General Methods for Analysis of Sequential "*n*-step" Kinetic Mechanisms: Application to Single Turnover Kinetics of Helicase-Catalyzed DNA Unwinding

Aaron L. Lucius, Nasib K. Maluf, Christopher J. Fischer, and Timothy M. Lohman
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

ABSTRACT Helicase-catalyzed DNA unwinding is often studied using "all or none" assays that detect only the final product of fully unwound DNA. Even using these assays, quantitative analysis of DNA unwinding time courses for DNA duplexes of different lengths, *L*, using "*n*-step" sequential mechanisms, can reveal information about the number of intermediates in the unwinding reaction and the "kinetic step size", *m*, defined as the average number of basepairs unwound between two successive rate limiting steps in the unwinding cycle. Simultaneous nonlinear least-squares analysis using "*n*-step" sequential mechanisms has previously been limited by an inability to float the number of "unwinding steps", *n*, and *m*, in the fitting algorithm. Here we discuss the behavior of single turnover DNA unwinding time courses and describe novel methods for nonlinear least-squares analysis that overcome these problems. Analytic expressions for the time courses, $f_{ss}(t)$, when obtainable, can be written using gamma and incomplete gamma functions. When analytic expressions are not obtainable, the numerical solution of the inverse Laplace transform can be used to obtain $f_{ss}(t)$. Both methods allow *n* and *m* to be continuous fitting parameters. These approaches are generally applicable to enzymes that translocate along a lattice or require repetition of a series of steps before product formation.

INTRODUCTION

The biological functions of many enzymes require the coupling of hydrolysis of nucleoside triphosphates (NTP or dNTP) to translocation of the enzyme along a linear protein or nucleic acid filament. Some of these so-called ''motor proteins'', such as the kinesin family of proteins, move along microtubules (Howard et al., 1989; Block et al., 1990; Vale and Fletterick, 1997). Others, such as DNA and RNA polymerases (Kornberg and Baker, 1992), helicases (Matson and Kaiser-Rogers, 1990; Lohman and Bjornson, 1996; Soultanas and Wigley, 2000; Patel and Picha, 2000), some nucleases (Kovall and Matthews, 1997), and some restriction enzymes (Szczelkun, 2002) translocate along linear nucleic acid filaments. A complete understanding of the molecular mechanisms by which these motor proteins function requires quantitative kinetic information (rates and rate constants) on the intermediate steps involved in the translocation process. Single molecule approaches are beginning to allow the study of intermediate steps in some of these processes (Wuite et al., 2000; Visscher et al., 1999; Dohoney and Gelles, 2001; Ha et al., 2002; Bianco et al., 2001), although ensemble studies can also often be used to obtain such information (Roman and Kowalczykowski, 1989; Eggleston et al., 1996; Taylor and Smith, 1980; Cheng et al., 2001; Ali and Lohman, 1997; Lucius et al., 2002; Dillingham et al., 2002; Jankowsky et al., 2000; Raney and Benkovic, 1995).

This manuscript focuses on the use of ensemble approaches and the methods of analyses needed to examine the kinetic mechanisms by which a class of nucleic acid enzymes, called helicases (Matson and Kaiser-Rogers, 1990; Lohman and Bjornson, 1996; Soultanas and Wigley, 2000; Patel and Picha, 2000) unwind duplex nucleic acids and translocate along DNA. The assays that are most generally used to study the kinetic behavior of these enzymes are ''all or none'' assays, meaning that they directly detect only the final product or step of the reaction (i.e., completely unwound single-stranded DNA), although assays that directly detect partially unwound DNA intermediates have also been developed (Roman and Kowalczykowski, 1989; Eggleston et al., 1996; Taylor and Smith, 1980; Cheng et al., 2001). However, even when ''all or none'' assays are used, one can often still obtain mechanistic information about the number of steps and the rate constants for the intermediate steps through quantitative analysis of the time courses of single turnover reactions (Ali and Lohman, 1997; Jankowsky et al., 2000; Lucius et al., 2002). Systematic studies of the dependence of the single turnover time course of DNA unwinding on duplex DNA length can yield estimates of the unwinding rate and the kinetic ''step size'', *m*, i.e., the average number of basepairs unwound between two successive rate limiting steps that are repeated during the unwinding cycle.

Submitted April 11, 2003, and accepted for publication July 15, 2003.

Address reprint requests to Timothy M. Lohman, Dept. of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110. Tel.: 314-362-4393; Fax: 314-362-7183; E-mail: lohman@biochem.wustl.edu.

0006-3495/03/10/2224/16 $2.00

"All or none" assays for helicase-catalyzed DNA unwinding

''All or none'' assays for studies of helicase-catalyzed unwinding of duplex nucleic acids involve the use of either radioactively labeled nucleic acid (Ali and Lohman, 1997;

Ali et al., 1999; Jankowsky et al., 2000; Lucius et al., 2002) or fluorescently labeled nucleic acid to monitor complete unwinding of the DNA (Bjornson et al., 1994; Houston and Kodadek, 1994; Cheng et al., 2001; A. L. Lucius and T. M. Lohman, unpublished results). These approaches are described in detail in the papers cited above; however, Fig. 1 outlines these approaches and shows examples of time courses of unwinding of a 40 bp DNA duplex by *Escherichia coli* RecBCD helicase.

In both types of experiments, the helicase is prebound to a duplex DNA substrate at equilibrium. In a chemical quenched-flow experiment (Fig. 1 *A*), the prebound complex is rapidly mixed with ATP and a large excess of protein trap (typically unlabeled DNA or heparin) to ensure that rebinding of free protein to the DNA does not occur during the ensuing reaction. This ensures that each helicase is involved in only a single round of DNA unwinding (i.e., reactions are single turnover with respect to the DNA, although multiple rounds of ATP hydrolysis occur). The reaction is then allowed to proceed for a specified amount of time (Δt) before a quenching solution (typically an excess of EDTA to chelate the free Mg^{2+}) is added to stop the reaction. The quenched sample is subsequently analyzed by nondenaturing gel electrophoresis to separate the product

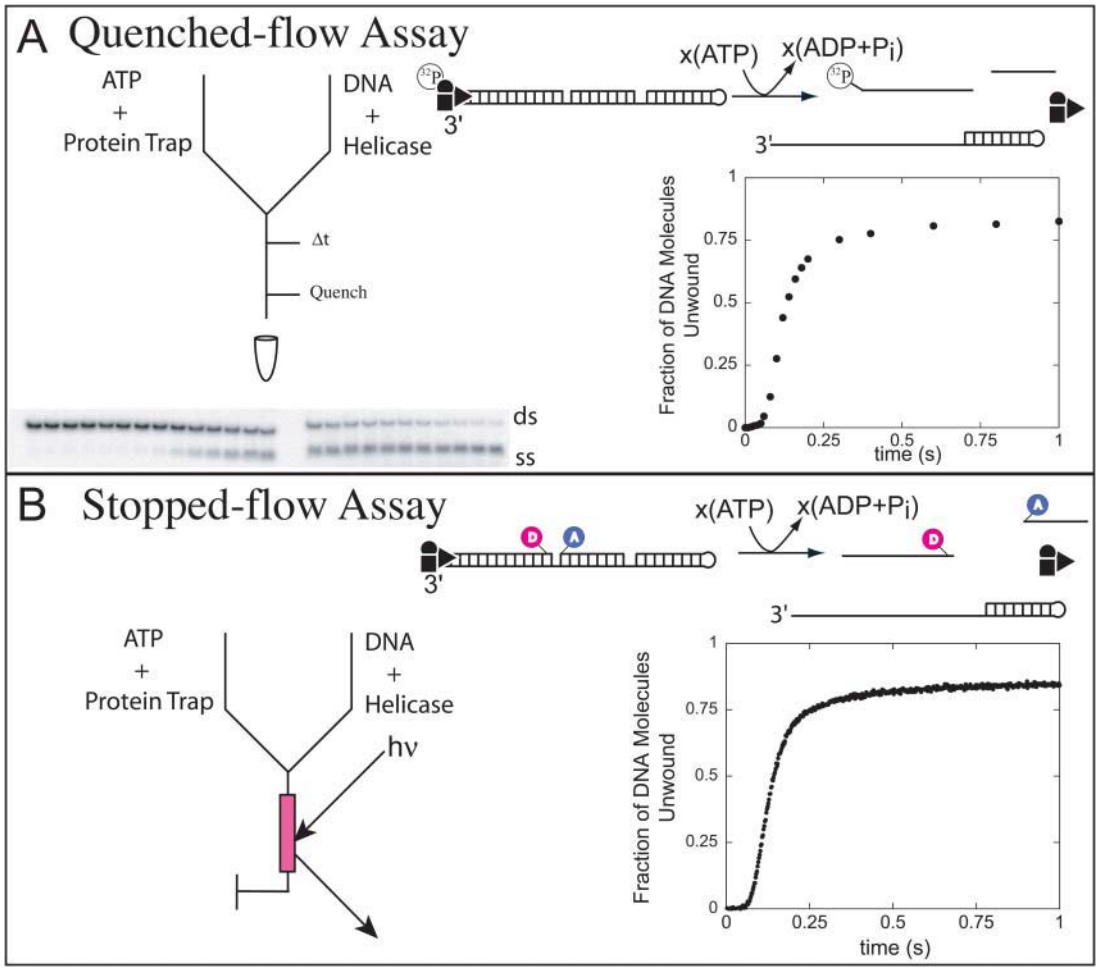

FIGURE 1 Schematic representation of an all or none assay using either chemical quenched-flow techniques or fluorescence stopped-flow techniques. The combined circle, square, and triangle represent the RecBCD helicase prebound to a blunt-ended DNA substrate. Upon addition of ATP, *x* number of ATP molecules are hydrolyzed and multiple steps are taken to completely unwind the duplex region. (*A*) DNA substrate is radiolabeled with ^{32}P at the 5′ end. The radiolabeled DNA substrate is incubated with helicase and allowed to achieve binding equilibrium. The helicase-DNA complex is loaded into one syringe of the quenched-flow apparatus, and ATP and a large excess of unlabeled DNA, to serve as a protein trap, are loaded into the other syringe. The samples are then rapidly mixed together and allowed to incubate for various times, Δt, before mixing with quench and expelling into a tube. Each reaction is run on a gel to separate ssDNA from dsDNA (denoted as ss and ds, respectively, in the *inset*); these bands are then quantitated and plotted in the graph in the inset. As described, the quenched-flow assay is both an all or none assay and a discontinuous assay in that each time point is an individual reaction. (*B*) DNA substrate is labeled with donor and acceptor FRET pairs. Again, the DNA substrate is incubated in one syringe with helicase and the other syringe contains ATP and protein trap. The samples are rapidly mixed together and the fluorescence signals from the fluorophores are observed. At time zero, the donor and acceptor are in close proximity, thus energy transfer between the two fluorophores is very efficient. In principle, upon complete unwinding, the fluorescence from the donor increases and the fluorescence from the acceptor should exhibit an anticorrelated decrease. In the inset is the donor fluorescence as a function of time.

(fully unwound ssDNA) from the fully native dsDNA substrate. One such experiment yields only one point in the time course and thus must be repeated ~20–25 times, varying Δt in each experiment, to obtain a full time course. Therefore, in addition to its ''all or none'' nature, the quenched-flow assay is also a discontinuous assay since each time point (see Fig. 1 *A*) is obtained from a separate experiment. In such experiments, the assumption is made that, at the time of quenching, all partially unwound DNA intermediates reanneal completely to re-form the native duplex DNA substrate, whereas any completely unwound ssDNA does not reanneal. Due to the length of time required to perform all of the experiments needed to describe an entire time course, this type of experiment is often performed at low DNA concentration (1 nM) so that no significant reannealing of the product DNA occurs during the course of the experiment. However, these experiments can be performed at high [DNA] by including a trap for ssDNA product (i.e., inclusion of a large excess of nonlabeled ssDNA that is complementary to one of the strands of the duplex DNA substrate) (Maluf et al., 2003).

In the fluorescence resonance energy transfer (FRET) analog of this experiment (Fig. 1 *B*), a DNA substrate is labeled with both a donor and an acceptor fluorophore. In the duplex DNA, the donor and acceptor fluorophores are in close proximity and thus the efficiency of FRET is high, whereas after complete unwinding of the DNA, the two complementary strands of ssDNA become separated, resulting in a complete loss of FRET. The resultant time course obtained from this experiment, in principle, will be identical to that determined in the quenched-flow experiment if the only change in fluorescence intensity of the donor and/or acceptor fluorophore is due to the change in FRET upon complete separation of the two fluorophores. Unlike the quenched-flow experiment, the FRET experiment can be performed as a continuous assay in a stopped-flow instrument. However, the assumption is still made that any partially unwound DNA intermediates from which the helicase dissociates before complete unwinding of the DNA will reanneal to re-form the duplex DNA substrate, resulting in no change in the FRET signal for that molecule.

The radioactive chemical quenched-flow assay is clearly more time consuming since each time point requires a separate experiment. This also sets a practical limit for the number of data points that are generally obtained for a single time course (typically 20–30). The fluorescence assay can be performed in a stopped-flow instrument, and thus one generally obtains 400–500 data points for a time course in a single experiment. The advantage of the chemical quenched-flow experiment is that one obtains a direct measure of the extent of DNA unwinding, whereas, in the FRET experiment, one must be able to correlate the extent of the fluorescence change with the known extent of DNA unwinding. Therefore, quantitative analysis of the fluorescence unwinding time course usually requires comparison with an otherwise identical experiment performed by chemical quenched-flow.

In the following sections, we first describe the general behavior of several ''*n*-step'' sequential mechanisms for helicase-catalyzed DNA unwinding, of increasing complexity. We then discuss the experimental approaches and analysis of a set of single turnover unwinding experiments needed to estimate the number of intermediate steps, n, required to unwind a duplex DNA of length, L, and the ''kinetic step size'', $m = L/n$. Finally, we introduce novel methods to simultaneously (globally) analyze, by nonlinear least-squares (NLLS) methods, a series of time courses obtained for a set of DNA substrates varying in duplex DNA length, L. These methods allow for a noninteger number of steps as well as a noninteger kinetic step size, thus enabling one to float these parameters continuously in a NLLS analysis. These methods involve use of gamma and incomplete gamma functions when a closed form expression can be obtained for the time course of unwinding, $f_{ss}(t)$, and Laplace transforms and numerical approaches to obtain the inverse Laplace transforms when a closed form expression for the time course is not obtainable. We note that many methods, including matrix methods and numerical methods, can be used if all that is sought is to simply simulate a kinetic time course for a given mechanism. Such methods are also suitable for NLLS analysis of reactions in which the number of steps in the pathway is defined and fixed. However, simultaneous NLLS analysis of a series of *n*-sequential step mechanisms in which the number of steps, n, is not fixed but changes for each substrate (DNA in our case) and is in fact an unknown parameter to be determined from the analysis, is much more suited to the method of Laplace transforms. These techniques have been applied recently to studies of RecBCD catalyzed DNA unwinding (Lucius et al., 2002), but are also applicable to the analysis of mechanistic studies of any enzyme-catalyzed reaction that requires repetition of a series of steps before product formation.

Single turnover kinetic time courses for helicase-catalyzed DNA unwinding determined from an "all or none" assay

Case 1 (Scheme 1): All unwinding steps the same, infinite processivity (P = 1)

The single turnover time courses for RecBCD helicase-catalyzed DNA unwinding obtained from either the ''all or none'' quenched-flow or FRET stopped-flow assays display clear lag phases for formation of fully unwound ssDNA product (see Fig. 1). Such a lag phase is expected if the helicase proceeds through multiple repeated steps to fully unwind the duplex DNA, and if the rate constants for each step are similar in magnitude (Gutfreund, 1995; Ali and

Lohman, 1997). The simplest mechanism displaying this behavior is the ''*n*-step'' sequential mechanism shown in Scheme 1,

$$HD \xrightarrow{k_U} I_1 \xrightarrow{k_U} I_2 \xrightarrow{k_U} \ldots \xrightarrow{k_U} I_{n-1} \xrightarrow{k_U} ssDNA, \quad \text{(Scheme 1)}$$

where the preformed helicase-DNA complex, *HD*, proceeds through a series of *n* repeated, irreversible steps ($n - 1$ intermediates, I_i) before yielding fully unwound ssDNA product. In Scheme 1, the rate constants for each step, k_U, are equal, and upon addition of ATP, the helicase unwinds *m* basepairs in each step. Previous treatments of *n*-step sequential mechanisms have only dealt with ''*n*'' irreversible steps with different rate constants for each step (Capellos and Bielski, 1972; Rodiguin and Rodiguina, 1964; Gutfreund, 1995). As written, Scheme 1 also assumes that once the helicase has initiated DNA unwinding, it does not dissociate until unwinding is completed (i.e., processivity, $P = 1$; see below).

We note that Scheme 1, and all of the schemes discussed in this manuscript, considers that all of the helicases initiate at time zero from the same point on the DNA, i.e., one end of the duplex, or a ss/ds-DNA junction. Thus, it is assumed that any helicase that unwinds a DNA molecule completely has proceeded through the same number of intermediates en route to product formation. This assumption needs to be kept in mind, since it is conceivable that some helicases may not initiate from the same point on the DNA substrate. For example, a helicase that requires a ssDNA flanking region (either 3′ or 5′ ssDNA) might initially be bound only to the ssDNA region, rather than the ss/ds-DNA junction. Upon addition of ATP, the helicase would then have to first translocate to the ss/ds-DNA junction before initiating DNA unwinding. If the ssDNA flanking region were long, then the steps required for translocation along the flanking ssDNA would also need to be considered in the analysis of the time course of DNA unwinding. Furthermore, since the ensemble population of helicases would be initially bound in some distribution (random?) along the ssDNA flanking region, all of these helicases would not reach the ss/ds-DNA junction synchronously. Although such a situation can be treated using the approaches described here, in this manuscript we only treat the case in which all of the helicases initiate DNA unwinding from the same start site.

There are several ways to obtain a general solution as a function of *n* for Scheme 1. The most direct method starts by solving the differential equations for a mechanism with one step ($n = 1$) to obtain an analytic solution for the time-dependent behavior of each species. One can then repeat the process for a two-step ($n = 2$) mechanism and so on until analytic solutions are obtained for $n = 1$, 2, 3, etc. These solutions can then be examined in an attempt to recognize a pattern and obtain a general analytic expression as a function of *n*. However, this is a tedious process even for the simplest mechanism.

We have used the method of Laplace transforms (see Appendix A) to solve the differential equations for the ''*n*-step'' sequential mechanisms described here. Any number of other methods, including matrix methods (Bujalowski and Jezewska, 2000; Arai et al., 1981) and numerical methods (Barshop et al., 1983), can be used if one is solely interested in simply simulating the time course of a defined mechanism in which the number of intermediates (steps) is known. However, these other methods are not well suited for NLLS analysis of an *n*-step sequential mechanism in which the number of steps, *n*, is to be determined from the analysis, or for simultaneous NLLS analysis of a series of substrates (in this case DNA) in which the number of steps, *n*, changes for each substrate. Matrix methods are not practical for global NLLS analysis of such systems because every change in the number of steps would require a new matrix of different dimensionality. However, a primary advantage of the Laplace transform method for the *n*-step sequential mechanisms discussed here is that it is relatively straightforward to obtain the Laplace transform for each intermediate species for a given ''*n*'', and to easily recognize emerging patterns in the Laplace transform as the number of steps, *n*, is changed (See Appendix A). A further advantage is that the Laplace transform describing ssDNA formation as a function of the Laplace variable, *s*, can always be obtained in analytic form as a continuous function of *n*. This leads to the other major advantage that one can readily use the numerical solution to the inverse Laplace transform in a NLLS fitting routine and therefore treat the number of steps *n* as a continuous fitting parameter even when an analytic solution in the time domain is not available. In this article, we demonstrate that the methods discussed here enable global NLLS analyses of a series of time courses, such that the number of steps, *n*, and the step size, *m*, can be floated continuously in the analysis, which was not possible using previous approaches (Ali and Lohman, 1997; Jankowsky et al., 2000).

As shown in Appendix A for Scheme 1, the Laplace transform, $F_{ss}(s)$, where *s* is the Laplace variable, of the fraction of ssDNA molecules formed in the time domain, $f_{ss}(t)$, is given in Eq. 1 (see Eq. A24 in Appendix A).

$$F_{ss}(s) = \frac{ssDNA'(s)}{[HD_0]} = \frac{k_U^n}{s(k_U + s)^n}. \quad (1)$$

The expression for the time dependence of the fraction of ssDNA molecules formed, $f_{ss}(t)$, is then obtained by taking the inverse Laplace transform of Eq. 1, yielding Eq. 2,

$$f_{ss}(t) = \mathcal{L}^{-1}(F_{ss}(s)) = \frac{ssDNA(t)}{[HD_0]} = \left(1 - \frac{\Gamma(n, k_U t)}{\Gamma(n)}\right) \quad (2)$$

(see Eq. A35 in Appendix A), where $\mathcal{L}^{-1}$ is the inverse Laplace transform operator, and $\Gamma(n)$ and $\Gamma(n, k_U t)$are the gamma function and incomplete gamma function, respectively (NB: some mathematics texts define the ratio, $\Gamma(n, k_U t)/\Gamma(n)$, as the incomplete gamma function, however, see

Appendix A for the definition used here). Note that in the general solution for $f_{ss}(t)$ given in Eq. 2, the number of steps, n, can be any positive value, including a noninteger. However, if n is an integer, then $f_{ss}(t)$ can be expressed as the series given in Eq. 3:

$$f_{ss}(t) = \left(1 - \sum_{r=1}^{n} \frac{(k_U t)^{r-1}}{(r-1)!} e^{-k_U t}\right). \quad (3)$$

Use of Eq. 2 has obvious advantages for NLLS analyses, since one can allow n to float continuously without restricting it to be an integer. Appendix A presents the derivations of these expressions as well as the expressions describing the time-dependent formation and disappearance of each intermediate species in Scheme 1.

In the following sections we also examine schemes that are more complex than Scheme 1. However, under some conditions, the solutions to these more complex schemes will have the same general form of Eqs. 1 and 2. In anticipation of this, we can write Eqs. 1 and 2 in the more general form given by Eq. 4, where k_U is replaced by an experimentally observed rate constant, k_{obs}, and n is replaced by an apparent number of steps, n_{app}. As we show below, each new scheme can then be described by Eq. 4 with a specific expression for k_{obs} and n_{app} applicable to that scheme.

$$f_{ss}(t) = \mathcal{L}^{-1}\left(\frac{k_{obs}^{n_{app}}}{s(k_{obs}+s)^{n_{app}}}\right) = \left(1 - \frac{\Gamma(n_{app}, k_{obs}t)}{\Gamma(n_{app})}\right). \quad (4)$$

Fig. 2 *A* shows a series of time courses for an ''all or none'' DNA unwinding assay simulated using Scheme 1 (Eq. 2) with values of $n = 1$–5. The time course for a single step reaction ($n = 1$) is a single exponential. However, the time courses for $n \geq 2$ display clear lag phases that increase in duration as the number of intermediates in the reaction increase. Fig. 2 *B* shows the time courses of formation and decay of the DNA substrate (*HD*), product (ssDNA), and the partially unwound intermediates, I_1, I_2, I_3, and I_4, simulated using Scheme 1 with $n = 5$. This demonstrates that the lag phase results from the fact that each intermediate state is significantly populated during the unwinding time course. It is important to note that if any one of the five rate constants in Scheme 1 is significantly slower than the other four (e.g., $k'_U \ll k_U$), then the observed time course of DNA unwinding will collapse to a single exponential with rate constant k'_U as given by Eq. 5:

$$f_{ss}(t) = 1 - e^{-k'_U t}. \quad (5)$$

Hence, the presence of a single slow step in the scheme yields a single exponential phase; the observation of a lag phase requires the presence of multiple steps with rate constants that are similar in magnitude.

Scheme 1 and the simulations in Fig. 2 indicate that in a single turnover helicase-catalyzed DNA unwinding

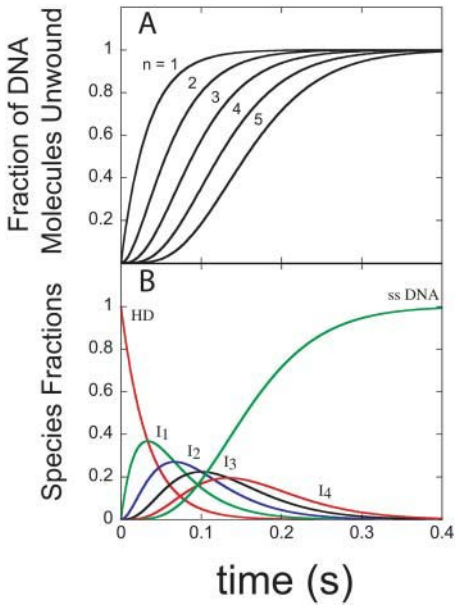

FIGURE 2 Simulated traces from Scheme 1 (Eq. 2). (*A*) Traces were simulated with an unwinding rate constant $k_U = 30\ s^{-1}$ and a number of steps $n = 1$, 2, 3, 4, and 5 as denoted in the panel. The plot illustrates the increase in the lag phase as a function of increasing number of steps. (*B*) The time dependence of the concentration of the helicase-DNA complex, HD (*red*), Intermediate I_1 (*green*), I_2 (*blue*), I_3 (*black*), I_4 (*red*), and ssDNA (*green*) from Scheme 1 with $k_U = 30\ s^{-1}$ and $n = 5$ steps.

reaction performed using an ''all or none'' assay, a lag phase will be observed if some rate-limiting step, with rate constant k_U, occurs repeatedly throughout the DNA unwinding reaction. Furthermore, the extent of this lag phase will increase as the number of bp, L, in the DNA duplex increases since the number of rate-limiting steps, n, will also increase in direct proportion to L. Therefore, if a series of experiments were performed as a function of duplex length, L, then analysis of the unwinding time courses can be used to estimate the number of steps required to unwind each DNA duplex and also the average number of basepairs unwound per rate limiting step, termed the kinetic step size, $m = L/n$ (Ali and Lohman, 1997).

Case 2 (Scheme 2): All unwinding steps the same, finite processivity ($0 < P < 1$)

The next case we consider is Scheme 2,

$$\underset{\downarrow k_d}{HD} \xrightarrow{k_U} \underset{\downarrow k_d}{I_1} \xrightarrow{k_U} \underset{\downarrow k_d}{I_2} \xrightarrow{k_U} \cdots \xrightarrow{k_U} \underset{\downarrow k_d}{I_{n-1}} \xrightarrow{k_U} ssDNA, \quad \text{(Scheme 2)}$$

which differs from Scheme 1 in that the helicase can dissociate from the DNA substrate, with rate constant, k_d, at each step during the course of unwinding. Scheme 2 introduces the concept of a finite processivity for DNA unwinding. Processivity, P, of DNA unwinding is a measure of the number of basepairs that are unwound by the helicase per DNA binding event (i.e., before the helicase

dissociates). It can be described, quantitatively, in either of two equivalent ways. First, processivity is the probability, P ($0 \leq P \leq 1$), that the helicase will proceed one step to unwind the next basepair (or m number of basepairs) rather than dissociate from the DNA. The processivity is also related to the average number of basepairs unwound (N) per binding event. As shown in Eq. 6 (McClure and Chow, 1980; Lohman and Bjornson, 1996; von Hippel and Delagoutte, 2001)

$$P = \frac{k_{\mathrm{U}}}{k_{\mathrm{U}} + k_{\mathrm{d}}} = e^{-(\mathrm{m/N})}, \tag{6}$$

where m is the DNA unwinding step size (bp unwound per step) (see Appendix B for a derivation of Eq. 6).

The differential equations that describe Scheme 2 can be solved following the same approaches detailed in Appendix A, yielding Eq. 7 for $F_{\mathrm{ss}}(s)$,

$$F_{\mathrm{ss}}(s) = \frac{k_{\mathrm{U}}^{\mathrm{n}}}{s(k_{\mathrm{U}} + k_{\mathrm{d}} + s)^{\mathrm{n}}}, \tag{7}$$

which is the Laplace transform of $f_{\mathrm{ss}}(t)$. The difference between the solutions for Scheme 2 (Eq. 7) versus Scheme 1 (Eq. 1) is that the rate constants k_{d} and k_{U} appear as a sum in the denominator in Eq. 7. This is always the case when dissociation can occur and thus we define an observed rate constant, k_{obs}, as in Eq. 8,

$$k_{\mathrm{obs}} = k_{\mathrm{U}} + k_{\mathrm{d}}, \tag{8}$$

which leads to Eq. 9:

$$F_{\mathrm{ss}}(s) = \left(\frac{k_{\mathrm{U}}}{k_{\mathrm{obs}}}\right)^{\mathrm{n}} \frac{k_{\mathrm{obs}}^{\mathrm{n}}}{s(k_{\mathrm{obs}} + s)^{\mathrm{n}}} = P^{\mathrm{n}} \frac{k_{\mathrm{obs}}^{\mathrm{n}}}{s(k_{\mathrm{obs}} + s)^{\mathrm{n}}}. \tag{9}$$

The inverse Laplace transform of Eq. 9 then yields Eq. 10 for $f_{\mathrm{ss}}(t)$:

$$f_{\mathrm{ss}}(t) = P^{\mathrm{n}} \mathcal{L}^{-1}\left(\frac{k_{\mathrm{obs}}^{\mathrm{n}}}{s(k_{\mathrm{obs}} + s)^{\mathrm{n}}}\right) = P^{\mathrm{n}}\left(1 - \frac{\Gamma(n, k_{\mathrm{obs}}t)}{\Gamma(n)}\right). \tag{10}$$

Note that with the exception of the processivity term, Eqs. 9 and 10 have the same form as Eqs. 1 and 2. For integer values of n, $f_{\mathrm{ss}}(t)$ can also be expressed as in Eq. 11:

$$f_{\mathrm{ss}}(t) = P^{\mathrm{n}}\left(1 - \sum_{\mathrm{r}=1}^{\mathrm{n}} \frac{((k_{\mathrm{obs}})t)^{\mathrm{r}-1}}{(r-1)!} e^{-(k_{\mathrm{obs}})t}\right). \tag{11}$$

In principle, an estimation of the processivity can be obtained if there is a detectable decrease in the total DNA unwinding amplitude with increasing duplex length, although there can also be other reasons for such decreases in amplitude.

Fig. 3 *A* shows a series of simulated time courses for Scheme 2, generated using Eq. 10 with $k_{\mathrm{U}} = 30\ \mathrm{s}^{-1}$, $k_{\mathrm{d}} = 1\ \mathrm{s}^{-1}$, (i.e., $k_{\mathrm{obs}} = 31\ \mathrm{s}^{-1}$) and n = 1–5. These show the expected lag phase that increases with increasing n. However, due to the finite dissociation rate constant, k_{d}, the final amplitude of DNA unwinding also decreases with n. Fig. 3 *B* shows a plot of the total unwinding amplitude ($f_{\mathrm{ss}}(\infty)$) as a function of n, along with a fit of these amplitudes to $f_{\mathrm{ss}}(\infty) = P^{\mathrm{n}}$. In principle, one can estimate the processivity, P, from such data.

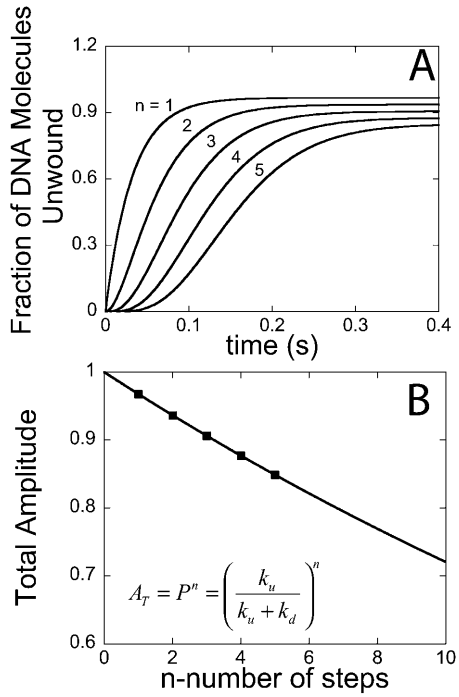

FIGURE 3 Simulated traces from Scheme 2 (Eq. 10). (*A*) Traces were simulated with an unwinding rate constant $k_{\mathrm{U}} = 30\ \mathrm{s}^{-1}$, a dissociation rate constant of $k_{\mathrm{d}} = 1\ \mathrm{s}^{-1}$, and a number of steps n = 1, 2, 3, 4, and 5. (*B*) The total amplitude ($A_{\mathrm{T}} = f_{\mathrm{ss}}(\infty)$) of each trace in *A* plotted as a function of the number of steps. Solid line is a nonlinear least-squares fit to $A_{\mathrm{T}} = P^{\mathrm{n}}$, where A_{T} is the total unwinding amplitude, P is the processivity expressed in Eq. 6, and n is the number of steps. The fit predicts a processivity of 0.9677, which is equal to $k_{\mathrm{U}}/(k_{\mathrm{U}} + k_{\mathrm{d}})$ or $(30\ \mathrm{s}^{-1})/(31\ \mathrm{s}^{-1})$.

Case 3 (Scheme 3): All unwinding steps the same, finite processivity, biphasic time course

A number of DNA helicases that have been studied exhibit biphasic time courses in single turnover DNA unwinding studies (Lucius et al., 2002; Ali and Lohman, 1997), such that a second, slower phase occurs in addition to the expected ''lag'' phase. This behavior has been interpreted as reflecting the presence of two types of helicase complexes that are prebound to the DNA substrate before the addition of ATP, as depicted in Scheme 3. One is a productively bound complex, $(HD)_{\mathrm{P}}$, that is poised to initiate DNA unwinding immediately. The other is a non-productively bound complex, $(HD)_{\mathrm{NP}}$, that must first isomerize to form a productive complex before initiating DNA unwinding.

$$\begin{array}{l}(HD)_{\mathrm{NP}}\\ \downarrow k_{\mathrm{NP}}\\ (HD)_{\mathrm{P}} \xrightarrow[\downarrow k_{\mathrm{d}}]{k_{\mathrm{U}}} I_1 \xrightarrow[\downarrow k_{\mathrm{d}}]{k_{\mathrm{U}}} I_2 \xrightarrow[\downarrow k_{\mathrm{d}}]{k_{\mathrm{U}}} \ldots \xrightarrow[\downarrow k_{\mathrm{d}}]{k_{\mathrm{U}}} I_{\mathrm{n}-1} \xrightarrow[\downarrow k_{\mathrm{d}}]{k_{\mathrm{U}}} ssDNA.\end{array}$$

(Scheme 3)

Scheme 3 was used to analyze the unwinding of DNA by *E. coli* UvrD helicase, using the approximate analytic solution shown in Eq. 12 (Ali and Lohman, 1997),

$$f_{\mathrm{ss}}(t) = A_{\mathrm{T}}\left[x\left\{1 - \sum_{\mathrm{r}=1}^{\mathrm{n}} \frac{(k_{\mathrm{obs}}t)^{\mathrm{r}-1}}{(r-1)!} e^{-k_{\mathrm{obs}}t}\right\} + (1-x)(1-e^{-k_{\mathrm{np}}t})\right], \tag{12}$$

where A_{T} is the total unwinding amplitude (Eq. 13), $k_{\mathrm{obs}} = (k_{\mathrm{U}} + k_{\mathrm{d}})$, and x is the fraction of productively bound complexes as defined in Eq. 14.

$$A_{\mathrm{T}} = \left(\frac{k_{\mathrm{u}}}{k_{\mathrm{u}} + k_{\mathrm{d}}}\right)^{\mathrm{n}}. \tag{13}$$

$$x = \frac{(HD)_{\mathrm{P}}}{(HD)_{\mathrm{NP}} + (HD)_{\mathrm{P}}} = \frac{[HD]_{\mathrm{P}}}{[HD]_0}. \tag{14}$$

The expression in Eq. 12 is an approximation that is strictly valid under conditions such that $k_{\mathrm{obs}} \gg k_{\mathrm{NP}}$, and for small numbers of steps, n.

Equation 15 is the Laplace transform of $f_{\mathrm{ss}}(t)$ for Scheme 3, with no assumptions:

$$\begin{aligned} F_{\mathrm{ss}}(s) &= \frac{k_{\mathrm{U}}^{\mathrm{n}}(k_{\mathrm{NP}} + sx)}{s(k_{\mathrm{NP}} + s)(k_{\mathrm{U}} + k_{\mathrm{d}} + s)^{\mathrm{n}}} \\ &= P^{\mathrm{n}} \frac{(k_{\mathrm{NP}} + sx)}{(k_{\mathrm{NP}} + s)} \frac{k_{\mathrm{obs}}^{\mathrm{n}}}{s(k_{\mathrm{obs}} + s)^{\mathrm{n}}}. \end{aligned} \tag{15}$$

The expression for $f_{\mathrm{ss}}(t)$ for Scheme 3 (with no assumptions) is obtained from the inverse Laplace transform of Eq. 15 and is given in Eq. 16. We now use Eq. 16 in place of Eq. 12 to analyze time courses based on Scheme 3 (Lucius et al., 2002):

$$\begin{aligned} f_{\mathrm{ss}}(t) = \left(\frac{k_{\mathrm{U}}}{k_{\mathrm{obs}}}\right)^{\mathrm{n}} \Bigg[&\left(1 - \frac{\Gamma(n, (k_{\mathrm{obs}})t)}{\Gamma(n)}\right) - e^{-k_{\mathrm{NP}}t}(1-x) \\ &\times \left(\frac{k_{\mathrm{obs}}}{k_{\mathrm{obs}} - k_{\mathrm{NP}}}\right)^{\mathrm{n}} \left(1 - \frac{\Gamma(n, (k_{\mathrm{obs}} - k_{\mathrm{NP}})t)}{\Gamma(n)}\right)\Bigg]. \end{aligned} \tag{16}$$

Equation 16, unlike Eq. 12, has the advantage of being a continuous function of the number of steps, n, and therefore is easily implemented into NLLS routines. For comparison with previously published equations, we also show Eq. 17, which is the series representation of Eq. 16, valid only for integer values of n, and therefore not useful for a NLLS analysis to determine n:

$$\begin{aligned} f_{\mathrm{ss}}(t) = \left(\frac{k_{\mathrm{U}}}{k_{\mathrm{obs}}}\right)^{\mathrm{n}} \Bigg[&\left(1 - \sum_{\mathrm{r}=1}^{\mathrm{n}} \frac{((k_{\mathrm{obs}})t)^{\mathrm{r}-1}}{(r-1)!} e^{-(k_{\mathrm{obs}})t}\right) \\ &- e^{-k_{\mathrm{NP}}t}(1-x)\left(\frac{k_{\mathrm{obs}}}{k_{\mathrm{obs}} - k_{\mathrm{NP}}}\right)^{\mathrm{n}} \\ &\times \left(1 - \sum_{\mathrm{r}=1}^{\mathrm{n}} \frac{((k_{\mathrm{obs}} - k_{\mathrm{NP}})t)^{\mathrm{r}-1}}{(r-1)!} e^{-(k_{\mathrm{obs}} - k_{\mathrm{NP}})t}\right)\Bigg]. \end{aligned} \tag{17}$$

Fig. 4 shows a series of time courses simulated using Scheme 3 (Eq. 16) for n = 1–5.

Case 4: Additional steps in the mechanism, not involved directly in DNA unwinding (infinite processivity)

As discussed above, the simplest mechanism that can produce a lag in the time course of production of ssDNA is Scheme 1 in which one step is repeated n times, and the rate constants for each repeated step are identical or similar in magnitude. The next level of complexity is a mechanism that contains additional steps beyond those associated with DNA unwinding. An example of such a mechanism is shown in Scheme 4,

$$\begin{array}{l} HD \xrightarrow{k_{\mathrm{c}}} HD_1^* \xrightarrow{k_{\mathrm{c}}} HD_2^* \ldots \xrightarrow{k_{\mathrm{c}}} HD_{\mathrm{h}}^* \xrightarrow{k_{\mathrm{U}}} \\ \quad I_1 \xrightarrow{k_{\mathrm{U}}} I_2 \xrightarrow{k_{\mathrm{U}}} \ldots \xrightarrow{k_{\mathrm{U}}} I_{\mathrm{n}-1} \xrightarrow{k_{\mathrm{U}}} ssDNA, \end{array} \qquad \text{(Scheme 4)}$$

in which a series of steps, with rate constant, k_{C}, occurs before the actual unwinding steps, but the rate constants for these steps are not sufficiently slow as to eliminate the presence of the lag phase. In fact, the actual position of the steps with rate constants, k_{C}, within Scheme 4 does not affect the final expression for the time course of formation of the final ssDNA product. Hence, the steps with rate constant, k_{C}, could also occur after the unwinding steps, as depicted in

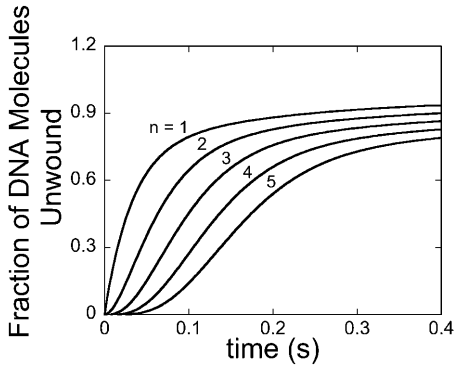

FIGURE 4 Simulated traces from Scheme 3 (Eq. 16). Traces were simulated with an unwinding rate constant $k_{\mathrm{U}} = 30\ \mathrm{s}^{-1}$, a dissociation rate constant $k_{\mathrm{d}} = 1\ \mathrm{s}^{-1}$, a nonproductive rate constant $k_{\mathrm{NP}} = 5\ \mathrm{s}^{-1}$, and a fraction of productively bound complexes $x = 0.8$, for n = 1, 2, 3, 4, and 5. These traces demonstrate a rapid lag phase followed by a slow second phase of unwinding.

Scheme 4a, or anywhere else within the sequential mechanism, if only the time course of the final ssDNA product is monitored.

For simplicity, we have not considered either dissociation of the helicase or the occurrence of any nonproductive complex in Scheme 4. The Laplace transform of the time dependence of ssDNA formation, $f_{ss}(t)$, for Scheme 4 is given in Eq. 18:

$$F_{ss}(s) = \frac{k_C^h k_U^n}{s(k_C + s)^h (k_U + s)^n}. \quad (18)$$

We have been unable to obtain a general analytic expression as a function of h for $f_{ss}(t)$, although one can obtain analytic expressions for specified integer values of h (i.e., $h = 1$, 2, or 3). However, even these analytic expressions are cumbersome; hence we do not include them here. On the other hand, as is generally true, one can obtain $f_{ss}(t)$ by solving for the inverse Laplace transform numerically (see below and Appendix C). A series of time courses simulated using Scheme 4 are shown in Fig. 5 ($k_U = 30\ s^{-1}$, $k_C = 10\ s^{-1}$, $n = 1$, 2, 3, 4, and 5, and $h = 2$). Comparison of these simulations with those generated using the simpler Scheme 1 (Fig. 2), indicates that inclusion of the slow steps, with rate constant, k_C, results in significantly less separation between time courses for increasing values of n (or equivalently, duplex lengths). Of course, in the limit where $k_C \ll k_U$, the time courses would become independent of n and display no dependence on duplex length. We note that for Scheme 4 (Fig. 5) when $n = 1$, a lag phase is still observed due to the presence of the two k_C steps, unlike for Scheme 1 (Fig. 2), where only a single exponential is observed for $n = 1$.

As stated above, an interesting feature of Scheme 4 is that the final expressions for $F_{ss}(s)$ and therefore $f_{ss}(t)$ are independent of the actual positions of the different k_C steps within Scheme 4. As a result, one cannot determine where these steps occur within the Scheme based only on the time course of formation of the final ssDNA product. Independent information about the time course of the intermediate species is needed to determine this.

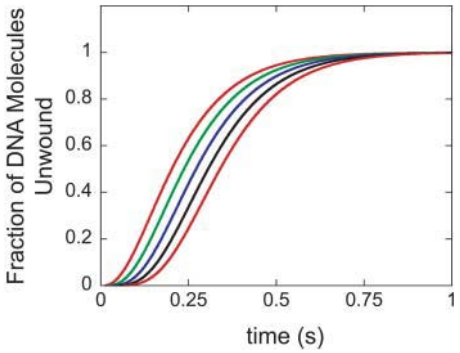

FIGURE 5 Simulated traces from Scheme 4 (Eq. 18). (*A*) Traces were simulated with an unwinding rate constant $k_U = 30\ s^{-1}$, $k_c = 10\ s^{-1}$, $n = 1$ (*red*), 2 (*green*), 3 (*blue*), 4 (*black*), and 5 (*red*), and $h = 2$. The plot illustrates that inclusion of a slow step results in significantly less separation between time courses for increasing values of n when compared to Fig. 2 where there is no slow step.

One can generalize Scheme 4 to the case in which a distribution of z different classes of steps exist, with each class of steps having rate constant k_i and each step within the class occurring n_i times, and with n_0 unwinding steps, each with unwinding rate constant, k_U. In this case, Eq. 18 can be generalized to Eq. 19:

$$F_{ss}(s) = \frac{k_U^{n_0}}{s(k_U + s)^{n_0}} \prod_{i=1}^{z} \frac{k_i^{n_i}}{(k_i + s)^{n_i}}. \quad (19)$$

For the case where a nonproductive binding mode occurs as in Scheme 3, then $F_{ss}(s)$ is given by Eq. 20:

$$F_{ss}(s) = \frac{(k_{NP} + sx)}{(k_{NP} + s)} \frac{k_U^{n_0}}{s(k_U + s)^{n_0}} \prod_{i=1}^{z} \frac{k_i^{n_i}}{(k_i + s)^{n_i}}. \quad (20)$$

Equations similar to Eqs. 19 and 20 have been described by Schnitzer and Block (1995). These authors also make the point that if a distribution of reaction rates occurs within a sequential scheme, then the Laplace transform, $F_{ss}(s)$, is simply the product of the terms in the Laplace domain describing each step, as can be seen in Eqs. 19 and 20. We note that for the cases we consider here, the presence of the $1/s$ term in Eqs. 19 and 20 arises from the fact that our expressions describe the formation of only the final ssDNA product.

As pointed out above, no matter where in the mechanism the steps with rate constant k_C occur, the expression for $f_{ss}(t)$ will be the same. However, all possible schemes will fall into three special cases. The first special case is when h is constant (independent of duplex length). The second special case is when $n = h$, i.e., each unwinding step is coupled to a different step (e.g., exhibiting a conformational change). The third special case is when both n and h are directly proportional to L, but with different proportionality constants, such that they occur with different frequencies during the unwinding scheme. These special cases are each considered below.

First special case: h is constant (independent of duplex length). The first special case can be described by either Scheme 4 or Scheme 4a and was used to analyze RecBCD-catalyzed DNA unwinding (Lucius et al., 2002):

$$HD \xrightarrow{k_U} I_1 \xrightarrow{k_U} I_2 \xrightarrow{k_U} \ldots \xrightarrow{k_U} I_{n-1} \xrightarrow{k_c} HD_1^* \xrightarrow{k_c} \ldots \xrightarrow{k_c} ssDNA. \quad \text{(Scheme 4a)}$$

The Laplace transform of $f_{ss}(t)$ for either Scheme 4 or Scheme 4a is given by Eq. 18.

The steps with rate constant k_C can be resolved only if $k_U \geq k_C$. In fact, in the limit as k_C approaches infinity, both

Scheme 4 and Scheme 4a (Eq. 18) collapse to Scheme 1 (Eq. 1). This can be shown by examining Eq. 18 in this limit, which yields Eq. 21,

$$\lim_{k_C \to \infty} \frac{k_C^h k_U^{L/m}}{s(k_C+s)^h (k_U+s)^{L/m}} = \frac{k_U^{L/m}}{s(k_U+s)^{L/m}}, \quad (21)$$

which is equivalent to $F_{ss}(s)$ for Scheme 1, with $n = L/m$. This makes sense, because if k_C is very fast relative to k_U, the step with rate constant k_C will not be observable, and Scheme 4 and Scheme 4a will collapse to Scheme 1.

However, if $k_C = k_U$, then the resulting time course can be described by Scheme 2 with n replaced by $n_{app} = n + h = L/m + h$. This can be illustrated by taking the limit of Eq. 18 as k_C approaches k_U, which yields Eq. 22:

$$\lim_{k_C \to k_U} \frac{k_C^h k_U^{L/m}}{s(k_C+s)^h (k_U+s)^{L/m}} = \frac{k_U^{(L/m)+h}}{s(k_U+s)^{(L/m)+h}} = \frac{k_{obs}^{n_{app}}}{s(k_{obs}+s)^{n_{app}}}. \quad (22)$$

The inverse Laplace transform of Eq. 22 yields Eq. 4, with n_{app} defined as in Eq. 23:

$$n_{app} = \frac{L}{m} + h. \quad (23)$$

This predicts that if a second set of repeating steps occurs in addition to those involved in DNA unwinding, then a plot of n_{app} versus duplex length, L, will be linear, but exhibit a positive y-intercept, h. In this case, the values of n_{app} are obtained by performing a nonlinear least-squares fit of the set of time courses using Scheme 1 (Eq. 4), constraining k_{obs} to be a global parameter (i.e., the same for all duplex lengths), whereas n_{app} is treated as a local parameter (i.e., different for each duplex length).

Second special case: $h = n$ (two repeating steps occur for every ''m'' basepairs unwound). Scheme 4b describes the case in which an additional step is linked to each unwinding step such that both steps are repeated an equal number of times during the unwinding time course.

$$HD \xrightarrow{k_U} \left(I_1 \xrightarrow{k_C} I_1^*\right) \xrightarrow{k_U} \left(I_2 \xrightarrow{k_C} I_2^*\right) \ldots \xrightarrow{k_U} \left(I_{n-1} \xrightarrow{k_C} ssDNA\right). \quad \text{(Scheme 4b)}$$

This might describe a mechanism in which each unwinding step is accompanied by a conformational change occurring with rate constant k_C.

The Laplace transform of $f_{ss}(t)$ for Scheme 4b is given by Eq. 24:

$$F_{ss}(s) = \frac{k_C^n k_U^n}{s(k_C+s)^n (k_U+s)^n} = \frac{k_C^{L/m} k_U^{L/m}}{s(k_C+s)^{L/m} (k_U+s)^{L/m}}. \quad (24)$$

For this case, it can be easily shown that by taking the limit as k_C approaches infinity, only the slowest step in each catalytic cycle will be observed (Eq. 24):

$$\lim_{k_C \to \infty} \frac{k_C^n k_U^n}{s(k_C+s)^n (k_U+s)^n} = \frac{k_U^n}{s(k_U+s)^n}. \quad (25)$$

This is an important result demonstrating that regardless of which repeating step occurs as the slowest step within each unwinding cycle, it will still report on the average number of basepairs unwound between two rate-limiting steps.

However, if $k_U = k_C$ in Eq. 25, then a determination of the true step size becomes more difficult. This can be shown by taking the limit of Eq. 24 as k_U approaches k_C yielding Eq. 26, with n_{app} defined in Eq. 27:

$$\lim_{k_U \to k_C} \frac{k_C^n k_U^n}{s(k_C+s)^n (k_U+s)^n} = \frac{k_C^{2n}}{s(k_C+s)^{2n}} = \frac{k_{obs}^{n_{app}}}{s(k_{obs}+s)^{n_{app}}} \quad (26)$$

$$n_{app} = 2n = L\left(\frac{2}{m}\right) = L\left(\frac{1}{\frac{1}{2}m}\right). \quad (27)$$

Therefore, if one were to fit time courses to Scheme 1 (Eq. 4) that were actually generated by a mechanism that has two repeating steps with equal rate constants associated with the unwinding of m bp, this would result in a twofold underestimate of the true step size, even though the time courses would be well described by this model. One way to potentially uncover such a situation would be to perform a series of unwinding experiments under a variety of solution conditions in the hope that the two repeating rate constants would be affected differentially and eventually become resolvable. Since one does not know a priori that a situation such as Scheme 4b might exist, it is always advisable to perform additional experiments as a function of solution conditions, such as salt, temperature, pH, [ATP], etc., to determine whether the apparent step size is affected by solution conditions.

As we have discussed, even though multiple elementary kinetic steps will be associated with the unwinding of each ''m'' basepairs, the analysis of the time course will generally detect only the slowest step that is repeated in the unwinding cycle. Furthermore, one does not know a priori what physical process is associated with this repeated rate-limiting step. In fact, although we have used the designation k_U, for unwinding, to denote this repeated step, one needs to keep in mind that the step could reflect other processes such as ATP binding, hydrolysis, product release, or a protein conformational change, rather than the actual step associated with DNA unwinding.

Third special case: Both n and h are directly proportional to L, but with different proportionality constants. A separate translocation step is repeated with a different frequency than the unwinding steps. The third special case of Scheme 4 occurs if the two different steps are repeated during the course of DNA unwinding, but with different frequencies

(e.g., two steps with different step sizes) as depicted in Scheme 4c:

$$HD \xrightarrow{k_U} I_1 \xrightarrow{k_U} I_2 \xrightarrow{k_U} I_3 \xrightarrow{k_T} I_3^* \xrightarrow{k_U} I_4 \xrightarrow{k_U} I_5 \xrightarrow{k_U} I_6 \xrightarrow{k_T} \ldots ssDNA. \quad \text{(Scheme 4c)}$$

In this case, one step with rate constant k_T (and step size m_2) occurs for every three steps with rate constant k_U (and step size m). The Laplace transform of $f_{ss}(t)$ for Scheme 4c is given in Eq. 28, where $n_2 = L/m_2$:

$$F_{ss}(s) = \frac{k_T^{n_2}\, k_U^{n}}{s(k_T + s)^{n_2}(k_U + s)^{n}} = \frac{k_T^{L/m_2}\, k_U^{L/m}}{s(k_T + s)^{L/m_2}(k_U + s)^{L/m}}. \quad (28)$$

This type of mechanism might reflect a case in which m basepairs are unwound for each of the n steps, followed by a translocation step with rate constant k_T, associated with a larger step size, m_2. The quantum inchworm model proposed for RecBC-catalyzed DNA unwinding (Bianco and Kowalczykowski, 2000) is an example of such a mechanism. The ability to detect either the smaller or the larger of the two step sizes depends strongly on both the difference between the two step sizes, m and m_2, and the relative magnitudes of the two rate constants, k_U and k_T. As pointed out above, unless the two rate constants are similar in magnitude, one will only be able to observe the slowest step in the repeating cycle. Therefore, if the translocation step is rate limiting, one will observe the step size associated with translocation and vice versa.

As a starting point, we can explore the behavior of n_{app} as a function of duplex length, L, when $k_U = k_T$. If $k_U = k_T$, then Eq. 28 reduces to Eq. 29, with n_{app}, defined by Eq. 30:

$$\frac{k_{obs}^{((L/m)+(L/m_2))}}{s(k_{obs} + s)^{((L/m)+(L/m_2))}} \quad (29)$$

$$n_{app} = L\left(\frac{1}{m} + \frac{1}{m_2}\right). \quad (30)$$

If $m_2 = m$, then a twofold underestimate of the unwinding step size will be obtained as discussed above (see Eq. 27). However, if $m_2 \gg m$, the time course will provide information only on m, the smaller of the two step sizes, as can be seen from Eq. 30. This analysis suggests that if k_U and k_T are equal, one will observe the smaller of the two step sizes; however, if one of the two rate constants is partially rate limiting, one will primarily observe the step size associated with the rate constant that is partially rate limiting. Again, since there is never a priori knowledge that a situation such as this might exist, it is advisable to examine whether n_{app} changes with solution conditions.

GENERAL METHODS FOR GLOBAL NONLINEAR LEAST-SQUARES ANALYSIS OF DNA UNWINDING TIME COURSES

Expressions for the time course of DNA unwinding that are continuous functions of *n* and *m*

To obtain the "best fit" values of the kinetic parameters, as well as n and m and their associated uncertainties, for a given kinetic mechanism it is necessary to use nonlinear least-squares methods to analyze the single turnover DNA unwinding time courses. It is certainly the case that numerical integration techniques can be used for NLLS fitting for complex kinetic models in which the number of steps is defined and fixed (Barshop et al., 1983). However, NLLS analysis of either a single DNA unwinding time course, or the simultaneous (global) analysis of a series of time courses obtained for a set of DNA substrates differing in duplex length, L, is not possible with the numerical integration packages that are currently available, since one is unable to float either the number of steps, n, or the kinetic step size, $m = L/n$. This limitation results from the fact that the currently available programs require one to include every intermediate explicitly in the "kinetic mechanism", and thus these programs can only consider one particular scheme with a specified number of steps, n. Therefore, for the "n-step" sequential mechanisms considered here, n must be constrained to be an integer in those programs. However, n is one of the parameters to be determined from a quantitative analysis of the time course, and is thus not generally known a priori. This problem is compounded when attempting to analyze a set of time courses for DNA duplexes of different lengths, L, since n will differ for each length, L. Therefore, in previous approaches, one attempted to fit the time course to a series of schemes involving progressively increasing numbers of steps, n, rather than allowing the number of steps to be determined by the NLLS analysis (i.e., allowing n to "float"). This is why early analyses of DNA unwinding time courses (Ali and Lohman, 1997) were performed by a "bootstrapping" method in which n was constrained to be an integer, followed by a NLLS analysis for that value of n, then n was changed to another integer value, and the NLLS process was repeated. The results of analyses of a series of experiments performed with several duplex lengths were then compared to find a set of values of n, which was consistent for the different duplex lengths (i.e., for which a constant step size, $m = L/n$, was achieved). For these reasons, expressions for $f_{ss}(t)$ that are continuous functions of n (and thus m) are required for rigorous NLLS analysis.

There are additional reasons for considering noninteger values of the number of steps, n, as well as the "step size", m. First, it is conceivable that the number of basepairs unwound per step is not an absolute constant, but varies within some range. Thus, the average DNA unwinding step size determined from studies of a series of DNA duplexes

might be expected to result in a noninteger value of m. A noninteger average step size could also arise if the DNA duplex lengths, L, used in the experimental determination of the kinetic step size are not all integer products of the step size (which of course is impossible to know a priori). Finally, an average noninteger step size is expected simply as a result of experimental uncertainty. Therefore, the ability to account for a noninteger number of steps simplifies any NLLS analysis, and allows one to consider more carefully the possible stochastic nature of the kinetic step size and its experimental uncertainties.

The analytic solutions for $f_{ss}(t)$ discussed above and previously (Ali and Lohman, 1997; Lucius et al., 2002), which are expressed as a series multiplied by an exponential, are limited in use to integer values of n. However, as shown above and in Appendix A, the general form of the solution for $f_{ss}(t)$ that is obtained using Laplace transform methods yields a ratio of gamma functions, which is a continuous function of n. For example, Eqs. 4, 10, and 16, which are the general solutions for Schemes 1, 2, and 3, respectively, are continuous functions of the number of steps, n, and therefore can be easily implemented for use in a NLLS fitting routine (Lucius et al., 2002).

As shown in Appendix A, it is relatively simple to obtain analytic expressions for the Laplace transform of $f_{ss}(t)$ (i.e., $F_{ss}(s)$), even for Schemes for which an analytic solution for $f_{ss}(t)$ is not obtainable. In fact, since one can obtain $f_{ss}(t)$ by obtaining the numerical value for the inverse Laplace transform of $F_{ss}(s)$, it is not even necessary to have the analytic solution for $f_{ss}(t)$. One can directly perform NLLS analyses on a time course or a series of time courses using Eq. 31:

$$f_{ss}(t) = A_T \mathcal{L}^{-1}(F_{ss}(s)). \tag{31}$$

In fact, by obtaining the inverse Laplace transform of $F_{ss}(s)$ in numerical form, Lucius et al. (2002) were able to perform NLLS analysis using complex models such as Scheme 4 (Eqs. 18 and 31), even though an analytic solution in the time domain could not be obtained that was continuous in both n and h. Using numerical solutions to the inverse Laplace transform, one can float both n and h in a NLLS fitting algorithm for a set of time courses for duplexes of different lengths. This is currently the preferred method for analysis of experimental data in our laboratory.

Methods for determining the kinetic step size, *m*

Single-turnover studies of helicase-catalyzed DNA unwinding can be used to estimate the ''kinetic step size'', m, i.e., the number of basepairs unwound between two successive rate limiting steps that are repeated in the unwinding cycle (Ali and Lohman, 1997; Jankowsky et al., 2000; Lucius et al., 2002). The expectation is that the number of ''steps'', n, that are repeated during unwinding should be directly proportional to duplex length, L, such that $n = L/m$. However, one should not simply perform an unwinding experiment with a DNA substrate of one duplex length and estimate m by dividing the duplex length, L, by the number of apparent steps, n_{app}, required to fit the time course to a defined mechanism. One must consider whether all of the steps identified occur within each repeated cycle of unwinding or whether some steps might either be distinct from the unwinding cycle (e.g., slow initiation steps, etc.), or be coupled to each unwinding step, although not involved directly in DNA unwinding (e.g., conformational changes that occur after each unwinding step). As discussed above, the presence of such additional steps, not involved in unwinding, would yield an underestimate of the true step size. It is also not recommended to estimate m based on experiments performed with one duplex length since this would yield unacceptably high uncertainties. The reader is referred to a detailed analysis of single turnover DNA unwinding time courses catalyzed by the *E. coli* RecBCD helicase (Lucius et al., 2002) using the techniques described here, in which it was concluded that two additional steps occur in the mechanism that are not involved in DNA unwinding (see Scheme 4).

APPENDIX A: LAPLACE TRANSFORM METHODS FOR OBTAINING SOLUTIONS FOR THE TIME DEPENDENCE OF FORMATION OF FULLY UNWOUND DNA ($F_{SS}(T)$) FOR A SINGLE-TURNOVER HELICASE-CATALYZED DNA UNWINDING REACTION OCCURRING BY "*N*-STEP" SEQUENTIAL MECHANISMS

Here we present the details for deriving the expressions for the time dependence of formation of fully unwound single stranded DNA, $f_{ss}(t)$, for single-turnover helicase-catalyzed DNA unwinding reactions. Since all of the ''n-step'' sequential mechanisms that we discuss in the text can be described by a series of linear, first order differential equations, we can use Laplace transform methods to solve for expressions for the time dependence of formation and dissipation of the various intermediates in each mechanism. In fact, this method can only be used for solutions of a series of linear first order differential equations. For more details, the reader can consult any general mathematics text (e.g., Kreyszig (1993)). Much of the mathematical manipulation presented here can be facilitated through use of any mathematical software package that allows symbolic manipulation of algebraic equations and matrices; however, we have generally used Mathematica (Wolfram Research, Champaign, IL) for this purpose.

One starts by writing the system of differential equations that describe the particular scheme under consideration. For example, the differential equations describing the time dependent formation and dissipation of the species in Scheme 1 are given by Eqs. A1–A6 for n number of steps:

$$\frac{d[HD]}{dt} = -[HD]k_U \tag{A1}$$

$$\frac{d[I_1]}{dt} = [HD]k_U - [I_1]k_U \tag{A2}$$

$$\frac{d[I_2]}{dt} = [I_1]k_U - [I_2]k_U \tag{A3}$$

$$\frac{d[I_j]}{dt} = [I_{j-1}]k_U - [I_j]k_U \tag{A4}$$

$$\frac{d[I_{n-1}]}{dt} = [I_{n-2}]k_U - [I_{n-1}]k_U \tag{A5}$$

$$\frac{d[ssDNA]}{dt} = [I_{n-1}]k_U. \tag{A6}$$

Application of the Laplace transform to this series of coupled differential equations results in a series of coupled algebraic equations, which can then be handled much more readily. The Laplace transform, $F(s)$, of a function, $f(t)$, is defined in Eq. A7,

$$F(s) = \mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}dt, \tag{A7}$$

and the inverse Laplace transform of $F(s)$ yields $f(t)$, as defined in Eq. A8,

$$f(t) = \mathcal{L}^{-1}[F(s)] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} F(s)e^{st}ds, \tag{A8}$$

where $i = \sqrt{-1}$, and c is chosen so that all of the singular points of $F(s)$ lie to the left of the line $\mathrm{Re}\{s\} = c$ in the complex s plane. Eqs. A7 and A8 are the strict definitions of the Laplace and the inverse Laplace transform, respectively, however, as will become clear below, one does not typically perform these integrations to solve the problem. From Eq. A7 it follows that

$$\mathcal{L}\left[\frac{df(t)}{dt}\right] = sF(s) - f(0), \tag{A9}$$

where $f(0)$ is $f(t)$at $t = 0$. It also follows from Eq. A7 that the inverse Laplace transform of the product of two Laplace transforms, $F(s)$ and $G(s)$, is equal to the convolution of the corresponding time-dependent functions, $f(t)$ and $g(t)$, as shown in Eq. A10:

$$\mathcal{L}^{-1}[F(s)G(s)] = \int_0^t f(u)g(t-u)\,du. \tag{A10}$$

Taking the Laplace transforms of Eqs. A1–A6 yields Eqs. A11–A16:

$$(s + k_U)HD'(s) = HD_0 \tag{A11}$$

$$-k_U HD'(s) + (s + k_U)I_1'(s) = 0 \tag{A12}$$

$$-k_U I_1'(s) + (s + k_U)I_2'(s) = 0 \tag{A13}$$

$$-k_U I_{j-1}'(s) + (s + k_U)I_j'(s) = 0 \tag{A14}$$

$$-k_U I_{n-2}'(s) + (s + k_U)I_{n-1}'(s) = 0 \tag{A15}$$

$$-k_U I_{n-1}'(s) + s(ssDNA'(s)) = 0, \tag{A16}$$

where $HD'(s)$, I_j'(s), and $ssDNA'$(s) represent the concentrations of substrate-enzyme complex, intermediate j, and product, respectively, as functions of the complex variable s. HD_0 is the initial concentration of the HD complex at $t = 0$, and we have assumed that the initial concentrations of all intermediates and the final $ssDNA$ product are zero.

Equations A11–A16 can be written in matrix form as in Eq. A17:

$$\begin{pmatrix} s+k_U & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ -k_U & s+k_U & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & -k_U & s+k_U & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -k_U & s+k_U & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -k_U & s+k_U & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k_U & s \end{pmatrix} \cdot \begin{pmatrix} HD'(s) \\ I_1'(s) \\ I_2'(s) \\ \vdots \\ I_j'(s) \\ \vdots \\ I_n'(s) \\ ssDNA'(s) \end{pmatrix} = \begin{pmatrix} HD_0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}. \tag{A17}$$

Equation A17 can then be row reduced to yield Eq. A18, where each term in the last column is the Laplace transform for the concentration of each species in Scheme 1 (see Eqs. A19–A24):

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} HD'(s) \\ I_1'(s) \\ I_2'(s) \\ \vdots \\ I_j'(s) \\ \vdots \\ I_n'(s) \\ ssDNA'(s) \end{pmatrix} = \begin{pmatrix} \frac{HD_0}{k_U+s} \\ \frac{k_U HD_0}{(k_U+s)^2} \\ \frac{k_U^2 HD_0}{(k_U+s)^3} \\ \vdots \\ \frac{k_U^j HD_0}{(k_U+s)^{j+1}} \\ \vdots \\ \frac{k_U^{n-1} HD_0}{(k_U+s)^n} \\ \frac{k_U^n HD_0}{s(k_U+s)^n} \end{pmatrix} \tag{A18}$$

$$HD'(s) = \frac{HD_0}{k_U + s} \tag{A19}$$

$$I_1'(s) = \frac{k_U HD_0}{(k_U + s)^2} \tag{A20}$$

$$I_2'(s) = \frac{k_U^2 HD_0}{(k_U + s)^3} \tag{A21}$$

$$I_j'(s) = \frac{k_U^j HD_0}{(k_U + s)^{j+1}} \tag{A22}$$

$$I_{n-1}'(s) = \frac{k_U^{n-1} HD_0}{(k_U + s)^n} \tag{A23}$$

$$ssDNA'(s) = \frac{k_U^n HD_0}{s(k_U + s)^n}. \tag{A24}$$

The time-dependent expressions for the concentrations of the initial enzyme-substrate complex, the final ssDNA product, and each intermediate are then obtained by taking the inverse Laplace transforms of Eqs. A19–A24. This can be accomplished using Eq. A8, or by consulting a table of known Laplace transforms or by using a mathematical software package such as Mathematica.

For example, Eq. A24 can be separated into two parts:

$$ssDNA'(s) = \frac{k_U^n HD_0}{s(k_U + s)^n} = \left(\frac{1}{s}\right)\left(\frac{k_U^n HD_0}{(k_U + s)^n}\right). \tag{A25}$$

The inverse Laplace transform of each part can be found in tables of Laplace transforms,

$$\mathcal{L}^{-1}\left[\frac{b}{(s-a)^n}\right] = b\frac{t^{n-1}e^{a\cdot t}}{\Gamma(n)}, \tag{A26}$$

where the gamma function, $\Gamma(n)$, is defined in Eq. A27:

$$\Gamma(n) = \int_0^\infty r^{n-1}e^{-r}dr. \tag{A27}$$

For integer values of n, the gamma function can be defined as in Eq. A28,

$$\Gamma(n) = (n-1)! \tag{A28}$$

and

$$\mathcal{L}^{-1}\left[\frac{1}{s}\right] = 1. \tag{A29}$$

Through application of Eq. A10, we obtain Eq. A30:

$$\mathcal{L}^{-1}[ssDNA'(s)] = \int_0^t k_U^n HD_0 \frac{u^{n-1}e^{-k_U u}}{\Gamma(n)}\,du. \tag{A30}$$

The integral in Eq. A30 can be rewritten as in Eq. A31,

$$\int_0^t u^{n-1}e^{-k_U u}\,du = \left(\frac{1}{k_U}\right)^n \int_0^{k_U\cdot t} z^{n-1}e^{-z}\,dz, \tag{A31}$$

and then further simplified as in Eq. A32,

$$\left(\frac{1}{k_U}\right)^n \int_0^{k_U\cdot t} z^{n-1}e^{-z}\,dz = \left(\frac{1}{k_U}\right)^n \left\{\int_0^\infty z^{n-1}e^{-z}\,dz - \int_{k_U\cdot t}^\infty z^{n-1}e^{-z}\,dz\right\}$$
$$= \left(\frac{1}{k_U}\right)^n \{\Gamma(n) - \Gamma(n, k_U t)\}, \tag{A32}$$

where $\Gamma(n,z)$ is the incomplete gamma function, defined in Eq. A33:

$$\Gamma(n,z) = \int_z^\infty r^{n-1}e^{-r}dr. \tag{A33}$$

Thus, the final expression for [$ssDNA(t)$] is given in Eq. A34:

$$ssDNA(t) = \frac{k_U^n HD_0}{\Gamma(n)}\left(\frac{1}{k_U}\right)^n \{\Gamma(n) - \Gamma(n, k_U t)\}. \tag{A34}$$

Upon rearranging, one obtains Eq. A35:

$$\frac{ssDNA(t)}{HD_0} = 1 - \frac{\Gamma(n, k_U t)}{\Gamma(n)}. \tag{A35}$$

If n is an integer, the ratio of the gamma functions in Eq. A35 can be replaced with the series shown in Eq. A36,

$$\frac{\Gamma(n,z)}{\Gamma(n)} = \sum_{r=1}^{n} \frac{(z)^{r-1}}{(r-1)!}e^{-z}, \tag{A36}$$

so that Eq. A35 can be rewritten as Eq. A37:

$$\frac{ssDNA(t)}{HD_0} = 1 - \sum_{r=1}^{n} \frac{(k_U t)^{r-1}}{(r-1)!}e^{-k_U t}. \tag{A37}$$

Using the same technique, the analytic solution for the time-dependent loss of the helicase-DNA complex and the formation and disappearance of the jth intermediate are given in Eqs. A38 and A39, respectively:

$$\frac{HD(t)}{HD_0} = \mathcal{L}^{-1}\left(\frac{1}{k_U + s}\right) = e^{-k_U t} \tag{A38}$$

$$\frac{I_j(t)}{HD_0} = \mathcal{L}^{-1}\left(\frac{k_U^j}{(k_U + s)^{j+1}}\right) = \frac{(k_U t)^j}{\Gamma(j+1)}e^{-k_U t}. \tag{A39}$$

For Scheme 2, the Laplace transforms and inverse Laplace transforms for each species are given in Eqs. A40–A42:

$$\frac{HD(t)}{HD_0} = \mathcal{L}^{-1}\left(\frac{1}{k_U + k_d + s}\right) = e^{-(k_U + k_d)t} \tag{A40}$$

$$\frac{I_j(t)}{HD_0} = \mathcal{L}^{-1}\left(\frac{k_U^j}{(k_U + k_d + s)^{j+1}}\right) = \frac{(k_U t)^j}{\Gamma(j+1)}e^{-(k_U + k_d)t} \tag{A41}$$

$$\frac{ssDNA(t)}{HD_0} = \mathcal{L}^{-1}\left(\frac{k_U^n}{s(k_U + k_d + s)^n}\right)$$
$$= \left(\frac{k_U}{k_U + k_d}\right)^n \left(1 - \frac{\Gamma(n, (k_U + k_d)t)}{\Gamma(n)}\right). \tag{A42}$$

For Scheme 3, the time-dependent expressions for the concentrations of each species are given in Eqs. A43–A46:

$$\frac{HD(t)}{HD_0} = \mathcal{L}^{-1}\frac{(1-x)}{(k_{NP} + s)} = (1-x)e^{-k_{NP}t} \tag{A43}$$

$$\frac{HD^*(t)}{HD_0} = \mathcal{L}^{-1}\frac{(k_{NP}+sx)}{(k_{NP}+s)(k_U+k_d+s)} = \frac{((k_U+k_d)x-k_{NP})}{k_U+k_d-k_{NP}}e^{-(k_U+k_d)t} + (1-x)\frac{k_{NP}}{k_U+k_d-k_{NP}}e^{-k_{NP}t} \quad \text{(A44)}$$

$$\frac{I_j(t)}{HD_0} = \mathcal{L}^{-1}\frac{k_U^j(k_{NP}+sx)}{(k_{NP}+s)(k_U+k_d+s)^{j+1}} = \frac{x(k_Ut)^j}{\Gamma(j+1)}e^{-(k_U+k_d)t} + (1-x)\left(\frac{k_{NP}}{k_U+k_d-k_{NP}}\right)\left(\frac{k_U}{k_U+k_d-k_{NP}}\right)^j e^{-k_{NP}t} \times\left(1-\frac{\Gamma(j+1,(k_U+k_d-k_{NP})t)}{\Gamma(j+1)}\right) \quad \text{(A45)}$$

$$\frac{ssDNA(t)}{HD_0} = \mathcal{L}^{-1}\frac{k_U^n(k_{NP}+sx)}{s(k_{NP}+s)(k_U+k_d+s)^n} = \left(\frac{k_U}{k_U+k_d}\right)^n\left[\left(1-\frac{\Gamma(n,(k_U+k_d)t)}{\Gamma(n)}\right) - e^{-k_{NP}t}(1-x)\left(\frac{k_U+k_d}{k_U+k_d-k_{NP}}\right)^n \times\left(1-\frac{\Gamma(n,(k_U+k_d-k_{NP})t)}{\Gamma(n)}\right)\right]. \quad \text{(A46)}$$

APPENDIX B: PROCESSIVITY OF DNA UNWINDING

The processivity, P, has been defined in terms of k_U, k_d (Eq. 6) as well as the average number of steps taken, N, (Eq. 6) (McClure and Chow, 1980). Derivations for these relationships have not been reported, thus we present them here.

The average value, $\langle x\rangle$, of any probability distribution function, $f(x)$, is given by Eq. B1:

$$\langle x\rangle = \frac{\int_{-\infty}^{\infty} xf(x)\,dx}{\int_{-\infty}^{\infty} f(x)\,dx}. \quad \text{(B1)}$$

Therefore, the average number of steps taken, $\langle N_{stp}\rangle$, is defined by Eq. B2. The lower limit of integration in Eq. B2 is zero since we consider only forward motion Eq. B2:

$$\langle N_{stp}\rangle = \frac{\int_0^{\infty} nP^n\,dn}{\int_0^{\infty} P^n\,dn}. \quad \text{(B2)}$$

The integration of Eq. B2 results in Eq. B3,

$$\langle N_{stp}\rangle = -\frac{1}{\ln(P)}, \quad \text{(B3)}$$

which upon rearrangement yields Eq. B4,

$$P = e^{-1/\langle N_{stp}\rangle}. \quad \text{(B4)}$$

Equation B4 can be expanded in a Taylor series to yield Eq. B5:

$$P = e^{-(1/\langle N_{stp}\rangle)} = \sum_{r=1}^{\infty}\frac{(-(1/\langle N_{stp}\rangle))^{r-1}}{(r-1)!} = 1-\frac{1}{\langle N_{stp}\rangle}+\frac{1^2}{2\langle N_{stp}\rangle^2}-\ldots \quad \text{(B5)}$$

In the limit of $(\langle N_{stp}\rangle) \gg 1$, this series can be approximated by only the first two terms

$$P \cong \frac{\langle N_{stp}\rangle - 1}{\langle N_{stp}\rangle}. \quad \text{(B6)}$$

To determine the average number of basepairs unwound, $\langle N_{bp}\rangle$, a similar integration can be performed by replacing n with L/m as in Eq. B7, if there are m basepairs unwound per step.

$$\langle N_{bp}\rangle = \frac{\int_0^{\infty}\frac{L}{m}P^{L/m}\,dL}{\int_0^{\infty}P^{L/m}\,dL} = -\frac{m}{\ln(P)}. \quad \text{(B7)}$$

Rearranging Eq. B7 and expanding the result in a series yields Eq. B8,

$$P \cong \frac{\langle N_{bp}\rangle - m}{\langle N_{bp}\rangle}, \quad \text{(B8)}$$

where again, the approximation is valid only when $(\langle N_{bp}\rangle) \gg m$.

APPENDIX C: NONLINEAR LEAST-SQUARES METHODS

We have used the program Conlin (Williams and Hall, 2000) for all NLLS analysis. The routine in Conlin used for the numerical inversion of the Laplace transform was purchased from Visual Numerics (Houston, TX) and is contained within the IMSL C Numerical Libraries. The routine is based on a modification of Weeks' method (Weeks, 1966) due to Garbow et al. (1988). To our knowledge, the only other commercially available package that implements the numerical inversion of the Laplace transform for the direct fitting of data is Scientist (MicroMath, St. Louis, MO). Mathematica can be used for determining the numerical inverse Laplace transforms; however, it does not contain a global NLLS analysis package. The gamma function representation of the analytic solution was implemented in Conlin using the routines for the gamma function and incomplete gamma function published in Numerical Recipes in C (Press, 1995).

The symbolic algebraic manipulations as well as the matrix manipulations described in Appendix A were performed using Mathematica. The time-dependent solutions were also found using the ''InverseLaplaceTransform'' command in Mathematica.

To determine if inversion of the Laplace transform by numerical methods could be used in our analysis, a set of data points were simulated using the analytic solution to Scheme 1 (Eq. 2). Data points were simulated for $n = 5$ steps, and $k_U = 50\ s^{-1}$. The simulated data points were fit to both the analytic solution as well as the numerical inversion of the Laplace transform. In the case of the fit to the analytic solution, the number of steps was set to $n = 5$ and the series in Eq. 3 was expanded, so that the only fitting parameter was the unwinding rate constant k_U. In the case of the fit to the numerical inversion, both the rate constant k_U and the number of steps were floated as fitting parameters. The solid line in Fig. 6 is the result from the fitting of the data to the numerical inverse Laplace transform, which yields a rate constant of $50\ s^{-1}$ and a number of steps $n = 5$. The difference between the sum of the squared residuals for the fit to the analytic solution and the numerical inversion is 2.0×10^{-19}, leading us to conclude that the fit using the numerical routine is identical to the fit using the analytic solution. Furthermore, fitting of the data to the numerical inverse Laplace transform is significantly less cumbersome,

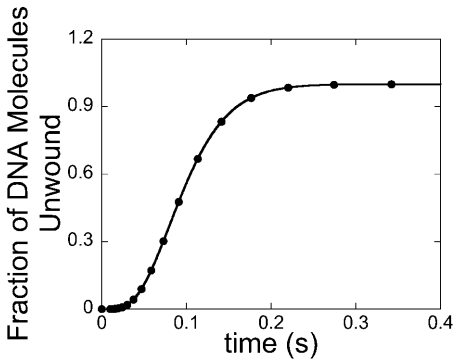

FIGURE 6 Simulation to determine if the numerical inversion of the Laplace transform can be used to perform nonlinear least-squares analysis. Data points (●) are simulated from Scheme 1 (Eq. 2) with $n = 5$ steps, and $k_U = 50\ s^{-1}$. The solid line is a nonlinear least-squares fit using the numerical inverse Laplace transform of Eq. 1. The best fit to the simulated data points predicts an unwinding rate constant $k_U = 50\ s^{-1}$, and the number of steps $n = 5$ by floating k_U and n as fitting parameters. The variance of the fit is 2.4×10^{-16}, which results in a percent error in the fit to be $1.6 \times 10^{-6}\%$.

given that finding the best model using the series solution results in having to fit the data several times to different expansions of the series until one finds the minimum in the sum of the squared residuals. Equally good agreement was obtained with the gamma function representations as expressed in Eq. 2.

All fitting models implemented in Conlin will be made available for download at http://biochem.wustl.edu/~lohman.

A. Lucius and N. K. Maluf received partial support from National Institutes of Health training grant T32 GM08492. C. Fischer received partial support from National Institutes of Health postdoctoral fellowship GM56105. This research was supported by National Institutes of Health grant GM45948.

REFERENCES

Ali, J. A., and T. M. Lohman. 1997. Kinetic measurement of the step size of DNA unwinding by Escherichia coli UvrD helicase. *Science*. 275:377–380.

Ali, J. A., N. K. Maluf, and T. M. Lohman. 1999. An oligomeric form of E. coli UvrD is required for optimal helicase activity. *J. Mol. Biol.* 293:815–833.

Arai, N., K. I. Arai, and A. Kornberg. 1981. Complexes of *Rep* protein with ATP and DNA as a basis for helicase action. *J. Biol. Chem.* 256:5287–5293.

Barshop, B. A., R. F. Wrenn, and C. Frieden. 1983. Analysis of numerical methods for computer simulation of kinetic processes: development of KINSIM—a flexible, portable system. *Anal. Biochem.* 130:134–145.

Bianco, P. R., L. R. Brewer, M. Corzett, R. Balhorn, Y. Yeh, S. C. Kowalczykowski, and R. J. Baskin. 2001. Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature*. 409:374–378.

Bianco, P. R., and S. C. Kowalczykowski. 2000. Translocation step size and mechanism of the RecBC DNA helicase. *Nature*. 405:368–372.

Bjornson, K. P., M. Amaratunga, K. J. M. Moore, and T. M. Lohman. 1994. Single-turnover kinetics of helicase-catalyzed DNA unwinding monitored continuously by fluorescence energy transfer. *Biochemistry*. 33:14306–14316.

Block, S. M., L. S. Goldstein, and B. J. Schnapp. 1990. Bead movement by single kinesin molecules studied with optical tweezers. *Nature*. 348:348–352.

Bujalowski, W., and M. J. Jezewska. 2000. Kinetic mechanism of the single-stranded DNA recognition by Escherichia coli replicative helicase DnaB protein. Application of the matrix projection operator technique to analyze stopped-flow kinetics. *J. Mol. Biol.* 295:831–852.

Capellos, C., and B. H. J. Bielski. 1972. Kinetic Systems: Mathematical Description of Chemical Kinetics in Solution. Wiley InterScience, New York.

Cheng, W., J. Hsieh, K. M. Brendza, and T. M. Lohman. 2001. E. coli Rep oligomers are required to initiate DNA unwinding in vitro. *J. Mol. Biol.* 310:327–350.

Dillingham, M. S., D. B. Wigley, and M. R. Webb. 2002. Direct measurement of single-stranded DNA translocation by PcrA helicase using the fluorescent base analogue 2-aminopurine. *Biochemistry*. 41:643–651.

Dohoney, K. M., and J. Gelles. 2001. χ-sequence recognition and DNA translocation by single RecBCD helicase/nuclease molecules. *Nature*. 409:370–374.

Eggleston, A. K., N. A. Rahim, and S. C. Kowalczykowski. 1996. A helicase assay based on the displacement of fluorescent, nucleic acid-binding ligands. *Nucleic Acids Res*. 24:1179–1186.

Garbow, B. S., G. Giunta, and J. N. Lyness. 1988. Software for an implementation of Weeks' method for the inverse Laplace transform problem. *ACM Trans. Math. Soft.* 14:163–170.

Gutfreund, H. 1995. Kinetics for the Life Sciences. Receptors, Transmitters and Catalysts. Cambridge University Press, Cambridge, England.

Ha, T., I. Rasnik, W. Cheng, H. P. Babcock, G. H. Gauss, T. M. Lohman, and S. Chu. 2002. Initiation and re-initiation of DNA unwinding by the Escherichia coli Rep helicase. *Nature*. 419:638–641.

Houston, P., and T. Kodadek. 1994. Spectrophotometric assay for enzyme-mediated unwinding of double-stranded DNA. *Proc. Natl. Acad. Sci. USA*. 91:5471–5474.

Howard, J., A. J. Hudspeth, and R. D. Vale. 1989. Movement of microtubules by single kinesin molecules. *Nature*. 342:154–158.

Jankowsky, E., C. H. Gross, S. Shuman, and A. M. Pyle. 2000. The DExH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature*. 403:447–451.

Kornberg, A., and T. A. Baker. 1992. DNA Replication. W.H. Freeman & Co., New York.

Kovall, R., and B. W. Matthews. 1997. Toroidal structure of lambda-exonuclease. *Science*. 277:1824–1827.

Kreyszig, E. 1993. Advanced Engineering Mathematics. John Wiley & Sons.

Lohman, T. M., and K. P. Bjornson. 1996. Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* 65:169–214.

Lucius, A. L., A. Vindigni, R. Gregorian, J. A. Ali, A. F. Taylor, G. R. Smith, and T. M. Lohman. 2002. DNA unwinding step-size of E. coli RecBCD helicase determined from single turnover chemical quenched-flow kinetic studies. *J. Mol. Biol.* 324:409–428.

Maluf, N. K., C. J. Fischer, and T. M. Lohman. 2003. A dimer of Escherichia coli UvrD is the active form of the helicase in vitro. *J. Mol. Biol.* 325:913–935.

Matson, S. W., and K. A. Kaiser-Rogers. 1990. DNA helicases. *Annu. Rev. Biochem.* 59:289–329.

McClure, W. R., and Y. Chow. 1980. The kinetics and processivity of nucleic acid polymerases. *Methods Enzymol.* 64:277–297.

Patel, S. S., and K. M. Picha. 2000. Structure and function of hexameric helicases. *Annu. Rev. Biochem.* 69:651–697.

Press, W.H. 1995. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, New York.

Raney, K. D., and S. J. Benkovic. 1995. Bacteriophage T4 Dda helicase translocates in a unidirectional fashion on single-stranded DNA. *J. Biol. Chem.* 270:22236–22242.

Rodiguin, N. M., and E. N. Rodiguina. 1964. Consecutive Chemical Reactions: Mathematical Analysis and Development. D. Van Nostrand, Princeton, NJ.

Roman, L. J., and S. C. Kowalczykowski. 1989. Characterization of the helicase activity of the *Escherichia coli* RecBCD enzyme using a novel helicase assay. *Biochemistry*. 28:2863–2873.

Schnitzer, M. J., and S. M. Block. 1995. Statistical kinetics of processive enzymes. *Cold Spring Harb. Symp. Quant. Biol.* 60:793–802.

Soultanas, P., and D. B. Wigley. 2000. DNA helicases: ''inching forward''. *Curr. Opin. Struct. Biol.* 10:124–128.

Szczelkun, M. D. 2002. Kinetic models of translocation, head-on collision, and DNA cleavage by type I restriction endonucleases. *Biochemistry*. 41:2067–2074.

Taylor, A., and G. R. Smith. 1980. Unwinding and rewinding of DNA by the RecBC enzyme. *Cell*. 22:447–457.

Vale, R. D., and R. J. Fletterick. 1997. The design plan of kinesin motors. *Annu. Rev. Cell Dev. Biol.* 13:745–777.

Visscher, K., M. J. Schnitzer, and S. M. Block. 1999. Single kinesin molecules studied with a molecular force clamp. *Nature*. 400:184–189.

von Hippel, P. H., and E. Delagoutte. 2001. A general model for nucleic acid helicases and their ''coupling'' within macromolecular machines. *Cell*. 104:177–190.

Weeks, W. T. 1966. Numerical inversion of Laplace transforms using Laquerre functions. *JACM*. 13:419–426.

Williams, D. J., and K. B. Hall. 2000. Monte Carlo applications to thermal and chemical denaturation experiments of nucleic acids and proteins. *Methods Enzymol.* 321:330–352.

Wuite, G. J., S. B. Smith, M. Young, D. Keller, and C. Bustamante. 2000. Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature*. 404:103–106.