3. J. A. Hoffmann, M. Lagueux, in *Comprehensive Insect Physiology, Biochemistry, and Pharmacology*, G. A. Kerkut, L. I. Gilbert, Eds. (Pergamon Press, Oxford, 1985), pp. 435–460.
4. V. M. Chavez, *et al.*, *Development* **127**, 4115 (2000).
5. W. S. Talbot, E. A. Swyryd, D. S. Hogness, *Cell* **73**, 1323 (1993).
6. M. Bender, F. B. Imam, W. S. Talbot, B. Ganetzky, D. S. Hogness, *Cell* **91**, 777 (1997).
7. M. Buszczak *et al.*, *Development* **126**, 4581 (1999).
8. T. Kozlova, C. S. Thummel, *Development* **129**, 1739 (2002).
9. T. Kozlova, C. S. Thummel, data not shown,
10. C. Antoniewski, B. Mugat, F. Delbac, J. A. Lepesant, *Mol. Cell. Biol.* **16**, 2977 (1996).
11. L. H. Frank, C. Rushlow, *Development* **122**, 1343 (1996).
12. M. L. Yip, M. L. Lamka, H. D. Lipshitz, *Development* **124**, 2129 (1997).
13. D. P. Kiehart, C. G. Galbraith, K. A. Edwards, W. L. Rickoll, R. A. Montague, *J. Cell. Biol.* **149**, 471 (2000).
14. L. Cherbas, X. Hu, I. F. Zhimulev, E. Belyaeva, P. Cherbas, *Development* **130**, 271 (2003).
15. A. Garen, L. Kauvar, J.-A. Lepesant, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5099 (1977).
16. Materials and methods are available as supporting material on *Science* Online.
17. R. Clifford, T. Schupbach, *Development* **115**, 853 (1992).
18. M. L. Lamka, H. D. Lipshitz, *Dev. Biol.* **214**, 102 (1999).
19. F. Schöck, N. Perrimon, *Genes Dev.* **17**, 597 (2003).
20. C. E. Roote, S. Zusman, *Dev. Biol.* **169**, 322 (1995).
21. M. Bownes, A. Shirras, M. Blair, J. Collins, A. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1554 (1988).
22. F. Schöck, N. Perrimon, *Dev. Biol.* **248**, 29 (2002).
23. J. T. Warren *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11043 (2002).
24. N. Perrimon, L. Engstrom, A. P. Mahowald, *Genetics* **111**, 23 (1985).
25. A. E. Oro, M. McKeown, R. M. Evans, *Development* **115**, 449 (1992).
26. N. Ghbeish *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3867 (2001).
27. L. von Kalm, D. Fristrom, J. Fristrom, *BioEssays* **17**, 693 (1995).
28. R. E. Ward, J. Evans, C. S. Thummel, *Genetics*, in press.
29. P. P. D'Avino, C. S. Thummel, *Dev. Biol.* **220**, 211 (2000).

30. We thank D. Brower, D. Hogness, A. Letsou, H. Lipshitz, F. Schöck, and the Bloomington Stock Center for fly stocks; D. Hogness for EcR antibodies; L. Cherbas for the *UAS-EcR-F645A* transformant; members of the Thummel lab, K. Clark, and A. Letsou for discussions and suggestions in the course of this work; and A. Bashirullah, A. Letsou, and R. Ward for critical comments on the manuscript. T.K. was a Research Associate and C.S.T. is an Investigator with the Howard Hughes Medical Institute.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1087419/DC1
Fig. S1
Materials and Methods
References

30 May 2003; accepted 21 August 2003
Published online 4 September 2003;
10.1126/science.1087419
Include this information when citing this paper.

Sequence-Dependent Pausing of Single Lambda Exonuclease Molecules

Thomas T. Perkins,[1]*‡ Ravindra V. Dalal,[2] Paul G. Mitsis,[3]† Steven M. Block[1,4]

Lambda exonuclease processively degrades one strand of duplex DNA, moving 5′-to-3′ in an ATP-independent fashion. When examined at the single-molecule level, the speeds of digestion were nearly constant at 4 nanometers per second (12 nucleotides per second), interspersed with pauses of variable duration. Long pauses, occurring at stereotypical locations, were strand-specific and sequence-dependent. Pause duration and probability varied widely. The strongest pause, GGCGAT TCT, was identified by gel electrophoresis. Correlating single-molecule dwell positions with sequence independently identified the motif GGCGA. This sequence is found in the left lambda cohesive end, where exonuclease inhibition may contribute to the reduced recombination efficiency at that end.

Exonucleases are integral components of many genetic repair and recombination pathways (*1*). Processive 5′→3′ exonucleases generate DNA intermediates with long 3′-overhangs involved in these pathways (*1*, *2*); in bacteriophage lambda, the Red α gene encodes the 24-kD subunit of such a nuclease (*3*). The structure of lambda exonuclease (λ-exo) consists of a toroidal homotrimer of subunits, with a tapered central channel sufficient to admit double-stranded DNA (dsDNA) at its entrance but only wide enough to pass single-stranded DNA (ssDNA) through its exit (*4*). A topological constraint may therefore underlie the high processivity of λ-exo (>3000 base pairs, where 1 bp corresponds to a helical rise of 0.338 nm along dsDNA) (*5*). λ-exo digestion requires Mg^{2+}, but does not require ATP or GTP, and proceeds along DNA at rates of 2 to 3.5 nm/s, as determined by bulk biochemical studies (*6*, *7*).

Here, we developed a novel high-resolution, single-molecule assay to study λ-exo motion by optical trapping nanometry. Besides measuring individual translocation rates, single-molecule studies provide insights into other processes that

[1]Department of Biological Sciences, [2]Department of Physics, and [4]Department of Applied Physics, Stanford University, Stanford, CA 94305, USA. [3]Praelux, Inc., Lawrenceville, NJ 08646, USA.

*Present address: JILA, National Institute of Standards and Technology and University of Colorado, Boulder, CO 80309–0440, USA.
†Present address: Amersham Biosciences, 800 Centennial Road, Piscataway, NJ 08855, USA.
‡To whom correspondence should be addressed. E-mail: tperkins@jila.colorado.edu

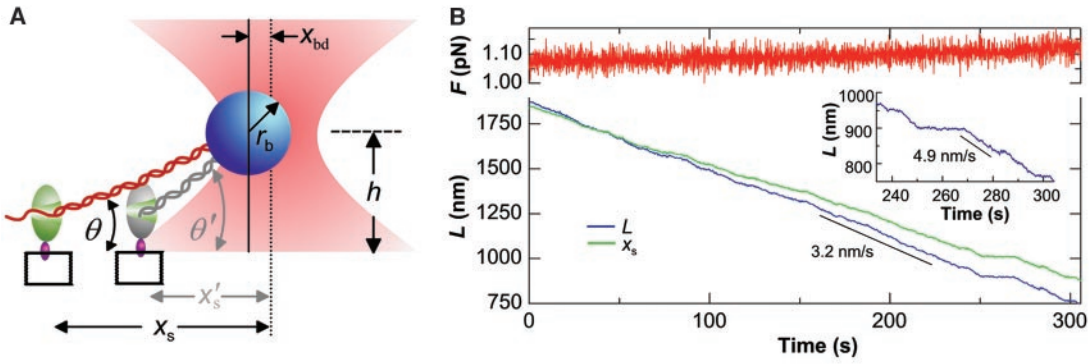


Fig. 1. (**A**) Experimental geometry of the stage-based optical force clamp (not to scale); bead displacement relative to the trap center (x_{bd}) was maintained by stage motions ($x_s \rightarrow x_s'$) that kept the load constant. The bead was positioned at a predetermined height above the coverslip; $h = 200$ nm + r_b, where r_b is the bead radius. (**B**) A representative record of λ-exo digestion in force-clamp mode over 5 min, during which the enzyme moved >1 μm. (Left) Force was clamped at 1.07 pN (red) and varied <5% (the small increase is due to changes in the vertical component of the force, which depends weakly on θ). (Right) Traces of stage motion (green) and the contour length, *L*, of the dsDNA remaining (blue), showing periods of steady movement interspersed with pauses. Inset, magnified view of DNA contour length over the interval 240 to 300 s that illustrates pauses of varying lengths. Line fits over the regions of uniform motion indicated (black lines) show the average velocity.

are obscured by bulk averages. For example, certain processive DNA enzymes pause for variable intervals, a behavior that may reflect off-pathway branches in the reaction cycle. Pausing behavior has been observed in single-molecule studies of transcriptional elongation by RNA polymerase (*8*), viral phage packaging by ϕ29 portal protein (*9*), and DNA unwinding by Rep helicase (*10*). The biological significance of pausing in such disparate enzyme systems is not yet well understood, however.

Our experimental assay consists of a surface-attached enzyme bound to a single DNA molecule (Fig. 1A) (*11*). A His-tagged version of λ-exo was preincubated with a 7.1-kbp substrate (derived from M13 viral DNA) in the presence of Ca^{2+}, which arrests the enzyme, and anchored stereospecifically with a His-specific antibody linkage to the cover glass surface inside a flow cell. Once Mg^{2+} was introduced into the buffer, digestion started and the enzyme moved processively. The DNA molecule was attached at its distal end to a small polystyrene bead, which formed a tether. The tethered bead was captured and held under tension with an optical trap, and its position was monitored by back focal plane interferometry (*12*). We implemented a stage-based force clamp using a computer-driven 3D piezoelectric stage that maintains the bead at a fixed offset from the trap center (*11*, *13*).

Motion of λ-exo was not constant along the substrate, but displayed frequent pauses (Fig. 1B) interspersed with periods of uniform movement. The locations and durations of these pauses could be accurately resolved by averaging data at 0.4 Hz under comparatively low load (1 pN) (Fig. 1B, inset). Higher loads, up to ~3 pN, had little effect on speeds (*14*).

When independent records of λ-exo motion were compared by eye, many of the longer pauses appeared at closely corresponding locations on the DNA substrate. A small vertical shift of individual traces sufficed to bring pauses at different DNA locations in separate records into alignment. To score pausing statistics and variations in digestion rate among individual λ-exo molecules, we focused on the region containing the dominant pause at ~900 nm (Fig. 2A) (*11*). The 900-nm pause and nearby weaker pauses were used as fiducial marks to align the corresponding locations for each record. To perform the alignment on 40 records, an average offset of −2.1 nm ± 12 nm (mean ± SD) was necessary; more than half (7 nm) of the offset variability (12 nm) is due to variation in the bead diameter (*11*). The alignment demonstrated that pausing was sequence-dependent and allowed us to compute the pause probability and duration for the most prominent pause locations (Table 1).

Among the possible explanations for pausing are changes in the local properties of the DNA structure, such as twists or bends (*15*); variations in local DNA energetics or stability, such as a series of A·T or G·C base pairs; and stereochemical interactions between the λ-exo enzyme and specific DNA sequences. If the structural or energetic properties of dsDNA alone dominated, pausing would be expected to occur at (or near) the same positions, regardless of the direction of digestion. To test this possibility, we engineered substrates with opposite polarity, so that digestion took place from a 5′-terminal phosphate on the complementary strand in the reverse direction. For such constructs, the dominant pause at 900 nm disappeared (Fig. 2B), and no significant correlation was found between pause locations in the forward and backward directions (*16*). This finding implies that pausing is strand-specific, as well as sequence-dependent, and eliminates a class of candidate mechanisms, including such symmetric measures as the G·C content. Notably, for the DNA region between 675 and 1200 nm (over which data were most plentiful), there were significantly more long pauses found for digestion of the sense strand (forwards) than for the nonsense strand (backwards) (*17*).

Is pausing a stochastic or a deterministic property of enzymatic activity? The wide variation from molecule to molecule in paus-

Table 1. Statistics for 11 pauses of various durations located between 600 and 1200 nm and identified by the pause-finding algorithm. Positions are reported as L_{nm} (±2.3 nm, SEM) and as L_{bp} (±7 bp, SEM); these are related through L_{bp} = 6406 bp − L_{nm}/0.338 nm/bp, where 6406 bp is the position of the digoxigenin-labeled primer. Pause probability (±SEM) is estimated from the ratio of traces pausing to all traces passing through a given site (*N*). Pause mean duration, $<\tau>$ (±SEM), is computed as the average of all nonzero pauses detected; relative pause strength, *RS* (±SEM), is computed as the strength of a given pause (strength = $<\tau>$ × probability) normalized to the pause at 900 nm. Pauses from different records that occurred within 5 nm of one another other (after realignment) were taken to represent the same location.

Location		Probability (%)	τ (s)	*RS*	*N*
L_{nm}	L_{bp}				
1191	2882	25 ± 9	3.5 ± 1.3	0.11 ± 0.06	24
1128	3069	21 ± 8	3.6 ± 0.9	0.10 ± 0.05	28
1050	3300	27 ± 8	4.6 ± 1.0	0.16 ± 0.06	33
1037	3339	21 ± 7	7.1 ± 1.3	0.19 ± 0.08	34
985	3491	19 ± 7	4.5 ± 1.5	0.11 ± 0.06	36
976	3517	17 ± 6	5.1 ± 2.0	0.11 ± 0.06	36
950	3597	25 ± 7	3.6 ± 0.5	0.12 ± 0.04	36
900	3742	69 ± 8	11.2 ± 1.5	1.00 ± 0.25	36
813	3999	14 ± 6	8.1 ± 3.9	0.15 ± 0.10	35
757	4167	23 ± 7	6.9 ± 2.6	0.20 ± 0.11	35
689	4368	39 ± 10	6.8 ± 2.7	0.34 ± 0.17	23

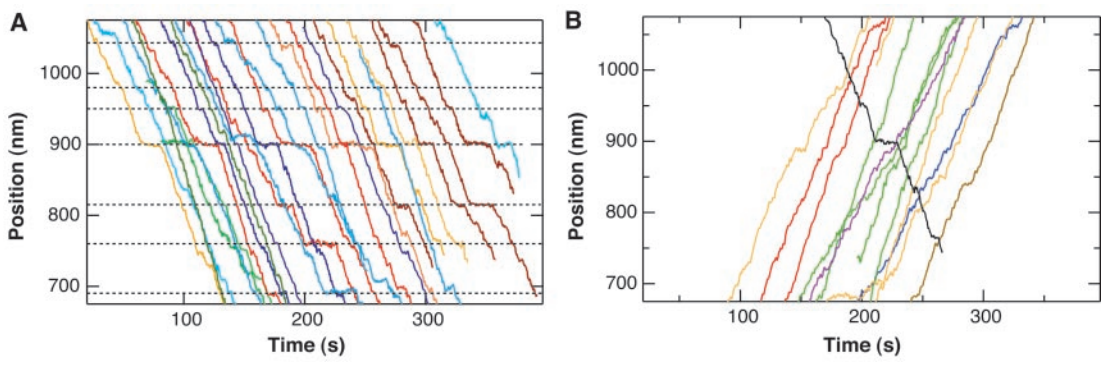


Fig. 2. Sequence specificity of pausing. (**A**) Records of digestion in the forward direction for 28 λ-exo molecules, aligned at the dominant pause at 900 nm. For clarity, only a selection of pause sites of various relative strengths (Table 1) are indicated (horizontal black lines). No alignment offset exceeded ±30 nm; the average offset was –2.1 ± 12 nm. (**B**) Records of digestion in the backward direction for 13 λ-exo molecules, displayed on the same scale as in (A). The dominant pause at 900 nm was not observed, nor were other common sites of pausing. A single trace showing digestion in the forward direction is overlaid for comparison (black). All trace origins were displaced horizontally by variable amounts of time to facilitate display of the records.

ing at characterized sites (Table 1) contrasts with the consistency in catalytic rates, reflected in the uniform speeds of digestion between pauses (Figs. 2A and 3A). This difference suggests that pausing is stochastic and is likely to represent an off-pathway state unrelated to any translocation events in the reaction cycle. The comparatively steady rate of λ-exo digestion also contrasts with the nearly fivefold variation in single-molecule speeds reported in experiments on RNA polymerase (*8*, *18*) and RecBCD (*19*, *20*). Is the propensity to pause a variable feature of each individual enzyme? It did not appear from an inspection of records that certain enzymes were more prone to pausing than others. Quantitatively, our data failed to display a correlation between the time spent at the dominant pause at 900 nm and at a pause at 950 nm (Fig. 3B), implying that pause durations at these two nearby sites are statistically independent. The histogram of durations for the pause at 900 nm is shown in Fig. 3C. Moreover, pausing appears to be independent of the surface immobilization; if bound enzymes paused more frequently, their average rates of digestion would be slower. The mean speed of single-molecule digestion (including pauses) was 3.3 ± 0.5 nm/s (10 ± 2 nucleotides (nt) per s) [mean ± SD; N = 40]), which agrees within error with the digestion rate of 12 nt/s reported in a recent biochemical study under similar buffer and temperature conditions (*7*).

To compare records with specific DNA sequences, we analyzed products of λ-exo digestion at steady state in the presence of an excess of 3′-^{32}P–labeled DNA by high-resolution polyacrylamide gel electrophoresis (*11*). In gel assays, the strand to be digested carried a 5′-phosphate, whereas digestion of the complementary strand was blocked by a 5′-OH group (*6*). Bands are produced at every base position, and pauses are characterized by darker bands; the density of any given band is related to the pause strength, which is the product of the duration and probability of pausing.

First, we set out to identify DNA sequences corresponding to the pauses identified in single-molecule records. Duplex DNA molecules (~300 bp) for several regions carrying pause sites were generated by polymerase chain reaction (PCR) (table S1), digested, and run out on gels. Analysis revealed a single prominent band at nucleotide 3,738 in the M13 sequence (Fig. 4A), corresponding to a computed position of 902 nm along the DNA; this matches up extremely well against the dominant pause found at 900 nm. We estimate our positional error, after drift removal and fiducial alignment, at approximately 2 nm (6 bp) (*21*). The digestion product displayed a single band, rather than a series of adjacent bands, which implied that λ-exo paused at a unique position. None of the other briefer, sequence-dependent pauses detected in single-molecule records were resolved by DNA gel analysis; this may be attributable to the comparatively weaker relative pause strengths at these locations (Table 1) and the relatively poor signal-to-noise ratio (*S/N*) afforded by gel analysis (*S/N* = 5 ± 0.5 for the dominant 900-nm pause).

Next, we sought to characterize the determinants responsible for the dominant pause. Oligonucleotides were designed in which 20-bp regions immediately before (FlipA) or after (FlipB) the pause site were exchanged from one strand to the other (Fig. 4D; table S2). Because single-molecule assays showed that pausing is strand-specific, we reasoned that swapping the full sequence specifying a pause from one strand to its complement would cause a new band to appear on gels digested in the backward direction, with concomitant loss of the band in the forward direction. Upon digestion, the FlipB substrate showed the expected behavior, and the pause site was successfully transferred to its corresponding position on the opposite strand, whereas for the FlipA substrate, the pause site was not transferred between strands (Fig. 4B). This finding restricts the sequence specifying a pause to a 20-bp region ahead of the active site of the enzyme. To further isolate the minimal sequence, a third set of oligonucleotides was designed that flipped the central 18 bp of the sequence (Flip18). This also caused the pause to exchange strands (Fig. 4C). Collectively, these results implicate a comparatively short region located up to 9 bp ahead of the enzyme active site, within the sequence 5′-GGCGATTCT-3′. Models of DNA interacting with the enzyme crystal structure suggest that 11 bp may be enclosed within the central channel (*22*), and nuclease protection assays suggest that the footprint of the enzyme spans ~13 to 14 bp (*6*). Our data therefore support a model where the DNA bases located within the central channel interact directly with the protein, as opposed to alternative possibilities, for example, hairpin formation in the ssDNA digestion product or the effect of upstream dsDNA sequences.

Does the sequence determining the dominant pause—or some sequence related to it—correlate with other pauses found in single-molecule assays? To investigate this question, we performed a cross-correlation analysis of the sequence in the DNA sense strand (the digested strand) against a function based on the average of single-molecule records, in an effort to identify the relevant motif (*11*). A histogram of dwell times was produced by binning aligned traces over a range of contour lengths from 590 to 1760 nm, with the use of a 5-nm window. A simple integer-scoring function was adopted to compare the running DNA sequence against a given candidate for the pause sequence. A trial oligomer of n nucleotides was selected for the candidate, with n ranging from 4 to 9 bases. This n-nucleotide oligomer was then compared with the substrate base sequence at all positions, with a score of +1 awarded for each exact base match, 0 for a mismatch, and an additional penalty of –1 assessed for any transversion (purine-pyrimidine interchange) altering the number of base-pair hydrogen bonds (i.e., for A↔C and G↔T). The running score was exponentiated (pause times are exponentially related to the underlying energetics) and smoothed. Finally, it was cross-correlated against the dwell time histogram and used to calculate a correlation coefficient, r. The value computed by this procedure supplies a numerical measure of how well the candidate sequence predicts the experimentally observed distribution of dwell times. Correlation coefficients for all possible n–mer sequences were computed and ranked (table S3).

For the 9-nt oligomer GGCGATTCT corresponding to the dominant pause sequence identified by gel analysis, the correlation was quite high, r = 0.451; less than 1.5% of all 262,144 possible 9-nt oligomers produced a correlation exceeding even half this value (*11*). Of 1024 possible 5-nt oligomers, the sequence GGCGA

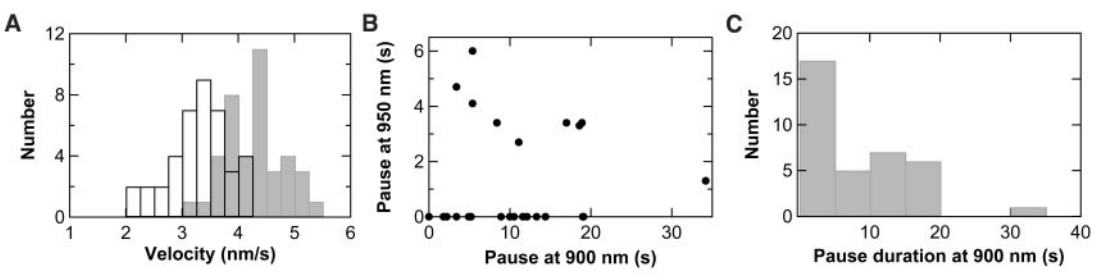
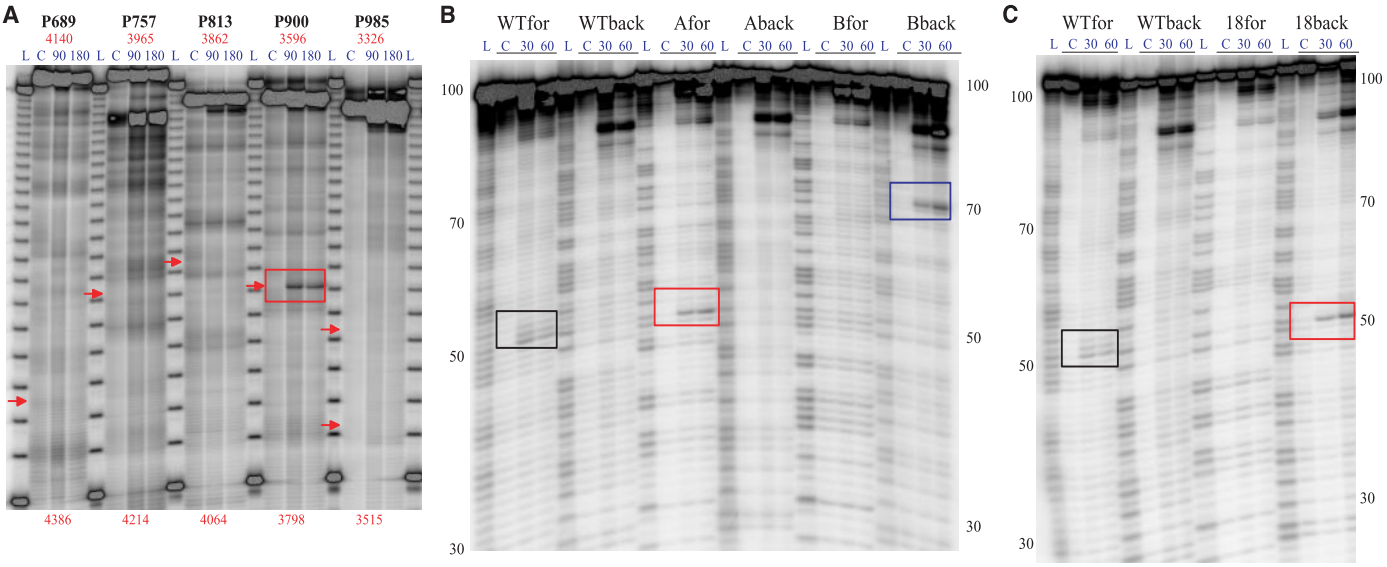


Fig. 3. Velocity and pause statistics. (**A**) Histogram of mean forward λ-exo digestion speeds, computed with (white bars) and without (gray bars) intervening pauses >1 s (N = 40). The average speeds with and without pauses were 3.3 ± 0.5 nm/s and 4.3 ± 0.5 nm/s, respectively. (**B**) Duration of the pause observed at 900 nm versus the duration of the pause at 950 nm for each molecule, showing a lack of correlation between times spent at these nearby locations (N = 35). (**C**) Histogram of durations for the dominant pause at 900 nm.

yielded the single largest correlation, $r = 0.296$. For each n-nucleotide oligomer size, we also examined the 1% of sequences producing the greatest correlation values, e.g., the top-ranked 2621 of all 9-nt oligomer sequences, etc. Among these best-correlated sets of sequences, we scored the most frequently occurring imbedded sequences (of length m, with $m \leq n$). The 5-bp motif GGCGA consistently appeared as either the first or the second most common imbedded sequence found in every set of 6-, 7-, 8-, and 9-nt oligomers, with no other imbedded sequences appearing consistently in the rankings. The same 5-bp motif was returned by a variant of this analysis, in which we substituted the pause strength at all sites identified by the pause-finding algorithm ($N = 124$) for the dwell-time histogram. This analysis identifies GGCGA as the 5-base motif statistically most correlated with pausing within the 1170-nm (~3.5-kb) interval studied (table S3). Of particular note, the correlation analysis relies solely on single-molecule data, without a priori knowledge of the site identified by gel assays. These complementary approaches independently suggest that determinants for pausing are contained within GGCGATTCT. The 5-nt oligomer GGCGA is only imperfectly correlated with dwell locations, however, which raises the likelihood that other sequence motifs may also contribute to pausing signals. Exploring alternative scoring functions, we found that pause locations failed to correlate with purine content ($r = 0.05$) or with the energetics of DNA stability ($r = 0.07$), and correlated only weakly with GC content alone ($r = 0.19$) (table S4).

What features of the consensus sequence GGCGA slow the enzyme? The initial triplet of the sequence, GGC, has a high melting temperature, second only to GCG among all 3-nt oligomers (*23*). However, thermal stability of the DNA alone is unlikely to supply the explanation, because of the following argument (and also owing to the failure of DNA stability and GC content to correlate well with pausing, as discussed). Gel analysis shows that the dominant pause occurs reproducibly at a single base and not over a distribution of several nearby bases. Because the enzyme cleaves mononucleotides, the most parsimonious explanation is that λ-exo translocation and DNA strand separation occur one base at a time. Based on the structural model (*4*), when the 5′-terminal G is positioned at the enzyme active site, located 15 to 18 Å from the



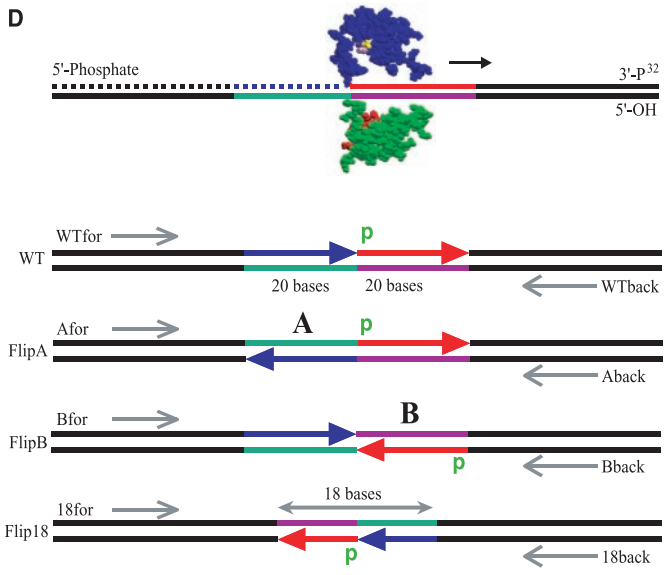


Fig. 4. Polyacrylamide gel analysis of pausing. (**A**) Exonuclease digestion of PCR products (~300 bp) flanking several pause sites identified in single-molecule records, designated P*nnn* atop each lane, where *nnn* is the nominal position, in nanometers. Numbers (red) above and below each lane specify the starting and ending bp of the M13 sequence. Lane designations (in blue): L, length standards in a 10-bp ladder; C, control lane without enzyme; 90, 180, digestion by enzyme for the times in seconds indicated. Arrows (red) show locations of pauses in single-molecule data. The dominant pause site at 900 nm corresponds to a strong band (lanes P900); other sites do not. (**B**) Gel analysis of the FlipA and FlipB constructs shown below and in table S2. Lane designations same as in (A), except that L is a purine ladder. The lane labeled WTfor displays the pause band at 900 nm (black box) for digestion in the forward direction only (compare with WTback). The bands in Afor and Aback show the identical pause location (red box). The band in Bback shows that the pause has been transferred to the backward strand (blue box) and is now absent from the forward strand (compare with Bfor); note that in its new orientation of the sequence, the location of the pause is offset. (**C**) Gel analysis of the Flip18 construct. Lanes WTfor and WTback are as in (B). Lanes 18for and 18back show that pause has been disrupted in the forward strand and successfully transferred to the backward strand (red box). (**D**) Flip-sequence experimental design. (Top) Cartoon showing enzyme orientation and strand digestion (dotted trace) starting from a unique 5′ phosphate (shown in the position appropriate for forward digestion); the pause region is displayed in several colors. Lower four drawings illustrate the flip sequences. Labels are as follows: pause base site, p (green); directions of digestion (black arrows); relative orientation of DNA sequences (red and blue arrows). WT, unmodified DNA sequence; FlipA and FlipB, sequences where 20 bp upstream or downstream of the pause site were swapped to the opposite strand, respectively; and Flip18, sequence where 18 bp on either side of the pause were swapped to the opposite strand.

central axis, ~2 to 5 terminal nucleotides would be unpaired at this point, i.e., they would be melted before the pause and not concomitant with it.

We instead favor a model where pausing is conferred by interactions between λ-exo residues and specific nucleotides in the DNA substrate. The unpaired bases located in the "frayed" segment of ssDNA postulated to exist between the central axis and the enzyme active site represent obvious candidates for such interactions. We speculate that residue Trp^{24} of the protein, located on helix αB lining the central channel and directly facing the catalytic site [residues 22 to 26; (*22*)], may intercalate between two adjacent guanosine bases in this segment and form a tight ring-stacking interaction. A similar stacking arrangement between DNA bases and aromatic amino acids has been postulated in the mechanism of translocation in helicase (*24*). The sequence GGCGA contains four purines that may potentiate stacking interactions and is sufficiently short to be entirely in single-stranded form inside the enzyme channel. Clearly, however, additional interactions between the ssDNA and the enzyme would be necessary to confer the observed degree of sequence specificity. A more complete understanding of such interactions awaits enzyme-DNA cross-linking studies or a cocrystal structure of λ-exo bound to a strong pause sequence, such as the one identified here.

We speculate that the propensity of exonuclease to pause at specific sequences may explain a longstanding puzzle about the inhomogeneity of recombination rates in bacteriophage lambda. When lambda recombination occurs in the absence of DNA replication—and is therefore dependent on the lambda Red system, of which λ-exonuclease is a principal component—the frequency of recombination is much lower at the left end of lambda than at the right (*25*). In addition, a strong, nested pause signal containing both GGCG and GGCGA is found within the first nine bases of left lambda cohesive end (cosNL), GGGCGGCGACCT (*26*). If the pausing of lambda exonuclease induced in vitro by these sites is reflected in its activity in vivo, then the slowing of digestion from the left end of lambda may singly, or in conjunction with lambda terminase (*27*), be able to account for the recombination deficit.

We anticipate that the improved spatiotemporal resolution achieved here, which made possible the nanometer-level identification of sequences, will not only facilitate further work on λ-exo, but also enable detailed studies of other DNA-based enzymes at the single-molecule level.

References and Notes

1. A. Kuzminov, *Microbiol. Mol. Biol. Rev.* **63**, 751 (1999).
2. S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, W. M. Rehrauer, *Microbiol. Rev.* **58**, 401 (1994).
3. C. M. Radding, *J. Mol. Biol.* **18**, 235 (1966).
4. R. Kovall, B. W. Matthews, *Science* **277**, 1824 (1997).
5. J. W. Little, *J. Biol. Chem.* **242**, 679 (1967).
6. P. G. Mitsis, J. G. Kwagh, *Nucl. Acids Res.* **27**, 3057 (1999).
7. K. Subramanian, W. Rutvisuttinunt, W. Scott, R. S. Myers, *Nucl. Acids Res.* **31**, 1585 (2003).
8. M. D. Wang *et al.*, *Science* **282**, 902 (1998).
9. D. E. Smith *et al.*, *Nature* **413**, 748 (2001).
10. T. Ha *et al.*, *Nature* **419**, 638 (2002).
11. Materials and methods are available as supporting material on *Science* Online.
12. K. Visscher, S. P. Gross, S. M. Block, *IEEE* (*Inst. Elect. Electron. Eng.*) *J. Sel. Top. Quant. Electr.* **2**, 1066 (1996).
13. A force clamp supplies improved resolution over open-loop systems and makes it unnecessary to apply corrections for the elastic compliance of the DNA-motor complex (*12*). Force clamps based on stage motion, as opposed to those that steer the optical trap by acousto-optic deflectors (*28*), are well suited to studying slow-moving, processive enzymes, because they afford a greater operating range (>10 μm) together with high precision (~2 nm) at the cost of reduced feedback loop closure time (~0.1 s).
14. The velocity was unchanged by a threefold increase in force from 1 pN (fig. S1A). This insensitivity to external load was observed for several molecules (fig. S1B) and implies that the mechanical translocation step of the reaction cycle is not rate-determining at low loads. A similar finding has been reported for RNA polymerase (*8*). Force-velocity analysis over a greater range was not possible because tethered complexes failed to support significant loads (>3 pN) for extended periods (~30 s).
15. E. N. Trifonov, *C.R.C. Crit. Rev. Biochem.* **19**, 89 (1985).
16. Comparing unaligned data sets, we found no significant correlation between the dwell time histograms for forward and backward directions of digestion (*11*). A cross-correlation analysis yielded a correlation coefficient $r = 0.050$ [−0.059, +0.159] over the region 605 to 1730 nm, where brackets denote the 95.4% statistical confidence interval determined by a bootstrap resampling analysis. The lack of a single dominant pause in the backward traces prevented us from making a robust backward alignment in that direction. Nevertheless, we attempted an alignment of backward traces, and cross-correlation between aligned forward and backward records yielded a correlation coefficient of $r = -0.002$ [−0.114, +0.110] over the region 595 to 1760 nm. All cross-correlations were computed over a region selected to contain a minimum of seven records in either direction (table S5).
17. The estimated numbers of long pauses (>5 s) moving forwards and backwards were 49 and 19, respectively (for pauses >10 s, the corresponding numbers were 19 and 2). Estimates of the number of forward pauses were scaled by a factor of 0.62 to correct for the ratio of the total distance traveled moving forwards to backwards by all molecules scored over this region (675 to 1200 nm).
18. R. J. Davenport, G. J. L. Wuite, R. Landick, C. Bustamante, *Science* **287**, 2497 (2000).
19. K. M. Dohoney, J. Gelles, *Nature* **409**, 370 (2001).
20. P. R. Bianco *et al.*, *Nature* **409**, 374 (2001).
21. The single-molecule traces were aligned to 900 nm, which corresponds to a peak in the distribution of shifts necessary to align the traces. The average offset was −2.1 nm. If we aligned the traces so the average offset was 0, then the dominant pause moved to 902.1 nm. In either case, the single-molecule data were within 2 nm of the gel-derived location.
22. R. A. Kovall, B. W. Matthews, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7893 (1998).
23. M. J. Doktycz, M. D. Morris, S. J. Dormady, K. L. Beattie, K. B. Jacobson, *J. Biol. Chem.* **270**, 8439 (1995).
24. P. Soultanas, D. B. Wigley, *Curr. Opin. Struct. Biol.* **10**, 124 (2000).
25. F. W. Stahl, K. D. McMilin, M. M. Stahl, Y. Nozu, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3598 (1972).
26. R. Wu, E. Taylor, *J. Mol. Biol.* **57**, 491 (1971).
27. Q. Yang, A. Hanagan, C. E. Catalano, *Biochemistry* **36**, 2744 (1997).
28. M. J. Lang, C. L. Asbury, J. W. Shaevitz, S. M. Block, *Biophys. J.* **83**, 491 (2002).

29. We thank K. Neuman and R. Knight for helpful discussions; E. Abbondanzieri for discussions and supplying his pause-finding algorithm; J. Einerson for extensive help with the assays; A. Engh, K. Herbert, and N. Padte for advice and help with sequencing gels; and U. Bali for the generous gift of lambda exonuclease. T.T.P. received support from a Burroughs Wellcome Fund Career Award in the Biomedical Sciences, a Princeton University Materials Institute Fellowship, and NIST. This work was supported by grants from the NIGMS GM 57035 (S.M.B.) and NHGRI HG 011821-01 (P.G.M.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1088047/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S5
References

16 June 2003; accepted 18 August 2003
Published online 28 August 2003;
10.1126/science.1088047
Include this information when citing this paper.

Salmonella SipA Polymerizes Actin by Stapling Filaments with Nonglobular Protein Arms

Mirjana Lilic,[1] Vitold E. Galkin,[2] Albina Orlova,[2] Margaret S. VanLoock,[2] Edward H. Egelman,[2] C. Erec Stebbins[1]*

Like many bacterial pathogens, *Salmonella* spp. use a type III secretion system to inject virulence proteins into host cells. The *Salmonella* invasion protein A (SipA) binds host actin, enhances its polymerization near adherent extracellular bacteria, and contributes to cytoskeletal rearrangements that internalize the pathogen. By combining x-ray crystallography of SipA with electron microscopy and image analysis of SipA-actin filaments, we show that SipA functions as a "molecular staple," in which a globular domain and two nonglobular "arms" mechanically stabilize the filament by tethering actin subunits in opposing strands. Deletion analysis of the tethering arms provides strong support for this model.

Salmonella spp. cause more than one billion new human infections each year that lead to more than three million deaths (*1*); they are classified as biodefense food and water safety threats, having been used in a documented biological attack (*2*). They use a protein se-