Supplementary Materials and Methods

Polymerase and expression vector

Nucleotide sequence of the plasmid used to construct the AviTag leg insertions (6379 bp). The polymerase gene (upper case) is cloned in a pBAD-HisC vector (lower case; Invitrogen). The C-terminal 6xHis tag is encoded (**CAT**)$_6$ at the 3'-end of the polymerase gene. The AviTag insertion made between amino acids K53-V54 divides the dinucleotide **GG** at 477-478 (underlined; numbered from 1 in the sequence), and the insertion between amino acids K229-F230 divides **GT** at 1005-1006.

aagaaaccaattgtccatattgcatcagacattgccgtcactgcgtcttttactgcctcttctcgctaaccaaaccggtaacc
ccgcttattaaaagcattctgtaacaaagcgggaccaaagccatgacaaaaacgcgtaacaaaagtgtctataatcacggcag
aaaagtccacattgattatttgcacggcgtcacactttgctatgccatagcatttttatccataagattagcggatcctacct
gacgctttttatcgcaactctctactgtttctccatacccgtttttgggctaacaggaggaattacatATGATTCTCGATAC
CGACTACATCACCGAGAACGGGAAGCCCGTGATAAGGGTCTTCAAGAAGGAGAACGGCGAGTTTAAAATCGAGTACGACAGAA
CCTTCGAGCCCTACTTCTACGCCCTTCTGAAGGACGATTCTGCGATAGAGGACGTCAAGAA**GG**TAACCGCAAAGAGGCACGGA
ACGGTTGTCAAGGTGAAGCGCGCCGAGAAGGTGCAGAAGAAGTTCCTCGGCAGGCCGATAGAGGTCTGGAAGCTCTACTTCAA
CCATCCTCAGGACGTCCCGGCGATTCGAGACAGGATACGCGCCCACCCCGCTGTCGTTGACATCTACGAGTACGACATACCCT
TCGCCAAGCGCTACCTCATCGACAAGGGCCTGATTCCGATGGAGGGCGACGAGGAGCTTACGATGCTCGCCTTCGCGATCGCA
ACCCTCTATCACGAGGGCGAGGAGTTCGGAACCGGGCCGATTCTCATGATAAGCTACGCCGACGGGAGCGAGGCGAGGGTGAT
AACCTGGAAGAAGATTGACCTTCCGTACGTTGACGTCGTCTCGACCGAGAAGGAGATGATTAAGCGCTTCCTCCGCGTCGTCA
GGGAGAAGGACCCCGACGTGCTCATCACCTACAACGGCGACAACTTCGACTTCGCCTACCTGAAGAAGCGCTGTGAGGAACTC
GGAATAAA***GT***TCACACTCGGCAGGGACGGGAGCGAGCCGAAGATACAGCGAATGGGCGACCGCTTTGCCGTTGAGGTGAAGGG
CAGGATTCACTTCGACCTCTACCCCGTCATAAGGCGCACGATAAACCTCCCGACCTACACCCTTGAGGCCGTTTACGAGGCCG
TCTTTGGAAAGCCCAAGGAGAAGGTTTACGCAGAGGAGATAGCGCAGGCCTGGGAGAGCGGGGAGGGCCTTGAAAGGGTTGCA
AGATACTCGATGGAGGACGCTAAGGTGACCTACGAGCTGGGAAGGGAGTTCTTCCCGATGGAGGCCCAGCTTTCGAGGCTTAT
AGGCCAGAGCCTCTGGGACGTCTCGCGCTCGAGCACCGGAAATTTGGTGGAGGCATTCCTCCTGCGGAAGGCCTACAAGAGGA
ACGAGCTCGCCCCAAACAAGCCCGACGAGAGGGAGCTCGCGAGACGGCGCGGGGGCTACGCTGGCGGGTACGTTAAGGAACCA
GAGCGGGGATTGTGGGACAACATTGTGTATTTAGACTTCCGCTCGTGGTATCCTTCAATCATCATAACCCACAACGTCTCGCC
GGATACCCTCAACCGCGAGGGCTGTAAAGAGTACGACGTCGCCCCTGAGGTTGGACACAAGTTCTGCAAGGACTTCCCCGGCT
TCATACCAAGCCTCCTGGGAGATTTGCTCGAGGAGGCGAGCAAGATAGAGCGGAAGATGAAGGCAACGGTTGACCCGCTGGAG
AAGAAACTCCTCGTGTACAGGCAGTGGCTTATAAAAATCCTCGCCAACAGCTTCTACGGCTACTACGGCTACGCCAAGGCCCG
GTGGTACTGCAAGGAGTGCGCCGAGAGCGTTACGGCCTGGGGAAGGGAGTATATAGAAATGGTTATCCGGGAACTCGAAGAAA
AATTCGGTTTTAAAGTTCTCTATGCCGATACAGACGGTCTCCATGCTACCATTCCCGGAGCAGACGCTGAAACAGTCAAGAAA
AAAGCAAAGGAGTTCTTAAAATACATTAATCCAAAACTGCCCGGCCTGCTCGAACTTGAGTACGAGGGCTTCTACGTGAGGGG
CTTCTTCGTCACGAAGAAGAAGTACGCTGTGATAGACGAGGAGGGCAAGATAACCACGAGGGGTCTTGAGATTGTGAGGCGCG
ACTGGAGCGAGATAGCGAAGGAGACCCAGGCCAGGGTCTTAGAGGCGATACTCAAGCACGGTGACGTCGAGGAGGCCGTTAGG
ATAGTCAAGGAAGTGACGGAAAAGCTGAGCAAGTATGAGGTCCCGCCCGAGAAGCTGGTAATCCACGAGCAGATAACGCGCGA
TTTGAGGGATTACAAAGCCACCGGCCCGCACGTTGCCGTTGCGAAGAGGCTCGCGGCGCGTGGAGTGAAAATCCGGCCCGGCA
CGGTGATAAGCTACATCGTCCTAAAGGGCTCTGGAAGGATAGGCGACAGGGCGATTCCAGCTGATGAGTTCGACCCGACGAAG
CACCGCTACGATGCGGAATACTACATCGAGAACCAGGTTCTCCCGGCGGTGGAGAGGATTCTAAAAGCCTTCGGCTATCGGAA
GGAGGATTTGCGCTACCAGAAGACGAAGCAGGTCGGCTCGGGCGCGTGGCTGAAGGTGAAGGGGAAGAAGGGTACCGAAGCTT
ACGTAGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGCCGTCGAC**CATCATCATCATCATCAT**TGAgtttaaacggtc
tccagcttggctgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaa
cagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatgg
tagtgtggggtctccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggccttt
cgttttatctgttgtttgtcggtgaacgctctcctgagtaggacaaatccgccgggagcggatttgaacgttgcgaagcaacg
gcccggagggtggcgggcaggacgcccgccataaactgccaggcatcaaattaagcagaaggccatcctgacggatggccttt
ttgcgtttctacaaactcttttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaat
gcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgcc
ttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaa
ctggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgct
atgtggcgcggtattatcccgtgttgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttg
agtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgat
aacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgt
aactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatgg

caacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggat
aaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggtc
tcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaacta
tggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatat
atactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaat
cccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatccttttttctgc
gcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactcttttt
ccgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaa
ctctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccg
ggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggggttcgtgcacacagcccagcttggag
cgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacag
gtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctg
tcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaac
gcggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataa
ccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcgg
aatagcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatctggtgcactctcagtacaatctgc
tctgatgccgcatagttaagccagtatacactccgctatcgctacgtgactgggtcatggctgcgccccgacacccgccaaca
cccgctgacgcgccctgacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtg
tcagaggttttcaccgtcatcaccgaaacgcgcgaggcagcagatcaattcgcgcgcgaaggcgaagcggcatgcataatgtg
cctgtcaaatggacgaagcagggattctgcaaaccctatgctactccgtcaagccgtcaattgtctgattcgttaccaattat
gacaacttgacggctacatcattcacttttttcttcacaaccggcacggaactcgctcgggctggccccggtgcatttttttaaa
tacccgcgagaaatagagttgatcgtcaaaaccaacattgcgaccgacggtggcgataggcatccgggtggtgctcaaaagca
gcttcgcctggctgatacgttggtcctcgcgccagcttaagacgctaatccctaactgctggcggaaaagatgtgacagacgc
gacggcgacaagcaaacatgctgtgcgacgctggcgatatcaaaattgctgtctgccaggtgatcgctgatgtactgacaagc
ctcgcgtacccgattatccatcggtggatggagcgactcgttaatcgcttccatgcgccgcagtaacaattgctcaagcagat
ttatcgccagcagctccgaatagcgcccttcccccttgcccggcgttaatgatttgcccaaacaggtcgctgaaatgcggctgg
tgcgcttcatccgggcgaaagaaccccgtattggcaaatattgacggccagttaagccattcatgccagtaggcgcgcggacg
aaagtaaacccactggtgataccattcgcgagcctccggatgacgaccgtagtgatgaatctctcctggcgggaacagcaaaa
tatcacccggtcggcaaacaaattctcgtccctgatttttcaccacccctgaccgcgaatggtgagattgagaatataacct
ttcattcccagcggtcggtcgataaaaaaatcgagataaccgttggcctcaatcggcgttaaacccgccaccagatgggcatt
aaacgagtatcccggcagcaggggatcattttgcgcttcagccatactttcatactcccgccattcagag

Polymerase AviTag™ constructs.

To insert an AviTag peptide at a single position in the polymerase, the entire plasmid vector (**above**) was first amplified in two separate PCR reactions (**Supplementary Figure S1**), which were then heteroduplexed prior to transformation into *E. coli*. The AviTag insertion is encoded by primers p1 and p3 (below) in their 5'-tails (red and green lines); the nucleotide sequences of the two tails are mutually complementary.

The procedure for constructing the K53-V54 insertion follows:
Primers (named according to **Supplementary Figure S1**) were from Integrated DNA Technologies Inc. and were PAGE-purified by the vendor prior to use.
p1:gctagatgcgccttcgtgccattcgattttctgagcttcgaagatgtcgttcagaccgctagacttcttgacgtcctctatcgcag,
p2:cttcttgacgtcctctatcgcag,
p3:tctagcggtctgaacgacatcttcgaagctcagaaaatcgaatggcacgaaggcgcatctagcgtaaccgcaaagaggcacgg,
p4:gtaaccgcaaagaggcacgg
The 25 uL PCR reactions contained plasmid DNA (8 pg/uL), 2 mM $MgSO_4$, 20 mM Tris-Cl pH 9.0, 50 mM KCl, 0.4 uM each of the two primers, 0.2 mM each dNTP, 80 units/mL of Taq DNA polymerase (Promega) and 20 units/mL of Pfu Ultra DNA polymerase (Stratagene). The thermal cycling schedule was 95ºC for 2 min, 10 cycles of (95 ºC for 15 sec, 58 ºC for 30 sec, 68 ºC for 6 min), 20 cycles of (95

ºC for 15 sec, 58 ºC for 30 sec, 68 ºC for 6 min extended by 5 sec each cycle), 68 ºC for 5 min. A 1-uL volume of Dpn I (20,000 units/mL, New England Biolabs) was added and both samples were incubated at 37 ºC for 60 min. The samples were resolved by electrophoresis in a 1.2% agarose gel (eGel, Invitrogen); the major product was the size of full-length plasmid DNA. The full-length product bands were purified and recovered in a 50 uL volume (Qiagen Gel Extraction Kit). The purified DNA products were mixed in equal volumes (4 uL of each in a 30 uL final volume containing 100 mM NaCl plus 20 mM Tris-Cl pH 8.0) and were heated at 99 ºC for 3 min followed by 2 cycles of (65 ºC for 5 min, 30 ºC for 15 min). Chemically-competent E. coli TOP 10 cells (Invitrogen) were transformed with 8 uL of the annealed sample and selected for ampicillin resistance. Clones were screened by transferring single colonies to 50 uL of water and amplifying in 25 uL reaction mixtures containing 1 uL of cell suspension, 5 mM $MgSO_4$, 20 mM Tris-Cl pH 9.0, 50 mM KCl, 0.4 uM each of diagnostic primers 5'-ccttctgaaggacgattctgcg and 5'-cgcttcaccttgacaaccg, and Taq DNA polymerase (20 units/mL); amplification conditions were 95 ºC for 2 min, 25 cycles of (95 ºC for 15 sec, 54 ºC for 30 sec, 68 ºC for 1 min), 68 ºC for 5 min. Samples were resolved by gel electrophoresis; clones containing the desired insert were identified as having a 180-bp amplicon, whereas negative clones had a 117-bp amplicon. Nine of eleven clones were positive by this PCR test. Two clones were confirmed by DNA sequencing. The first was the desired sequence while the second had an unintended mutation and was discarded.

The same procedure was used for inserting the second AviTag peptide between K229-F230, starting either from the same parent gene as above (to construct a single leg insertion at K229) or from the K53-V54 insert obtained above (to add the second leg at K229). Primers for the second insertion were: p1:gctagatgcgccttcgtgccattcgattttctgagcttcgaagatgtcgttcagaccgctagactttattccgagttcctcacagcg, p2: ctttattccgagttcctcacagcg,
p3:tctagcggtctgaacgacatcttcgaagctcagaaaatcgaatggcacgaaggcgcatctagcttcacactcggcagggacgg, p4: ttcacactcggcagggacgg, with the diagnostic primers ttcgctgtatcttcggctcg and tacctgaagaagcgctgtgag. For both the single and the double leg constructs, the same high yield of positive clones was obtained (~90%) and individual clones were confirmed by sequencing.

Processivity of Therminator™ Polymerase

In order to assess whether the engineered complexes display enhanced processivity, it was first necessary to determine processivity in the Therminator (exo^-) parent polymerase under comparable conditions. As explained below, we found that Therminator processivity is less than about 20 nucleotides, which demonstrates that processivity is greatly enhanced in the complexes.

To determine processivity, we used a gel assay similar to that of Davidson et al (Nucleic Acids Research 31:4702 2003). In our experiments, polymerase was first pre-incubated with a tagged primer-template before initiating DNA synthesis by addition of unlabeled dNTPs plus a DNA trap. The trap binds free polymerases, thereby arresting the size distribution of first-round primer extension products for subsequent analysis by gel electrophoresis. Whereas we detected DNA products using a single fluorescent label on the primer, Davidson incorporated multiple labels using ^{32}P-dGTP. This difference in labeling methods affects the subsequent analysis of processivity (discussed below). Davidson observed that Taq(exo^-) polymerase incorporated 1-25 nucleotides before dissociating from the template, which suggests a processivity value somewhere between 1-25 nucleotides; a Taq derivative engineered

to bind thioredoxin displayed a set of much longer products characteristic of increased processivity (Figure 3 of Davidson, reproduced here in Supplementary Figure S5 A).

In our assay of Therminator(exo-) processivity, primers were extended hundreds of nucleotides in the absence of inhibitor DNA, but only about 1-50 nucleotides in the presence of the largest tested amount of inhibitor (Supplementary Figure S6). This size range, 1-50 nucleotides, places an upper limit on processivity, since additional inhibitor might have depressed fragment length even further. While this *qualitative* result of 1-50 nucleotides is sufficient for contrasting the very different processivities of Therminator and engineered complexes, we suggest that the profile of decreasing band intensities observed in the gel lanes offers an opportunity to *quantify* processivity by applying a simple model of polymerase action. As described below, this analysis gave processivities of < 20 nucleotides for Therminator (our data) and 4.6 nucleotides for Taq (Davidson data).

Exponential decay. Band intensities of end-labeled primer extension products are expected to decrease exponentially with fragment size based on the dual assumptions i) that the probability of polymerase dissociation is uniform along the template and ii) that 5'-3' exonuclease activity is negligible compared to the polymerization rate. For example, a 10% dissociation probability (d = 0.1) would leave 90% of polymerases bound after incorporating the first nucleotide n=1, 81% at n=2, 73% at n=3, 66% at n=4 and so on,

$$B_n = (1-d)^n, \quad (1)$$

where B_n is the band intensity of labeled primers extended n nucleotides. Equation (1) describes a normalized exponential distribution, which can be written equivalently in terms of the natural base e and decay constant k

$$B_n = e^{-kn}. \quad (2)$$

An equation of form (2) is obtained in fitting data to an exponential function. The two equations are identical for corresponding values of k and d, obtained by equating (1) with (2) and solving for each variable in terms of the other.

$$(1-d)^n = e^{-kn}$$

$$d = (-1 + e^k) / e^k \quad (3)$$

$$k = -\ln(1 - d) \quad (4)$$

Thus, if we specify k, we can calculate d, and vice versa. Equations (1) and (2) give identical plots for corresponding values of d and k (not shown). Moreover, equation (3) shows that the dissociation probability d approaches 1, as expected, when the exponential decay constant k goes to infinity (Supplementary Figure S7). As implemented here, d is understood as the probability of polymerase dissociation per added nucleotide

$$d = \frac{(probability\ of\ dissociation\ per\ unit\ time)}{(probability\ of\ nucleotide\ addition\ per\ unit\ time)}.$$

Since the dissociation probability in reality varies by position along the template (indicated by uneven band intensities in Supplementary Figures S5 and S6), the decay constant k obtained in a least-squares fit to band intensities is an average value related to the averaged probability of polymerase dissociation along the template.

Calculating processivity. Processivity is normally understood as the expected number of nucleotides added before the polymerase dissociates. As such, in order to determine processivity, we just need to know the mean value of the fitted exponential function. The mean value of a normalized exponential function $f = e^{-kn}$ is just the inverse of the decay constant

processivity = distribution mean = 1/k.

While this relationship is often explained in terms of Poisson statistics (http://mathworld.wolfram.com/ExponentialDistribution.html), we derive it here in context of the processivity assay.

Processivity is the average number of nucleotides added to a primer strand before the polymerase dissociates. In a set of first-round primer extension products, processivity P may be calculated as the ratio

$$P = \frac{(total\ number\ of\ nucleotides\ incorporated\ into\ all\ extended\ primer\ strands)}{(total\ number\ of\ extended\ primer\ strands)} = \frac{N}{R} \quad (5)$$

Case: Labeled primer (Therminator). Relative values for both N and R are conveniently determined by fitting the profile of band intensities to an exponential model. When the primer is tagged with a single label, the number of DNA fragments in a given band is directly related to band intensity. In an exponential model, the normalized band intensity B_n for primers extended n nucleotides is

$$\boldsymbol{B_n = e^{-k*n}}$$

Thus, the total number R of extended primer strands is just the sum of intensities of all bands B_n comprising the lane profile of fragment sizes n = 0 to n_{max}

$$R = \int_0^{n_{max}} B_n\, dn = \int_0^{n_{max}} e^{-k*n}\, dn = -\frac{-1+ e^{-k*n_{max}}}{k} \ .$$

The number of nucleotides incorporated into primers extended n nucleotides is just n times the number of primers ($n \cdot B_n$), and the total number N of nucleotides incorporated into all extended primers is just the sum over all bands B_n

$$N = \int_0^{n_{max}} n * B_n\, dn = \int_0^{n_{max}} n * e^{-k*n} dn = -\frac{-1+e^{-k*n_{max}}+e^{-k*n_{max}}*k*n_{max}}{k^2}.$$

Processivity is the quotient (Eq 5)

$$P = \frac{N}{R} = \frac{-\frac{-1 + e^{-k*n_{max}} + e^{-k*n_{max}} * k * n_{max}}{k^2}}{-\frac{-1 + e^{-k*n_{max}}}{k}} = \frac{-1 + e^{-k*n_{max}} + e^{-k*n_{max}} * k * n_{max}}{k * (-1 + e^{-k*n_{max}})} \quad (6)$$

Evaluated at n_{max} = 100 and k = 2, equation (6) gives P = 0.50000000, and at k=1/5 gives P = 4.99999979, which suggests that P = 1/k as expected.

Mathematically, in the limit where maximum fragment size n_{max} goes to infinity (i.e., an infinitely long gel lane) and where k > 0, the processivity expression (Eq 6) simplifies to

$$P = \frac{1}{k},$$

which confirms the inverse relationship between processivity and the exponential decay constant. Applying this analysis to Therminator, we estimate a processivity of 20 nucleotides (Supplementary Figure 5B). The approach could be improved by combining the primer extension products from several different primer sequences in a single lane in order to average out the effects of specific sequences on the probability of polymerase dissociation and provide smoother profiles.

Case: Multiple, sequence-specific base labels (Taq). Analysis is more complicated for the case of Taq polymerase (Davidson) because of the different labeling method. As discussed above, the Davidson assay incorporates multiple labels from ^{32}P-dGTP into the primer extension products. The labeling pattern along the synthesized strand thus depends on the specific template sequence used in the experiment. Therefore, in order to fit band intensity profiles to an exponential model, it is first necessary to correct for the specific number of G nucleotides along the template. In the template of Davidson, the sequence of the first 20 incorporated nucleotides and the cumulative number of potential labeling sites would have been

```
5'-TGTTGTGTGGAATTGTGAGC... (primer strand sequence)
   01112233455555667788... (cumulative number of G label sites),
```

where primer extension products incorporating only the first nucleotide (T) would go undetected. We applied exponential analysis to the data of Davidson by first correcting the intensity profile for the number of labels. This normalizes the intensities to what they would be if there were only a single label per DNA strand. For example, intensities between the second and third bands were left unchanged because they each have only 1 potential label, while intensities between the fifth and sixth bands were divided by 2, and the seventh to eighth bands by 3, and so on; corrections to each nominal band size n were applied to the continuous profile (i.e. pixel intensities) in the range n $\mp$ 0.5. The corrected intensity profile more accurately reflects the number of primer strands in each band, yielding a processivity for Taq polymerase of 4.6 nucleotides (Supplementary Figure 5B).

Case: High processivity.
The exponential analysis as presented is useful only in cases of low processivity. In the competitive assay used here, all polymerases, regardless of their processivity, are expected to generate exponential decay profiles of band intensities. However, whereas low processivity results in steep profiles decaying

over a short size range of detectable (intense) bands, high processivity would produce shallow profiles dispersing the DNA over a broad range of fragment lengths. Highly dispersed and weak in intensity, the individual bands may be difficult to detect. Moreover, in cases of extreme processivity or short reaction times, primer extension products would appear as a cluster of high-molecular weight bands as seen in experiments with the engineered Taq-thioredoxin polymerase (Supplementary Figure 5, lanes 5 and 9). With the Taq-thioredoxin polymerase of Davidson, in addition to the observed band cluster, there is also an exponentially-decaying component of low-molecular weight bands; we attribute the latter to a subclass of low-processivity polymerases that failed to bind thioredoxin.

Supplemental Figure Legends

Figure S1. The approach of Chiu et al (Chiu, J., March, P.E., Lee, R., Tillett, D. 2004. Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 hours. Nucleic Acids Res 32: e174) that was used for inserting the AviTag legs.

Figure S2. $DOPA_3$-PEG compounds used for coating ITO-coated coverglass with a non-stick PEG film derivatized with biotin. Details are given in Materials and Methods.

Figure S3. Chamber used in immobilization experiments. Details are given in Materials and Methods.

Figure S4. Demonstration that φ-29 DNA polymerase is capable of strand-displacement synthesis, while Taq DNA polymerase is not. Primed M13 DNA (2 nM) was mixed in buffer C (Materials and Methods) with 200 uM of each of the 4 dNTPs, 5 mM MgCl2 and either Taq or φ-29 DNA polymerase (New England Biolabs) in 10 uL volumes. After incubation for 90 min at 55ºC (Taq) or 37ºC (φ-29), samples were mixed with 10 uL of 0.2% sodium dodecylsulfate and 5 uL of 50% glycerol, heated at 65ºC for 5 min, and resolved by electrophoresis in a 1.5% agarose gel. The gel was stained with SYBR Gold™ and photographed. Lane 1: M13 DNA control. Lanes 2-3: Taq polymerase at 20 and 400 unit/mL. Lanes 4-5: φ-29 polymerase at 20 and 400 unit/mL.

Figure S5. Processivity determination of Taq DNA polymerase (Davidson et al, Nucleic Acids Research 31:4702 2003). (A) Figure 3A of Davidson shows decreasing band intensities at longer fragment lengths. Bands comprising longer products are seen when thioredoxin is added to polymerases engineered with a thioredoxin binding domain (lanes 5 and 9).

(B) Quantitative estimation of Taq processivity from the band intensity profile of lane 7. The profile (blue line) was obtained from the displayed image originating in a pdf file of the Davidson publication using Image J software (http://rsb.info.nih.gov/ij/). The band density profile, originally tabulated by pixel position in the lane, was rescaled to units of nucleotides (n) by visually counting individual bands; band intensities (B_n) were mapped to an 8-bit gray scale and corrected for sequence-specific labeling stoichiometry as explained in the Supplementary text. The corrected band intensity profile was fit to an exponential trendline, giving decay constant $k = 0.22$ and processivity $P = 1/k = 4.6$ nucleotides.

.

Figure S6. Processivity determination by primer extension in the presence of competing inhibitor oligonucleotide. (A) Polymerase (2 nM) was mixed with 5'-labeled primed M13 template DNA (10nM) in Reaction Buffer (10 mM Tris-Cl pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 0.1% Tween-20 and 0.1 mM EDTA) in an 8 uL sample. Samples were held in capped tubes at 20°C for at least 5 minutes before use. Reactions were initiated, incubated and terminated, one sample at a time, as follows. Each capped tube was pre-warmed in a 52°C thermal cycler heat block for 30 seconds. Polymerization was initiated by adding 2 uL of Reaction Buffer containing all 4 dNTPs (final concentration, 200 uM each) plus varying amounts of the inhibitor oligonucleotide (final concentrations 0, 100, 630 or 4000 nM replicated in lanes 1..4 and 5..8). After this addition, the tubes were incubated 60 seconds in the heat block and reactions were terminated by adding 1 uL of 100 mM EDTA. Samples were mixed with 10 uL of gel loading buffer (LI-COR #829-04459) and 4 uL volumes were analyzed by electrophoresis in a 25 cm long polyacrylamide-urea gel (6.5% KB^{PLUS}, LI-COR #827-05607) on a LI-COR 4200 DNA analyzer. The positions of primers extended by 3 or 50 nucleotides was determined by visually counting bands from the primer front (bottom dark band) in an expanded image of that shown here. The primed template was prepared by annealing a fluorophore-labeled primer (83 nM of IR700-5'-ACCGCCAGCCATTGCAACAGGA, LI-COR Inc., primer T_m = 64.4°C calculated at www.idtdna.com) plus M13 template (80 nM of M13 ssDNA, New England Biolabs) in 60 uL of Reaction Buffer (above) using temperature profile 80°C for 60 sec, 70°C for 10 sec, ramp 70°C to 50°C at 0.1°C /sec. The inhibitor oligonucleotide was prepared by annealing a primer (40 uM of 5'-ACACCACCATCAGCCCCGTGGAT) and a template (40 uM of 5'-CTTTACTTCTTATGTTAGTTATCCACGGGGCTGATGGTGGTGT, both from IDT, primer T_m = 65.1°C calculated at www.idtdna.com) in 100 uL of Reaction Buffer in 60 sec temperature steps of 90, 80, 75, 70, 65 and 60°C. Annealed primed template and inhibitor oligo preparations were stored at -20°C prior to use.

B) The band intensity profile of the sample containing the largest amount of inhibitor (lane 4) was obtained from the image and fit to an exponential function as explained in the Supplementary text. The fitted decay constant k = 0.05 gave a processivity of P = 1/k = 20 nt.

Figure S7. Plot of polymerase dissociation probability d as a function of the decay constant k (Equation 3, Supplementary text). The dissociation probability approaches 1 as k goes to infinity.

Movie M1. Mpeg movie of labeled DNA synthesized by immobilized complexes is included with the Supplemental Data package. The sample was prepared as described in Figure 8. The pixel dimension is 0.27 um and the exposure time was 80 msec per frame.

Figure S1

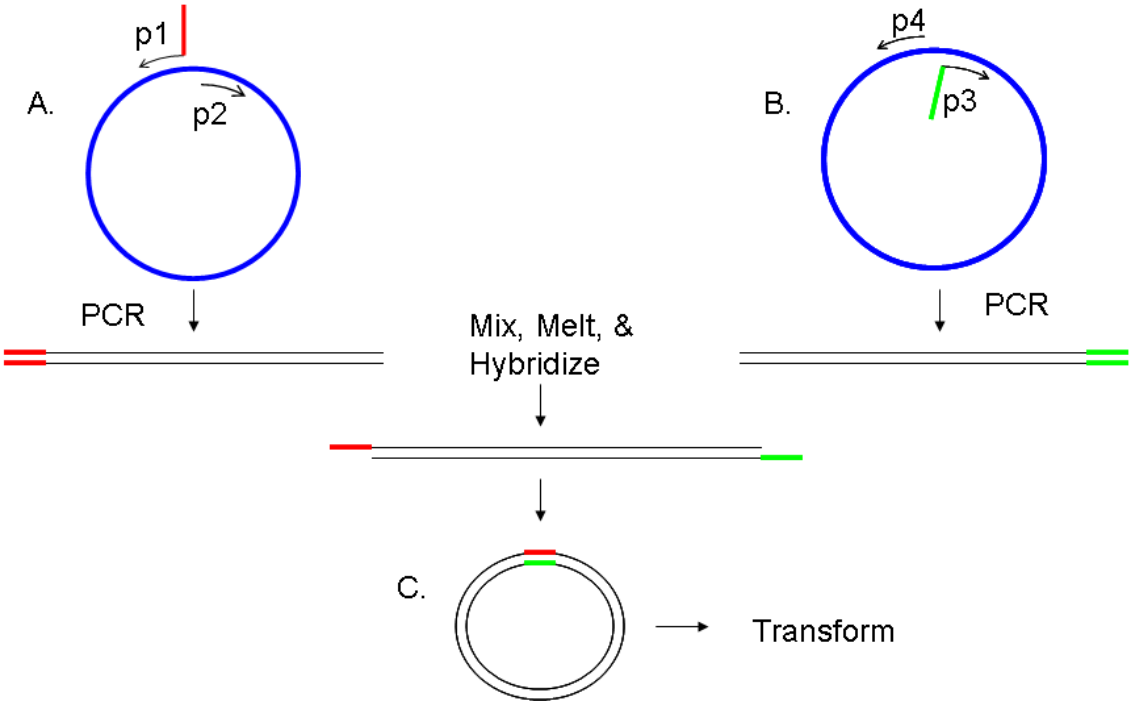


Figure S2

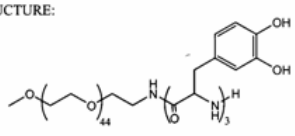
NERITES

Product Specification Sheet

Date: July 19, 2005

PRODUCT NAME:
mPEG-DOPA (tri)

CHEMICAL STRUCTURE:



PRODUCT ANALYSIS:
Prepared from mPEG-NH_2 2000 (Nektar 2M2V0D11)
DOPA content (UV-vis): 3.6
Appearance: light yellow powder

PACKAGE & STORAGE:
The product is shipped in a glass bottle filled with argon and the bottle should be placed in the freezer (< -4°C) upon arrival. The bottle must be allowed to warm up to room temperature before opening the lid to avoid water condensation. It is recommended to refill the bottle with inert gas, such as nitrogen or argon, before restocking the bottle in the freezer.

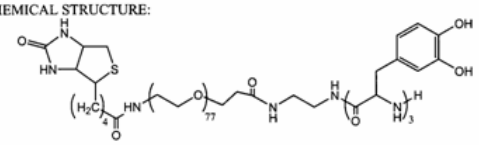
NERITES

Product Specification Sheet

Date: July 19, 2005

PRODUCT NAME:
Biotin-PEG-DOPA (tri)

CHEMICAL STRUCTURE:



PRODUCT ANALYSIS:
Prepared from Biotin-PEG-NHS 3,400 (Nektar OH4M0F02)
DOPA content (UV-vis): 2.9
Biotin (NMR): confirmed
Appearance: light beige powder

PACKAGE & STORAGE:
The product is shipped in a glass bottle filled with argon and the bottle should be placed in the freezer (< -4°C) upon arrival. The bottle must be allowed to warm up to room temperature before opening the lid to avoid water condensation. It is recommended to refill the bottle with inert gas, such as nitrogen or argon, before restocking the bottle in the freezer.

Figure S3

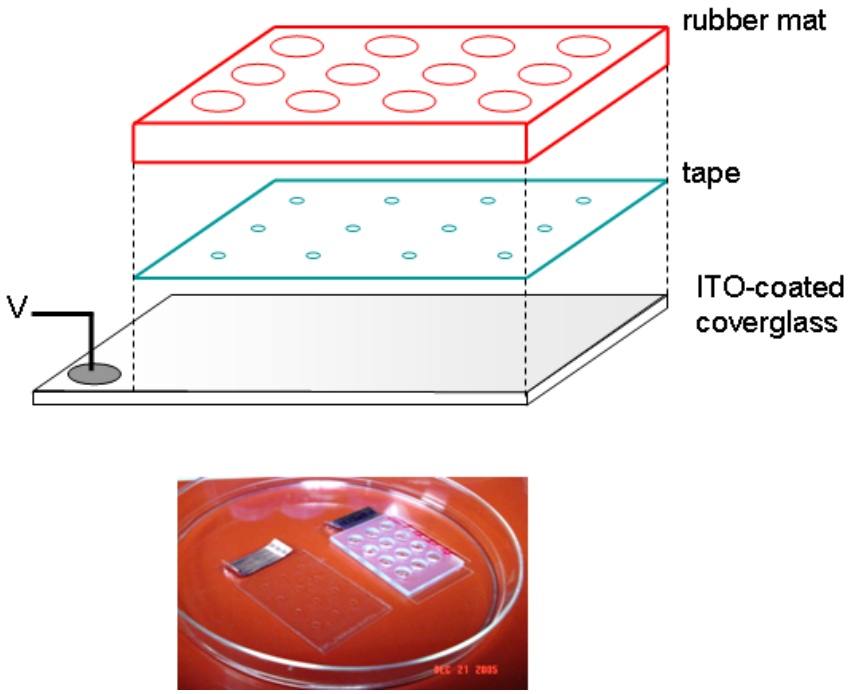


Figure S4

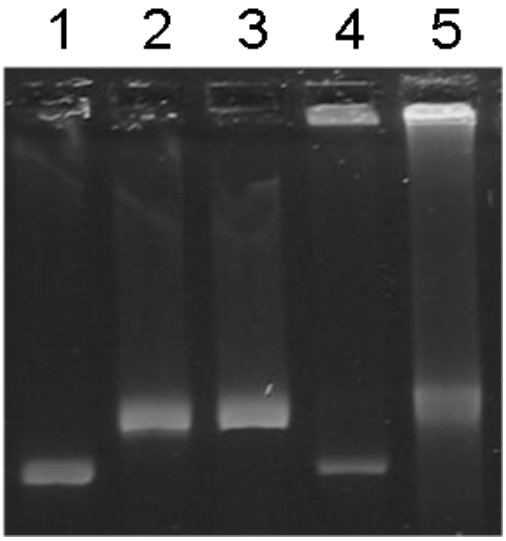


Figure S5

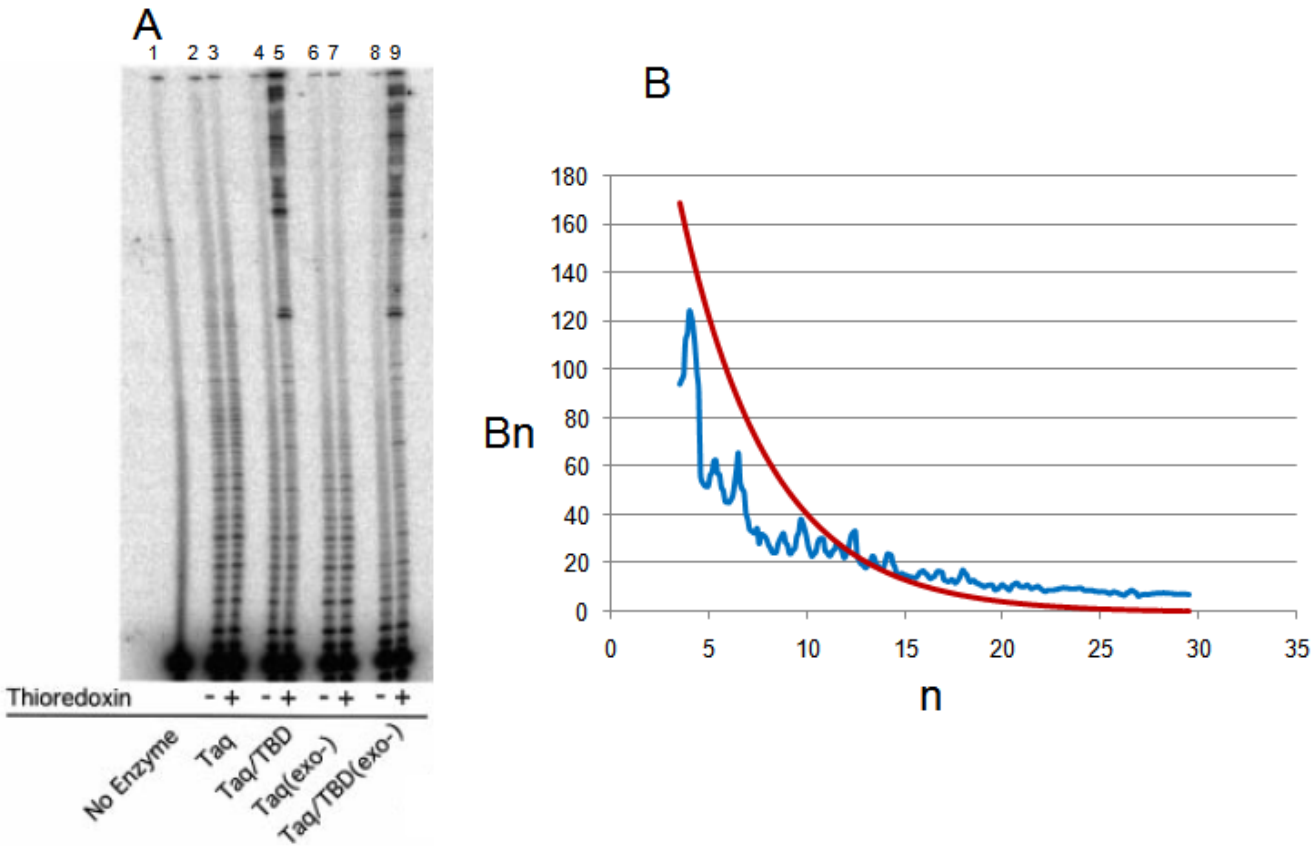

A
1 2 3 4 5 6 7 8 9
Thioredoxin - + - + - + - +
No Enzyme
Taq
Taq/TBD
Taq(exo-)
Taq/TBD(exo-)
B
Bn
n
180
160
140
120
100
80
60
40
20
0
0 5 10 15 20 25 30 35

Fig S6

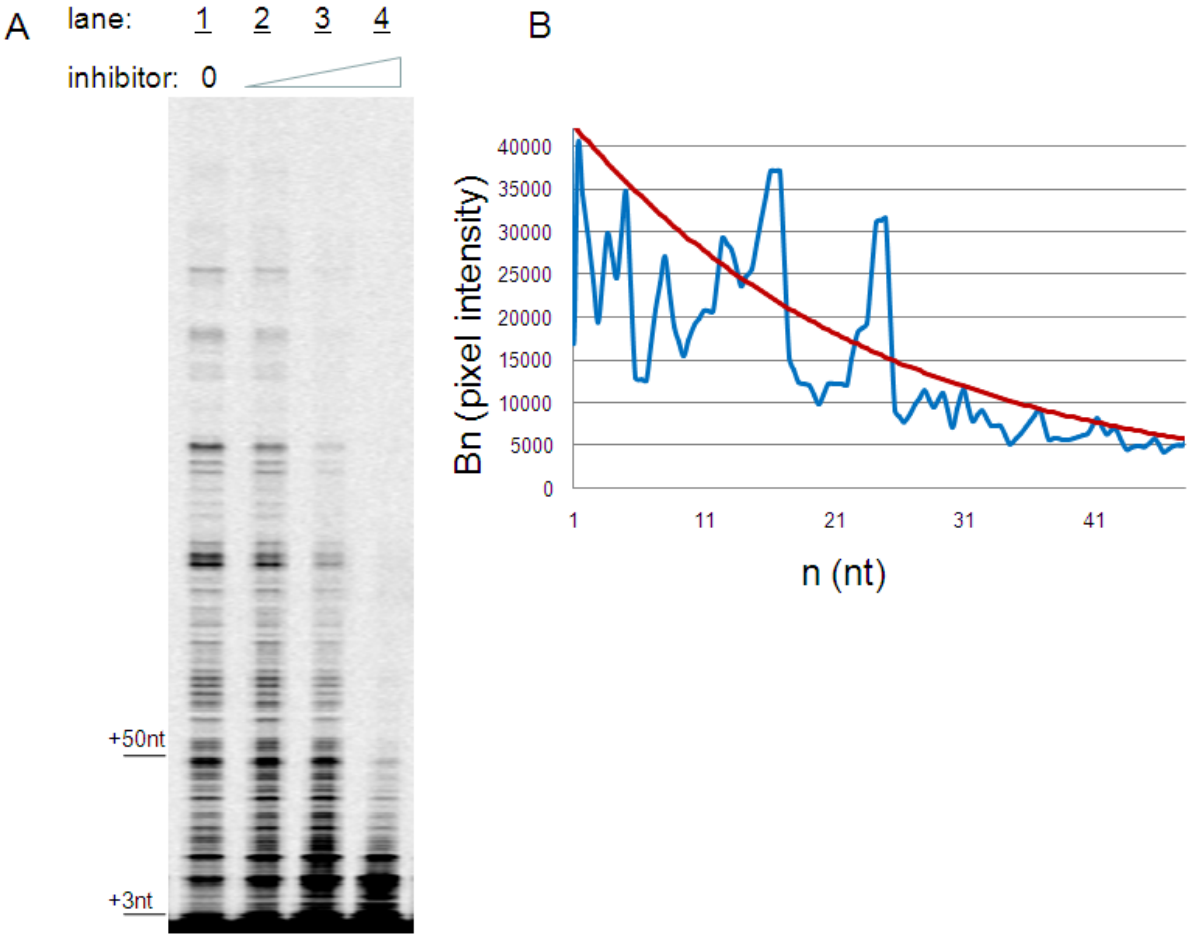
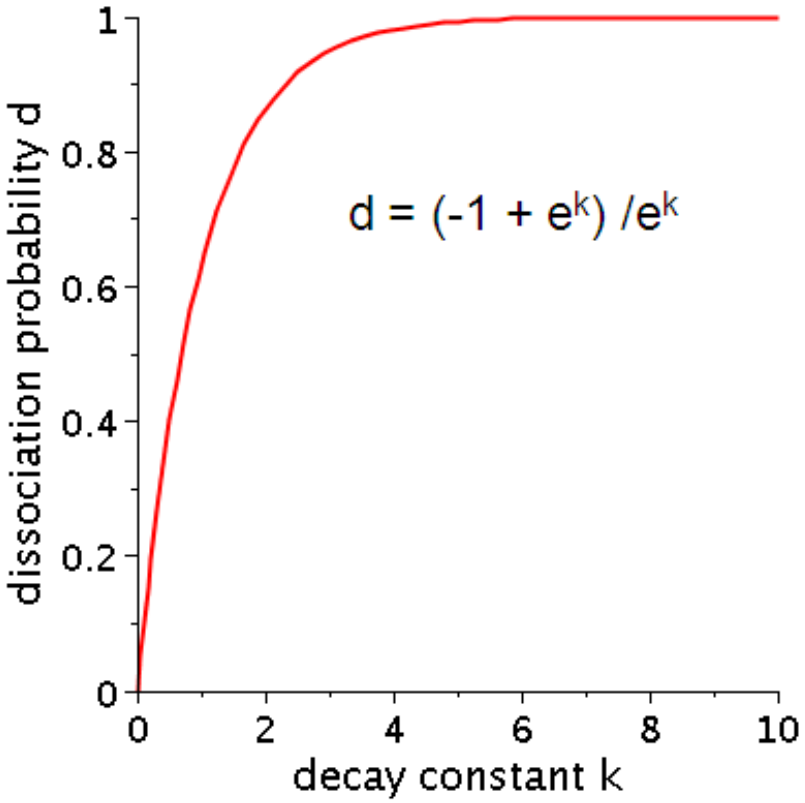

A
lane: 1 2 3 4
inhibitor: 0
+50nt
+3nt
B
Bn (pixel intensity)
40000
35000
30000
25000
20000
15000
10000
5000
0
1
11
21
31
41
n (nt)

Fig S7


dissociation probability d
1
0.8
0.6
0.4
0.2
0
d = (-1 + e^k) /e^k
0
2
4
6
8
10
decay constant k