Supporting Information

Fazal et al. 10.1073/pnas.1514028112

SI Materials and Methods

Substrate RNA. Except where otherwise noted, all PNPase and RNase R data were collected using an RNA substrate with the sequence 5′-GGCCUGACUAGAGUCCUUGGCGAACCGGUGUUUGACGUCCAGGAAUGUCAAAUCCGUGGCGUGACCUAUUCCGCACCGCUGCGCGUUAAACUGCGUCUGGUGAUCUAUGAGCGCGAAGCGCCGGAAGGCACCGUAAAAGACAUUAAAGAACAAGAAGAAAAAAAAAAAA-3′.

This RNA was generated by in vitro transcription of a DNA template, which was obtained by PCR of the pALB3 plasmid using the primers RNase R forward: 5′-AAAAAGAGTATAATACGACTCACTATAGGCCTGACTAGAGTCCTTGGCGA-3′ and 166 reverse: 5′-TTTTTTTTTTTTCTTCTTGTTCTTTAATGTCTT-3′.

Substrate Complement: DNA. For assays performed on DNA–RNA hybrids, the strand complementary to the enzyme substrate carried a 5′ biotin tag, and was chemically synthesized (Integrated DNA Technologies). The sequence was as follows (bases complementary to the substrate RNA are underscored): 5′-TTTCGCCCATGTAGACTTCTTGTTCTTTAATGTCTTTTACGGTGCCTTCCGGCGCTTCGCGCTCATAGATCACCAGACGCAGTTTAACGCGCAGCGGTGCGGAATAGGTCACGCCACGGATTTGACATTCCTGGACGTCAAACACCGGTTCGCCAAGGACTCTAGTCAGGGCTCCAAGTAATGAGGGCTACGGTCGAACA-3′.

Substrate Complement: RNA. For assays performed using dsRNA, the substrate complement strand was obtained by ligating a chemically synthesized 20-nt RNA oligomer (Integrated DNA Technologies) carrying a 5′ biotin tag to a 187-nt RNA sequence produced by in vitro transcription. The DNA template used to generate the 187-nt sequence was obtained by PCR of pALB3 using the primers RNAHand forward: 5′-AAAAAGAGTATAATACGACTCACTATAGGTCTTGTTCTTTAATGTCTTTTACGGT-3′ and RNAHand reverse: 5′-TGTTCGACCGTAGCCCTCATTACTTGTATCGGCCTGACTAGAGTCCTTGGCGA-3′. After the 20-nt and 187-nt RNA sequences were annealed to a DNA splint, the two components were joined by T4 RNA ligase 2. The splint had the sequence 5′-AGACATTAAAGAACAAGACCTCTACATGGGCGAAATTCCG-3′. The final RNA product (207 nt) was as follows (bases complementary to the substrate RNA are underscored): 5′-CGGAAUUUCGCCCAUGUAGAGGUCUUGUUCUUUAAUGUCUUUUACGGUGCCUUCCGGCGCUUCGCGCUCAUAGAUCACCAGACGCAGUUUAACGCGCAGCGGUGCGGAAUAGGUCACGCCACGGAUUUGACAUUCCUGGACGUCAAACACCGGUUCGCCAAGGACUCUAGUCAGGCCGAUACAAGUAAUGAGGGCUACGGUCGAACA-3′.

We note that in the dsRNA construct the sequence for the double-stranded region is frame-shifted by two base pairs relative to that of the DNA–RNA hybrid. Consequently, the substrate RNA in the former case carried a 14-nt-long 3′ overhang [sequence $AG(A)_{12}$], whereas the substrate RNA in the latter case carried a 5′ GG overhang in addition to a 12-nt-long 3′ overhang (sequence A_{12}).

DNA–RNA Hybrid: AU50 Construct. Certain experiments used a modified RNA substrate in which the first 48 bp of the DNA–RNA hybrid encountered by PNPase were replaced with a randomly generated sequence with high A+U content. The first 50 nt encountered by the ribonuclease in this construct included 35 nt that were either A or U. The sequence of this substrate (AU50) was as follows, with the randomly generated component underscored: 5′-GGCCUGACUAGAGUCCUUGGCGAACCGGUGUUUGACGUCCAGGAAUGUCAAAUCCGUGGCGUGACCUAUUCCGCACCGCUGCGCGUUAAACUGCGUCUGGUGAUCUAAUUGUUUUGAUACUUUGGUCAUAUUUCUGUUGUAGGAGUGAAUUCACUAAAAAAAAAAAA-3′.

As before, this sequence was generated by in vitro transcription of a DNA template which was obtained by PCR of pALB3 using primers RNase R forward and 48-ATrich reverse: 5′-TTTTTTTTTTTTAGTGAATTCACTCCTACAACAGAAATATGACCAAAGTATCAAAACAATTAGATCACCAGACGCAGTTTAACGCGC-3′.

The RNA was hybridized to a DNA complement carrying a 5′ biotin tag (Integrated DNA Technologies). The sequence was as follows (bases complementary to the substrate RNA are underscored): 5′-CCCATGTAGAAGTGAATTCACTCCTACAACAGAAATATGACCAAAGTATCAAAACAATTAGATCACCAGACGCAGTTTAACGCGCAGCGGTGCGGAATAGGTCACGCCACGGATTTGACATTCCTGGACGTCAAACACCGGTTCGCCAAGGACTCTAGTCAGGCCATGAGGGCTACGGTCGAACA-3′.

RNase R: 1,500-bp dsRNA Constructs. Three copies of the 155-bp pALB3 sequence were inserted, in a tandem series, into a pUC19 plasmid between the EcoRI and HindIII restriction sites. The final sequence of the full 483-bp insert was as follows, with restriction sites (adjoining the 155 bp repeats) underscored: 5′-AATTCGGCCTGACTAGAGTCCTTGGCGAACCGGTGTTTGACGTCCAGGAATGTCAAATCCGTGGCGTGACCTATTCCGCACCGCTGCGCGTTAAACTGCGTCTGGTGATCTATGAGCGCGAAGCGCCGGAAGGCACCGTAAAAGACATTAAAGAACAAGAGAATTCGGCCTGACTAGAGTCCTTGGCGAACCGGTGTTTGACGTCCAGGAATGTCAAATCCGTGGCGTGACCTATTCCGCACCGCTGCGCGTTAAACTGCGTCTGGTGATCTATGAGCGCGAAGCGCCGGAAGGCACCGTAAAAGACATTAAAGAACAAGAAAGCTTGGCCTGACTAGAGTCCTTGGCGAACCGGTGTTTGACGTCCAGGAATGTCAAATCCGTGGCGTGACCTATTCCGCACCGCTGCGCGTTAAACTGCGTCTGGTGATCTATGAGCGCGAAGCGCCGGAAGGCACCGTAAAAGACATTAAAGAACAAGAA-3′.

The plasmid was used to generate DNA templates from which new substrate and substrate-complement RNAs could be transcribed. The DNA template for the substrate RNA was generated by PCR using primers 19rep-motor-NEW forward: 5′-AAAAAGAGTATAATACGACTCACTATAGGCAGTGTTATCACTCATGGTTATGGCAGCAC-3′ and 19rep-motor reverse: 5′-TTTTTTTTTTTTCTGCTATGACCATGATTACGCCAAGCT-3′.

The DNA template for the substrate-complement RNA was obtained by PCR using primers 19rep-teth forward: 5′-AAAAAGAGTATAATACGACTCACTATAGGGTCGCTCTAATTTAGCTCTCGTGTCGCCCATGTAGAGGGCTATGACCATGATTACGCCAAGCT-3′ and 19rep-teth-NEW reverse", 5′-TGTTCGACCGTAGCCCTCATGGCAGTGTTATCACTCATGGTTATGGCAGCAC-3′.

The resulting RNAs were designed to yield a 1,500-bp dsRNA construct when hybridized. In this construct, a 26-bp dsRNA spacer separates the 12-nt polyadenine 3′ overhang from the pALB3 insert sequences.

RNase R data also were acquired for an alternative 1,500-bp construct carrying a single modified copy of the 155-bp pALB3 sequence in which a 25-bp segment was moved internally to a new location. To generate this modified construct, a DNA insert with the following sequence was synthesized chemically (Integrated DNA Technologies) and placed between the EcoRI and HindIII restriction sites in pUC19: 5′-ATCCTGACTAGAGTCCTTG-GCGAACCGGTGTTTGACGTCACCGCTGCGCGTTAAA-CTGCGTCTGCAGGAATGTCAAATCCGTGGCGTGACC-TATTCCGCGTGATCTATGAGCGCGAAGCGCCGGAAG-GCACCGTAAAAGACATTAAAGAACAAGA-3′.

As above, this plasmid was used to generate DNA templates from which substrate and substrate-complement RNAs could be transcribed. The DNA template for the substrate RNA was generated by PCR using the primers pUC19-motor forward: 5′-AAAA-AGAGTATAATACGACTCACTATAGGCCCAGTGCTGCA-ATGATACCGCGAG-3′ and 166 reverse. The template for the substrate-complement RNA was obtained by PCR using primers pUC19-teth forward: 5′-AAAAAGAGTATAATACGACTCAC-TATAGGGTCGCTCTAATTTAGCTCTCGTGTCGCCCATGT-AGAGGTCTTGTTCTTTAATGTCTTTTACGGTGCCTTCCG-3′ and pUC19-teth reverse: 5′-TGTTCGACCGTAGCCCTCATG-GCCCAGTGCTGCAATGATACCGCGAG-3′.

DNA Handles. The 3,057-bp and 2,250-bp digoxygenin-labeled DNA handles were derived by PCR from plasmid M13mp18 using primers L_dig_3306 for the 3,057-bp handle: 5′-TTAAGGCTTCAAAAC-CTCCCGCAAG-3′, L_dig_4114long for the 2,250-bp handle: 5′-AAGGATTCTAAGGGAAAATTAATTAATAGCGACG-3′, and P_absc_splintanneal: 5′-TCCACCGATTATGTCCGTACGCGCT/idSp/ATGTGCTGCAAGGCGATTAAGTTGG-3′.

The resulting product carrying a 25-nt 5′ overhang was annealed and ligated to a DNA oligomer (Integrated DNA Technologies) of sequence 5′-AAGCGCGTACGGACATAATCGGTGGATGTTCGA-CCGTAGCCCTCAT-3′ to generate a 20-nt 3′ overhang. For DNA–RNA hybrid experiments, the DNA complementary to the RNA substrate was present in the reaction buffer and was simultaneously annealed and ligated to the handle.

RNase R experiments performed on 1,500-bp dsRNA constructs used a second biotin-labeled 1,044-bp DNA handle, which was derived by PCR from the pRL732 plasmid using primers 732-Handle forward: 5′-CACGAGAGCTAAATTAGAGCGACCC/idSp/TATCATCCCTTACCGTGGTTCCTGGC-3′ and 732-Handle reverse: 5′-CTCAGACAGCGGGTTGTTCTGG-3′.

Hybridization and Storage of Tether Components. The RNA substrate, its complement, and the DNA handles were annealed and hybridized in a solution of 80% formamide, 40 mM Pipes (pH 6.5), 400 mM NaCl, and 1 mM EDTA. Formamide was used because it reduces the thermal stability of double-stranded polynucleotides and allows annealing to be performed at lower temperatures (1). Components were brought to a temperature of 85 °C for 10 min, reduced to a temperature of 62 °C for 90 min, then to 52 °C for 90 min, and finally to 10 °C over the course of 10 min. Tethers were precipitated overnight in isopropanol at −20 °C, resuspended in TE buffer (Thermo Fischer Scientific), aliquoted, flash-frozen in liquid nitrogen, and stored at −80 °C.

Data Collection and Analysis. The optical trapping instrument has been described previously (2). Positional data were acquired at 2 kHz using custom software (LabVIEW), filtered at 1 kHz using an eight-pole low-pass Bessel filter, and boxcar averaged over a 20-point window to provide positional feedback to an active force clamp at a rate of 100 Hz. This rate is sufficient to measure speeds up to ∼1,000 nt/s. Uncertainties in force, caused by variations in bead size and systematic calibration errors, were estimated at 15%. Data were analyzed using Igor Pro software (WaveMetrics).

Calculation of Enzyme Velocity. Velocities of individual proteins were determined by dividing the observed change in extension by the corresponding time. Pause-free velocities for RNase R were obtained by limiting this analysis to motion over a comparatively pause-free region located 16–25 nm from the initial starting extension. This approach was deemed preferable to one that attempted to detect and remove pauses from the individual records.

Identification of Pauses in RNase R Records. To estimate the locations of sequence-dependent pauses by RNase R on the 155-bp dsRNA construct, each record initially was truncated to include only ∼1 s of positional data before the start of processive unwinding and ∼0.5 s after the cessation of enzyme motion. A histogram of occupancy versus position (bin size, 0.1 nm) was constructed for each trace, with peaks corresponding to enzyme pausing. The baseline starting position for each record was identified by fitting a Gaussian function to the first peak present (corresponding to the initial ∼1 s before motion) to allow alignment of traces. Once aligned, the data were summed for all records, with the contribution of each trace weighted by its integrated occupancy over the entire record (excluding the contribution of the initial and final peaks) to minimize the influence of outliers. The resulting positional occupancy histogram displayed three well-defined peaks, corresponding to pause positions for RNase R, the centers of which were identified by Gaussian fits. This analysis returned peaks at the same positions when performed using an unweighted sum of the individual records, and the calculation was robust against minor changes in histogram bin size.

A similar calculation was performed for records of RNase R motion over the 1,500-bp dsRNA construct carrying tandem repeats of the 155-bp sequence. Here, the individual contributions of records were weighted by the integrated occupancy over the first 20 nm traversed because not all enzymes moved beyond this location. To adjust for possible bias in the positional occupancy arising from only a subset of motors advancing to a given position, the occupancy at each position in the cumulative histogram was normalized relative to the number of enzymes that unwound the RNA up to that point.

Periodogram Analyses. To investigate stepping behavior, records were truncated to include only 0.5 s of positional data before processive motion, and the terminal backtracks of 12 ± 1 nt observed in records of PNPase (see main text) were excluded from the analysis. We calculated the positional separations of all points relative to each other within each record (i.e., pairwise distances), and histogrammed these positions (bin size, 0.001 nm) to determine the distribution of pairwise positions. These data then were summed across equivalent single-molecule records, with each record weighted by its respective number of pairwise separations, to ensure that each record contributed appropriately to the analysis and to minimize the influence of any potential outliers. A Fourier transform of this distribution was performed to identify peaks in spatial frequency. Each peak was fit to a Gaussian function to identify its center, and the SD of a peak was taken to represent an estimate of the positional error. Given the levels of signal and noise, we estimate that our system would not detect steps smaller than three nucleotides.

1. Casey J, Davidson N (1977) Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucleic Acids Res* 4(5): 1539–1552.

2. Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, Block SM (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* 438(7067): 460–465.

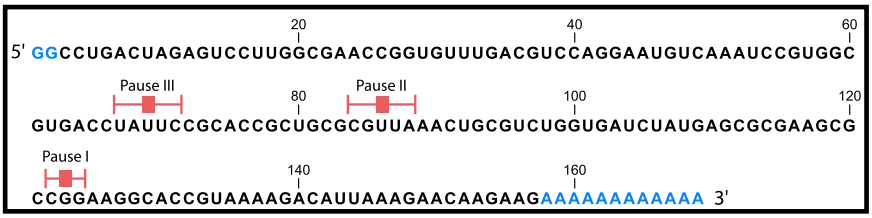

Fig. S1. Pause locations for RNase R on the 155-bp construct. Filled boxes (red) above the sequence indicate the locations of the three pauses (pause 1, pause II, and pause III) along with positional uncertainties (error bars). Two nucleotides at the 5′ end and the 12 nt at the 3′ end (colored blue) represent overhang sequences that do not hybridize with the complementary strand.

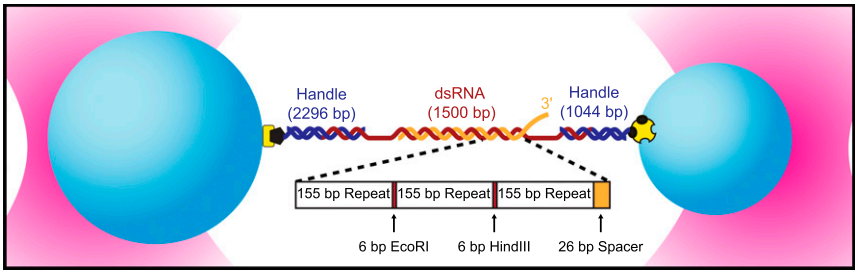

Fig. S2. Single-molecule assay—the 1,500-bp dsRNA construct. As in Fig. 1 in the main text, a double-stranded segment of RNA is placed under tension between two polystyrene beads (light blue) that are held in separate optical traps (pink). One strand of RNA (the digested strand, orange) carries a 3′, adenine-rich terminal overhang that serves as a loading site for RNase R. The complementary strand (red) is hybridized at both ends to DNA handles that are coupled to separate beads via biotin–avidin and digoxygenin–antibody linkages. The RNA immediately adjacent to the 3′ overhang consists of a 26-bp spacer element followed by three copies of the 155-bp sequence used previously. Tandem repeats (indicated) are separated by six-base pair restriction sites (red).

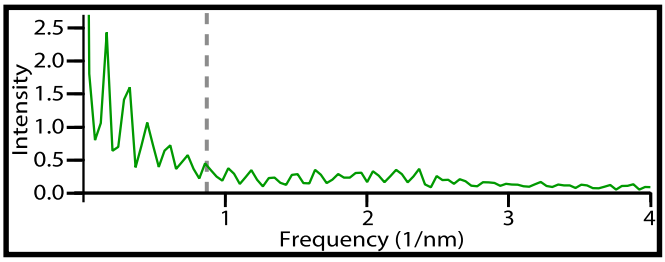

Fig. S3. Periodogram analysis of records of RNase R movement (green curve; n = 13). Analysis was performed on data collected from the 155-bp dsRNA construct. Unlike the corresponding periodograms for PNPase, no strong peak is evident at a position indicating steps of six or seven nucleotides (vertical gray dashed line, positioned as in Fig. 3*B* in the main text). The two peaks in the distribution at 0.16 nm^{-1} and 0.31 nm^{-1} correspond to the larger ~6-nm and ~3-nm separations found between sequence-dependent pauses, as reported earlier (Fig. 2*B* in the main text).

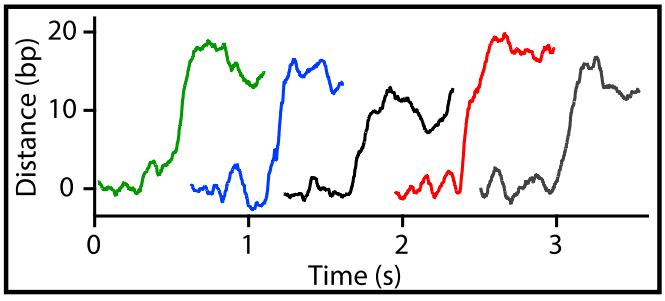

Fig. S4. Five representative single-molecule records of mutant PNPase (ΔKH–ΔS1) activity on 155-bp dsRNA constructs. Enzymes moved an average distance of 14 ± 1 bp (mean ± SE). Individual traces are displayed in different colors and are offset horizontally.