Supplementary Information

Diffusive ssDNA surveillance and bidirectional exoribonucleolytic digestion by RNase H2 ensure thorough clearance of RNA–DNA hybrids

Yi Zhao[1], Yuting Zhang[1], Yanan Li[1,2], Bingkai Cheng[1], Xia Zhang[1], Xintian Ma[1], Lishuang Chen[1], Shaojuan Wu[1], Jinfeng Chen[1], Cong Liu[3], Chao Liu[4], Wei Li[4], and Bo Sun[1]✉

[1]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [2]College of Life Sciences, Xinyang Normal University, Xinyang, China. [3]Interdisciplinary Research Center on Biology and Chemistry, State Key Laboratory of Chemical Biology, Shanghai Academy of Natural Sciences (SANS), Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, China. [4]Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China.

✉e-mail: sunbo@shanghaitech.edu.cn

This file includes

Supplementary Figures S1 – S14

Supplementary Table S1

Supplementary References S1 – S3

Supplementary Figures

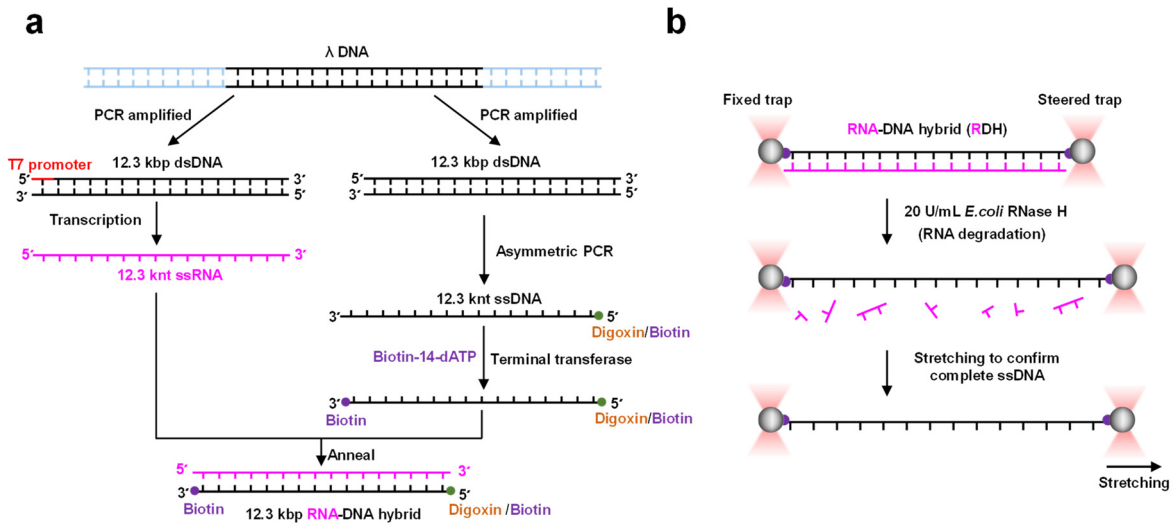

Figure S1. Preparation of the RDH substrate and ssDNA generation for OT assays. **(a)** Schematic of the construction of the RNA–DNA hybrid (RDH) template used in the OT assays. **(b)** Schematic of the workflow for generating ssDNA via RNase H-mediated degradation of the hybrid. In the *E. coli* RNase H channel, RNase H degrades the RNA strand of the hybrid under a force of 10 pN. Following degradation, the resulting ssDNA molecule is verified by the force–extension analysis[1].

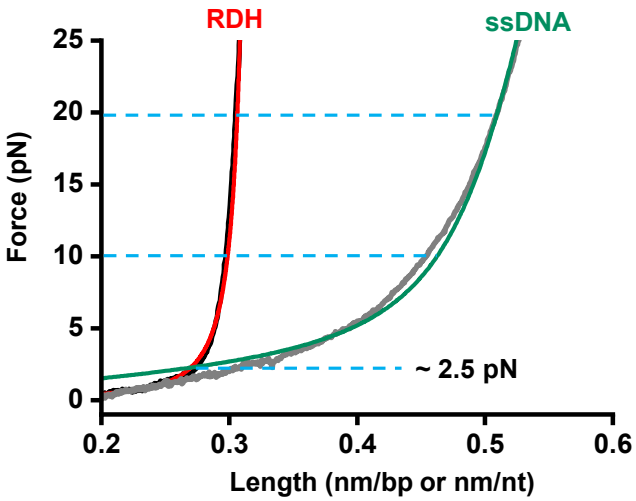

Figure S2. Force–extension characterization of the RDH and ssDNA substrates. Normalized force–extension curves for the RDH template (black) and ssDNA template (grey) in reaction buffer are shown. The two curves are fitted with the modified Worm-Like Chain (mWLC) model (red) and the Freely-Jointed Chain (FJC) model (green), respectively[2,3]. Based on these characterizations, the hybrid-to-ssDNA transition mediated by RNase H2 at forces over ~2.5 pN would cause an increase in the tether length.

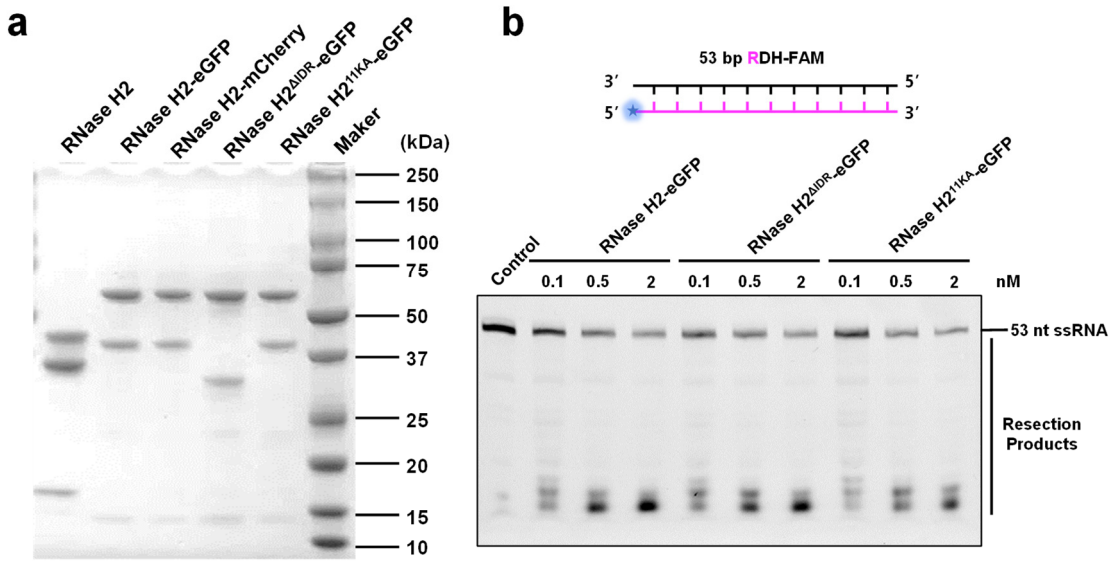

Figure S3. Purification and characterization of the fluorescently labeled RNase H2. **(a)** Representative gel electrophoresis images of purified recombinant proteins: RNase H2-eGFP, RNase H2-mCherry, RNase H2$^{\Delta IDR}$-eGFP, and RNase H2^{11KA}-eGFP. **(b)** Representative gel image showing cleavage of RDH substrates by RNase H2 variants. Specifically, 0.1 nM, 0.5 nM, and 2 nM of each RNase H2 variant were incubated with 10 nM of the 53 bp RDH substrate (sequences provided in **Table S1**) in cleavage buffer (25 mM Tris-HCl pH 8.0, 50 mM KCl, 10 mM $MgCl_2$, 0.1 mg/mL BSA, 1 mM DTT) at 25 °C for 10 minutes. Reactions were terminated by the addition of formamide gel loading buffer supplemented with 20 mM EDTA, followed by heating at 95 °C for 10 minutes. These findings verified that these labeled RNase H2 retained the nuclease activity toward RDHs.

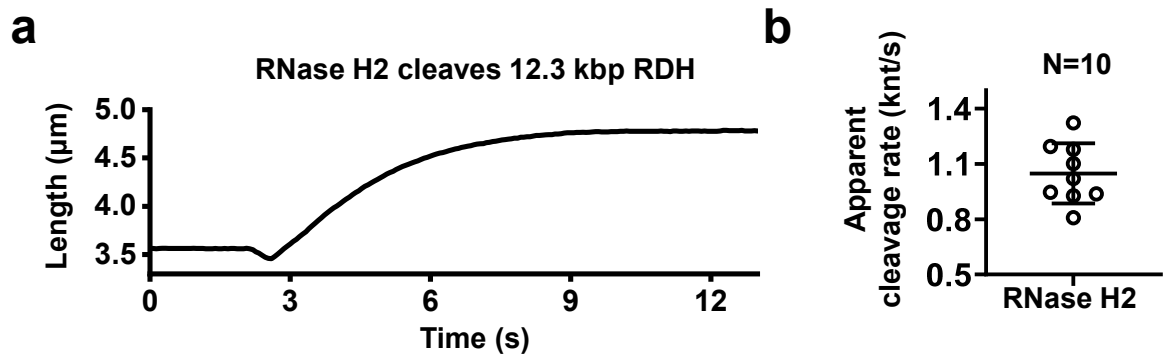

Figure S4. RNase H2 cleavage activity is not affected by Cy3 labeling of the RNA strand. (a) A representative real-time extension trajectory of a 12.3 kbp unlabeled RDH substrate cleaved by 10 nM wild-type RNase H2 under a constant force of 10 pN. **(b)** Quantification of the apparent cleavage rate for the cleavage of the unlabeled RDH substrate by wild-type RNase H2. Data are shown as mean ± SD, with N=10 independent tethers.

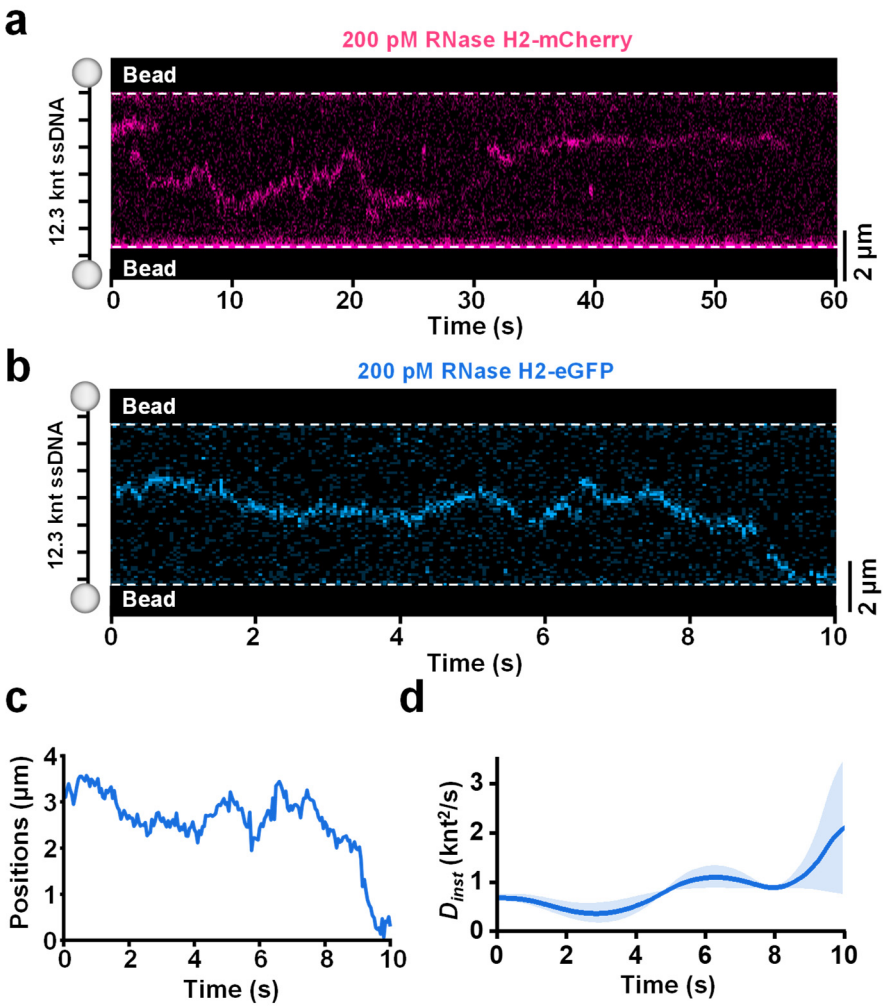

Figure S5. Analysis of the instantaneous diffusion coefficient of RNase H2 along ssDNA. (a) A representative kymograph depicting the 1D diffusion of RNase H2-mCherry (200 pM) along a single ssDNA molecule held under 10 pN. **(b)** A representative kymograph showing the 1D diffusion of RNase H2-eGFP (200 pM) along ssDNA under 10 pN. **(c)** A representative real-time trajectory of RNase H2-eGFP derived from the kymograph in **(b)**. **(d)** The time-resolved instantaneous diffusion coefficient (***D*$_{inst}$**) obtained from DeepDiffusion analysis of the trajectory shown in **(c)**, illustrating the dynamic fluctuations in diffusion speed. Data are shown as mean ± SD.

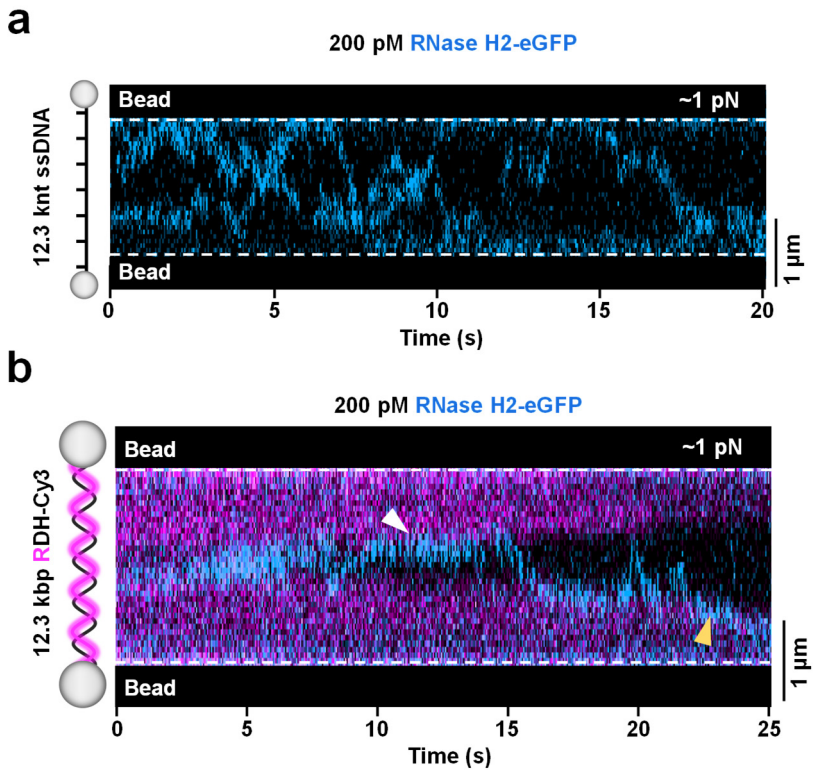

Figure S6. RNase H2 retains 1D diffusion and bidirectional exoribonucleolytic cleavage at low forces. (a) A representative kymograph showing the 1D diffusion of RNase H2-eGFP along ssDNA under a constant force of ~1 pN. **(b)** A representative kymograph showing the bidirectional exoribonucleolytic cleavage of an RDH substrate by RNase H2-eGFP at a low force (~1 pN). White and yellow arrowheads indicate two distinct cleavage directions.

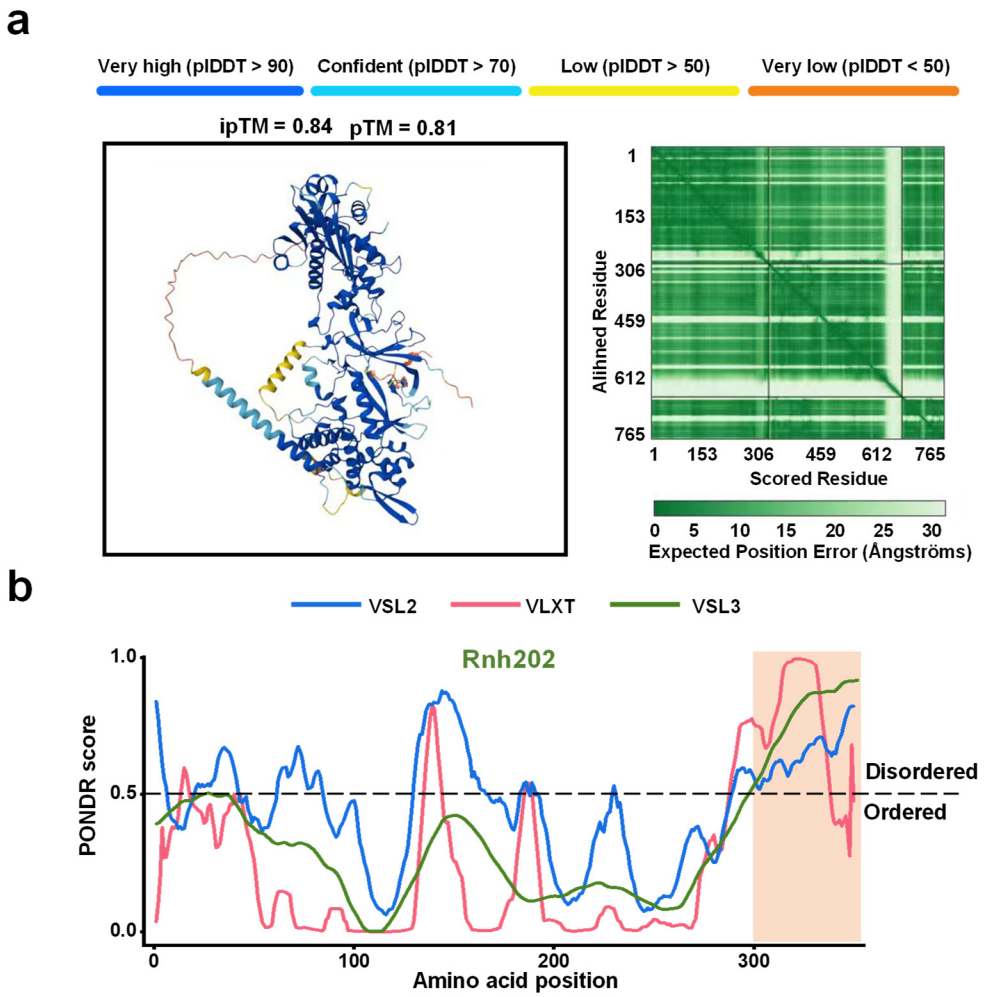

Figure S7. Prediction of the RNase H2 IDR. (a) Protein structure prediction of the RNase H2 heterotrimer generated by AlphaFold3. The AlphaFold Predicted Aligned Error (PAE) heatmap is shown on the right, depicting the expected distance error for each pair of residues and thereby providing an estimate of positional uncertainty within the predicted structure. **(b)** PONDR disorder scores for the Rnh202 subunit of RNase H2 predicted using multiple independent computational models.

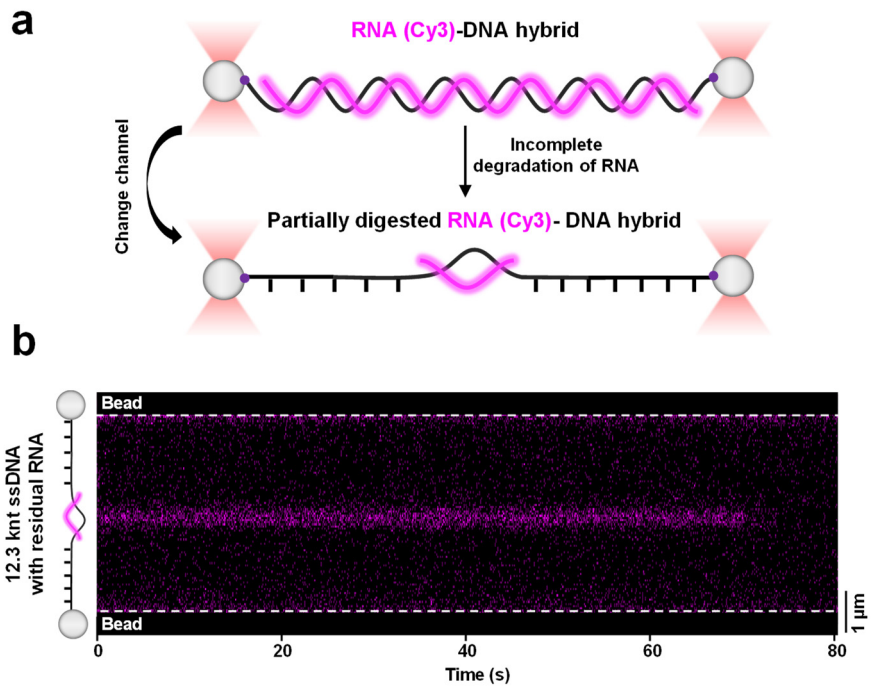

Figure S8. Generation and characterization of the ssDNA template with partial residual RNA fragments. (a) Schematic of the experimental procedure for preparing 12.3 knt ssDNA with residual RNA-Cy3. A single 12.3 kbp RDH-Cy3 tether is transferred into a channel containing 5 U/mL *E. coli* RNase H to achieve incomplete digestion of the RNA strand. The resulting template consists of a long ssDNA backbone sparsely decorated with residual Cy3-labeled short RNA fragments. **(b)** A representative kymograph of the generated template showing the residual Cy3-RNA signal (magenta) on the ssDNA tether under a constant force of 10 pN. In the absence of RNase H2, these residual RNA fragments persist stably, with no detectable signal loss over a period exceeding 1 minute.

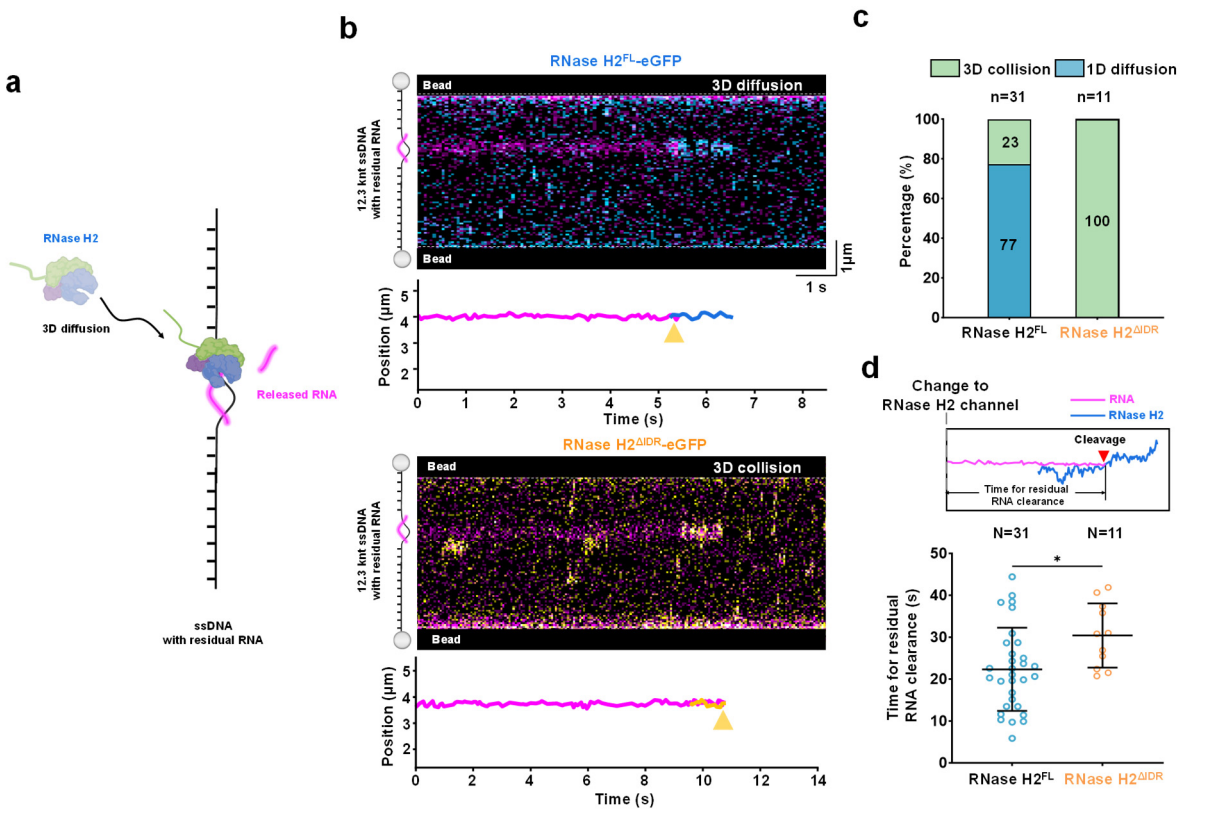

Figure S9. RNase H2 recognizes and cleaves residual RNA via 3D diffusion. (a) Schematic model illustrating that, in addition to ssDNA-based 1D diffusion, RNase H2 can directly access residual RNA fragments through 3D diffusion from solution and subsequently completes RNA cleavage. **(b)** Representative kymographs showing an RNase $H2^{FL}$-eGFP molecule (blue) or RNase $H2^{\Delta IDR}$-eGFP (orange) arriving from solution at a residual RNA-Cy3 site (magenta) on ssDNA, followed by rapid disappearance of the RNA-Cy3 signal within seconds. Yellow arrowheads mark the moment of RNA cleavage. Corresponding real-time trajectories of RNA-Cy3 and RNase H2-eGFP are shown below. **(c)** Stacked bar chart showing the percentage of residual RNA cleavage events mediated by 1D diffusion versus 3D collision for RNase $H2^{FL}$ and RNase $H2^{\Delta IDR}$. **(d)** Top: Schematic illustrating the measurement of the time required for residual RNA clearance. Timing begins upon transferring the ssDNA tether containing residual RNA fragments into the RNase H2 channel and ends when the RNA fluorescence signal is completely eliminated by the enzyme. Bottom: Scatter plot comparing the residual RNA clearance time between RNase $H2^{FL}$ and RNase $H2^{\Delta IDR}$. Data are presented as mean ± SD. Unpaired two-tailed t-test , *$p < 0.05$.

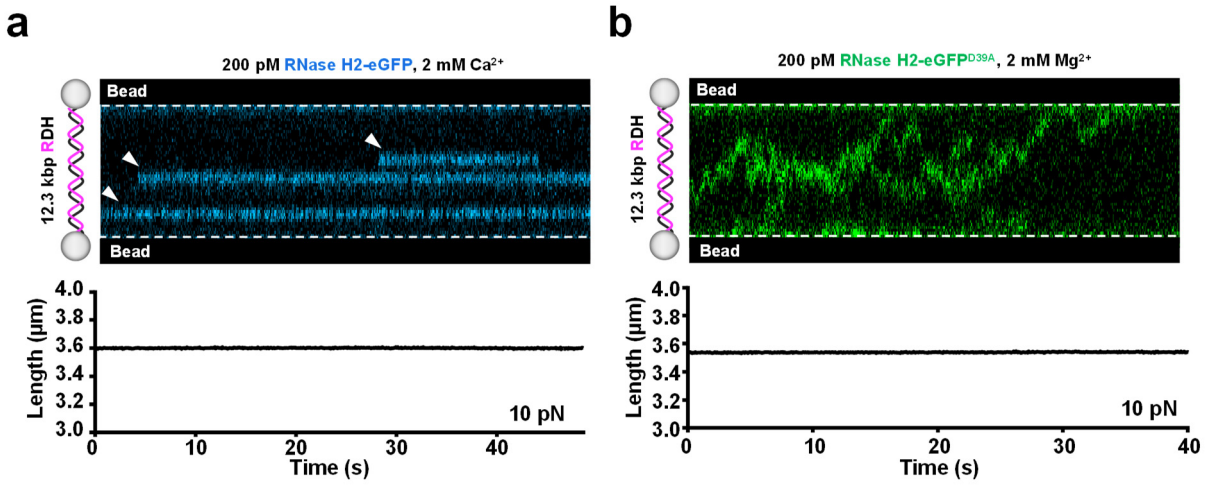

Figure S10. RNase H2 binds to RDH but fails to cleave under non-catalytic conditions. (a) A representative kymograph showing the binding of RNase H2-eGFP (200 pM) to a 12.3 kbp RDH molecule in the presence of 2 mM Ca^{2+}, with no translocation observed. The length of the examined molecule remained unchanged. **(b)** A representative kymograph of the catalytically inactive mutant RNase $H2^{D39A}$-eGFP (200 pM) bound to the RDH substrate in the presence of 2 mM Mg^{2+}. The mutant diffuses along the substrate but no increase in tether length is observed, indicating a complete loss of cleavage activity.

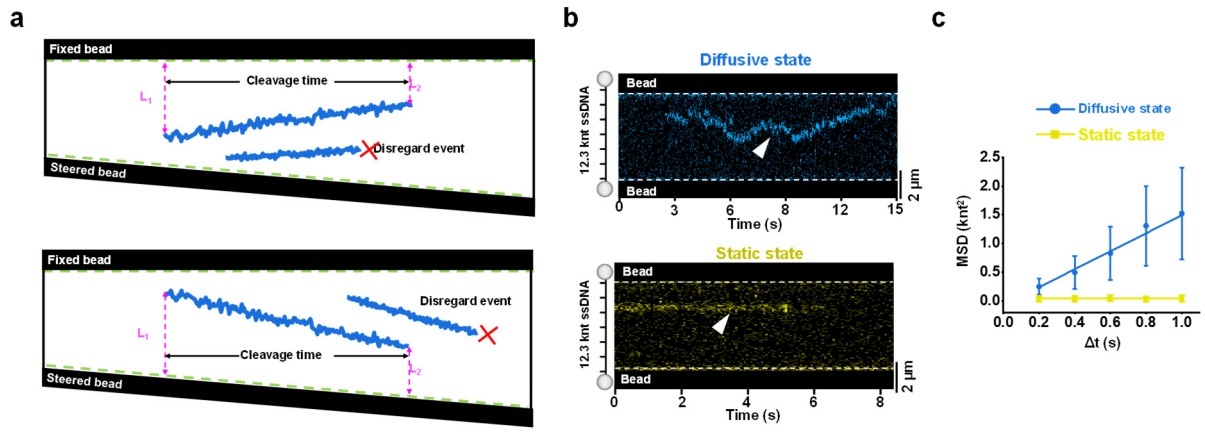

Figure S11. Measurements of cleavage processivity and rate, and MSD-based distinction between static binding and 1D diffusion. (a) The cleavage distance of RNase H2 is measured from the change in distance between the protein and either the fixed or mobile bead, depending on the cleavage direction (L_1-L_2). Displacement (μm) was converted to kbp using the RDH force–extension curve (**Fig. S2**), and the cleavage rate was calculated as distance divided by time. To ensure the accuracy of the data, adjacent cleavage events exhibiting significant temporal or spatial overlap were strictly excluded. As illustrated in the schematic, when another trajectory with distinct cleavage characteristics is present in the adjacent region above or below the target event, and the two events significantly overlap in time or space, the adjacent event is considered an interfering event and discarded (indicated by the red cross), thereby ensuring the purity of the single-event data included in the statistical analysis. **(b)** Two representative and illustrative kymographs showing distinct visual features of a protein in a diffusive state (top, blue) and a static state (bottom, yellow) on ssDNA. The diffusive state is characterized by continuous positional shifts of the fluorescence signal along the DNA over time, whereas the static state exhibits a fixed position with no apparent movement. These kymographs are provided for illustration purposes only and do not represent specific experimental data. **(c)** MSD–time plots comparing the diffusive (blue) and static (yellow) states. For the diffusive state, the MSD increases linearly with time, a hallmark of 1D Brownian diffusion. For the static state, the MSD remains flat and near zero, consistent with the absence of translational motion. Data are shown as mean ± SD.

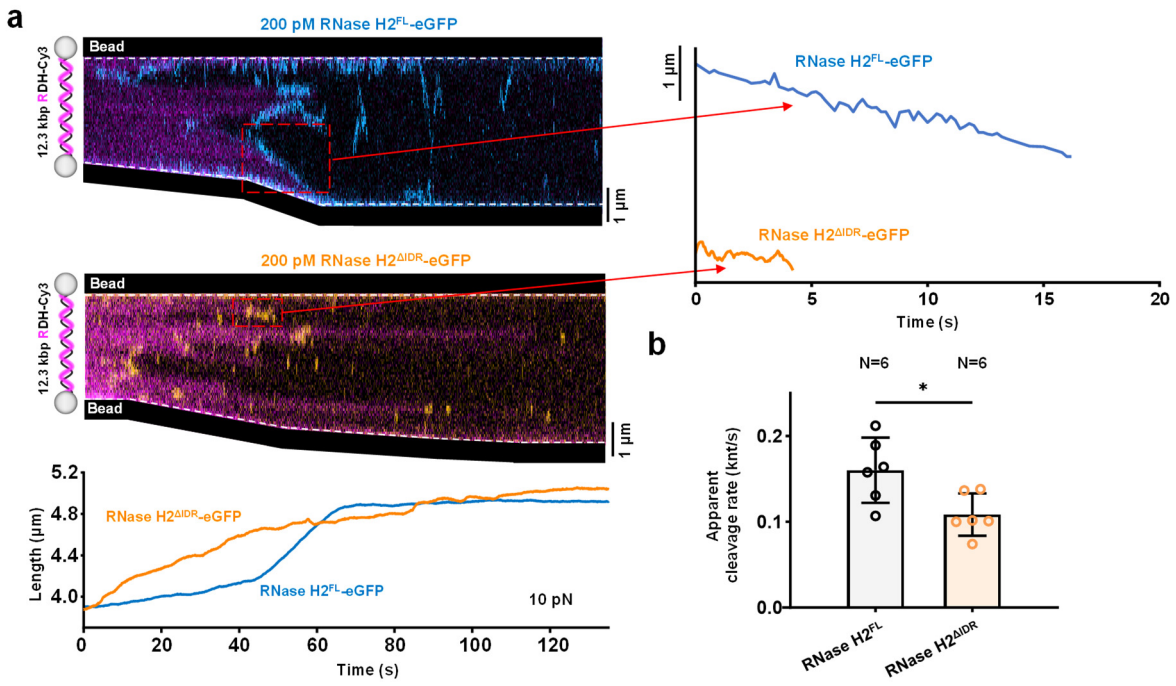

Figure S12. Cleavage of a 12.3 kbp RDH substrate by the RNase H2$^{\Delta IDR}$. (a) Representative kymographs showing the cleavage of a 12.3 kbp RDH-Cy3 substrate by RNase H2^{FL}-eGFP (top) and RNase H2$^{\Delta IDR}$-eGFP (bottom) (200 pM) under a constant force of 10 pN. The corresponding tether length traces are shown below the kymographs. The right panels show representative real-time trajectories of individual cleavage events for RNase H2^{FL}-eGFP (blue) and RNase H2$^{\Delta IDR}$-eGFP (orange), respectively. **(b)** Scatter plot comparing the apparent cleavage rate of RNase H2^{FL} or RNase H2$^{\Delta IDR}$. Data are presented as mean ± SD. N denotes the number of independent tethers. *p < 0.05. Unpaired two-tailed t-test.

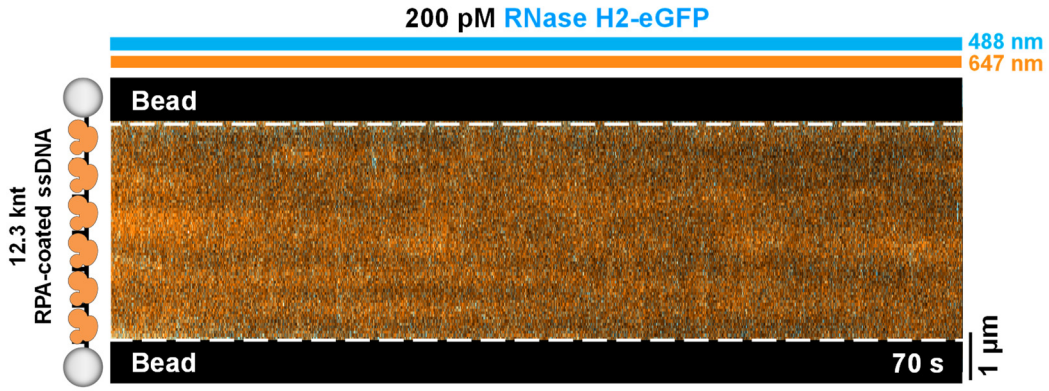

Figure S13. The response of RNase H2 to RPA-coated ssDNA. A representative kymograph showing the uniform coating of RPA-LD655 (orange) on ssDNA. RNase H2 loading on the ssDNA was prohibited.

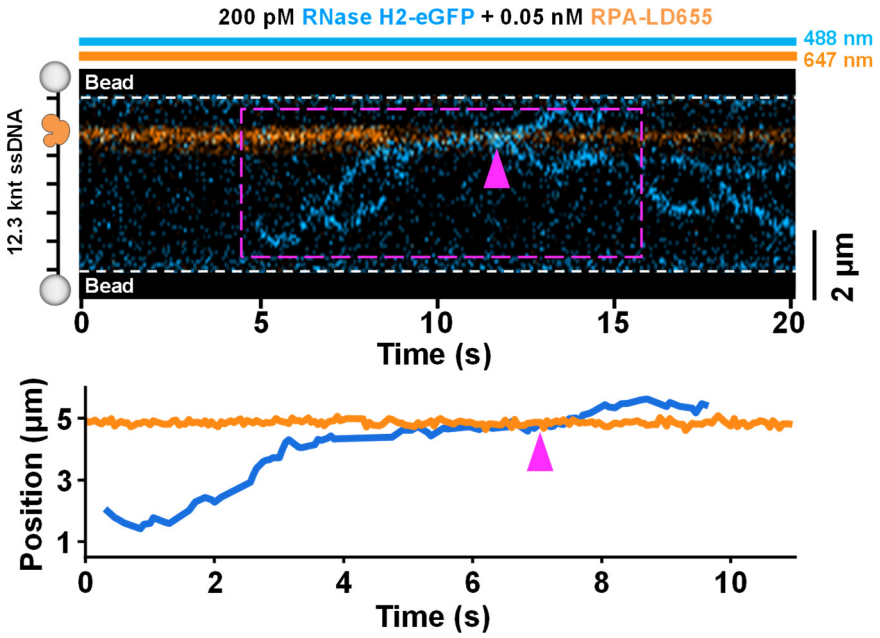

Figure S14. RNase H2 rarely bypasses RPA upon collision. A representative kymograph showing a diffusing RNase H2-eGFP molecule (blue) encountering sparsely bound RPA (orange) on ssDNA. Magenta triangles mark representative collision events between RNase H2 and RPA. Corresponding real-time trajectories of RPA-LD655 and RNase H2-eGFP derived from the magenta-boxed region are shown below. In 5 out of 55 detected collision events, RNase H2 was found to bypass ssDNA-bound RPA.

Supplementary Table 1. The sequences of primers and oligonucleotides used.

Template	Name	Sequence
Primers for 12.3 kbp dsDNA used in ssDNA generation	Forward	5′-AGCCTTTGCCTCGCTATACA-3′
	Reverse	5′-CAGCATAAGCGGCTACATGA-3′
Primers for 12.3 kbp dsDNA used in ssRNA generation	Forward	5′-TAATACGACTCACTATAGGAGCCTTTGCCTCGCTATACA-3′
	Reverse	5′-CAGCATAAGCGGCTACATGA-3′
Primers for end-labeled 12.3 knt ssDNA	Reverse1	5′-biotin-CAGCATAAGCGGCTACATGA-3′
	Reverse2	5′-digoxigenin-CAGCATAAGCGGCTACATGA-3′
53 bp RDH-FAM	53 nt ssDNA	5′-TGTAGTTGCCGTCGTCCTTGAAGAAGATGGTGATCTCCTGGACGTAGCCTTCG-3′
	53 nt ssRNA	5′-FAM-CGAAGGCUACGUCCAGGAGAUCACCAUCUUCUUCAAGGACGACGGCAACUACA-3′

Supplementary References

1. Li, Y. et al. RPA-ssDNA co-phase separation facilitates RAD51 enrichment during homologous recombination. Nucleic Acids Res 54, gkag586 (2026).
2. Smith, S.B., Cui, Y. & Bustamante, C. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. Science 271, 795-9 (1996).
3. Wang, M.D., Yin, H., Landick, R., Gelles, J. & Block, S.M. Stretching DNA with optical tweezers. Biophys J 72, 1335-46 (1997).