doi:10.1016/S0022-2836(02)01067-7 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 324, 409–428

JMB

DNA Unwinding Step-size of *E. coli* RecBCD Helicase Determined from Single Turnover Chemical Quenched-flow Kinetic Studies

Aaron L. Lucius[1], Alessandro Vindigni[1], Razmic Gregorian[1] Janid A. Ali[1], Andrew F. Taylor[2], Gerald R. Smith[2] and Timothy M. Lohman[1]*

[1]*Department of Biochemistry and Molecular Biophysics Washington University School of Medicine, 660 S. Euclid Ave. Box 8231, St. Louis, MO 63110 USA*

[2]*Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. North, Seattle, WA 98109 USA*

**Corresponding author*

The mechanism by which *Escherichia coli* RecBCD DNA helicase unwinds duplex DNA was examined *in vitro* using pre-steady-state chemical quenched-flow kinetic methods. Single turnover DNA unwinding experiments were performed by addition of ATP to RecBCD that was pre-bound to a series of DNA substrates containing duplex DNA regions ranging from 24 bp to 60 bp. In each case, the time-course for formation of completely unwound DNA displayed a distinct lag phase that increased with duplex length, reflecting the transient formation of partially unwound DNA intermediates during unwinding catalyzed by RecBCD. Quantitative analysis of five independent sets of DNA unwinding time courses indicates that RecBCD unwinds duplex DNA in discrete steps, with an average unwinding "step-size", $m = 3.9(\pm 1.3)$ bp step^{-1}, with an average unwinding rate of $k_U = 196(\pm 77)$ steps s^{-1} ($mk_U = 790(\pm 23)$ bp s^{-1}) at 25.0 °C (10 mM $MgCl_2$, 30 mM NaCl (pH 7.0), 5% (v/v) glycerol). However, additional steps, not linked directly to DNA unwinding are also detected. This kinetic DNA unwinding step-size is similar to that determined for the *E. coli* UvrD helicase, suggesting that these two SF1 superfamily helicases may share similar mechanisms of DNA unwinding.

Keywords: helicase; motor protein; recombination; ATPase; kinetics; mechanism

Introduction

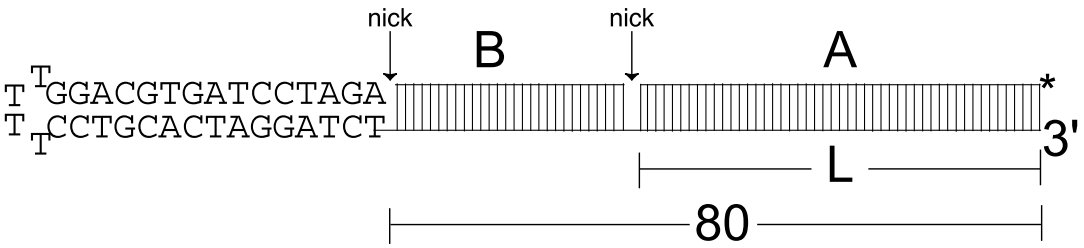
DNA helicases are motor proteins that use nucleoside triphosphate (NTP) binding and/or hydrolysis to unwind duplex DNA, thus forming the single-stranded (ss) DNA intermediates required for processes such as DNA replication, recombination and repair. In addition to unwinding duplex DNA, these enzymes must also translocate along the DNA filament.[1–4] Enzymes with helicase or putative helicase activity have been characterized based on amino acid sequence conservation and grouped into superfamilies, with the SF1 and SF2 superfamilies containing the largest numbers.[5] Many DNA helicases function as oligomeric enzymes, particularly dimers or hexamers,[1,3,6,7] although mechanisms have also been proposed for monomeric helicases.[8,9] High resolution crystal structures have been reported for two SF1 helicases bound to DNA, the *Bacillus stearothermophilus* PcrA[9] and *Escherichia coli* Rep proteins.[10]

The *E. coli* RecBCD enzyme is essential for genetic recombination and has multiple activities, including double-stranded (ds) and ssDNA exonuclease, ssDNA endonuclease, DNA-dependent ATPase, and helicase activities.[11,12] Through a combination of these activities RecBCD produces ssDNA onto which it loads RecA protein.[13] These

Present addresses: A. Vindigni, International Centre for Genetic Engineering and Biotechnology, Area Science Park, Padriciano 99, I-34012 Trieste, Italy; R. Gregorian, The Lewin Group, 9302 Lee Hwy, Suite 500, Fairfax, VA 22031, USA; J. A. Ali, Infinity Pharmaceuticals, Inc., 650 Albany Street, Boston, MA 02118, USA.

Abbreviations used: ss, single-stranded; ds, double-stranded; NLLS, non-linear least-squares; SSR, sum of the squared residuals.

E-mail address of the corresponding author: lohman@biochem.wustl.edu



	Length (L)	Sequences of Oligonucleotide "A"
I	24	5′-CCATGGCTCCTGAGCTAGCTGCAG
II	30	5′-CCATGGCTCCTGAGCTAGCTGCAGTAGCCT
III	40	5′-CCATGGCTCCTGAGCTAGCTGCAGTAGCCTAAAGGATGAA
IV	48	5′-CCATGGCTCCTGAGCTAGCTGCAGTAGCCTAAAGGATGAAACTAGGAT
V	60	5′-CCATGGCTCCTGAGCTAGCTGCAGTAGCCTAAAGGATGAAACTAGGATCTTATGCTCCAT
VI	80	5′-CCATGGCTCCTGAGCTAGCTGCAGTAGCCTAAAGGATGAAACTAGGATCTTATGCTCCATGGATACGTCGAGTCGCATCC

Figure 1. Schematic representation of the double nicked DNA substrates. Each DNA substrate is composed of three separate oligodeoxynucleotides, a constant 114 nucleotide bottom strand containing a hairpin and two top strands, A and B. The two nicks are indicated. Oligodeoxynucleotide A is radiolabeled at its 5′ end with ^{32}P (denoted by an asterisk). The sequences and lengths, *L*, in nucleotides, of the series of A oligodeoxynucleotides used to form the different DNA substrates are indicated. The oligodeoxynucleotide designated as VI is the full 80 nucleotide sequence of oligonucleotides A plus B. The complete sequence for the bottom strand including the hairpin is given in Materials and Methods.

RecA protein–ssDNA filaments are essential precursors to the formation of joint molecule intermediates in homologous recombination. The enzyme consists of three distinct polypeptides, RecB, RecC, and RecD, the masses of which sum to 330 kDa. Both the RecB and RecD polypeptides are members of the SF1 helicase superfamily,[5] although helicase activity has been demonstrated only for the isolated RecB polypeptide.[14] Although the RecD polypeptide is not required for helicase activity, the RecBC enzyme displays a slower rate of unwinding and a lower affinity for DNA,[15,16] suggesting a role for RecD in DNA unwinding by RecBCD (A.F.T. & G.R.S., unpublished data).

A variety of evidence indicates that RecBCD initiates unwinding of duplex DNA at a dsDNA end. The enzyme shows no detectable nuclease activity on circular dsDNA; potent nuclease activity is seen only with linear ds or ssDNA.[17] Genetic results indicate that an important intracellular substrate for RecBCD is linear dsDNA,[18] and the helicase activity of the purified enzyme has a strong preference for linear dsDNA with blunt or nearly blunt ends.[19,20] Analysis of partially unwound DNA by electron microscopy showed that the enzyme initiates unwinding at the end of dsDNA and can unwind >20 kb of DNA with high processivity at rates of ~ 350 bp s^{-1}.[21] Depending on the solution conditions, the rate of unwinding can approach 500 bp s^{-1} at 25 °C with sufficient processivity to unwind DNA with an average length of 30,000 bp.[22,23] More recently, single molecule experiments measured the rate of RecBCD unwinding to be $502(\pm 243)$ bp s^{-1} at 23 °C and $972(\pm 172)$ bp s^{-1} at 37 °C.[24] In the absence of ATP, which is required for processive DNA unwinding, the enzyme binds with high affinity to a dsDNA end, binding most tightly to those with short ssDNA extensions.[19,25] In this initiation complex the RecB polypeptide can undergo UV-induced crosslinking to the DNA strand with the 3′ end and the RecC and RecD polypeptides can be crosslinked to the DNA strand with the 5′ end.[26] Binding of RecBCD to the duplex end results in a perturbation of the conformation of ~ 5–6 bp in an ATP-independent, but Mg^{2+}-dependent reaction, as determined from chemical protection experiments[27] using $KMnO_4$.

DNA unwinding by a processive helicase must involve the cycling of the enzyme through a number of steps resulting in the unwinding of some number of base-pairs during each cycle, which we refer to as the unwinding step-size. On the basis of pre-steady-state single turnover kinetic studies of DNA unwinding, Ali & Lohman[28] determined that *E. coli* UvrD (helicase II), an SF1 helicase involved in methyl-directed mismatch repair[29] and nucleotide excision repair,[30,31] unwinds DNA with a kinetic step-size of $4(\pm 1)$ bp. Here we describe the use of such pre-steady-state chemical quenched-flow kinetic approaches to estimate the kinetic step-size for DNA unwinding by the *E. coli* RecBCD helicase.

Results

DNA substrate design

The DNA substrates used in our single turnover kinetic studies, referred to as "double nicked" substrates, were designed based on DNA substrates used by Taylor & Smith[25] and are shown schematically in Figure 1. These linear DNA substrates contain a DNA hairpin with a T_4 loop at one end and a blunt-ended duplex at the other. This design was used so that RecBCD would initiate DNA unwinding only at one end of the DNA substrate, the

single blunt end. In these DNA substrates, the length of the DNA duplex stem region that closes the hairpin is 15 bp. These substrates are composed of three separate oligodeoxynucleotides, a constant "bottom" strand containing the hairpin and two "top" strands, "A" and "B", that when annealed to the bottom strand form a fully base-paired duplex (with the exception of the T_4 loop in the hairpin), but which contain two nicks, as depicted in Figure 1. In the studies reported here, the top A strand was radioactively labeled at its 5′-end with ^{32}P (denoted as an asterisk in Figure 1) and the release of this DNA strand was used to monitor DNA unwinding. The length of the duplex region to be unwound, *L*, is therefore equal to the length of oligodeoxynucleotide A, which was varied by changing the lengths of the A and B oligodeoxynucleotides, while maintaining the sum of their lengths at 80 nt. The nucleotide sequences of the A DNA strands, designated I–V, are given in Figure 1. We note that the 80 base oligodeoxynucleotide listed in Figure 1 was not used in DNA unwinding experiments; its sequence is given solely to indicate the complete sequence of the DNA substrate.

The second oligodeoxynucleotide B was included in these substrates so that there would be a sufficiently long duplex region (at least 35 bp including the 15 bp region that closes the hairpin) separating the hairpin end of the DNA substrate from oligodeoxynucleotide A that is to be unwound. This design was used for two reasons. First, to minimize any potential interference with unwinding by a second RecBCD enzyme that might bind to the hairpin end of the DNA substrate, and second, to maintain a region of duplex DNA (at least 35 bp) beyond the first oligonucleotide (A) in the event that RecBCD needs to make contacts with those regions of the DNA, as suggested for the RecBC enzyme.[15] However, as discussed below, we have observed that the single turnover kinetic time courses for unwinding are unaffected if strand B is missing from the DNA substrate.

To test whether initiation of DNA unwinding occurs only at one end of the DNA substrate, we first determined the length for the duplex stem region closing the hairpin that is required to prevent RecBCD from initiating unwinding at the hairpin. We examined three different "double hairpin" oligodeoxynucleotides depicted schematically in the lower inset of Figure 2(b). These DNA molecules contain hairpin regions at both ends, but differing in the number of base-pairs, *N*, that form the duplex stem closing the hairpin. The length of the central duplex region between the two nicks in these DNA molecules is 40 bp. A ^{32}P label was incorporated at the 5′ end of one oligonucleotide (designated by an asterisk in Figure 2(b)). Chemical quenched-flow experiments were performed as described in Materials and Methods by pre-incubating RecBCD (20 nM) with each DNA substrate (2 nM) in buffer M at 25.0 °C and initiating the reactions by addition of 10 mM ATP

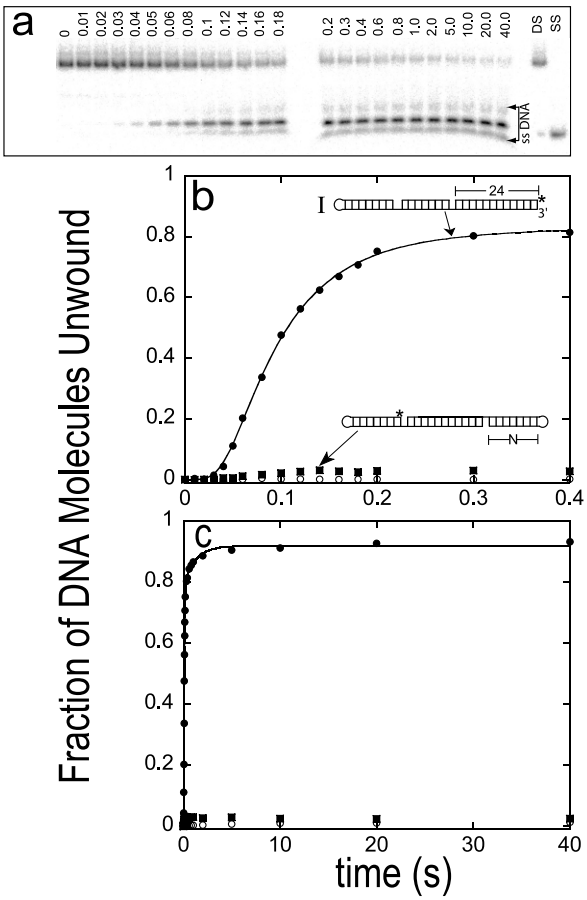


Figure 2. Single turnover kinetic time-course for RecBCD-catalyzed unwinding of a 24 bp duplex shows a distinct lag phase. Rapid chemical quenched-flow experiments were performed by pre-incubating 20 nM RecBCD with 2 nM DNA substrate I (Figure 1) in buffer M at 25 °C. Reactions were initiated by addition of an ATP (5 mM final concentration) solution, which also contained unlabeled 16 bp hairpin DNA (5 μM final concentration) in buffer M at 25 °C. The reactions were quenched by addition of 0.4 M EDTA in 10% glycerol (see Materials and Methods). (a) Phosphorimager image of a 10% non-denaturing polyacrylamide gel after electrophoresis of the ^{32}P end-labeled DNA (substrate I) samples resulting from the rapid quenched-flow experiments. The time (in seconds) after the addition of ATP is indicated for each lane. The lane marked DS shows where the duplex DNA substrate migrates in the absence of protein, and the lane marked SS shows the same substrate after boiling. The three bands, indicated as ssDNA, represent the unwound ssDNA products. (b) Unwinding time courses plotted as the fraction of duplex DNA molecules unwound for: (●) DNA substrate I ($L = 24$ bp), (■) a double hairpin substrate with a 40 bp duplex region and $N = 15$ bp closing the two hairpin regions. (○) A control experiment in which 10 μM of the 16 bp hairpin DNA trap was pre-mixed with 2 nM DNA substrate I in buffer M containing ATP and then rapidly mixed with 20 nM RecBCD in buffer M (ATP concentration after mixing was 5 mM). (c) The same single turnover DNA unwinding time-course of RecBCD-catalyzed unwinding of DNA substrate I showing data to 40 seconds.

(5 mM after mixing) and a large excess of a 16 bp hairpin trap (unlabeled). The results of the experiment with $N = 15$ bp are shown in Figure 2(b) and indicate that RecBCD cannot initiate unwinding at a hairpin that is closed by duplex regions of $N = 15$ bp. However, substantial and rapid unwinding by RecBCD can occur on a DNA double hairpin with $N = 5$ bp (data not shown). The ability of RecBCD to initiate unwinding at a nick if the stem region is only 5 bp is consistent with the observation that RecBCD is able to distort a 5–6 bp region upon binding to a blunt end duplex in a reaction that requires Mg^{2+}, but not ATP.[27] For this and other reasons discussed below, we used a 15 bp duplex stem to close the hairpin regions of the DNA substrates used in our kinetic studies ($N = 15$ bp).

Single-turnover kinetic time courses of DNA unwinding catalyzed by RecBCD enzyme

Single turnover kinetic studies of RecBCD-catalyzed DNA unwinding of the DNA substrates shown in Figure 1 were performed as described in Materials and Methods in a rapid chemical quenched-flow apparatus. Briefly, RecBCD was incubated with a ^{32}P-labeled DNA substrate in buffer M followed by rapid mixing of the pre-formed RecBCD–DNA substrate complex with an equal volume (30 μl) of buffer M containing 10 mM ATP and a large excess of a 16 bp DNA hairpin (unlabeled) at 25.0 °C. Therefore, the final ATP concentration at the start of the reaction was 5 mM, with 10 mM $MgCl_2$. The excess 16 bp DNA hairpin was included to serve as a protein trap to prevent initiation of DNA unwinding by any free RecBCD or re-binding of any RecBCD that dissociates during the course of unwinding. Therefore, these kinetic time courses measure a single round of DNA unwinding by RecBCD complexes that are pre-bound to the DNA substrate in a competent initiation complex. The experiments were performed at low DNA substrate concentrations (1 nM final concentration after mixing) to prevent re-annealing of the unwound DNA substrate during the course of the experiment. The gel assay used to monitor DNA unwinding is an "all or none" assay, since it detects only fully unwound DNA molecules. However, quantitative analysis of the time courses can yield information about intermediates formed during the course of DNA unwinding.[28]

Figure 2 shows the results from a single turnover experiment performed by pre-incubating 20 nM RecBCD with 2 nM DNA substrate I (Figure 1) and monitoring the time course of complete unwinding of the 24 bp duplex region ($L = 24$ bp). Under the conditions used in our unwinding experiments (10 mM Mg^{2+}, 5 mM ATP), the nuclease activity of RecBCD will preferentially degrade the unlabeled bottom strand (from 3′ to 5′), while the ^{32}P-labeled top strand will be cleaved only infrequently.[32] A control experiment in which the bottom strand was labeled with ^{32}P at its 5′ end showed that ~85% of the bottom strand DNA was degraded during unwinding (data not shown), consistent with published results.[32]

Figure 2(a) shows an image of a 10% non-denaturing polyacrylamide gel showing separation of the duplex DNA substrate from the unwound product, ^{32}P-labeled 24 base oligodeoxynucleotide A. The duplex DNA substrate ($t = 0$ time-point) shows the presence of two bands. The faster migrating major band reflects the complete DNA substrate, as shown in Figure 1, whereas the slower migrating minor band reflects a small fraction of the DNA substrate which is missing the B oligonucleotide, as shown by control experiments (data not shown). The three fastest migrating bands shown in Figure 2(a) reflect fully unwound "ssDNA product". This conclusion is supported by the observation that the time-course of formation is the same for each of these three bands. The middle band, representing the majority of the ssDNA product is the full-length oligodeoxynucleotide A. The fastest moving band(s) in the ssDNA product triplet reflects ssDNA oligodeoxynucleotide A that has been partially degraded. The time-course of unwinding determined from experiments performed under conditions where the RecBCD nuclease activity is inhibited ([ATP] in excess over the [Mg^{2+}]) does not show any ssDNA degradation, but exhibits the same rate and step-size of DNA unwinding (data not shown). The slowest migrating band in the ssDNA product triplet reflects full-length oligonucleotide A that possesses some stable secondary structure (possibly residual base-pairing) such that it migrates slower than the bulk of the oligodeoxynucleotide A. This was confirmed by boiling the reaction samples before gel electrophoresis, which showed that the slowest migrating band in the ssDNA product triplet co-migrated with the major oligodeoxynucleotide A band.

On the basis of the above discussion, data of the type shown in Figure 2(a) were analyzed to obtain the time-course of DNA unwinding by including all three of the fastest migrating bands in our quantitation of "unwound ssDNA". In further support of this, the DNA unwinding time courses obtained by including all three fastest migrating bands as unwound ssDNA agree quantitatively with the time courses obtained using a fluorescence-stopped-flow unwinding assay, which is blind to any nuclease activity (A.L.L., unpublished results). However, exclusion of the slowest migrating band of the triplet in our quantitative analysis resulted in no statistically significant difference in either the DNA unwinding time-course or the fitting parameters determined from the time-course. Exclusion of the fastest migrating band reduced the amplitude of DNA unwinding by ~5%.

Figure 2(b) shows the fraction of DNA molecules unwound as a function of time. The unwinding reaction is biphasic, displaying an initial "lag phase" in which ~70% of the total DNA molecules

$$(R \bullet D)_{NP}$$
$$\uparrow\downarrow k_{NP}$$
$$(R \bullet D)_L \xrightarrow{k_U} I_{(L-m)} \xrightarrow{k_U} \ldots I_{(L-im)} \ldots \xrightarrow{k_U} ssDNA$$

Scheme 1.

are unwound within the first 100 ms, followed by a much slower second phase (evident in Figure 2(c)) in which an additional ~20% of the DNA molecules are unwound. The distinct lag phase in the first ~30 ms is not due to re-annealing of the fully unwound DNA, since the half-life for re-annealing is >2.5 hours at the 1 nM DNA concentrations used in these experiments.[33] Furthermore, the fact that the RecBCD nuclease activity degrades the bottom strand from the 3′ to 5′ direction also prevents re-annealing of the unwound DNA. We attribute the slower second phase to DNA unwinding by a fraction of RecBCD that is initially bound in a non-productive mode that must undergo a slow isomerization step, with rate constant k_{NP}, before fast unwinding can proceed (see Scheme 1). As shown below, the presence of the lag phase and the resulting rate of DNA unwinding are independent of the RecBCD concentration as long as the RecBCD is pre-mixed with the DNA substrate in the presence of Mg^{2+}. When RecBCD is not pre-incubated with the DNA substrate, but rather added with the ATP to initiate the unwinding reaction, a much slower unwinding time-course is observed due to the fact that the rate of RecBCD binding to the DNA partially limits the observed rate of formation of fully unwound DNA (data not shown).

The following control experiment was performed to test the efficacy of the 16 bp hairpin as a trap for free RecBCD protein: 10 mM ATP, 10 μM 16 bp hairpin, and 2 nM ^{32}P-labeled DNA substrate were pre-mixed in buffer M in one loop of the quenched-flow apparatus and then rapidly mixed with 20 nM RecBCD in buffer M in the other loop. As shown in Figure 2(b) (open circles), a final hairpin trap concentration of 5 μM, in the presence of 1 nM DNA substrate is sufficient to prevent initiation of unwinding by free RecBCD under the conditions of our experiments.

We considered whether the slower second phase of the unwinding reaction shown in Figure 2 might reflect unwinding of a fraction of the DNA substrate that is lacking the B oligonucleotide (see Figure 1). To examine this we performed single turnover DNA unwinding experiments with the $L = 60$ bp duplex (substrate V in Figure 1) and compared this with unwinding of the same substrate, but lacking the 20 nucleotide long B strand. Similar experiments were also performed with the $L = 24$ bp substrate in the presence and absence of the 56 nucleotide long B strand (substrate I in Figure 1). In each case, the single turnover unwinding time courses were unaffected by the absence of the B DNA strand, as shown in Figure 3. Thus, the second slow phase of these unwinding time courses is not due to a fraction of the DNA substrate that is missing the B oligodeoxynucleotide. Furthermore, the resulting DNA unwinding time-course of any of these DNA substrates is unaffected under our standard conditions even if it is lacking the B strand.

Effects of duplex DNA length on RecBCD-catalyzed DNA unwinding

In studies of single turnover DNA unwinding catalyzed by *E. coli* UvrD helicase, Ali & Lohman[28] observed a lag phase in the time-course. This lag phase increased with increasing duplex length due to the fact that complete unwinding of the DNA required a series of "*n*" sequential steps, resulting in the transient formation of a significant population of partially unwound DNA intermediates during the course of the unwinding reaction. If the lag phase observed in our single turnover studies of RecBCD-catalyzed DNA unwinding has the same origin, then the extent of the lag should also increase with increasing duplex length. We therefore examined RecBCD-catalyzed unwinding of the series of DNA substrates, depicted in Figure 1, with duplex lengths, $L = 24$, 30, 40, 48 and 60 bp.

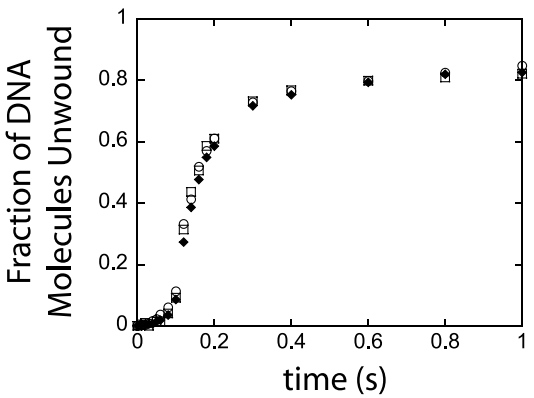


Figure 3. The single turnover kinetic time-course of RecBCD-catalyzed unwinding of oligodeoxynucleotide A of the double nicked DNA substrate is unaffected by the absence of oligonucleotide B and the absence of trap. RecBCD (20 nM) was pre-incubated with DNA substrate V ($L = 60$ bp) (2 nM) in the (□) presence and (○) absence of oligonucleotide B and (◆) in the absence of the 16 bp DNA hairpin trap and the reaction was initiated by addition of ATP (5 mM final concentration) (buffer M, 25.0 °C).

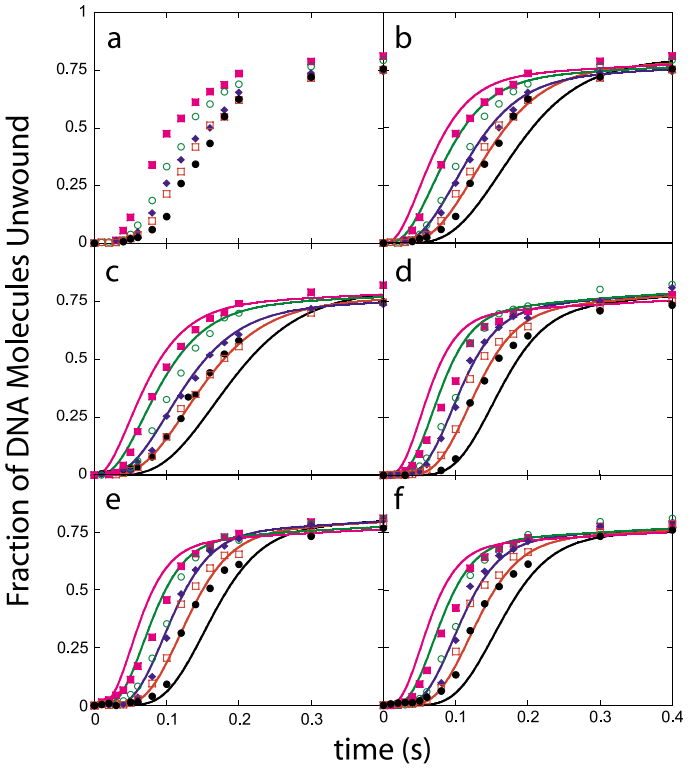


Figure 4. Single turnover RecBCD-catalyzed DNA unwinding time-courses as a function of DNA duplex length are not well described by Scheme 1. Five sets of time courses, each determined independently on separate days with ^{32}P-labeled DNA substrates that were freshly prepared, are shown in (a)–(f) ((a) and (b) show the same data). RecBCD (20 nM) was pre-incubated with each DNA substrate (2 nM) and unwinding was initiated with ATP (5 mM final concentration) containing unlabeled 16 bp hairpin DNA trap (5 μM final concentration) (buffer M at 25.0 °C) as described in Materials and Methods. DNA substrate I ($L = 24$ bp) (▦), substrate II ($L = 30$ bp) (○), substrate III ($L = 40$ bp) (◆), substrate IV ($L = 48$ bp) (□), and substrate V ($L = 60$ bp) (●). (a) Results from one series of experiments illustrate the increase in the lag as a function of duplex length. The five independent data sets in (b)–(f) were each analyzed globally. (b) All five time courses from (a) were globally analyzed using Scheme 1 (equations (6) and (8)) yielding: $k_U = 38.4(\pm 9)\ s^{-1}$, $m = 8.4(\pm 1.3)$ bp step^{-1}, $x = 0.84(\pm 0.02)$, and $k_{NP} = 0.8(\pm 0.2)\ s^{-1}$ (variance of the fit equal to $3.0(\pm 0.2) \times 10^{-3}$). The continuous curves are the simulated time courses based on these best-fit parameters. (c) Independent, replicate data set analyzed as in (b) yielding: $k_U = 33.2(\pm 4.4)\ s^{-1}$, $m = 9.3(\pm 1.1)$ bp step^{-1}, $x = 0.83(\pm 0.01)$, and $k_{NP} = 1.2(\pm 0.1)\ s^{-1}$ (variance of the fit equal to $3.4(\pm 0.2) \times 10^{-3}$). (d) Independent, replicate data set, analyzed as in (b), yielding: $k_U = 62.8(\pm 23)\ s^{-1}$, $m = 5.7(\pm 0.8)$ bp step^{-1}, $x = 0.81(\pm 0.01)$, and $k_{NP} = 1.8(\pm 0.3)\ s^{-1}$ (variance of the fit equal to $3.8(\pm 0.3) \times 10^{-3}$). (e) Independent, replicate data set, but performed in the presence of *E. coli* SSB protein (1 μM final concentration), and analyzed as in (b), yielding: $k_U = 60.7(\pm 9.4)\ s^{-1}$, $m = 4.91(\pm 0.9)$ bp step^{-1}, $x = 0.84(\pm 0.01)$, and $k_{NP} = 1.4(\pm 0.2)\ s^{-1}$ (variance of the fit equal to $3.5(\pm 0.2) \times 10^{-3}$). (f) Independent, replicate data set, but performed in the presence of *E. coli* SSB protein (1 μM final concentration), and analyzed as in (b), yielding: $k_U = 69(\pm 22)\ s^{-1}$, $m = 5.3(\pm 0.9)$ bp step^{-1}, $x = 0.80(\pm 0.01)$, and $k_{NP} = 2.0(\pm 0.3)\ s^{-1}$ (variance of the fit equal to $3.7(\pm 0.3) \times 10^{-3}$). All parameters are summarized in Table 1.

Figure 4(a) shows the single turnover DNA unwinding time courses resulting from experiments performed in buffer M by pre-incubating 20 nM RecBCD with 2 nM DNA and initiating the reaction by addition of ATP, as described in Materials and Methods. Each unwinding time-course shown in Figure 4(a) displays a lag phase, and the extent of the lag phase increases with increasing duplex length. This indicates that RecBCD-catalyzed DNA unwinding is a multi-step process and that a rate-limiting step (or steps) is repeated n times in order to completely unwind each DNA. In order to determine the reproducibility of these data, we compared the results from five independent sets of DNA unwinding experiments. These experiments are shown in Figure 4(b)–(f) ((a) and (b) show the same data) and each set was performed on a different day, with different ^{32}P-labeled duplex DNA samples, which were labeled independently. The reproducibility of the time courses from these five sets of independent experiments is quite good. Although the reproducibility of these experiments is even better when the experiment is repeated with the same DNA samples, we analyzed these five sets of independent data in order to better assess how the variability among data sets would affect our ability to estimate the kinetic parameters for DNA unwinding and the kinetic step-size.

The data in Figure 4(e) and (f) are from experiments in which 2 μM *E. coli* SSB protein (1 μM final reaction concentration) was pre-incubated with the RecBCD and the DNA substrate. These time courses are also the same, within our uncertainty, as those performed in the absence of SSB protein (a)–(d), thus the presence of *E. coli* SSB protein has no effect on the RecBCD-catalyzed DNA unwinding of these short DNA substrates.

The amplitude of each unwinding time-course is the same, independent of duplex DNA length. This indicates that no significant dissociation of RecBCD occurs during the course of DNA unwinding, consistent with the fact that the processivity of DNA unwinding by RecBCD is very high under

Table 1. Kinetic parameters determined from NLLS fitting of the experimental data to Scheme 1

Figure 4	k_U (s^{-1})	k_C	k_{NP} (s^{-1})	m (bp $step^{-1}$)	h	mk_U (bp s^{-1})	Variance ($\times 10^{-3}$)
b	38.4 ± 9	NA	1.7 ± 0.7	8.4 ± 1.3	0	322 ± 10	3.0 ± 0.2
c	33.2 ± 4.4	NA	1.6 ± 0.5	9.3 ± 1.1	0	309 ± 7	3.4 ± 0.2
d	62.80 ± 23	NA	1.6 ± 0.7	5.7 ± 0.8	0	358.4 ± 13	3.8 ± 0.3
e	60.7 ± 9.4	NA	2.25 ± 1.2	4.91 ± 0.9	0	358.97 ± 7	3.5 ± 0.2
f	69 ± 22	NA	2.82 ± 1.3	5.3 ± 0.9	0	367.9 ± 16	3.7 ± 0.3
Average	53 ± 16	NA	2 ± 0.5	6.7 ± 2.0	0	343 ± 26	

these conditions.[22] However, it is evident from Figure 4 that ~10–12% of the DNA molecules are not unwound in these single turnover experiments, independent of duplex length. This is not the result of some fraction of the DNA being free of RecBCD, since independent binding studies (data not shown) indicate that all of the DNA substrate is bound by RecBCD at these concentrations before addition of ATP. The ~10% of the DNA that remains duplex does not result from re-annealing of the ssDNA products, since inclusion of a 1000-fold excess of *E. coli* SSB protein does not affect the total unwinding amplitude (Figure 4(e) and (f)). Furthermore, as discussed above, upon ^{32}P labeling of the bottom strand we observed that most of the bottom strands were completely degraded by the nuclease activity of RecBCD during unwinding (data not shown). Finally, if the 16 bp DNA hairpin trap is not included in the reaction, the remaining 8–12% of the DNA still remains unwound (Figure 3), indicating that the inability to unwind ~10% of the DNA is not due to an inability to re-bind RecBCD.

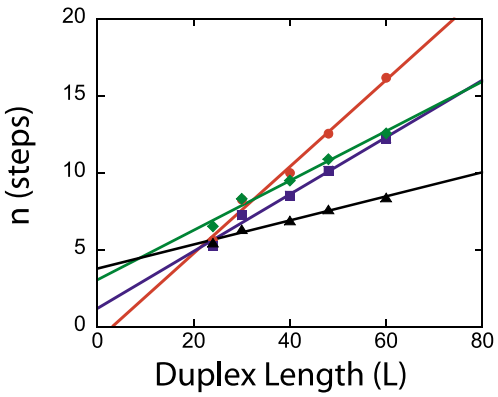


Figure 5. Scheme 2 with two to three additional h steps needed, so that the number of sequential unwinding steps, n, is directly proportional to the duplex length, L, which is unwound by RecBCD. Each independent set of data shown in Figure 4 was globally analyzed by NLLS analysis using Scheme 2 (equations (7) and (8)), with h constrained to be 0, 1, 2, or 3, to obtain estimates of A_T, k_U, k_C, k_{NP}, x, and n (=L/m). In this analysis, the parameters, k_U, k_C, k_{NP}, x, and h were constrained to be identical for each substrate, while A_T, and n were allowed to float in the NLLS analysis for each substrate. The number of unwinding steps, n, each with rate constant, k_U, obtained from the NLLS analysis of each duplex length in the five independent sets of data were averaged and plotted *versus* duplex length, L, for Scheme 2 with $h = 0$ (▲), $h = 1$ (◆), $h = 2$ (■), and $h = 3$ (●). Each line can be described by the equation, $n = L/m + b$, where n is the number of steps, L is the duplex length, m is the step size and b is the y-intercept. (▲) Scheme 2 constraining $h = 0$ yields $m = 12.8$ bp $step^{-1}$, and $b = 3.8$. (◆) Scheme 2 constraining $h = 1$, yields $m = 6.2$ bp $step^{-1}$, and $b = 3.1$. (■) Scheme 2 constraining $h = 2$, yields $m = 5.4$ bp $step^{-1}$, and $b = 1.2$. (●) Scheme 2 constraining $h = 3$, yields $m = 3.6$ bp $step^{-1}$, and $b = -0.9$.

We can offer two possible explanations for this result. Either ~10% of the DNA molecules are in a state that is incapable of being unwound or that some fraction (~10%) of the RecBCD–DNA complexes can form a tightly bound "dead end" complex that cannot initiate DNA unwinding, at least on the time-scale of our single turnover studies. Although we cannot differentiate between these two possibilities, this latter explanation is consistent with a recent single molecule study of RecBCD-catalyzed DNA unwinding,[34] which reported that some fraction of the DNA molecules bound by RecBCD could not be unwound.

Quantitative analysis of the DNA unwinding time courses using Scheme 1

As shown in studies of the *E. coli* UvrD helicase,[28] the single turnover time courses for DNA unwinding can be analyzed to estimate the number of steps, n, required to unwind each duplex of length, L. The expectation is that the number of steps, n, should be directly proportional to the duplex length, L, and therefore, the kinetic step-size, m, for DNA unwinding can be estimated as $m = L/n$.

Our first attempt to analyze these time courses made use of the simplest scheme, Scheme 1, which assumes that a pre-bound helicase–DNA complex exists at equilibrium in two states, $(RD)_L$, which is competent to initiate DNA unwinding upon addition of ATP, and $(RD)_{NP}$, which must first isomerize with rate constant k_{NP} to form the $(RD)_L$ state before it can initiate DNA unwinding. In Scheme 1, the $(RD)_L$ state unwinds the DNA in a series of n repeating steps, with an average of m base-pairs being unwound per step with rate constant, k_U. We note that the k_U term in Scheme 1 does not necessarily reflect the actual unwinding rate constant, but rather represents the rate-limiting step for unwinding of the m base-pairs within each repeated unwinding cycle. Due to the high processivity of RecBCD, we have not explicitly included dissociation of RecBCD from the DNA once unwinding has been initiated. Equivalent

$$\begin{array}{c}(R\bullet D)_{NP}\\ \uparrow\downarrow k_{NP}\\ (R\bullet D)_L \xrightarrow{k_C} (R\bullet D)_L^1 \xrightarrow{k_C} \ldots (R\bullet D)_L^h \xrightarrow{k_U} I_{(L-m)} \xrightarrow{k_U} \ldots . I_{(L-im)} \ldots \xrightarrow{k_U} ssDNA\end{array}$$

Scheme 2.

expressions for the time dependence of formation of ssDNA product, based on Scheme 1, are given in equations (2) and (5) in Materials and Methods. However, as discussed in Materials and Methods, it is actually more convenient to directly analyze the unwinding time courses using equations (6) and (8), since these allow for a non-integer step size, m, and thus enable us to float the step-size directly in a non-linear least-squares (NLLS) analysis, even for complex models. Identical kinetic parameters are obtained when determined from NLLS analysis using either equation (5) or equations (6) and (8).

Each of the five independent sets of data (Figure 4(b)–(f)) was analyzed by simultaneous NLLS fitting of all five time courses ($L = 24$, 30, 40, 48, 60 bp) to equations (6) and (8). In these global analyses, the parameters, k_U, k_{NP}, m, and x, the fraction of productively bound RecBCD complexes, were constrained to be identical for each substrate, while A_T, the total unwinding amplitude, was allowed to float for each substrate. The simulated time courses describing the best global fits of these data to equations (6) and (8) are shown in Figure 4(b)–(f) and the kinetic parameters determined from these analyses are given in Table 1. As is clear from Figure 4, the quality of these fits is poor, suggesting that the simple mechanism in Scheme 1 is inadequate to describe these data.

In order to assess the mechanism in Scheme 1 further, we considered the relationship between the numbers of steps, n, and the duplex length, L, as determined from the NLLS fitting of the data in Figure 4 to equations (6) and (8). To determine this relationship, each independent set of time courses was reanalyzed using Scheme 1 by constraining the parameters, k_U, k_{NP}, and x to be identical for each substrate, while A_T, and n were allowed to

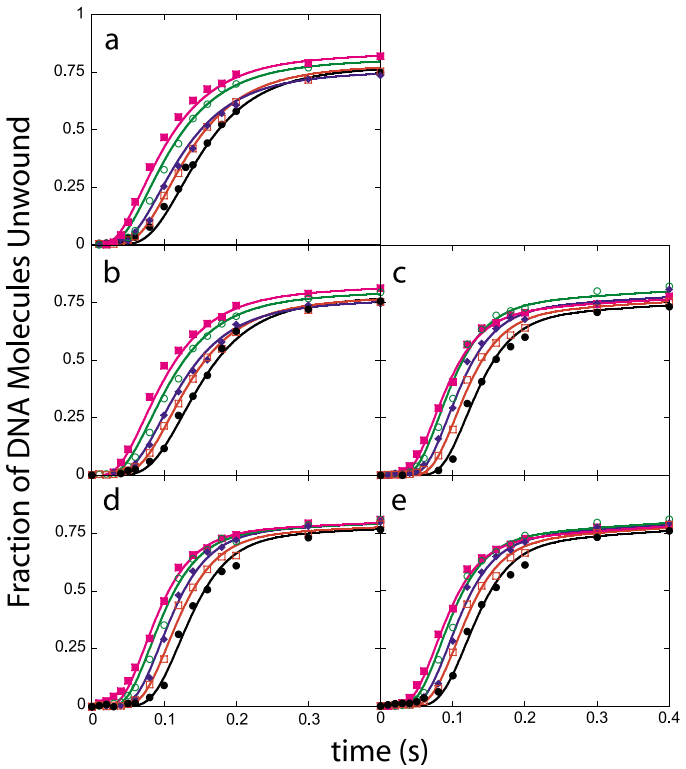


Figure 6. Single turnover RecBCD-catalyzed DNA unwinding time courses as a function of DNA duplex length are well described by Scheme 2 with $h = 2$ or 3. The data in (a)–(e) are the identical data shown in Figure 4(b)–(f). (■) DNA substrate I ($L = 24$ bp); (○) DNA substrate II ($L = 30$ bp); (◆) DNA substrate III ($L = 40$ bp); (□) DNA substrate IV ($L = 48$ bp); (●) DNA substrate V ($L = 60$ bp). All five independent data sets were globally analyzed by non-linear least-squares methods using Scheme 2 (equations (7) and (8)) yielding the following parameters: (a) $k_U = 72(\pm 25)\ s^{-1}$, $m = 8.8(\pm 3.3)$ bp step^{-1}, $k_C = 20(\pm 5)\ s^{-1}$, $h = 1.5(\pm 0.3)$ steps, $x = 0.84(\pm 0.02)$, and $k_{NP} = 0.8(\pm 0.2)\ s^{-1}$ (variance of the fit equal to $6.5(\pm 0.4) \times 10^{-4}$). (b) $k_U = 121(\pm 45)\ s^{-1}$, $m = 6.1(\pm 2.1)$ bp step^{-1}, $k_C = 23(\pm 3)\ s^{-1}$, $h = 1.7(\pm 0.2)$ steps, $x = 0.83(\pm 0.01)$, and $k_{NP} = 1.2(\pm 0.1)\ s^{-1}$ (variance of the fit equal to $2.8(\pm 0.2) \times 10^{-4}$). (c) $k_U = 377(\pm 144)\ s^{-1}$, $m = 2.23(\pm 1)$ bp step^{-1}, $k_C = 42(\pm 8)\ s^{-1}$, $h = 2.8(\pm 0.4)$ steps, $x = 0.81(\pm 0.01)$, and $k_{NP} = 1.8(\pm 0.3)\ s^{-1}$ (variance of the fit equal to $3.9(\pm 0.3) \times 10^{-4}$). (d) $k_U = 221(\pm 90)\ s^{-1}$, $m = 3.6(\pm 1.9)$ bp step^{-1}, $k_C = 38(\pm 7)\ s^{-1}$, $h = 2.5(\pm 0.4)$ steps, $x = 0.84(\pm 0.01)$, and $k_{NP} = 1.4(\pm 0.2)\ s^{-1}$ (variance of the fit equal to $3.6(\pm 0.3) \times 10^{-4}$). (e) $k_U = 258(\pm 97)\ s^{-1}$, $m = 3.5(\pm 1.4)$ bp step^{-1}, $k_C = 47(\pm 9)\ s^{-1}$, $h = 3.2(\pm 0.5)$ steps, $x = 0.80(\pm 0.01)$, and $k_{NP} = 2.0(\pm 0.3)\ s^{-1}$ (variance of the fit equal to $3.9(\pm 0.3) \times 10^{-4}$). All parameters are summarized in Table 2. The continuous curves are the simulated time courses based on Scheme 2 and the best-fit parameters.

Table 2. Kinetic parameters determined from NLLS fitting of the experimental data to Scheme 2

Figure 6	k_U (s^{-1})	k_C (s^{-1})	k_{NP} (s^{-1})	m (bp $step^{-1}$)	h (steps)	mk_U (bp s^{-1})	Variance ($\times 10^{-4}$)
a	72 ± 25	20 ± 5	0.8 ± 0.2	8.8 ± 3.3	1.5 ± 0.3	653 ± 36	6.5 ± 0.4
b	121 ± 45	23 ± 3	1.2 ± 0.1	6.1 ± 2.1	1.7 ± 0.2	719 ± 29	2.8 ± 0.2
c	377 ± 142	42 ± 8	1.8 ± 0.3	2.2 ± 1	2.8 ± 0.4	841 ± 35	3.9 ± 0.3
d	221 ± 90	38 ± 7	1.4 ± 0.2	3.6 ± 1.9	2.5 ± 0.4	773 ± 24	3.6 ± 0.3
e	258 ± 97	47 ± 9	2.0 ± 0.3	3.5 ± 1.4	3.2 ± 0.5	893 ± 38	3.9 ± 0.3
Average	210 ± 120	34 ± 12	1.4 ± 0.5	5 ± 3	2.4 ± 0.	800 ± 67	
Fit of averaged data points	196 ± 77	29 ± 3	1.1 ± 0.2	3.9 ± 1.3	2.0 ± 0.2	790 ± 23	2.1 ± 0.2

float for each substrate (L/m was replaced with n in equation (6)). The number of steps, n, required to fit each duplex length for all five data sets was then averaged and plotted *versus* L in Figure 5 (plot for Scheme 2 with $h = 0$). This plot indicates that n, as determined from analysis using Scheme 1 (equations (6) and (8)), is not directly proportional to L, but rather shows a positive intercept. We have shown through simulations (A.L.L. *et al.*, unpublished results) that the observation of a positive intercept from such an analysis (i.e. the fact that n is not directly proportional to L) is an indication that additional steps must occur in the mechanism, but that these steps are not part of the repeated unwinding cycles.

Quantitative analysis of the DNA unwinding time courses using Scheme 2

We next considered the more complex Scheme 2, in which an additional "h" steps with rate constant, k_C, occur within the mechanism, although not within the repeated unwinding cycles. We note that although Scheme 2 shows the steps with rate constant k_C at the beginning of the mechanism, in principle these steps could occur anywhere within the mechanism, although not within the repeated unwinding cycles defined by rate constant, k_U (A.L.L. *et al.*, unpublished results). We reanalyzed the five sets of unwinding data using Scheme 2, constraining $h = 1$, 2, or 3, and determined the number of steps, n, each with rate constant, k_U, required to obtain the best NLLS fit to the data. Figure 5 shows plots of the average value of n from analysis of all five data sets (Figure 4(b)–(f)) as a function of L, for $h = 0$, $h = 1$, $h = 2$ and $h = 3$. As seen in Figure 5, as h increases, the value of the intercept decreases, eventually becoming negative for Scheme 2 with $h = 3$. This indicates that Scheme 2 with two or three additional steps with rate constant, k_C, is required to describe the unwinding time courses, in order that the number of unwinding steps, n, is directly proportional to duplex length, L (i.e. $n = L/m$). Furthermore, the value of k_C is not dependent upon [ATP] (A.L.L., unpublished data), consistent with our conclusion that these steps do not occur as part of the repeated unwinding cycle.

Each of the five independent sets of data (Figure 4(b)–(f)) was re-analyzed by global NLLS fitting of the five time courses ($L = 24$, 30, 40, 48, and 60 bp) to Scheme 2 using equations (7) and (8) to obtain estimates of A_T, k_U, k_C, k_{NP}, x, m and h. In this analysis, the parameters, k_U, k_C, k_{NP}, x, m and h, were constrained to be identical for each of the five DNA substrates, while A_T, the total amplitude, was allowed to float for each substrate. In the global fitting, we also assumed that x, the fraction of productively bound RecBCD complexes, is independent of duplex length for a given protein concentration. The simulated curves describing the best global fits of these data to Scheme 2 (equations (7) and (8)) are shown overlaid on each independent data set in Figure 6. Comparison of these fits to those in Figure 4 indicates that Scheme 2 provides a much better description of the DNA unwinding time courses. The kinetic parameters determined from these five independent analyses are summarized in Table 2. Although each of the five sets of independent time courses shows good reproducibility, one sees that the values of the fitted parameters listed in Table 2 vary considerably among the independent data sets. For example, m ranges from 2.2(±1) bp $step^{-1}$ to 8.8(±3.3) bp $step^{-1}$, and k_U varies from 72(±25) s^{-1} to 377(±142) s^{-1}, although the product, mk_U is much better constrained, ranging from 653(±36) s^{-1} to 893(±38) s^{-1}. On the other hand,

Table 3. Kinetic parameters determined from NLLS fitting of a set of time courses simulated using Scheme 3, but fit to Scheme 2

Simulation	k_U (s^{-1})	k_C (s^{-1})	k_{NP} (s^{-1})	m (bp $step^{-1}$)	h (steps)	mk_U (bp s^{-1})
1	179 ± 60	42 ± 5	1.3 ± 0.2	5.4 ± 1.6	2.2 ± 0.4	962 ± 41
2	136 ± 42	51 ± 7	1.0 ± 0.2	7.7 ± 2.3	2.9 ± 0.5	1046 ± 44
3	355 ± 113	38 ± 6	1.2 ± 0.3	2.5 ± 0.8	2.11 ± 0.5	888 ± 37
4	386 ± 110	26 ± 5	0.9 ± 0.4	2.24 ± 0.6	1.37 ± 0.5	866 ± 37
5	88 ± 30	60 ± 11	1.6 ± 0.3	11 ± 4	3.0 ± 0.7	997 ± 42
Average	229 ± 133	43 ± 13	1.2 ± 0.3	5.8 ± 3.7	2.3 ± 0.6	952 ± 75
Fit of averaged data	250 ± 66	37 ± 3	1.2 ± 0.1	4.0 ± 0.9	2.1 ± 0.2	984 ± 84
Parameters used in data simulation	200	35	1	5.0	2	1000

the average values for the parameters determined from analysis of the five sets of independent unwinding data, along with their 68% confidence limits, are: $k_U = 210(\pm 120)\ s^{-1}$, $m = 5.0(\pm 3.0)$ bp step^{-1}, $k_C = 34(\pm 12)\ s^{-1}$, $k_{NP} = 1.4(\pm 0.5)\ s^{-1}$, $x = 0.82(\pm 0.02)$, and $h = 2.4(\pm 0.7)$ steps, with $mk_U = 800(\pm 67)$ bp s^{-1}.

The observed variation among the values of the fitted parameters determined from the independent analyses of the five sets of DNA unwinding time courses is expected, even for data of this quality, and has several origins (see Discussion and Table 3). First of all, it reflects the fact that these five data sets were determined completely independently and that the number of data points that can be obtained, reasonably, from a set of quenched-flow experiments is limited (~25–30 points per duplex length). Secondly, it reflects the fact that we are fitting these data to a complex model containing seven parameters, in which there is significant parameter correlation between k_U and m, and also between k_C and h. As expected (see Discussion and Materials and Methods), the best estimate of the parameters and their uncertainties is obtained from a NLLS analysis of the averages of the five independent sets of time courses.

On the basis of the above results, a final NLLS fit of the averaged data set to Scheme 2 (equations (7) and (8)) was performed by constraining k_U, k_C, k_{NP}, x, m and h to be identical for all duplexes, independent of length, while allowing only the total unwinding amplitude A_T to float for each duplex. The parameters obtained from this fit are: $k_U = 196(\pm 77)\ s^{-1}$, $k_C = 29(\pm 3)\ s^{-1}$, $k_{NP} = 1.1(\pm 0.2)\ s^{-1}$, $m = 3.93(\pm 1.34)$ bp step^{-1}, $h = 2.04(\pm 0.2)$ steps, and $mk_U = 790(\pm 23)$ bp step^{-1}, and are given in Table 2. The simulated time courses based on these parameters and Scheme 2 are shown overlaid on the averaged set of data in Figure 7(a). We consider these parameters to be the best estimates from the fit of the data to Scheme 2. These parameters are in excellent agreement with those determined from analysis of unwinding of a series of fluorescent DNA substrates using stopped-flow techniques (A.L.L., unpublished experiments). Attempts were also made to fit these data by allowing the two additional h steps to have rate constants that are different from one another (i.e. k_{C1}, k_{C2}). However, when this was done, the values of k_{C1} and k_{C2} always floated to values that were identical to each other within less than 1% (data not shown). Therefore, we conclude that it is most appropriate to constrain these additional two steps to have identical rate constants, k_C.

To further assess our estimate of the DNA unwinding step-size, m, obtained from analysis of the data in Figure 7(a), we performed a "grid-search" on these data (see Materials and Methods). In this procedure, we systematically constrained the step-size to have values from 1 bp step^{-1} to 10 bp step^{-1}, and then optimized the global NLLS fit of the time courses in Figure 7(a) to Scheme 2 (equations (7) and (8)) by allowing the remaining parameters to float for each step-size. The sum of the squared residuals (SSR) obtained from the analysis for each value of m is plotted in Figure 7(b). The absolute minimum of this grid-search occurs at a step-size of 3.97 bp step^{-1}, which is the same as the step-size obtained from the non-constrained global analysis of the data in Figure 7(a) as reported above ($m = 3.93(\pm 1.34)$ bp step^{-1}).

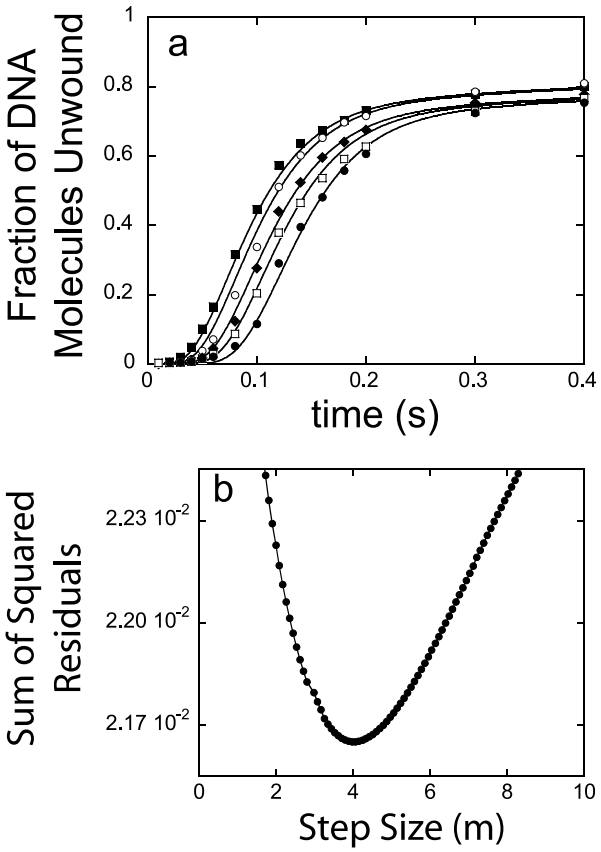


Figure 7. Estimation of the DNA unwinding step-size from analysis of the averages of the time courses for each of the five duplex DNA lengths from the five independent sets of single turnover kinetic time courses for RecBCD-catalyzed DNA unwinding. DNA substrates are described in Figure 1: (■) DNA substrate I ($L = 24$ bp), (○) substrate II ($L = 30$ bp), (◆) substrate III ($L = 40$ bp), (□) substrate IV ($L = 48$ bp), (●) substrate V ($L = 60$ bp). The data points shown for each DNA substrate represent the average of the data points for the five independent time courses shown in Figure 4(b)–(f). (a) All five time courses were globally analyzed using Scheme 2 (equations (7) and (8)) yielding: $k_U = 196(\pm 77)\ s^{-1}$, $m = 3.93(\pm 1.34)$ bp step^{-1}, $k_C = 29(\pm 3)\ s^{-1}$, $h = 2.04(\pm 0.2)$ steps, $x = 0.83(\pm 0.01)$, and $k_{NP} = 1.1(\pm 0.2)\ s^{-1}$ (variance of the fit equal to $2.1(\pm 0.2) \times 10^{-4}$). (b) SSR plotted *versus* step-size (m), showing a minimum at $m = 3.97$. Each SSR was determined by constraining the step size to have a constant value between 1 and 10 and then optimizing the fit using Scheme 2 with $h = 2$.

Effect of RecBCD concentration on the single turnover DNA unwinding time-course

The experiments described above were all performed with RecBCD in excess over the DNA substrate. We have also examined the effect of [RecBCD] (1–40 nM pre-incubation concentrations) on the single turnover kinetic time-course for unwinding of the 24 bp DNA substrate (2 nM substrate I in Figure 1). The resulting time courses are shown in Figure 8(a). Each time-course shows the expected lag phase, with the final extent of DNA unwinding increasing with increasing RecBCD concentration, reflecting an increase in the concentration of pre-formed RecBCD–DNA complexes that are competent to initiate DNA unwinding. Simultaneous (global) analysis of the set of six experiments shown in Figure 8(a) to Scheme 2 (equations (7) and (8)) was performed to obtain a single set of values for k_U, k_{NP}, and x that provide the best fit to all of the time courses. In this analysis, the unwinding step-size, m, the rate constant k_C, and h were constrained to be independent of [RecBCD], with values of $m = 3.9$ bp step^{-1}, $k_C = 29$ s^{-1}, and $h = 2$, based on the values obtained from the fit of the averaged data as a function of duplex length (Table 2). The resulting NLLS fits yield: $k_U = 215(\pm 5)$ s^{-1}, $k_{NP} = 1.4(\pm 0.1)$ s^{-1}, and $x = 0.82(\pm 0.01)$. The simulated time courses using Scheme 2 (equations (7) and (8)) and these best-fit parameters are shown overlaid on the data in Figure 8(a).

Figure 8(b) shows a plot of the total unwinding amplitude as a function of the ratio, [RecBCD]total/[DNA]total. The unwinding amplitude increases linearly with total [RecBCD], reaching a plateau value of ~90% of the DNA molecules unwound with an apparent breakpoint occurring at a [RecBCD]total/[DNA]total molar ratio of 1.6(±0.1). These data indicate that the binding of RecBCD to the ends of the DNA substrate is stoichiometric, i.e. the affinity of RecBCD for the DNA is too high to measure accurately under these conditions. One can only estimate an upper limit of $K_d < 0.04$ nM, for the apparent equilibrium dissociation constant for RecBCD binding to the blunt-end of the DNA substrate under these conditions (buffer M, 10 mM $MgCl_2$, 30 mM NaCl, 5% glycerol, 25.0 °C). This is consistent with previous estimates of the affinity of RecBCD for the blunt end of a DNA duplex,[23,25] although the previous estimates were made under slightly different solution conditions. The unwinding rate constants determined from fits of the individual time courses for each RecBCD concentration are plotted in Figure 8(c) and have values of $k_U = 217(\pm 10)$ steps s^{-1}, independent of RecBCD concentration.

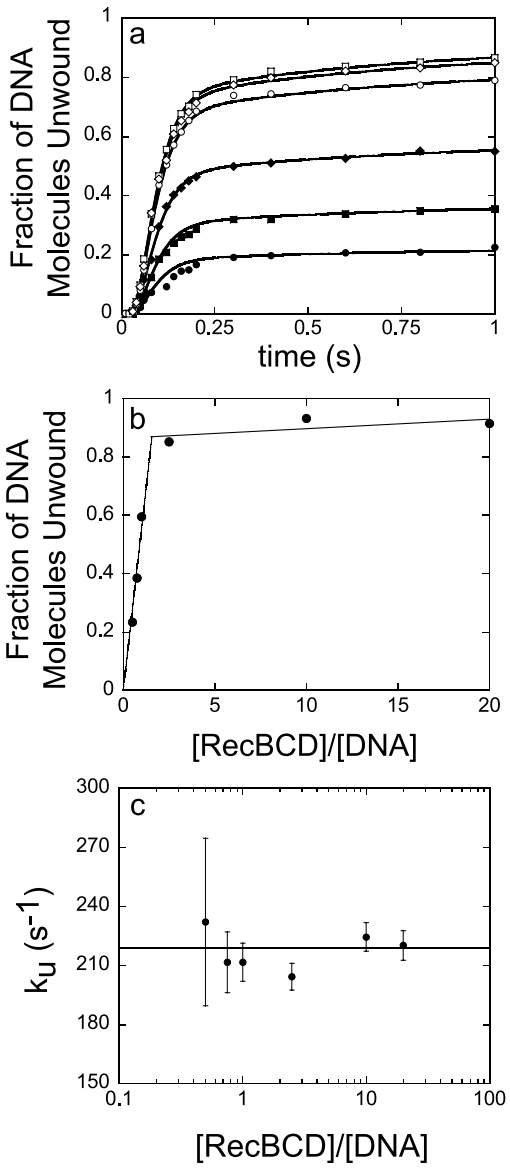


Figure 8. Dependence of the single turnover DNA unwinding time-course on RecBCD concentration. Single turnover DNA unwinding experiments were performed as described in Materials and Methods and in the legend to Figure 2 using the 24 bp DNA substrate I. (a) A series of experiments was performed by pre-incubating RecBCD at the following concentrations with 2 nM DNA: (●) 1 nM, (■) 1.5 nM, (◆) 2 nM, (○) 5 nM, (□) 20 nM, (◇) 40 nM RecBCD and initiating the reaction by addition of ATP (5 mM final concentration) containing the unlabeled 16 bp hairpin (5 μM final concentration). Experiments were performed in buffer M (30 mM NaCl, 10 mM $MgCl_2$) at 25.0 °C. All six time courses were analyzed globally by NLLS using Scheme 2 (equations (7) and (8)), and constraining $m = 3.9$ bp step^{-1}, $k_C = 29$ s^{-1}, and $h = 2$, to obtain estimates of $k_U = 215(\pm 5)$ s^{-1}, $k_{NP} = 1.4(\pm 0.1)$ s^{-1}, and $x = 0.82(\pm 0.01)$. (b) The total unwinding amplitude of each DNA unwinding reaction is plotted *versus* the ratio of [RecBCD]/[DNA]. The intersection of the two lines is 1.6(±0.1) [RecBCD]/[DNA]. (c) Values of k_U obtained from the individual fits of each time-course in (a) to Scheme 2 (equations (7) and (8)), constraining $m = 3.9$ bp step^{-1}, $k_C = 29$ s^{-1}, and $h = 2$, are plotted *versus* [RecBCD]/[DNA], with the horizontal line drawn through 217 steps s^{-1}.

Discussion

Using rapid chemical quenched-flow methods, we have examined the pre-steady-state

a

$$(R \bullet D)_{NP}$$
$$\uparrow\downarrow k_{NP}$$
$$(R \bullet D)_L \xrightarrow{k_C} (R \bullet D)^1_L \xrightarrow{k_C} (R \bullet D)^2_L \xrightarrow{k_U} I_{(L-m)} \xrightarrow{k_U} \ldots I_{(L-im)} \ldots \xrightarrow{k_U} ssDNA$$

b

$$(R \bullet D)_{NP}$$
$$\uparrow\downarrow k_{NP}$$
$$(R \bullet D)_L \xrightarrow{k_U} I_{(L-m)} \xrightarrow{k_U} \ldots I_{(L-im)} \ldots \xrightarrow{k_U} I^1_{(L-nm)} \xrightarrow{k_C} I^2_{(L-nm)} \xrightarrow{k_C} ssDNA$$

Scheme 3.

ATP-dependent kinetics of RecBCD-catalyzed unwinding of a series of short duplex DNA substrates, differing in duplex length, under single turnover conditions (i.e. a single round of DNA unwinding). Such experiments, in contrast to multiple turnover DNA unwinding experiments, can yield information about the kinetic steps that are directly involved in RecBCD-catalyzed DNA unwinding and translocation of the helicase along the DNA. The inclusion of a "trap" for free RecBCD protein is an important component of these experiments, since it eliminates the need to consider the kinetics of RecBCD binding to the DNA substrate in the quantitative analysis of the DNA unwinding time courses. The single turnover DNA unwinding time courses displayed a distinct lag phase that increased as a function of duplex DNA length, indicating that RecBCD requires multiple catalytic cycles to completely unwind each DNA duplex and that the intermediate states of partially unwound DNA produced during this process are populated significantly during the course of the reaction. Although the assay used to monitor DNA unwinding detects directly only fully unwound DNA and thus is an all or none assay, quantitative analysis of the unwinding time courses can yield information about the number of kinetic steps required by RecBCD to unwind each duplex DNA[28] (A.L.L. *et al.*, unpublished results).

Minimal kinetic mechanism of RecBCD-catalyzed DNA unwinding

Our detailed quantitative analysis of the time courses of RecBCD-catalyzed DNA unwinding performed as a function of duplex DNA length, L, supports the minimal n-step sequential kinetic mechanism shown in Scheme 3(a) and (b), which is the same as Scheme 2 with $h = 2$. In this scheme, the pre-bound RecBCD can exist in two states in the presence of Mg^{2+}, $(RD)_L$, which is competent to initiate DNA unwinding upon addition of ATP, and $(RD)_{NP}$, which is a non-productive state that first must isomerize to form the $(RD)_L$ state before it can initiate DNA unwinding. The evidence for these two states is based primarily on the observation that all time courses display biphasic character. Upon addition of ATP, the $(RD)_L$ state can initiate DNA unwinding through a series of repeated rate-limiting steps, each with rate constant $k_U = 196(\pm 77)\ s^{-1}$, and between every two rate-limiting steps an average of $m = 3.9(\pm 1.3)$ bp are unwound. We refer to this value of $m = 3.9(\pm 1.3)$ bp as the kinetic step-size for DNA unwinding. The simplest and most conservative interpretation of this kinetic step-size is that an average of $3.9(\pm 1.3)$ bp are unwound with rate constant k_U between every two rate-limiting steps. We also emphasize that the repeated step with rate constant k_U simply reflects the rate-limiting step that is repeated within the unwinding cycle. It may not reflect the actual DNA unwinding step, which may be much faster than k_U, but rather could reflect a slower conformational change in the protein, release of ADP or inorganic phosphate, etc.

The kinetic step-size of DNA unwinding estimated here should not be confused with the number of base-pairs unwound per ATP molecule hydrolyzed, although it is possible that the two are equivalent. Nor is the kinetic unwinding step-size necessarily the same as the physical translocation step-size (see further discussion below). As yet, we do not know the efficiency of the coupling of ATP hydrolysis to RecBCD-catalyzed DNA unwinding under our conditions. Steady-state studies of ATP hydrolysis during DNA unwinding by RecBCD[16,35] estimated a macroscopic coupling of two to three ATP molecules per base-pair unwound, averaged over the course of unwinding several thousand base-pairs. However, more detailed pre-steady-state kinetic experiments will

be needed to determine which steps in the process are coupled to ATP binding and hydrolysis. In fact, it is possible that the actual process of base-pair melting is not coupled to either ATP binding or hydrolysis. The footprinting experiments of Farah & Smith[27] indicate that the binding of RecBCD to a duplex DNA blunt end results in approximately 5–6 bp becoming sensitive to potassium permanganate. This perturbation of the duplex end is dependent upon Mg^{2+}, but does not require ATP. Therefore, ATP binding/hydrolysis may not be involved in the actual unwinding reaction, but rather may be needed for translocation, or possibly to increase the rate of an otherwise slow step in the cycle. In this regard, it is certainly interesting that the perturbation size estimate of ~5–6 bp upon binding of RecBCD to the end of the duplex is similar to our estimate of a kinetic step-size of 3–5 bp unwound.

Interestingly, the simplest scheme that might describe such an unwinding mechanism, i.e. Scheme 1, does not describe the time courses for RecBCD-catalyzed DNA unwinding quantitatively. Rather, two additional steps, with rate constant, k_C, are required somewhere within the n-step sequential mechanism. However, based on our current information, we cannot determine either what these steps represent or at what point these two steps occur within Scheme 3, aside from the fact that they must occur after the $(RD)_L$ state and before the formation of completely unwound ssDNA product. Since the value of $k_C = 29(\pm 3)\ s^{-1}$ is independent of both duplex length and [ATP] (A.L.L., unpublished experiments), it seems most likely that these two steps occur either at the beginning or even the end of the scheme, as shown in the two alternate forms of Scheme 3; however, our analysis does not allow us to conclude even this much. Attempts to fit the time courses to a model in which these two additional steps have different rate constants (e.g. k_{C1} and k_{C2}), always result in a fit where the additional steps float to very similar values. Hence, the simplest mechanism to consider is that these two steps have the same rate constant, k_C. Further experiments are needed to resolve the identity of these two steps, as well as their position within Scheme 3. On a comparative note, it is interesting that the time courses for the single turnover kinetic studies of *E. coli* UvrD helicase-catalyzed DNA unwinding could be described by the simpler Scheme 1,[28] whereas the single turnover time courses for DNA unwinding by the vaccinia virus NPH-II RNA helicase required inclusion of one additional [ATP]-independent step.[36]

Comments on the use of chemical quenched-flow techniques to determine a kinetic step-size for helicase-catalyzed DNA unwinding

Here, we have used chemical quenched-flow methods to obtain single turnover kinetic time courses for RecBCD-catalyzed DNA unwinding of radioactively labeled DNA substrates. As is evident from our discussion, the analysis of this data to obtain a minimal kinetic mechanism that can quantitatively describe the time courses as well as to estimate values for the rate constants and kinetic step-size was not simple. It is therefore worth commenting on the general utility of such experiments and the cautions one must take in their quantitative analysis to determine a kinetic step-size. We first state the obvious point that a quenched-flow experiment yields a finite number of data points. In general, our experiments were designed to obtain ~25–30 points for each time-course. How well can one estimate kinetic parameters from a series of such experiments performed as a function of DNA duplex length? Although there is no general answer to this question, it clearly becomes easier to exclude kinetic mechanisms when more points are obtained for each time-course and as more duplex lengths are examined. Furthermore, the uncertainties associated with any fitted kinetic parameters also decrease. However, the ability to estimate a kinetic unwinding step-size also depends on the complexity of the kinetic mechanism needed to describe the unwinding time courses. For example, fewer experiments are needed to determine a step-size within the same degree of uncertainty if the data can be described by the more simple Scheme 1 rather than by Scheme 3. In fact, when additional slow steps are present in a mechanism (e.g. Scheme 3), the observed increase in the lag phase per additional step, n, is smaller than when all steps are equivalent. That is, when additional slow steps are present, the unwinding time courses for different duplex lengths become less distinguishable. In the limit when these steps become very slow and thus completely rate limiting, one will only observe a single exponential time-course, with no lag, independent of duplex length, as observed for vaccinia NPH-II in the presence of Mg^{2+}.[36]

In our studies, we performed five independent sets of RecBCD-catalyzed DNA unwinding experiments on five DNA substrates with duplex lengths of 24, 30, 40, 48 and 60 bp. Upon overlaying the data for the five different experiments in Figure 4 we conclude that they are the same within our experimental error. However, global analysis of each independent set of time courses using Scheme 2 yielded quite varied estimates of the kinetic DNA unwinding step size, m: these being $6.1(\pm 2.1)$, $3.6(\pm 1.9)$, $3.5(\pm 1.4)$, $2.2(\pm 1)$ and $8.8(\pm 3.3)$ bp step^{-1} (see Table 1). This large variability in the determinations of the step-size is due to the limited information in each data set, the complexity of the model (seven parameters), and the parameter correlation that exists between k_U and n and between k_C and h. Clearly, if we had performed and analyzed only one set of experiments, then a much higher uncertainty would be associated with both the kinetic parameters and the step-size.

These conclusions are supported by two pieces of evidence. The correlation matrix for a given fit indicates that the unwinding rate constant, k_U, and

the step-size, m, are negatively correlated (correlation coefficient $> 90\%$), while the additional step rate constant, k_C, and the value of h are positively correlated (correlation coefficient $> 90\%$). We have also demonstrated this through analysis of a series of simulated time courses. In this exercise, we simulated a perfect set of data for a known model and then introduced small amounts of random error in these data. We then determined the variability in the fitted parameters when a limited set of replicas were available. Simulated time courses (22 points each) were generated using Scheme 3 for duplex lengths that were the same as those used in our experiments (NB Scheme 3 is identical to Scheme 2 with $h = 2$). We then incorporated 4% Gaussian error randomly to each point and subsequently fit these data to Scheme 2 (equations (7) and (8)). The 4% Gaussian error was chosen based on the fact that we observe $\sim 4\%$ error between our experimental data points and the best fit determined from our non-linear least-squares fitting. The values of the fitted parameters for both the simulated (Table 3) and experimental data (Table 2) showed similar variability. Furthermore, as expected, upon performing a NLLS fit to the average of the five simulated sets of data, we obtained parameters that were much closer to those used to generate the perfect data set. Naturally, if one simulates perfect data and performs a NLLS fit on the perfect data, the exact parameters are obtained. Furthermore, if one averages enough independent sets of data with 4% random Gaussian noise, analysis of these averaged data will eventually yield the original set of parameters. As such, we conclude that our best estimate of the parameters for the kinetic mechanism for RecBCD catalyzed DNA unwinding comes from the NLLS fitting of the averaged data set (Figure 7(a)) because these data points are the closest approximation to the "true" unwinding time courses. Therefore, we report the unwinding step-size for RecBCD as $3.9(\pm 1.3)$ bp step^{-1} from the fitting of the average of the five data sets, under these solution conditions (buffer M, 25 °C), although it will be of interest to examine further any potential effects due to changes in solution conditions. These results point to the need to perform multiple sets of experiments when using chemical quenched-flow techniques to determine a step-size from a complex mechanism.

Rates and processivity of RecBCD-catalyzed DNA unwinding

In addition to an estimate of the DNA unwinding step-size, our pre-steady-state kinetic studies provide estimates of DNA unwinding rates. The average DNA unwinding rate under saturating ATP concentrations is $k_U = 196(\pm 77)$ steps s^{-1} ($mk_U = 790(\pm 23)$ bp s^{-1}), (see Table 2). In assessing these values, it is important to note that any of our estimates of k_U, in units of steps s^{-1}, is negatively correlated to our estimate of the unwinding step-size, m, for that data set; thus, the individual estimates of k_U and m have larger uncertainties than their product, mk_U, which represents the average unwinding rate (bp s^{-1}). The rates determined here compare well with those determined from previous studies, although the solution conditions differ somewhat, and these differences can potentially influence both the rate and processivity of DNA unwinding.[22,23] Using electron microscopy, Taylor & Smith[21] estimated a rate of DNA unwinding of ~ 300 bp s^{-1} at 37 °C in 5 mM Mops (pH 7.0), 100 mM NaCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 5 mM ATP. Using a fluorescence assay monitoring the binding of *E. coli* SSB protein to the ssDNA formed upon unwinding, Roman & Kowalczykowski[22,23] examined both the rates and processivities of RecBCD unwinding under a variety of conditions and showed that these are sensitive to temperature, [NaCl] and [$MgCl_2$]. Under conditions that most closely resemble those used in our study, they report rates of $250(\pm 20)$ bp s^{-1},[23] $420(\pm 84)$ bp s^{-1} (25 °C, 25 mM Tris-acetate (pH 7.5),[22] 1 mM magnesium acetate, 1 mM DTT, 1 mM ATP (with an ATP regenerating system)), and $502(\pm 243)$ bp s^{-1} determined in single molecule experiments[24] (23 °C, 41 mM $NaHCO_3$ pH 8.0, 35 mM DTT, 13% (w/v) sucrose, 2 mM magnesium acetate). At 25 °C (30 mM NaCl), a processivity of $P = 0.999967$ was also measured,[22] corresponding to the unwinding of an average of $(1 - P)^{-1} = 30(\pm 3)$ kb before dissociation of the helicase. Since we have used relatively short DNA substrates (24–60 bp), our experiments cannot be used to measure such a high processivity. However, our single turnover experiments show that the total unwinding amplitude remains constant with increasing duplex length, consistent with such a high processivity (Figure 7(a)).

Our experiments also bear on the form of the RecBCD enzyme required for helicase activity. In experiments where the total RecBCD and DNA concentrations are both 2 nM (before mixing with ATP) (Figure 8(a) and (b)), we observe that $\sim 50\%$ of the DNA molecules are unwound within the initial lag phase, while another $\sim 10\%$ is unwound in the second slower phase. Since these experiments were performed with a trap for free RecBCD protein added with the ATP, then no more than 50% of the DNA molecules could be unwound if a dimer of RecBCD hetero-trimers were required for helicase activity. Therefore, these results indicate that a single hetero-trimer of RecBCD has helicase activity, consistent with previous conclusions.[25]

DNA unwinding step-size *versus* translocation step-size

We have noted above that the unwinding kinetic step-size of 3–5 bp determined here for the RecBCD helicase is not necessarily equivalent to a physical step-size for translocation of the enzyme along the DNA. In one extreme limit, a series of smaller, rapid translocation steps could occur, followed by a slow rate limiting step (e.g. a protein

conformational change) every 3–5 bp. It is also possible that physical translocation could occur with a larger step-size, or in two or more steps with different step-sizes. In fact, a recent study has suggested that the RecBC helicase (minus the D subunit) can bypass ssDNA gaps as large as ~23 nt within a DNA duplex,[15] which is approximately equivalent to the length of duplex DNA protected in "footprint" experiments when RecBCD is bound to a duplex end.[26,27] This has been interpreted to suggest that the RecBC helicase can translocate in steps of ~23 bp, although since the ssDNA gaps in the DNA substrates used in those studies are quite flexible, the actual distance over which the enzyme might translocate could be considerably smaller than a contour length equivalent to 23 bp (~78 Å). Although it has not been established that the mechanisms of DNA unwinding are identical for the RecBCD and RecBC helicases, nonetheless, the fact that the DNA footprint of RecBCD, and possibly the translocation step-size, may be much larger than the unwinding step-size (3–5 bp) suggests that RecBCD may unwind DNA by a type of "inch-worm" mechanism as previously proposed.[15,27] Our results, if combined with those of Bianco & Kowalczykowski,[15] might suggest that the RecBCD helicase could undergo two separate translocation events. The leading DNA binding site of the enzyme might translocate with a large step-size, followed by translocation of the trailing (unwinding) DNA binding site of the helicase in a series of multiple smaller steps corresponding to the 3–5 bp unwinding step-size measured here.

We have considered the possibility that the additional two steps in Scheme 3, with rate constant k_C, might reflect this larger translocation step. However, we find that our kinetic time courses can be described using only one class of ATP-dependent repeating steps, with rate constant, k_U, that result in the unwinding of an average of 3–5 bp step^{-1}. Both the number of steps with rate constant k_C and the value of k_C are independent of duplex length and [ATP] (A.L.L. *et al*., unpublished results). Furthermore, our data are not described well by a model where a step with rate constant k_U repeats with step-size m_1 and a step with rate constant k_C repeats with a different step-size m_2. However, although our results do not present any direct evidence for a large translocation step-size, our experiments would not detect such a translocation step if its rate were much faster than the unwinding rate; hence, we cannot rule out the possibility that translocation may occur as a step separate from the unwinding step.

Related to this issue, Goldmark & Linn[17] characterized the DNA products from both the dsDNA and ssDNA exonuclease activities of the RecBCD enzyme and found that the average size of the products was ~4.5 nt, with 90% of the product distribution ranging from 3 nt to 6 nt with no products shorter than 3 nt. This fragment size is similar to the unwinding step-size that we measured for RecBCD. This could indicate that the nuclease site within the RecBCD enzyme (at least partly on the B subunit[37]) also moves in average steps of ~3–5 nucleotides and suggests that the unwinding step-size that we report here may correspond to a physical step-size taken by the enzyme during one phase of the translocation process.

RecBCD and UvrD helicases unwind DNA with similar step sizes

The unwinding step-size of 3–5 bp reported here for RecBCD-catalyzed DNA unwinding is the same, within our uncertainties, as the unwinding step-size estimated for *E. coli* UvrD helicase by Ali & Lohman.[28] It also appears to be the same as the unwinding step-size of ~6 bp estimated for the SF2 vaccinia virus NPH-II RNA helicase[36] using the same methodology developed by Ali & Lohman.[28] These results suggest that at least some aspects of the unwinding mechanisms of these helicases may be similar. An oligomeric form of the *E. coli* UvrD protein is required for optimal helicase activity *in vitro*[6] and recent experiments indicate that a dimer is required for DNA helicase activity (N. K. Maluf, C. J. Fischer & T.M.L., unpublished results). UvrD unwinds with 3′ to 5′ polarity,[38,39] although it can also initiate unwinding at nicks and blunt ends at high protein concentration.[39–41] Only two of the three RecBCD subunits (RecB[42] and RecD[43]) possess SF1 superfamily helicase motifs. Interestingly, the RecB protein alone can form oligomers and exhibits limited helicase activity with apparent 3′ to 5′ polarity,[14] the same unwinding polarity observed for the RecBC enzyme.[15] However, RecBCD and UvrD differ in the DNA substrates required to initiate DNA unwinding *in vitro*. Whereas RecBCD unwinds blunt-ended DNA readily, although it binds more tightly to DNA with a short 5′ ssDNA tail,[25] UvrD initiates poorly on a blunt-ended DNA[41] and shows optimal unwinding of a substrate with a 40 nucleotide 3′ ssDNA tail.[6] These differences suggest that RecBCD interacts with both single strands of the unwound duplex as well as the duplex,[26,27] whereas UvrD appears to interact with the duplex and only one of the single strands at the ss/dsDNA junction (the unwound 3′ single strand). However, both enzymes appear to unwind by active mechanisms in which the enzyme binds to duplex DNA (at ss/dsDNA junctions) and separates the two strands of the double helix. Therefore, although the unwinding step-size is similar for RecBCD and UvrD, other aspects of their mechanisms clearly differ.

Materials and Methods

Buffers

Buffers were prepared with reagent grade chemicals using twice distilled water, that was subsequently

deionized with a Milli-Q purification system (Millipore Corp., Bedford, MA). Buffer C is 20 mM KPO_4 (pH 6.8), 0.1 mM DTT, 0.1 mM EDTA, 10% (v/v) glycerol. Buffer M is 20 mM Mops-KOH, 10 mM $MgCl_2$, 30 mM NaCl, 5% (v/v) glycerol (pH 7.0 at 25 °C).

Proteins

E. coli RecBCD protein (fraction D) was purified as described[19,25,44] and stored at 4 °C in buffer C. RecBCD concentration was determined spectrophotometrically using an extinction coefficient of $\varepsilon_{280} = 4.5 \times 10^5\ M^{-1}\ cm^{-1}$ for the hetero-trimer in buffer C; all protein concentrations reported refer to the RecBCD hetero-trimer. This extinction coefficient was determined from comparison of the absorption spectra of several aliquots of denatured RecBCD protein in buffer C plus 6 M guanidinium–HCl with the absorption spectra of aliquots of the native RecBCD protein in buffer C (but replacing the 0.1 mM DTT with 1 mM 2-mercaptoethanol). The extinction coefficient of the RecBCD in 6 M guanidinium–HCl was calculated from the extinction coefficients of the individual aromatic amino acids (56 Trp, 75 Tyr, 110 Phe) in 6 M guanidinium–HCl as described.[45] The average of three determinations on three separate RecBCD stocks yielded $\varepsilon_{280} = 4.5(\pm 0.2) \times 10^5\ M^{-1}\ cm^{-1}$ for the extinction coefficient of the native RecBCD protein in buffer C. The specific activity (ATP-dependent ds exonuclease) is ~240,000 units mg^{-1}.[46]

E. coli SSB protein (kindly provided by Dr Alex Kozlov) was purified as described[47] with the addition of a dsDNA-cellulose column to remove a minor exonuclease contaminant. SSB protein concentration was determined spectrophotometrically in 10 mM Tris, 0.1 mM Na_3EDTA, 0.20 M NaCl (pH 8.1) using and an extinction coefficient of $\epsilon_{280} = 1.13 \times 10^5\ M^{-1}$ (tetramer) cm^{-1}.

Nucleic acids and DNA substrate design

The oligodeoxynucleotides used here were synthesized and purified by electroelution of the DNA from denaturing polyacrylamide gels.[48] The concentrations of the DNA strands were determined by spectrophotometric analysis of the mixture of mononucleotides resulting after phosphodiesterase I digestion of the DNA in 100 mM Tris–HCl, 3 mM $MgCl_2$(pH 9.2 at 25 °C).[49] A 16 bp hairpin DNA with the sequence (5′-GACTCGTTACCTGAGT-T_4-ACTCAGGTAACGAGTC) was used to trap any free RecBCD or RecBCD that dissociated during the course of unwinding. In the DNA substrates used for our DNA unwinding studies, the length of the DNA duplex region that closes the hairpin is 15 bp. The substrates used in our studies (Figure 1) are composed of three oligodeoxynucleotides, a constant bottom strand containing the hairpin, and two top strands, labeled A and B in Figure 1, that when annealed to the bottom strand form a fully base-paired duplex (with the exception of the T_4 loop in the hairpin), but which contains two nicks. Oligonucleotide A was labeled at its 5′ end with ^{32}P (denoted by an asterisk in Figure 1) using T_4 polynucleotide kinase (US Biochemical Corp., Cleveland, OH) and purified as described.[48] The sequences of the A oligonucleotides are given in Figure 1. A stock solution of each DNA substrate (0.1 μM DNA molecules) was prepared by mixing (5′-^{32}P)-labeled A oligonucleotide with an equal concentration of the B oligonucleotide and a 25% excess of the bottom strand in (buffer M), followed by heating to 90 °C in a one liter water bath for five minutes, followed by slow cooling to ~25 °C. An excess of the bottom strand was added to insure complete annealing of the radiolabeled strand. The sequence of the bottom strand to which all substrates given in Figure 1 were annealed, is: 5′-AGATCC-TAGTGCAGGTTTCCTGCACTAGGATCTGGATGCG-ACTCGACGTATCCATGGAGCATAAGATCCTAGTTT-CATCCTTTAGGCTACTGCAGCTAGCTCAGGAGCCA-TGG.

Chemical quenched-flow kinetics of DNA unwinding

Rapid chemical quenched-flow experiments were carried out using a three syringe, pulsed quenched-flow apparatus (KinTek RQF-3, University Park, PA). All experiments were performed at 25 °C, using RecBCD dialyzed into the final buffer conditions (buffer M) on the day of use. RecBCD protein at twice the final concentration was pre-mixed with the DNA substrate (2 nM) in buffer M containing 6 μM bovine serum albumen that had been dialyzed into buffer M. This mixture was incubated on ice for 20 minutes and then loaded in one loop (30 μl) of the quenched-flow apparatus. The other loop (30 μl) contained ATP (twice the final concentration) and a large excess of non-radioactive 16 bp DNA hairpin (10 μM). These samples were pre-incubated in the loops for three minutes at 25 °C prior to performing the quenched-flow experiment. Increasing the incubation time to five minutes had no effect on the kinetic time-course (data not shown). The excess non-radioactive 16 bp DNA hairpin serves as a trap to prevent initiation of unwinding by any free RecBCD present at the start of the reaction or re-initiation of unwinding by any RecBCD that might dissociate from the ^{32}P-labeled DNA substrate during the course of the unwinding reaction. An identical time-course was observed when a higher concentration (25 μM final concentration) of the 16 bp hairpin trap was used. Furthermore, no unwinding of the ^{32}P-labeled DNA substrate occurred when excess 16 bp hairpin (10 μM) was pre-incubated with the ^{32}P-labeled DNA substrate (2 nM) and ATP, followed by addition of RecBCD helicase, insuring that we were monitoring single turnover unwinding time courses. The final ATP concentration in all experiments was 5 mM.

DNA unwinding reactions were initiated by rapidly mixing the two solutions in the quenched-flow instrument for a period of time, Δt, after which the reaction was quenched by the addition of 0.4 M EDTA and 10% glycerol (final concentrations: 0.3 M EDTA and 7% glycerol). Each time-course was constructed from a series of reactions, performed by varying Δt from 1 ms to 40 seconds. The zero time-point was determined in an identical manner as described above, but in the absence of ATP. Each quenched sample was stored on ice (for <one hour) until all samples were collected and then the samples were subjected to non-denaturing gel electrophoresis in a 10% (w/v) polyacrylamide gel to separate the duplex from the unwound ssDNA. A solution of 10% (w/v) SDS (1.25% final) was added to each sample to denature any protein that remained bound to the substrate after quenching. We have verified that no re-annealing of the unwound DNA occurred between quenching the reaction and running of the gel due to the low (1 nM) DNA concentrations used (data not shown). (NB The concentration of DNA after quenching is 0.3 nM.) After quenching, only duplex and fully

unwound DNA exists in the sample; hence, this is an all or none assay.

The radioactivity in each band was quantitated using a Molecular Dynamics Storm 840 system (Molecular Dynamics, Sunnyvale, CA) and the fraction of duplexes unwound, $f_{ss}(t)$, was calculated using equation (1) as described:[28]

$$f_{ss}(t) = \frac{\dfrac{C_S(t)}{C_S(t)+C_D(t)} - \dfrac{C_{S,0}}{C_{S,0}+C_{D,0}}}{1 - \dfrac{C_{S,0}}{C_{S,0}+C_{D,0}}} \quad (1)$$

where $C_S(t)$ corresponds to the radioactive counts for the unwound ssDNA (sum of the three fastest migrating bands (see Figure 1)), and $C_D(t)$ corresponds to the radioactive counts for the native, duplex DNA (sum of the two slowest migrating bands, consisting of fully duplexed DNA and duplex DNA minus the oligonucleotide (see Results)). $C_{S,0}$ and $C_{D,0}$ are the corresponding quantities at $t = 0$.

To examine the reproducibility of our experiments, we performed five independent sets of experiments with the series of DNA substrates that vary in duplex length. For each panel in Figures 4 and 6 the series of five DNA substrates (I–V in Figure 1) was end-labeled with ^{32}P and single turnover DNA unwinding experiments were performed on each substrate on a single day. Each time-course shown in Figure 7 represents the average of the five independent time courses determined for each duplex length.

Analysis of DNA unwinding kinetics

Single turnover DNA unwinding time courses were initially analyzed using the n-step sequential mechanism shown in Scheme 1.[28] In this scheme, $(RD)_L$ represents pre-formed productive RecBCD–DNA complexes, which can initiate DNA unwinding rapidly upon addition of ATP. DNA unwinding proceeds in steps by unwinding an average of m bp step^{-1} with rate constant k_U to form the partially unwound DNA intermediates, I_{L-m}, I_{L-2m}, I_{L-3m}, etc., before eventually forming fully unwound ssDNA. In Scheme 1, L is the total number of base-pairs in the duplex DNA to be unwound, n is the number of steps required to fully unwind the L bp, m is the step-size (bp unwound per step $= L/n$). The subscripts indicated for each intermediate, i, refer to the number of intact base-pairs remaining to be unwound in each partially unwound duplex. Scheme 1 also incorporates the fact that some fraction of the DNA is initially bound to RecBCD in non-productive complexes, $(RD)_{NP}$, that must first slowly isomerize, with net rate constant k_{NP}, to form the productive $(RD)_L$ complexes before initiation of DNA unwinding can occur. RecBCD could, in principle, dissociate with rate constant, k_d, at each intermediate step along the unwinding pathway. However, we have not included any dissociation steps in our analysis, since RecBCD is known to be highly processive under these conditions,[22] consistent with the fact that we observe no decrease in unwinding amplitude as a function of increasing duplex length. However, we do include an amplitude term (A_T) to account for the fact that we observe unwinding of only 88–92% of the DNA molecules.

The analytic expression describing the time-dependent formation of the fraction of unwound ssDNA molecules, $f_{ss}(t)$, for Scheme 1 is given in equation (2):

$$f_{ss}(t) = A_T\left[\left(1 - \sum_{r=1}^{n}\frac{(k_U t)^{r-1}}{(r-1)!}e^{-k_U t}\right) - e^{-k_{NP}t}(1-x) \times \left(\frac{k_U}{k_U - k_{NP}}\right)^n \left(1 - \sum_{r=1}^{n}\frac{((k_U - k_{NP})t)^{r-1}}{(r-1)!} \times e^{-(k_U - k_{NP})t}\right)\right] \quad (2)$$

where x is the fraction of productively bound complexes given by equation (3) and A_T is the total unwinding amplitude for a given time-course:

$$x = \frac{(RD)_L}{(RD)_L + (RD)_{np}} \quad (3)$$

However, one limitation of equation (2) is that it is only useful for integer values of n and thus also for integer values of L/m. Since one goal of our analysis is to determine the kinetic step-size, m, it is useful to allow the value of m to float to non-integer values, since this would greatly facilitate NLLS of the time courses and the estimation of uncertainties in the fitted values of m.

To overcome the limitations of equation (2), we have implemented two modifications (A.L.L. *et al.*, unpublished results). The first is to make the substitution shown in equation (4) for the product of each Taylor series multiplied by an exponential:

$$\sum_{r=1}^{n}\frac{(z)^{r-1}}{(r-1)!}e^{-z} = \frac{\Gamma(n,z)}{\Gamma(n)} \quad (4)$$

where $\Gamma(n, z)$ is the incomplete gamma function of n and z, and $\Gamma(n)$ is the gamma function of n.[50] In the case of Scheme 1, equation (2) can be rewritten as equation (5):

$$f_{ss}(t) = A_T\left[\left(1 - \frac{\Gamma(n,(k_U)t)}{\Gamma(n)}\right) - e^{-k_{NP}t}(1-x) \times \left(\frac{k_U}{k_U - k_{NP}}\right)^n \left(1 - \frac{\Gamma(n,(k_U - k_{NP})t)}{\Gamma(n)}\right)\right] \quad (5)$$

In fact, equation (5) represents the general solution for $f_{ss}(t)$, for Scheme 1 that is valid for all values of n, including non-integer values, whereas equation (2) is valid only for integer values of n. However, this substitution is only useful when a closed form expression for $f_{ss}(t)$ is obtainable, this is not the case for the more complex mechanism given by Scheme 2 which contains h additional steps, each with rate constant, k_C. Therefore, we have used the more general method of Laplace transforms to solve the system of differential equations describing the mechanism under consideration (e.g. Schemes 1 and 2, or even more complex schemes) (A.L.L. *et al.*, unpublished results). This method yields a solution for $F_{ss}(s)$, where $F_{ss}(s)$ is the Laplace transform of $f_{ss}(t)$. The advantage of this approach is that one can always obtain a closed-form expression for $F_{ss}(s)$. The resulting expression for $F_{ss}(s)$ for Scheme 1 is given in equation (6):

$$F_{ss}(s) = \frac{k_U^{L/m}(k_{NP} + sx)}{s(k_{NP} + s)(k_U + s)^{L/m}} \quad (6)$$

whereas the resulting expression for $F_{ss}(s)$ for Scheme 2

is given in equation (7):

$$F_{ss}(s) = \frac{k_C^h k_U^{L/m}(k_{NP} + sx)}{s(k_C + s)^h(k_{NP} + s)(k_U + s)^{L/m}} \quad (7)$$

Note that Scheme 2 is identical to Scheme 1 when $h = 0$, and that equation (7) reduces to equation (6) when $h = 0$. One can then obtain $f_{ss}(t)$ by taking the inverse Laplace transform of $F_{ss}(s)$, as indicated in equation (8), where $\mathscr{L}^{-1}$ indicates the inverse Laplace transform operator:

$$f_{ss}(t) = A_T(\mathscr{L}^{-1}[F_{ss}(s)]) \quad (8)$$

For complex models (e.g. Scheme 2), where it is difficult to obtain a closed form expression for $f_{ss}(t)$, the inverse Laplace transform can be obtained using numerical integration packages available in various programs such as Mathematica (Wolfram Research, Champaign IL), IMSL C numerical Libraries (Visual Numerics, Houston, TX), or Scientist (Micromath, St. Louis, MO) (A.L.L. *et al.*, unpublished results). As such, one can perform a non-linear least-squares fit to the data directly using equation (8). This approach allows for non-integer values of n, and thus also $m = L/n$, and therefore one can treat the step-size, m, as a fitting parameter that can take on any value. The time courses given in Figure 4 were globally fit to Scheme 1 using equation (6) to yield estimates of k_U, k_{NP}, m, and x and the time courses shown in Figures 6–8 were globally fit to Scheme 2 using equation (7) to yield estimates of k_U, k_{NP}, m, h, k_C and x.

The NLLS analyses of the time courses given in Figure 4 to yield the information in Figure 5 were performed by replacing L/m in equation (7) with n. In this analysis, h in equation (7) was constrained to equal 0, 1, 2 or 3, while n was allowed to float for each duplex length. The number of steps, n, obtained from the fits of all five data sets in Figure 4 was then averaged for each duplex length, and the average n is plotted *versus* duplex length in Figure 5 for each constrained value of h.

The grid-search calculations resulting in Figure 7(b) were performed for a series of step sizes, m, ranging from 1 to 10 (in intervals of 0.1). In each calculation, m was constrained and the resulting SSR was obtained after optimization of the global non-linear least-squares fit to the series of time courses shown in Figure 7(a). This was then repeated for the indicated range of step sizes.

All non-linear least-squares analyses were performed using Conlin,[51] kindly provided by Dr Jeremy Williams. Fitting models were written in the programming language C and compiled with the Microsoft C++ 6.0 compiler on Windows NT 4.0 Workstation. The routine in Conlin used for the numerical inversion of the Laplace transform was purchased from Visual Numerics Incorporated (Houston, TX) and is contained within the IMSL C Numerical Libraries. All uncertainties reported represent 68% confidence limits (±1 standard deviation). The uncertainties on the parameters obtained from non-linear least-squares fitting were determined by performing a 50 cycle Monte Carlo simulation[52] using the built in routine available in Conlin.

Estimation of parameter uncertainties from fitting of simulated kinetic data

Individual NLLS analysis of each of the five independent sets of experimental time courses reported here, using Scheme 2, yielded values of the kinetic parameters that varied considerably among the five data sets. In order to assess the origin of this variability, we simulated a set of perfect unwinding time courses for duplexes of lengths 24, 30, 40, 48, and 60 bp using Scheme 2 with $k_U = 200\ s^{-1}$, $k_C = 35\ s^{-1}$, $k_{NP} = 1\ s^{-1}$, $x = 0.8$, $m = 5$ bp step^{-1}, and $h = 2$ steps. To each perfect datum, 4% random Gaussian error was added using the method described by Di Cera.[53] We chose 4% error because this represents the average deviation between the experimental data points and the simulated time-course obtained from the best fits shown in Figure 6 (i.e. the square root of the variance of each global fit is ~4% (see Table 2)). This was repeated to generate five independent sets of "noisy" simulated time courses. A NLLS analysis was then performed on each simulated data set using equation (7) to obtain a set of fitted kinetic parameters for each data set. The parameters obtained from each of the five NLLS fits are given in Table 3, along with those obtained from a NLLS fit to the time courses resulting from the average of the five simulated data sets. The results of this exercise indicate that the fitting of the averaged set of data yields parameter estimates that are closer to the values of the true input parameters than those obtained from the average of the parameters obtained from the individual fits. Furthermore, the observed variability in the parameters obtained from the NLLS fits to each set of simulated data (Table 3) mirrors very closely the variability observed from the NLLS fits to each independent set of experimental time courses. This indicates that each of the five experimental time courses falls within the distribution of time courses expected for data with ~4% random error in each point.

Acknowledgements

We thank G. Waksman, N. K. Maluf, C. Fischer, K. Brendza, J. Wong, & D. Levine for comments on the manuscript and T. Ho for synthesis and purification of the DNA substrates. We also thank D. J. Williams and Andi Moeglich for their help in implementing Conlin. This research was supported in part by NIH grants GM45948 (T.M.L.) and GM32194 and GM31693 (G.R.S.), NIH postdoctoral fellowship (F32 GM17014 to R.G.) and NIH training grant (5 T32 GM08492) for partial support of A.L.

References

1. Lohman, T. M. & Bjornson, K. P. (1996). Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* **65**, 169–214.
2. Matson, S. W. & Kaiser-Rogers, K. A. (1990). DNA helicases. *Annu. Rev. Biochem.* **59**, 289–329.
3. Patel, S. S. & Picha, K. M. (2000). Structure and function of hexameric helicases. *Annu. Rev. Biochem.* **69**, 651–697.
4. von Hippel, P. H. & Delagoutte, E. (2001). A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell*, **104**, 177–190.
5. Gorbalenya, A. E. & Koonin, E. V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419–429.

6. Ali, J. A., Maluf, N. K. & Lohman, T. M. (1999). An oligomeric form of *E. coli* UvrD is required for optimal helicase activity. *J. Mol. Biol.* **293**, 815–833.
7. Cheng, W., Hsieh, J., Brendza, K. M. & Lohman, T. M. (2001). *E. coli* Rep oligomers are required to initiate DNA unwinding *in vitro*. *J. Mol. Biol.* **310**, 327–350.
8. Mechanic, L. E., Hall, M. C. & Matson, S. W. (1999). *Escherichia coli* DNA helicase II is active as a monomer. *J. Biol. Chem.* **274**, 12488–12498.
9. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. (1999). Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75–84.
10. Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M. & Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell*, **90**, 635–647.
11. Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol. Rev.* **58**, 401–465.
12. Smith, G. R. (1990). RecBCD enzyme. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J., eds), pp. 78–98, Springer, Berlin.
13. Anderson, D. G. & Kowalczykowski, S. C. (1997). The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner. *Cell*, **90**, 77–86.
14. Boehmer, P. E. & Emmerson, P. T. (1992). The RecB subunit of the *Escherichia coli* RecBCD enzyme couples ATP hydrolysis to DNA unwinding. *J. Biol. Chem.* **267**, 4981–4987.
15. Bianco, P. R. & Kowalczykowski, S. C. (2000). Translocation step size and mechanism of the RecBC DNA helicase. *Nature*, **405**, 368–372.
16. Korangy, F. & Julin, D. A. (1994). Efficiency of ATP hydrolysis and DNA unwinding by the RecBC enzyme from *Escherichia coli*. *Biochemistry*, **33**, 9552–9560.
17. Goldmark, P. J. & Linn, S. (1972). Purification and properties of the recBC DNase of *Escherichia coli* K-12. *J. Biol. Chem.* **247**, 1849–1860.
18. Thaler, D. S. & Stahl, F. W. (1988). DNA double-chain breaks in recombination of phage λ and of yeast. *Annu. Rev. Genet.* **22**, 169–197.
19. Taylor, A. F. & Smith, G. R. (1985). Substrate specificity of the DNA unwinding activity of the recBC enzyme of *Escherichia coli*. *J. Mol. Biol.* **185**, 431–443.
20. Telander-Muskavitch, K. & Linn, S. (1982). A unified mechanism for the nuclease and unwinding activities of the recBC enzyme of *Escherichia coli*. *J. Biol. Chem.* **257**, 2641–2648.
21. Taylor, A. & Smith, G. R. (1980). Unwinding and rewinding of DNA by the RecBC enzyme. *Cell*, **22**, 447–457.
22. Roman, L. J., Eggleston, A. K. & Kowalczykowski, S. C. (1992). Processivity of the DNA helicase activity of *Escherichia coli* recBCD enzyme. *J. Biol. Chem.* **267**, 4207–4214.
23. Roman, L. J. & Kowalczykowski, S. C. (1989). Characterization of the helicase activity of the *Escherichia coli* RecBCD enzyme using a novel helicase assay. *Biochemistry*, **28**, 2863–2873.
24. Bianco, P. R., Brewer, L. R., Corzett, M., Balhorn, R., Yeh, Y., Kowalczykowski, S. C. & Baskin, R. J. (2001). Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature*, **409**, 374–378.
25. Taylor, A. F. & Smith, G. R. (1995). Monomeric RecBCD enzyme binds and unwinds DNA. *J. Biol. Chem.* **270**, 24451–24458.
26. Ganesan, S. & Smith, G. R. (1993). Strand-specific binding to duplex DNA ends by the subunits of the *Escherichia coli* RecBCD enzyme. *J. Mol. Biol.* **229**, 67–78.
27. Farah, J. A. & Smith, G. R. (1997). The RecBCD enzyme initiation complex for DNA unwinding: enzyme positioning and DNA opening. *J. Mol. Biol.* **272**, 699–715.
28. Ali, J. A. & Lohman, T. M. (1997). Kinetic measurement of the step size of DNA unwinding by *Escherichia coli* UvrD helicase. *Science*, **275**, 377–380.
29. Lahue, R. S., Au, K. G. & Modrich, P. (1989). DNA mismatch correction in a defined system. *Science*, **245**, 160–164.
30. Caron, P. R., Kushner, S. R. & Grossman, L. (1985). Involvement of helicase II (*uvrD* gene product) and DNA polymerase I in excision mediated by the uvrABC protein complex. *Proc. Natl Acad. Sci. USA*, **82**, 4925–4929.
31. Husain, I., van Houten, B., Thomas, D. C., Abdel-Monem, M. & Sancar, A. (1985). Effect of DNA polymerase I and DNA helicase II on the turnover rate of UvrABC excision nuclease. *Proc. Natl Acad. Sci. USA*, **82**, 6774–6778.
32. Dixon, D. A. & Kowalczykowski, S. C. (1993). The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell*, **73**, 87–96.
33. Bjornson, K. P., Amaratunga, M., Moore, K. J. M. & Lohman, T. M. (1994). Single-turnover kinetics of helicase-catalyzed DNA unwinding monitored continuously by fluorescence energy transfer. *Biochemistry*, **33**, 14306–14316.
34. Dohoney, K. M. & Gelles, J. (2001). χ-Sequence recognition and DNA translocation by single RecBCD helicase/nuclease molecules. *Nature*, **409**, 370–374.
35. Roman, L. J. & Kowalczykowski, S. C. (1989). Characterization of the adenosinetriphosphatase activity of the *Escherichia coli* RecBCD enzyme: relationship of ATP hydrolysis to the unwinding of duplex DNA. *Biochemistry*, **28**, 2873–2881.
36. Jankowsky, E., Gross, C. H., Shuman, S. & Pyle, A. M. (2000). The DExH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature*, **403**, 447–451.
37. Yu, M., Souaya, J. & Julin, D. A. (1998). The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **95**, 981–986.
38. Matson, S. W. (1986). *Escherichia coli* helicase II (urvD gene product) translocates unidirectionally in a 3′ to 5′ direction. *J. Biol. Chem.* **261**, 10169–10175.
39. Runyon, G. T. & Lohman, T. M. (1989). *Escherichia coli* Helicase II (UvrD) protein can completely unwind fully duplex linear and nicked circular DNA. *J. Biol. Chem.* **264**, 17502–17512.
40. Runyon, G. T., Bear, D. G. & Lohman, T. M. (1990). *Escherichia coli* Helicase II (UvrD) protein initiates DNA unwinding at nicks and blunt ends. *Proc. Natl Acad. Sci. USA*, **87**, 6383–6387.
41. Runyon, G. T. & Lohman, T. M. (1993). Kinetics of *Escherichia coli* helicase II-catalyzed unwinding of

fully duplex and nicked circular DNA. *Biochemistry*, **32**, 4128–4138.
42. Finch, P. W., Storey, A., Chapman, K. E., Brown, K., Hickson, I. D. & Emmerson, P. T. (1986). Complete nucleotide sequence of the *Escherichia coli* recB gene. *Nucl. Acids Res.* **14**, 8573–8582.
43. Finch, P. W., Storey, A., Brown, K., Hickson, I. D. & Emmerson, P. T. (1986). Complete nucleotide sequence of recD, the structural gene for the α subunit of Exonuclease V of *Escherichia coli*. *Nucl. Acids Res.* **14**, 8583–8594.
44. Amundsen, S. K., Taylor, A. F., Chaudhury, A. M. & Smith, G. R. (1986). *recD*: the gene for an essential third subunit of exonuclease V. *Proc. Natl Acad. Sci. USA*, **83**, 5558–5562.
45. Lohman, T. M., Chao, K., Green, J. M., Sage, S. & Runyon, G. (1989). Large-scale purification and characterization of the *Escherichia coli* rep gene product. *J. Biol. Chem.* **264**, 10139–10147.
46. Eichler, D. C. & Lehman, I. R. (1977). On the role of ATP in phosphodiester bond hydrolysis catalyzed by the RecBC deoxyribonuclease of *Escherichia coli*. *J. Biol. Chem.* **252**, 499–503.
47. Lohman, T. M., Green, J. M. & Beyer, S. (1986). Large-scale overproduction and rapid purification of the *E. coli ssb* gene product. Expression of the *ssb* gene under P_L control. *Biochemistry*, **25**, 21–25.
48. Wong, I., Chao, K. L., Bujalowski, W. & Lohman, T. M. (1992). DNA-induced dimerization of the *Escherichia coli* Rep helicase allosteric effects of single-stranded and duplex DNA. *J. Biol. Chem.* **267**, 7596–7610.
49. Holbrook, J. A., Capp, M. W., Saecker, R. M. & Record, M. T., Jr (1999). Enthalpy and heat capacity changes for formation of an oligomeric DNA duplex: interpretation in terms of coupled processes of formation and association of single-stranded helices. *Biochemistry*, **38**, 8409–8422.
50. Kreyszig, E. (1993). *Advanced Engineering Mathematics*, Wiley, New York.
51. Williams, D. J. & Hall, K. B. (2000). Monte Carlo applications to thermal and chemical denaturation experiments of nucleic acids and proteins. *Methods Enzymol.* **321**, 330–352.
52. Straume, M. & Johnson, M. L. (1992). Monte Carlo method for determining complete confidence probability distributions of estimated model parameters. *Methods Enzymol.* **210**, 117–129.
53. Di Cera, E. (1992). Use of weighting functions in data fitting. *Methods Enzymol.* **210**, 68–87.

Edited by D. E. Draper

The manuscript was first submitted June 2001, with a revised version received 23 September 2002

(Received 23 September 2002; accepted 26 September 2002)