t\cdot s}$ has never been clarified.

In this paper, we examined carefully the relations between $f_{s\cdot s}$, $f_{t\cdot s}$ and $\mathscr{F}$, based on the FP method recently proposed by us to investigate the true fidelity of exo^--DNAP or exo^+-DNAP. We conclude that for exo^--DNAP, the steady-state assay can measure $\mathscr{F}$ with

very high precision, while the transient-state assay offers a rough estimate of $\mathscr{F}$ (with no order of magnitude difference), just by measuring the specificity constant (k_{cat}/K_m or k_{pol}/K_d).

For exo^+-DNAP, however, the situation is more complicated. The steady-state assay or the transient-state assay can still be used to measure the initial discrimination $\mathscr{F}_{ini}$, as done for exo^--DNAP (so the above cited estimates of f_{ini} are reliable). But either method fails to directly measure the effective excision rate and the effective elongation rate, and thus in principle they can not characterize the proofreading efficiency $\mathscr{F}_{pro}$. So the widely cited estimates $f_{pro} \sim 10^{2\sim3}$ are very suspicious. Our analysis shows that these kinetic assays may be valid to measure $\mathscr{F}_{pro}$ only if the involved rate constants satisfy some conditions. If there are no further evidences to support these conditions, the assays *per se* may largely underestimate $\mathscr{F}_{pro}$. Even when all these conditions are met, there were still quite different strategies in using these assays which often give inconsistent results. In this paper, we have shown definitely that there is only one proper way to estimate $\mathscr{F}_{pro}$ by using the transient-state assay.

How can one estimate the true fidelity unambiguously? A possible way is to dissect the reaction mechanism, *i.e.* measuring the rate constants of each step by transient-state experiments [17, 23, 34–38], and then calculate the effective rate according to Eq.(7) and Eq.(8). This is a perfect approach but needs heavy work. Theoretically, there is another approach, a single-molecule assay based on the FP analysis, to directly measure the effective rates. Simply put, in the framework of the first-passage theory, each effective rate required to calculate the true fidelity can be interpreted as the reciprocal of the residence time at the corresponding state in the stochastic dNTP incorporation process or the terminal excision process at the single-DNA level. This might be done, at least in principle, by analyzing the stochastic trajectories obtained in the single-molecule experiments. We do not discuss it here and leave the details in SM Sec.V. We hope that this may inspire future single-molecule assays on DNAP fidelity.

Last but not least, we have focused on the first-order (nearest) neighbor effect in this paper and just mentioned the higher-order neighbor effects which may also be important to DNAP fidelity. As shown in Sec.III C 2, the second-order (next-nearest) neighbor effect may even be more significant than the first-order effect in the proofreading of human mitochondrial DNAP. Although this conclusion is based on the rough estimates given there, we believe that it's somewhat universal. The proofreading efficiency $\mathscr{F}_i$ of exo^+-DNAP consists of two factors, $\overline{r}_{R_{i-2}R_{i-1}W_i}/\overline{k}_{R_{i-1}W_iR_{i+1}}$ and $\overline{r}_{R_{i-1}W_iR_{i+1}}/\overline{k}_{W_iR_{i+1}R_{i+2}}$, representing the first-order and the second-order proofreading efficiency respectively. These two factors are both dependent on the stability of the primer-template duplex. For naked dsDNA duplex, numerous experiments have shown that a penultimate mismatch leads to much lower stability than a terminal mismatch [39]. This implies that a penultimate mismatch may more significantly disturb the base stacking of the primer-template conjunction in the polymerase domain and thus the forward translocation of DNAP will be slower and the *Pol*-to-*Exo* transfer of the primer terminal will be faster, if compared with the terminal mismatch. More explicitly, it's very likely that $\overline{r}_{R_{i-1}W_iR_{i+1}} > \overline{r}_{R_{i-2}R_{i-1}W_i}$ and $\overline{k}_{W_iR_{i+1}R_{i+2}} < \overline{k}_{R_{i-1}W_iR_{i+1}}$. In such cases, the second-order factor may be larger than the first-order factor.

We hope that the analysis and the suggestions presented in this paper will urge serious experimental reexaminations on the conventional kinetic assays of DNAP fidelity and offer some inspirations to single-molecule experimentalists to conduct more accurate fidelity analysis.

V. ACKNOWLEDGMENTS

The authors thank the financial support by National Natural Science Foundation of China (No.11675180,11774358), the CAS Strategic Priority Research Program (No.XDA17010504), Key Research Program of Frontier Sciences of CAS (No.Y7Y1472Y61), Research Fund of Wenzhou Institute CAS (No.WIUCASYJ2020004,WIUCASQD2020009).

APPENDIX A: BASICS OF THE FP METHOD

For the minimal reaction scheme (Fig.1), the master equations for replicating a given template $X_1X_2...X_L$ (in the direction $1 \to L$) are shown below (here we only discuss the first-order neighbor effects as an example),

$$\begin{aligned}
\frac{d}{dt}P^{X_1...X_L}_{\alpha_1} &= \sum_{\alpha_2}\overline{r}^{X_1X_2}_{\alpha_1\alpha_2}P^{X_1X_2...X_L}_{\alpha_1\alpha_2} - \sum_{\alpha_2}\overline{k}^{X_1X_2}_{\alpha_1\alpha_2}P^{X_1...X_L}_{\alpha_1} \\
\frac{d}{dt}P^{X_1...X_i...X_L}_{\alpha_1...\alpha_i} &= \overline{k}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}P^{X_1...X_{i-1}...X_L}_{\alpha_1...\alpha_{i-1}} + \sum_{\alpha_{i+1}}\overline{r}^{X_iX_{i+1}}_{\alpha_i\alpha_{i+1}}P^{X_1...X_iX_{i+1}...X_L}_{\alpha_1...\alpha_i\alpha_{i+1}} \\
&- \left(\overline{r}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i} + \sum_{\alpha_{i+1}}\overline{k}^{X_iX_{i+1}}_{\alpha_i\alpha_{i+1}}\right)P^{X_1...X_i...X_L}_{\alpha_1...\alpha_i}, \quad 2 \le i \le L-2 \\
\frac{d}{dt}P^{X_1...X_{L-1}X_L}_{\alpha_1...\alpha_{L-1}} &= \overline{k}^{X_{L-2}X_{L-1}}_{\alpha_{L-2}\alpha_{L-1}}P^{X_1...X_{L-2}...X_L}_{\alpha_1...\alpha_{L-2}} - \left(\overline{r}^{X_{L-2}X_{L-1}}_{\alpha_{L-2}\alpha_{L-1}} + \sum_{\alpha_L}\overline{k}^{X_{L-1}X_L}_{\alpha_{L-1}\alpha_L}\right)P^{X_1...X_{L-1}X_L}_{\alpha_1...\alpha_{L-1}} \\
\frac{d}{dt}P^{X_1...X_L}_{\alpha_1...\alpha_L} &= \overline{k}^{X_{L-1}X_L}_{\alpha_{L-1}\alpha_L}P^{X_1...X_{L-1}X_L}_{\alpha_1...\alpha_{L-1}}
\end{aligned} \tag{A1}$$

Here $P^{X_1...X_i...X_L}_{\alpha_1...\alpha_i}$ is the probability of the primer with the sequence $\alpha_1...\alpha_i$ at time t. $\overline{k}$ and $\overline{r}$ are the nucleotide incorporation rate and the excision rate respectively. In these equations, we have assumed that the first unit α_1 and the last unit α_L of the primer chain cannot be excised (the boundary condition), and α_1 may be R or W with the given probability p_{α_1} (the initial condition).

We are concerned only about the sequence distribution of the final products $P^{X_1...X_L}_{\alpha_1...\alpha_L}(t \to \infty)$, which can be given by integrating Eq.(A1)

$$\begin{aligned}
-p_{\alpha_1} &= \sum_{\alpha_2}\overline{r}^{X_1X_2}_{\alpha_1\alpha_2}\Gamma^{X_1X_2...X_L}_{\alpha_1\alpha_2} - \sum_{\alpha_2}\overline{k}^{X_1X_2}_{\alpha_1\alpha_2}\Gamma^{X_1...X_L}_{\alpha_1} \\
0 &= \overline{k}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}\Gamma^{X_1...X_{i-1}...X_L}_{\alpha_1...\alpha_{i-1}} + \sum_{\alpha_{i+1}}\overline{r}^{X_iX_{i+1}}_{\alpha_i\alpha_{i+1}}\Gamma^{X_1...X_iX_{i+1}...X_L}_{\alpha_1...\alpha_i\alpha_{i+1}} \\
&- \left(\overline{r}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i} + \sum_{\alpha_{i+1}}\overline{k}^{X_iX_{i+1}}_{\alpha_i\alpha_{i+1}}\right)\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i}, \quad 2 \le i \le L-2 \\
0 &= \overline{k}^{X_{L-2}X_{L-1}}_{\alpha_{L-2}\alpha_{L-1}}\Gamma^{X_1...X_{L-2}...X_L}_{\alpha_1...\alpha_{L-2}} - \left(\overline{r}^{X_{L-2}X_{L-1}}_{\alpha_{L-2}\alpha_{L-1}} + \sum_{\alpha_L}\overline{k}^{X_{L-1}X_L}_{\alpha_{L-1}\alpha_L}\right)\Gamma^{X_1...X_{L-1}X_L}_{\alpha_1...\alpha_{L-1}} \\
P^{X_1...X_L}_{\alpha_1...\alpha_L}(t \to \infty) &= \overline{k}^{X_{L-1}X_L}_{\alpha_{L-1}\alpha_L}\Gamma^{X_1...X_{L-1}X_L}_{\alpha_1...\alpha_{L-1}}
\end{aligned} \tag{A2}$$

Here $\int_0^{+\infty}\frac{d}{dt}P^{X_1...X_i...X_L}_{\alpha_1...\alpha_i}dt = 0$ for any $1 < i < L$. $\Gamma_x = \int_0^{\infty}P_x dt$ is precisely the average residence time at the state x (see Appendix A of Ref.[20]). One can solve these equations numerically to obtain $P^{X_1...X_L}_{\alpha_1...\alpha_L}(t \to \infty)$ and directly give the site-specific fidelity $\mathscr{F}_i$.

$\mathscr{F}_i$ can be numerically computed for any given kinetic parameters. It can also be calculated analytically, though approximately, under some restrictive conditions on the parameters which seem quite reasonable for real DNA replications. That is, (a) $\overline{k}^{X_iX_{i+1}}_{R_iR_{i+1}} \gg \overline{k}^{X_iX_{i+1}}_{R_iW_{i+1}}$. (b) $\overline{k}^{X_iX_{i+1}}_{W_iW_{i+1}} \approx 0s^{-1}$. (c) $\overline{k}^{X_iX_{i+1}}_{R_iR_{i+1}} \gg \overline{r}^{X_{i-1}X_i}_{R_{i-1}R_i}, \overline{r}^{X_{i-1}X_i}_{W_{i-1}R_i}$. Under such conditions, one can get the approximate expressions of the true fidelity, as given by Eq.(1)-(5). Details can be found in SM Sec.I A.

APPENDIX B: THE UNIQUE REDUCTION OF THE MULTI-STEP SCHEME TO THE MINIMAL SCHEME

The FP method can also be applied to more complex reaction schemes such as the multi-step scheme in Fig.3. These schemes can be reduced uniquely to the minimal scheme Fig.1 and hence the true fidelity F can be calculated by Eq.(1)-(5). Here we take the scheme Fig.3(B) as an example to illustrate the reduction procedure.

Similar to Appendix A, one can write the master equations for all the states in the reaction scheme of replicating a template $X_1...X_L$ and integrate these equations to give the sequence distribution of the final products $P^{X_1...X_L}_{\alpha_1...\alpha_L}(t \to$

∞). Here we only show the integrated equations for the variables $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,2}, ..., \Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N}, \Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p}$ ($\Gamma^{X_1...X_i...X_L}_{\alpha_1,...\alpha_i,x}$ corresponds to $E_x \cdot D_i$ in Fig.3(B), $x = 2...N, p$)

$$\begin{aligned}
0 &= k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,1}\Gamma^{X_1...X_{i-1}...X_L}_{\alpha_1...\alpha_{i-1},p} + r^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,2}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,3} - \left(k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,2} + r^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,1}\right)\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,2} \\
0 &= k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,j-1}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,j-1} + r^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,j}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,j+1} - \left(k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,j} + r^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,j-1}\right)\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,j}, 2 < j < N \\
0 &= k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N-1} - \left(k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N} + r^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}\right)\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N}
\end{aligned} \tag{B1}$$

and

$$0 = k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N} + \sum_{\alpha_{i+1}} r^{X_iX_{i+1}}_{\alpha_i\alpha_{i+1},1}\Gamma^{X_1...X_{i+1}...X_L}_{\alpha_1...\alpha_{i+1},2} - \sum_{\alpha_{i+1}} k^{X_iX_{i+1}}_{\alpha_i\alpha_{i+1},1}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p} \tag{B2}$$

Here $k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,x}, r^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,x}$ correspond to $k_{x,i}, r_{x,i}$ in Fig.3(B). According to Eq.(B1), the variables $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,2}...\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N}$ can be solved as functions of $\Gamma^{X_1...X_{i-1}...X_L}_{\alpha_1...\alpha_{i-1},p}$. There is a similar system of equations for $\Gamma^{X_1...X_{i+1}...X_L}_{\alpha_1...\alpha_{i+1},2}...\Gamma^{X_1...X_{i+1}...X_L}_{\alpha_1...\alpha_{i+1},N}$, and $\Gamma^{X_1...X_{i+1}...X_L}_{\alpha_1...\alpha_{i+1},2}$ can also be solved as a function of $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p}$. So Eq.(B2) can be rewritten as

$$0 = k^{*X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}\Gamma^{X_1...X_{i-1}...X_L}_{\alpha_1...\alpha_{i-1},p} - \sum_{\alpha_{i+1}} k^{*X_iX_{i+1}}_{\alpha_i\alpha_{i+1}}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p} \tag{B3}$$

with k^* given by Eq.(7). Comparing Eq.(B3) with Eq.(A2), one can identify k^* as the effective incorporation rate ($k^{*X_{i-1}X_i}_{\alpha_{i-1}\alpha_i} = \overline{k}^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i}$, if $\overline{r} = 0$). No approximations such as the steady-state assumption or the quasi-equilibrium assumption are needed in the reduction procedure. It should be noted that the above reduction is unique in the sense that other reduction procedure eliminating $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p}(i = 1, ..., L)$ will lead to equations much different from Eq.(A2). For example, the reduced integrated equation with retained $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N}$ is

$$\begin{aligned}
0 = &\frac{k'^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}k^{X_{i-2}X_{i-1}}_{\alpha_{i-2}\alpha_{i-1},N}}{\sum_{\alpha_i} k'^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}}\Gamma^{X_1...X_{i-1}...X_L}_{\alpha_1...\alpha_{i-1},N} - \left(r'^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1} + k^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N}\right)\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N} \\
&+ \frac{k'^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}\left[\sum_{\alpha_i}\left(r'^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N}\right)\right]}{\sum_{\alpha_i} k'^{X_{i-1}X_i}_{\alpha_{i-1}\alpha_i,N-1}}
\end{aligned} \tag{B4}$$

k'_{N-1}, r'_{N-1} is given by Eq.(7). The form of Eq.(B4) is totally different from Eq.(B3) and Eq.(A2), so no effective incorporation rates can be properly defined and Eq.(2) and Eq.(A2) cannot be used to calculate the fidelity.

Similarly, the complex scheme Fig.3(A) can also be mapped to Fig.1 by the same logic. One can write the complete system of the integrated equations, and eliminate $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,2}, ..., \Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,N}$ (correspoding to $E_2 \cdot D_i, ..., E_N \cdot D_i$) while retain $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p}$ (corresponding to $E_p \cdot D_i$). This procedure leads to exactly the same equations as Eq.(A2) (with only $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i}$ replaced by $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p}$), which defines the effective incorporation rate and the effective excision rate rigorously (Eq.(7)(8)) and thus the probability distribution of the final products $P^{X_1...X_L}_{\alpha_1...\alpha_L}(t \to \infty)$ can be calculated in terms of these effective rates. Other reduction procedures which eliminates $\Gamma^{X_1...X_i...X_L}_{\alpha_1...\alpha_i,p}$ will result in equations totally different from Eq.(A2) and thus no effective rates can be properly defined.

[1] T. A. Trautner, M. N. Swartz, and A. Kornberg, Proc. Natl. Acad. Sci. U.S.A. **48**, 449 (1962).
[2] Z. W. Hall and I. R. Lehman, J. Mol. Biol. **36**, 321 (1968).
[3] D. F. Lee, J. Lu, S. Chang, J. J. Loparo, and X. S. Xie, Nucleic Acids Res. **44**, e118 (2016).
[4] A. M. de Paz, T. R. Cybulski, A. H. Marblestone, B. M. Zamft, G. M. Church, E. S. Boyden, K. P. Kording, and K. E. Tyo, Nucleic Acids Res. **46**, e78 (2018).
[5] K. Clayton and W. Branscomb, J. Biol. Chem. **254**, 1902 (1979).

[6] J. G. Bertram, K. Oertell, J. Petruska, and M. F. Goodman, Biochemistry **49**, 20 (2010).
[7] J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **71**, 4135 (1974).
[8] J. Ninio, Biochimie **57**, 587 (1975).
[9] A. R. Fersht, *Enzyme Structure and Mechanism* (W.H.Freeman & Co Ltd., 1985), 2nd ed.
[10] A. R. Fersht, Proc. R. Soc. London, Ser. B **187**, 397 (1974).
[11] W. M. Kati, K. A. Johnson, L. F. Jerva, and K. S. Anderson, J. Biol. Chem. **267**, 25988 (1992).
[12] K. A. Johnson, Annu. Rev. Biochem. **62**, 685 (1993).
[13] A. R. Fersht, J. W. Knill-Jones, and W. C. Tsui, J. Mol. Biol. **156**, 37 (1982).
[14] A. R. Fersht, Proc. Natl. Acad. Sci. U.S.A. **76**, 4946 (1979).
[15] B. M. Wingert, E. E. Parrott, and S. W. Nelson, Biochemistry **52**, 7723 (2013).
[16] A. K. Vashishtha and R. D. Kuchta, Biochemistry **54**, 240 (2015).
[17] M. J. Donlin, S. S. Patel, and K. A. Johnson, Biochemistry **30**, 538 (1991).
[18] A. Bębenek and I. Ziuzia-Graczyk, Curr. Genet. **64**, 985 (2018).
[19] P. Gaspard, Phys. Rev. E **96**, 042403 (2017).
[20] Q. S. Li, P. D. Zheng, Y. G. Shu, Z. C. Ou-Yang, and M. Li, Phys. Rev. E **100**, 012131 (2019).
[21] Y. S. Song, Y. G. Shu, X. Zhou, Z. C. Ou-Yang, and M. Li, J. Phys.: Condens. Matter **29**, 025101 (2017).
[22] R. Lamichhane, S. Y. Berezhna, J. P. Gill, E. Van Der Schans, and D. P. Millar, J. Am. Chem. Soc. **135**, 4735 (2013).
[23] I. Wong, S. S. Patel, and K. A. Johnson, Biochemistry **30**, 526 (1991).
[24] A. A. Johnson and K. A. Johnson, J. Biol. Chem. **276**, 38097 (2001).
[25] A. A. Johnson and K. A. Johnson, J. Biol. Chem. **276**, 38090 (2001).
[26] K. A. Johnson, The Enzymes **20**, 1 (1992).
[27] M. F. Goodman, S. Creighton, L. B. Bloom, J. Petruska, and T. A. Kunkel, Crit. Rev. Biochem. Mol. Biol. **28**, 83 (1993).
[28] J. M. Dahl, A. H. Mai, G. M. Cherf, N. N. Jetha, D. R. Garalde, A. Marziali, M. Akeson, H. Wang, and K. R. Lieberman, J. Biol. Chem. **287**, 13407 (2012).
[29] K. R. Lieberman, J. M. Dahl, A. H. Mai, M. Akeson, and H. Wang, J. Am. Chem. Soc. **134**, 18816 (2012).
[30] K. R. Lieberman, J. M. Dahl, A. H. Mai, A. Cox, M. Akeson, and H. Wang, J. Am. Chem. Soc. **135**, 9149 (2013).
[31] K. R. Lieberman, J. M. Dahl, and H. Wang, J. Am. Chem. Soc. **136**, 7117 (2014).
[32] J. A. Morin, F. J. Cao, J. M. Lázaro, J. R. Arias-Gonzalez, J. M. Valpuesta, J. L. Carrascosa, M. Salas, and B. Ibarra, Nucleic Acids Res. **43**, 3643 (2015).
[33] Z. Ren, Nucleic Acids Res. **44**, 7457 (2016).
[34] Y. C. Tsai and K. A. Johnson, Biochemistry **45**, 9675 (2006).
[35] V. Purohit, N. D. Grindley, and C. M. Joyce, Biochemistry **42**, 10200 (2003).
[36] C. M. Joyce, O. Potapova, A. M. DeLucia, X. Huang, V. P. Basu, and N. D. Grindley, Biochemistry **47**, 6103 (2008).
[37] Y. Santoso, C. M. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J. P. Torella, N. D. Grindley, and A. N. Kapanidis, Proc. Natl. Acad. Sci. U.S.A. **107**, 715 (2010).
[38] J. Hohlbein, L. Aigrain, T. D. Craggs, O. Bermek, O. Potapova, P. Shoolizadeh, N. D. Grindley, C. M. Joyce, and A. N. Kapanidis, Nat. Commun. **4**, 1 (2013).
[39] J. SantaLucia and D. Hicks, Annu. Rev. Biophys. Biomol. Struct. **33**, 415 (2004).