

Dissecting the multistep reaction pathway of an RNA enzyme by single-molecule kinetic “fingerprinting”

Shixin Liu*, Gregory Bokinsky*, Nils G. Walter†, and Xiaowei Zhuang*^{‡§¶}

Departments of *Chemistry and Chemical Biology and †Physics, and ‡Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138; and †Department of Chemistry, Single-Molecule Analysis Group, University of Michigan, Ann Arbor, MI 48109

Edited by Robert J. Silbey, Massachusetts Institute of Technology, Cambridge, MA, and approved April 4, 2007 (received for review November 30, 2006)

Single-molecule FRET is a powerful tool for probing the kinetic mechanism of a complex enzymatic reaction. However, not every reaction intermediate can be identified via a distinct FRET value, making it difficult to fully dissect a multistep reaction pathway. Here, we demonstrate a method using sequential kinetic experiments to differentiate each reaction intermediate by a distinct time sequence of FRET signal (a kinetic “fingerprint”). Our model system, the two-way junction hairpin ribozyme, catalyzes a multistep reversible RNA cleavage reaction, which comprises two structural transition steps (docking/undocking) and one chemical reaction step (cleavage/ligation). Whereas the docked and undocked forms of the enzyme display distinct FRET values, the cleaved and ligated forms do not. To overcome this difficulty, we used Mg^{2+} pulse-chase experiments to differentiate each reaction intermediate by a distinct kinetic fingerprint at the single-molecule level. This method allowed us to unambiguously determine the rate constant of each reaction step and fully characterize the reaction pathway by using the chemically competent enzyme–substrate complex. We found that the ligated form of the enzyme highly favors the docked state, whereas undocking becomes accelerated upon cleavage by two orders of magnitude, a result different from that obtained with chemically blocked substrate and product analogs. The overall cleavage reaction is rate-limited by the docking/undocking kinetics and the internal cleavage/ligation equilibrium, contrasting the rate-limiting mechanism of the four-way junction ribozyme. These results underscore the kinetic interdependence of reversible steps on an enzymatic reaction pathway and demonstrate a potentially general route to dissect them.

fluorescence resonance energy transfer | hairpin ribozyme | reaction kinetics | ribozyme

Enzymatic reactions often involve multiple kinetic steps such as substrate binding, folding of the enzyme–substrate complex, catalytic chemistry, and product dissociation. Because of the difficulty associated with the isolation of each intermediate species of the reaction, it is a challenging task to determine the entire set of microscopic rate constants that constitute such a multistep reaction scheme. Monitoring the reaction of a single molecule can potentially alleviate this problem. As an example, FRET has been exploited to probe conformational changes of single molecules in real time, making it well suited for monitoring structural intermediates (1–3). However, chemical reaction intermediates typically differ by only one or a few covalent bonds and often cannot be distinguished directly by the distance-sensitive probing based on FRET. Here, we demonstrate a kinetic “fingerprint” strategy to overcome this difficulty by using an RNA enzyme as a model system.

The hairpin ribozyme catalyzes a reversible site-specific RNA cleavage reaction (4, 5). The enzyme consists of two helix–loop–helix domains, with the cleavage site located within the substrate strand that makes up half of loop A (Fig. 1A). The two domains may be connected by a two-way helical junction (2WJ) in the minimal form or by a four-way helical junction (4WJ) in its natural form [supporting information (SI) Fig. 7]. The reaction mechanism of the hairpin ribozyme has been explored by

ensemble kinetic and structural studies (6–15) and single-molecule experiments (16–21). The enzyme catalyzes the cleavage reaction in several steps (Fig. 1B): (i) the extended (undocked) enzyme–substrate complex U_L first folds (docks) into a catalytically active state D_L ; (ii) cleavage of the substrate then occurs, generating the docked enzyme–product complex D_C ; (iii) the complex then unfolds (undocks), generating U_C ; and (iv) finally release of the cleavage product yields P_R . The conformational transitions (docking/undocking) can be readily monitored with FRET, but the cleaved and ligated forms of the ribozyme do not give distinct FRET values, making it difficult to determine the rate constants for every reaction step. Although it may be helpful to use modified, chemically inert analogs to assist quantitative analysis of the kinetic scheme (16, 19) as is often done in mechanistic studies (22), this approach has the caveat that analogs may convey significantly different properties to the enzyme–substrate complex than the wild-type substrate (20).

In this work, we dissected the multistep reaction pathway of the 2WJ hairpin ribozyme and determined the rate constant of each reaction step without the use of modified analogs. It was achieved by combining single-molecule FRET with multiple sequential buffer exchanges such that each of the five reaction states (U_L , D_L , D_C , U_C , and P_R) can be distinguished by a distinct FRET time sequence, which we refer to as a kinetic fingerprint. We find that the overall reaction is rate-limited by the structural transitions and the internal chemistry equilibrium between cleavage and ligation. This rate-limiting mechanism is different from that described previously for a 4WJ ribozyme where the overall reaction rate is primarily limited by the internal cleavage kinetics (20), suggesting an important effect of the interdomain junction on the reaction kinetics.

Results

Multistep Reaction Observed at the Single-Molecule Level. To monitor the hairpin ribozyme reaction via FRET, we attached a donor/acceptor pair (Cy3/Cy5) to the ends of the two helix–loop–helix domains of the RNA (Fig. 1A). Individual molecules were immobilized on the surface via streptavidin–biotin linkage to facilitate single-molecule detection. In the absence of Mg^{2+} , the molecules displayed a FRET value of ≈ 0.2 . When a buffer containing 12 mM Mg^{2+} was added, the FRET level exhibited an instantaneous, but small, shift to 0.30, which was followed by stochastic transitions to FRET ≈ 0.8 (more precisely 0.76) (Fig. 1C and D and SI Fig. 8). These two FRET levels (0.3 and 0.8)

Author contributions: S.L. and X.Z. designed research; S.L. performed research; S.L., G.B., and N.G.W. contributed new reagents/analytic tools; S.L. analyzed data; and S.L., N.G.W., and X.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: 2WJ, two-way helical junction; 4WJ, four-way helical junction; HMM, hidden Markov modeling; fps, frames per second.

[¶]To whom correspondence should be addressed. E-mail: zhuang@chemistry.harvard.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0610597104/DC1.

© 2007 by The National Academy of Sciences of the USA

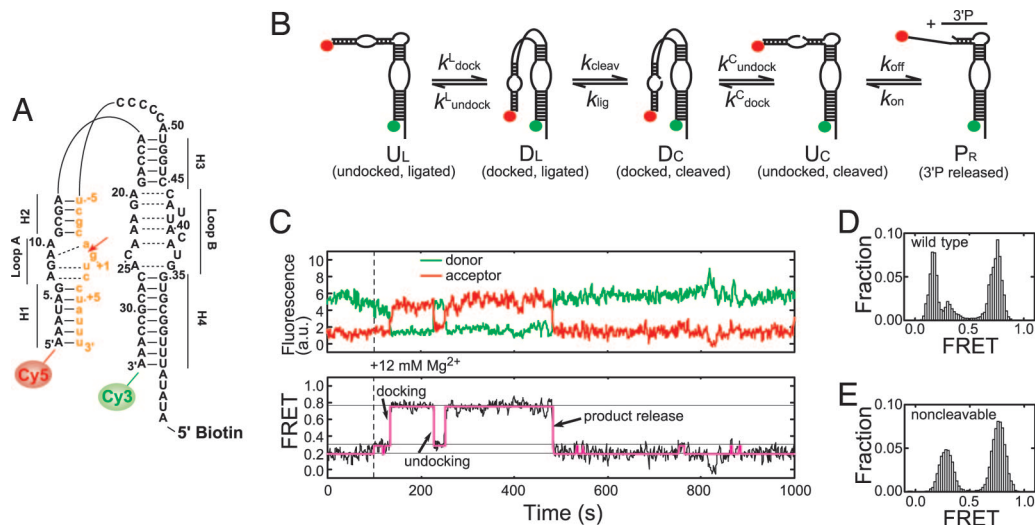


Fig. 1. Multistep catalytic reaction by the hairpin ribozyme. (A) The 2WJ hairpin ribozyme used in this study. An AC₅ linker was added to the ribozyme to keep substrate or the 5' product linked to the ribozyme. It has been shown that the ribozyme with AC₅ linker behaves similarly to the nicked 2WJ ribozyme (11, 17). Substrate is colored in orange. The cleavage site is indicated by a red arrow. Biotin and fluorophores Cy3 and Cy5 were attached as indicated. (B) Multistep cleavage reaction scheme. (C) Representative time trace of Cy3 and Cy5 fluorescence and the corresponding FRET values of a ribozyme molecule undergoing cleavage. Mg²⁺ (12 mM) was added at 100 s. The trace was recorded at 1 frame per second (fps), and the HMM analysis was applied to the FRET trace (pink line). (D) Histogram of FRET values obtained from many time traces as shown in C, but counting only the part after Mg²⁺ addition. The three peaks at FRET = 0.18, 0.30, and 0.76 represent product-released (P_R), undocked (U_L and U_C), and docked (D_L and D_C) states, respectively. (E) FRET histogram for ribozymes complexed with noncleavable substrate (with the native 2'-OH of A-1 substituted with 2'-OME) in the presence of 12 mM Mg²⁺.

have previously been identified as the undocked and docked state, respectively; mutations and changes in ionic conditions destabilizing the docked state perturb the transition kinetics between the two levels in a predictable manner (16, 17, 19). Eventually, molecules underwent cleavage and release of the 3' cleavage product (3'P) to yield the product-released state P_R, characterized by a stable FRET ≈ 0.2 (more precisely 0.18) level (Fig. 1 C and D and SI Fig. 8). When a high concentration of 3'P was added to the buffer to prevent product dissociation or when a ribozyme with a noncleavable substrate was probed (Fig. 1 E), the FRET trajectories in the presence of Mg²⁺ only displayed stochastic switching between FRET levels of 0.3 and 0.8, without exhibiting the 0.2 FRET level, demonstrating that the transition to 0.2 FRET indeed indicates product release. The FRET = 0.2, 0.3, and 0.8 states are all relatively long-lived and can be readily distinguished by visual inspection. To remove potential human bias in determining the transitions between these states, we also used a previously developed hidden Markov modeling (HMM) analysis (ref 23; the HMM software was downloaded from <http://bio.physics.uiuc.edu/HaMMMy.html>) to locate the FRET transitions (purple line in Fig. 1C). The HMM analysis agrees well with visual inspection except for the occasional and very brief excursions from the FRET = 0.2 (or 0.3) level to FRET = 0.3 (or 0.2), which are most likely false identification of transitions caused by finite single-to-noise ratio of the FRET signal.

It is therefore evident that the five states along the reaction pathway (U_L, D_L, D_C, U_C, and P_R) degenerate into three distinct FRET levels (0.8 for D_L and D_C, 0.3 for U_L and U_C, and 0.2 for P_R). This degeneracy makes it difficult to deduce all individual rate constants of the reaction pathway. For instance, the internal cleavage and ligation rate constants (k_{cleav} and k_{lig}) cannot be readily determined as the D_L and D_C states cannot be differentiated by FRET. Even the rate constants k_{undock}^L and k_{undock}^C cannot be readily determined as the lifetime of the FRET = 0.8 state is the convolution of several elementary rate constants (k_{undock}^L , k_{undock}^C , k_{lig} , and k_{cleav}). To resolve the degeneracy and convolution, we designed single-molecule kinetic fingerprinting experiments to al-

low all reaction intermediates to be differentiated and the rate constant of each reaction step to be determined.

Rate Constants for Product Binding and Dissociation. The product-bound, undocked ribozyme (U_C) displayed a FRET of 0.3, whereas the product-released state (P_R) exhibited a FRET of 0.2. Although stable, long-lived FRET = 0.2 and 0.3 levels were clearly distinguishable from each other, rapid transitions between the two would be more difficult to probe. To determine the product binding and dissociation kinetics, we thus attached the fluorescence quencher dabcy1 to the 3' end of 3'P (3'P-D), designed to quench the Cy5 fluorescence when 3'P was bound (Fig. 2 A and B). After *in situ* generation of P_R using a sufficient incubation time (1 h) of enzyme–substrate complexes with Mg²⁺ so that most complexes (≈90%) underwent cleavage and populated state P_R, we added a subsaturating concentration (500 nM) of 3'P-D. Binding and dissociation of 3'P-D caused the Cy5 fluorescence to fluctuate between low (quenched) and high (unquenched) values (Fig. 2B). The transitions between quenched and unquenched states were determined by using two different automated analysis methods (threshold fitting and HMM) (for details of the analysis methods, see SI Fig. 9). Both methods generated nearly identical results. The dwell-time distributions of these two states can be fit by first-order kinetics (Fig. 2 C and D), from which we deduced the dissociation and binding rate constant $k_{\text{off}} = 2.6 \pm 0.1 \text{ s}^{-1}$ and $k_{\text{on(obs)}} = 2.8 \pm 0.1 \text{ s}^{-1}$ at 500 nM of 3'P-D, respectively. The binding rate increased linearly with the 3'P-D concentration (data not shown), the slope of which yielded a bimolecular binding rate $k_{\text{on}} = 5.6 \pm 0.2 \mu\text{M}^{-1}\text{s}^{-1}$ and a corresponding equilibrium dissociation constant $K_{\text{d}} = 0.46 \pm 0.04 \mu\text{M}$.

Docking Rate Constants for Ligated and Cleaved Forms of the Ribozyme. The docking rate constants can be deduced from the lifetimes of the undocked states U_L and U_C, both of which can be readily identified in FRET trajectories. To determine k_{dock}^L , the ribozymes were first placed in buffer lacking Mg²⁺. Upon addition of Mg²⁺ with a home-built buffer-exchange apparatus

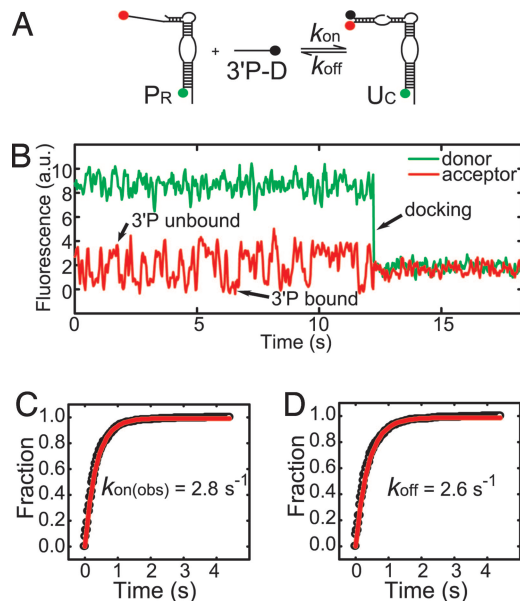


Fig. 2. Binding and dissociation of the cleavage product 3'P. (A) Schematic of experimental principle. 3'P with fluorescence quencher dabcyI (3'P-D) was added to the immobilized and cleaved ribozyme in the presence of Mg^{2+} . The green, red, and black dots indicate Cy3, Cy5, and dabcyI, respectively. (B) Representative time traces of Cy3 and Cy5 fluorescence of a single ribozyme molecule in the presence of 500 nM 3'P-D. The trace was recorded at the frame rate of 33 fps. In the undocked state, 3'P-D stochastically bound to and dissociated from the ribozyme, leading to repetitive quenching and de-quenching of the Cy5 signal, whereas the signal from the more distant Cy3 stayed large. In the docked state, 3'P-D remained bound, quenching signals from Cy3 and Cy5, now both in proximity to the quencher. (C) Cumulative dwell-time histogram of the 3'P unbound state in the presence of 500 nM 3'P-D. The histogram (circles) is fit to a single exponential (red line), yielding an observed binding rate constant of 2.8 s^{-1} . (D) Cumulative dwell-time histogram of the 3'P bound state (circles), which is fit to a single exponential (red line), yielding a dissociation rate constant of 2.6 s^{-1} .

with a dead time of $\approx 1 \text{ s}$ or less, FRET immediately shifted from 0.2 to 0.3 and then to 0.8, indicating docking. As the undocked state was catalytically inactive, the dwell time of the first FRET = 0.3 state before transition to FRET = 0.8 represented the lifetime of the U_L state (Fig. 3A). From 1,032 trajectories, we constructed a cumulative distribution of lifetimes that was fit well to first-order kinetics, yielding $k_{\text{dock}}^L = 0.013 \pm 0.001 \text{ s}^{-1}$ (Fig. 3C).

To measure k_{dock}^C , state P_R was again generated *in situ* by letting cleavage to take place for 1 h in a Mg^{2+} -containing buffer that lacked 3'P, and then $30 \mu\text{M}$ 3'P was added (Fig. 3B). At this concentration, the binding rate of 3'P was calculated to be $168 \pm 12 \text{ s}^{-1}$ (from $k_{\text{on}} = 5.6 \pm 0.2 \mu\text{M}^{-1}\text{s}^{-1}$), and the binding equilibrium was estimated to be 65 ± 7 , much in favor of the 3'P bound state. Indeed, an immediate shift from 0.2 to 0.3 was observed in the FRET trajectories upon addition of 3'P (Fig. 3B), confirming rapid binding. The large binding equilibrium of 3'P allowed us to attribute the time between the 3'P addition and the first docking transition entirely to the lifetime of the U_C state. The lifetime distribution obtained from 873 trajectories were fit to first-order kinetics, yielding $k_{\text{dock}}^C = 0.012 \pm 0.002 \text{ s}^{-1}$ (Fig. 3D).

Equilibrium Constants for Docking and the Internal Chemistry Step. In contrast to the straightforward determination of the docking rate constants, the undocking rate constants cannot be readily extracted from the lifetimes of the FRET = 0.8 state, which represents a mix of both ligated (D_L) and cleaved (D_C) forms of

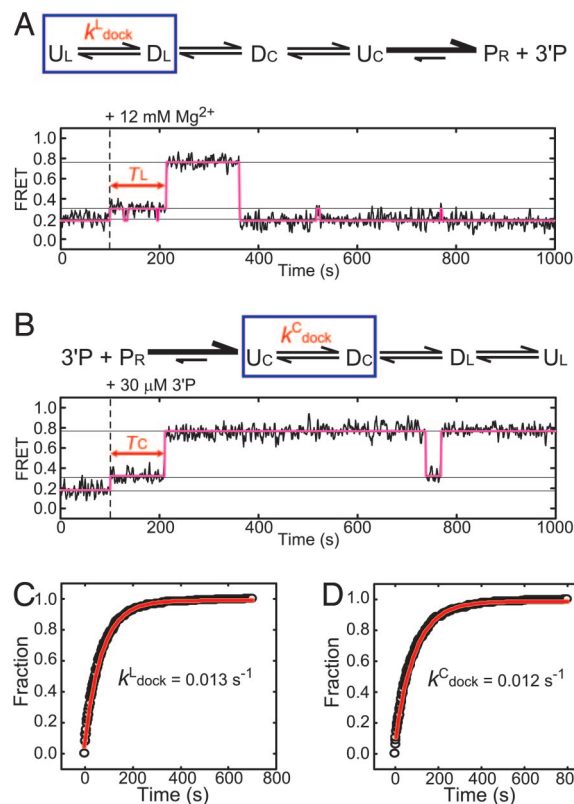


Fig. 3. Docking kinetics of the ligated and cleaved forms of the ribozyme. (A) Reaction scheme to determine k_{dock}^L with Mg^{2+} added at 100 s to induce docking, shown together with a representative single-molecule FRET time trace (1 fps) and HHM analysis (pink line). k_{dock}^L is determined from the time (T_L) between Mg^{2+} addition and docking. (B) Reaction scheme to determine k_{dock}^C with 3'P added at 100 s to induce docking, shown with a representative FRET time trace (1 fps) and HHM analysis (pink line). k_{dock}^C is determined from the time (T_C) between 3'P binding and the first docking event. (C) The cumulative histogram of T_L (circles), which is fit to a single exponential (red line), yielding a docking rate constant $k_{\text{dock}}^L = 0.013 \text{ s}^{-1}$. (D) The cumulative histogram of T_C (circles), which is fit to a single exponential (red line), yielding a docking rate constant $k_{\text{dock}}^C = 0.012 \text{ s}^{-1}$.

the docked states. The determination of the cleavage and ligation rate constants, k_{cleav} and k_{lig} , is also not straightforward, because D_L and D_C are not directly distinguishable by FRET. As a step toward resolving these difficulties, we determined the population of molecules in the U_L , D_L , D_C , and U_C states at equilibrium.

Equilibrium was obtained in a buffer containing 12 mM Mg^{2+} and saturating concentration of 3'P ($30 \mu\text{M}$) such that the state P_R was not significantly populated. After sufficient time (1 h), the populations of U_L , D_L , D_C , and U_C reached equilibrium. The FRET trajectories showed repeated docking and undocking transitions between 0.8 (D_L and D_C) and 0.3 (U_L and U_C) FRET levels (SI Fig. 10). To further distinguish U_L from U_C and D_L from D_C at equilibrium, we used sequential buffer exchanges to remove and replenish Mg^{2+} at specific times and generated characteristic FRET time sequences (fingerprints) for these four states (Fig. 4A). At time t_0 , an unfolding buffer (10 mM EDTA, no Mg^{2+} , no 3'P) was added to force undocking and 3'P release. Then 100 s later at t_1 , a folding buffer containing 12 mM Mg^{2+} but no 3'P was added to trigger docking. The trajectories were monitored for 1,500 s after t_1 . Because $k_{\text{dock}}^L = 0.013 \text{ s}^{-1}$, the probability that a ligated molecule does not dock within 1,500 s would be negligibly small. Molecules originally in state U_L (FRET = 0.3) at equilibrium would stay in U_L after t_0 (signified

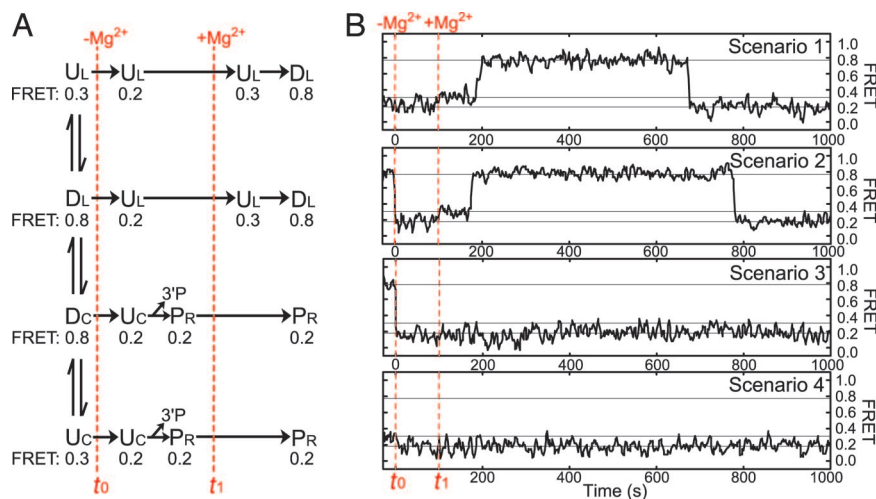


Fig. 4. Equilibrium constants for docking and the internal chemistry. (A) Schematic of the sequential buffer exchange experiment. Equilibrium among state U_L , D_L , D_C , and U_C was first reached by placing molecules in a solution containing 12 mM Mg^{2+} and 30 μM 3'P for > 1 h. At time t_0 , an unfolding buffer containing no Mg^{2+} or 3'P but 10 mM EDTA was added to force undocking and product release. Then a folding buffer containing 12 mM Mg^{2+} but no 3'P was added at t_1 to trigger docking. Molecules in state U_L , D_L , D_C , and U_C at t_0 showed distinct kinetic fingerprints, allowing them to be clearly distinguished. (B) Representative FRET time traces (1 fps) for each of the four scenarios. The first and second buffer exchanges were at 0 and 100 s, respectively. Scenarios 1, 2, 3, or 4 correspond to a molecule at state U_L , D_L , D_C , or U_C at t_0 , respectively.

by a slight shift to FRET = 0.2 caused by removal of Mg^{2+}). After Mg^{2+} was replenished at t_1 , these molecules would eventually dock to D_L (FRET = 0.8), as exemplified by the trace in scenario 1 of Fig. 4B. Molecules originally in D_L would undock to U_L at t_0 . After Mg^{2+} was added at t_1 , these molecules would eventually dock to D_L (scenario 2). Molecules originally in D_C would undock to U_C and release 3'P rapidly at t_0 , generating P_R . After t_1 , these molecules would stay in P_R (scenario 3). Molecules originally in U_C would release 3'P at t_0 and stay in P_R after t_1 (scenario 4). Thus, these characteristic fingerprints allowed us to clearly differentiate molecules that were in states U_L , D_L , D_C , or U_C at time t_0 . From a total of 824 molecules, we counted 21 in state U_L , 591 in D_L , 47 in D_C , and 165 in U_C . The equilibrium constants for docking were determined to be $K_{dock}^L = k_{dock}^L/k_{undock}^L = P_{D_L}/P_{U_C} = 28 \pm 6$ and $K_{dock}^C = k_{dock}^C/k_{undock}^C = P_{D_C}/P_{U_C} = 0.28 \pm 0.06$, where P_X denotes the population of molecules in state X . The equilibrium constant for the internal chemistry reaction is $K_{int} = k_{lig}/k_{cleav} = P_{D_L}/P_{D_C} = 13 \pm 2$.

Notably, the hairpin ribozyme exhibited heterogeneous undocking kinetics as shown (16, 17, 19). However, the majority ($\approx 60\%$) of molecules were catalytically highly active with stably docked states, whereas only a minor fraction appeared to dock into unstable docked states with rapid undocking kinetics (16, 17, 19). Single-molecule time traces allowed us to distinguish and isolate the major population, and our analyses in this work were focused on this population (SI Fig. 10).

Undocking Rate Constants for Ligated and Cleaved Forms of the Ribozyme. Now that the docking rate constants and the docking equilibrium constants have been determined unambiguously for the ligated and cleaved forms of the ribozyme, the undocking rate constants can be deduced readily. We obtained $k_{undock}^L = k_{dock}^L/K_{dock}^L = 0.00045 \pm 0.00015$ s $^{-1}$, and $k_{undock}^C = k_{dock}^C/K_{dock}^C = 0.043 \pm 0.015$ s $^{-1}$. Remarkably, the undocking rate constant for the cleaved form of the ribozyme was 100-fold faster than that of the ligated form.

Rate Constants for Internal Cleavage and Ligation. To determine k_{lig} and k_{cleav} , we revisited the cleavage time traces as exemplified in Figs. 1C and 5A. Because 3'P dissociated rapidly ($k_{off} = 2.6$ s $^{-1}$) from the enzyme with 3'P absent from the buffer, U_C would be

extremely short-lived and rapidly evolve into a stable FRET = 0.2 level representing state P_R . Thus product release was manifested as a direct transition from FRET = 0.8 to FRET = 0.2 level in our 1-s time resolution traces, evident by both visual inspection and HMM analysis (Fig. 5A). All FRET = 0.3 states observed in our trajectories corresponded to U_L . All FRET = 0.8 levels, except for the last one, started with state D_L and also ended with D_L , whereas the very last FRET = 0.8 event, right before adopting the P_R state, started with D_L and ended with D_C . Because most trajectories showed only one FRET = 0.8 event due to slow undocking, we focused on the dwell time T of the last docked state (Fig. 5A). We denoted the probability that T adopts a value between t and $t + \Delta t$ as $p_{last}(t)\Delta t$. This probability is related to the normalized population of molecules that enter state U_C (or equivalently P_R) between t and $t + \Delta t$, i.e., $p_{last}(t)\Delta t = \Delta P_{U_C}(t) = k_{undock}^C P_{D_C}(t)\Delta t$, if we assume that all molecules are in state D_L at $t = 0$. The equations governing the kinetics of these last FRET = 0.8 events are:

$$\frac{dP_{D_L}(t)}{dt} = -(k_{undock}^L + k_{cleav})P_{D_L}(t) + k_{lig}P_{D_C}(t),$$

$$\frac{dP_{D_C}(t)}{dt} = -(k_{undock}^C + k_{lig})P_{D_C}(t) + k_{cleav}P_{D_L}(t),$$

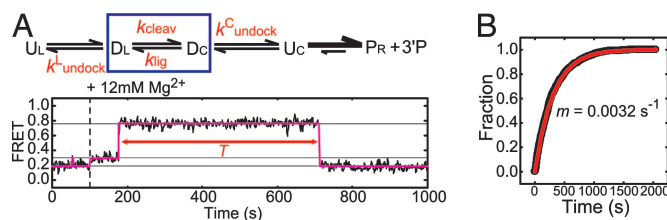


Fig. 5. Dwell-time analysis of the last docking event before 3'P release. (A) Reaction scheme with Mg^{2+} added at 100 s to induce docking and cleavage, shown with a representative FRET time trace (1 fps). The HMM analysis is shown in a pink line. The dwell time T of the last FRET = 0.8 event before releasing 3'P was analyzed. (B) The cumulative histogram of T (circles) is fit to a single exponential (red line) with a decay constant of $m = 0.0032$ s $^{-1}$, where m is a function of k_{undock}^L , k_{undock}^C , k_{lig} , and k_{cleav} .

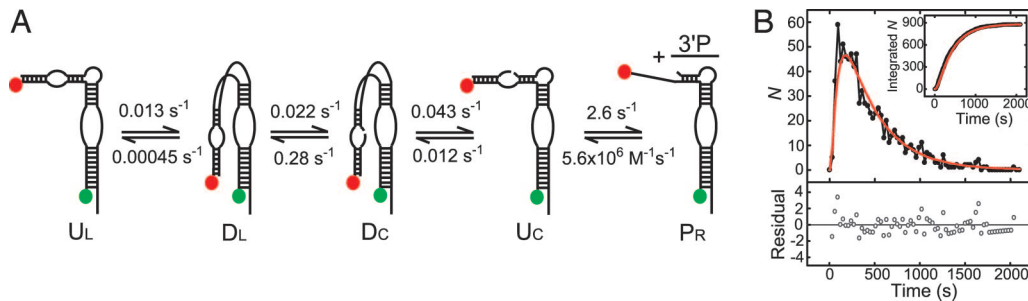


Fig. 6. The overall cleavage kinetics catalyzed by the 2WJ hairpin ribozyme. (A) Summary of the rate constants along the reaction pathway for the major (stably docked) population of the enzyme determined in this work. (B) (Upper) Number of molecules transitioned to the P_R state per 30-s interval (N) as a function of time with Mg^{2+} being added at $t = 0$. Red curve indicates data simulated ($N_{sim,i}$) using the kinetic rate constants determined above without any adjustable parameter and black curve indicates data obtained experimentally ($N_{exp,i}$). (Inset) Integrated number of molecules accumulated in the P_R state as a function of time (simulated curve: red; experimental curve: black). (Lower) The residual plot showing the difference between the simulated and experimental data. A standardized residual is calculated as $(N_{exp,i} - N_{sim,i})/\sqrt{N_{sim,i}}$.

$$\frac{dP_{U_C}(t)}{dt} = k_{undock}^C P_{D_C}(t).$$

Here, P_X denotes the normalized population of molecules in state X and the initial conditions for solving the above equations are $P_{D_L}(0) = 1$, $P_{D_C}(0) = 0$, and $P_{U_C}(0) = 0$. Note that P_{U_L} does not enter the equation because U_L is not populated during this process. However, undocking of the D_L state is competing with cleavage to impact the dwell time so that k_{undock}^L will affect the final dwell-time distribution. Solving the above equations yielded

$$P(t) = \int_0^t P_{last}(\tau) d\tau = 1 - e^{-mt},$$

where

$$m = \frac{k - \sqrt{k^2 - 4(k_{undock}^L k_{lig} + k_{undock}^C k_{cleav} + k_{undock}^L k_{undock}^C)}}{2},$$

define k

$$= k_{undock}^L + k_{cleav} + k_{lig} + k_{undock}^C.$$

This is an approximate solution with $<5\%$ error (see *SI Text*).

From 831 cleavage trajectories, we derived the cumulative dwell-time distribution $P(t)$ of the last FRET = 0.8 event. Indeed, the distribution can be quantitatively fit to a single exponential function with $m = 0.0032 \pm 0.0001 \text{ s}^{-1}$ (Fig. 5B). From the above expression of m and the already determined values of k_{undock}^L , k_{undock}^C , and k_{lig}/k_{cleav} , we obtained $k_{cleav} = 0.022 \pm 0.006 \text{ s}^{-1}$ and $k_{lig} = 0.28 \pm 0.10 \text{ s}^{-1}$.

Overall Cleavage Kinetics. The above experiments allowed us to determine the rate constants of each individual step along the reaction pathway of the 2WJ hairpin ribozyme (Fig. 6A). From these rate constants, we simulated the overall cleavage time courses [number of molecules transitioned to the P_R state per unit time (red curve in Fig. 6B)] and total number of molecules accumulated in the P_R state (red curve in Fig. 6B Inset) after the addition of Mg^{2+} , without any adjustable parameter. These predictions agreed quantitatively well with the experimentally determined cleavage time courses (for 878 molecules) (black curves in Fig. 6B). The reduced χ^2 value χ_r^2 determined from the residue plot (Fig. 6B Lower) was calculated to be 0.96, very close to 1 and indicating a good agreement.

Discussion

We have designed single-molecule kinetic fingerprinting experiments to dissect the rate constants for each elementary step on the multistep reaction pathway catalyzed by the 2WJ hairpin ribozyme. Through this system, we have demonstrated that single-molecule kinetic fingerprinting by FRET is a powerful tool to characterize complex multistep reaction pathways.

The docking and undocking rate constants for the intact (ligated) 2WJ ribozyme–substrate complex were found to be $k_{dock}^L = 0.013 \text{ s}^{-1}$ and $k_{undock}^L = 0.00045 \text{ s}^{-1}$, respectively. The corresponding equilibrium is substantially biased toward docking ($k_{dock}^L/k_{undock}^L = 28$). Interestingly, the same or similar ribozymes with a noncleavable substrate (carrying a chemistry-blocking 2'-OMe A-1 instead of 2'-OH A-1 at the cleavage site) has a comparable docking rate constant (0.018 s^{-1}), but undocks 20-fold faster (0.01 s^{-1}) (9, 16, 17, 19). These results indicate that the structural dynamics derived from a nonreactive substrate analog do not faithfully represent those of the unmodified substrate. What may be the structural basis for this difference? Crystal structures have been solved for the hairpin ribozyme with noncleavable substrate analog harboring 2'-OMe A-1, with a substrate harboring the unmodified 2'-OH A-1 in the context of an inactivating G8 mutation of the ribozyme, with a vanadate mimicking the transition state, and with cleavage products (13–15). Molecular dynamics simulations have been used to simulate the structure with the unmodified 2'-OH A-1 substrate (24). From these structures, the ribose of 2'-OMe A-1 is proposed to be in the 2'-endo pucker conformation, twisting the 2'-O and G+1 phosphate further away from A38 than in the 2'-exo conformations proposed for the 2'-OH A-1, the transition state analog, and the product structures. As a result, at least one net hydrogen bond is lost between the A38 ribose and the active site, which is part of the loop–loop docking interaction network. This difference provides a likely mechanism by which 2'-OMe modification of A-1 may compromise docking stability. In contrast to the undocking rate constant of the ribozyme–substrate complex, its docking rate constant is essentially unaffected by the 2'-OMe modification of A-1, consistent with our previous result that the native tertiary contacts are largely absent from the docking transition state (17). Upon cleavage the docking equilibrium changes dramatically, in a manner similar to the 4WJ ribozyme (20). Whereas the docking rate remains largely unchanged ($k_{dock}^C = 0.012 \text{ s}^{-1}$), the undocking rate is accelerated by 100-fold ($k_{undock}^C = 0.043 \text{ s}^{-1}$). These docking and undocking rates are similar to those of the ribozyme complexed with a nonligatable 3'P analog (with a 5'-Me replacing the native 5'-OH) (data not shown).

The internal chemistry equilibrium for the 2WJ ribozyme is substantially biased toward ligation with $K_{\text{int}} = k_{\text{lig}}/k_{\text{cleav}} = 13$, which is significantly higher than a previously determined value ($K_{\text{int}} = 2.5$) based on the docking and undocking rate constants from the noncleavable substrate and product analogs (19). The chemical kinetics of the 2WJ ribozyme found here ($k_{\text{cleav}} = 0.022 \text{ s}^{-1}$, $k_{\text{lig}} = 0.28 \text{ s}^{-1}$ at pH 7.5) are comparable to those determined for the 4WJ form ($k_{\text{cleav}} = 0.01 \text{ s}^{-1}$, $k_{\text{lig}} = 0.35 \text{ s}^{-1}$ at pH 7.5) (20), suggesting similar local environments of the active sites. The two forms of the ribozyme, however, exhibit substantially different conformational dynamics. Compared with the 4WJ form, the docking rate constant of the 2WJ ribozyme in the ligated state and the undocking rate constant of its cleaved state are two or more orders of magnitude slower. The acceleration of docking and undocking for the 4WJ is likely caused by the intrinsic structural dynamics of the 4WJ, which facilitates the transition between the docked and undocked states (18). As a result, the rate-limiting mechanisms are substantially different for the 2WJ and 4WJ ribozymes. For the 2WJ ribozyme, undocking of the cleaved ribozyme–product complex is much slower than ligation, so internal cleavage and ligation reach equilibrium before undocking takes place. As a result, the overall cleavage reaction is rate-limited by docking and undocking ($k_{\text{dock}}^{\text{c}}$ and $k_{\text{undock}}^{\text{c}}$) and the internal equilibrium (K_{int}). In the case of the 4WJ ribozyme, rapid docking of the intact enzyme–substrate complex and undocking of the cleaved complex conspire to make the relatively slow bond scission rate-limiting (20). These differences underscore the important role of junction dynamics in the enzymatic reaction kinetics.

Interestingly, the hairpin ribozyme appears to be a meticulously optimized enzyme: stable docking of the intact ribozyme–substrate complex gives the enzyme ample time to complete cleavage. In the meantime, the potential disadvantage of inefficient product release caused by stable docking is overcome by the cleavage-induced acceleration of undocking of the enzyme–product complex. While this property is imbedded in the loop–loop interactions of both the 2WJ and 4WJ ribozymes, the presence of a 4WJ in the natural form of the enzyme appears to improve both properties by specifically accelerating both dock-

ing kinetics of the intact enzyme–substrate complex and undocking kinetics of the cleaved state.

Materials and Methods

RNA Preparation. The Cy5-labeled RNA strand (RzA) was purchased from the W. M. Keck Foundation Biotechnology Resource Laboratory (Yale University, New Haven, CT). Cy3 was coupled to the 3' end of RzA postsynthetically as described (16). The biotin-labeled RNA strand (RzB) and the 3' product 3'P were purchased from Dharmacon (Lafayette, CO). 3'P with the fluorescence quencher dabcy1 (3'P-D) was purchased from Trilink Biotechnologies (San Diego, CA). RNA was deprotected according to the manufacturers' protocols and then purified by denaturing PAGE and C8 RP-HPLC as described (9). To assemble the ribozyme, we annealed RzA and RzB strands by heating to 90°C for 1 min and then slowly cooling to room temperature over 45 min in annealing buffer (50 mM Tris-HCl, pH 7.5/50 mM NaCl/1 mM EDTA).

Single-Molecule FRET Measurements. The annealed biotinylated ribozyme was bound to a streptavidin-coated quartz slide surface via biotin–streptavidin interaction. The donor (I_D) and acceptor (I_A) fluorescence signals of optically resolved single molecules (characterized by single-step photobleaching) were detected on a total internal reflection fluorescence setup as described (25). The FRET value, defined as $I_A/(I_A + I_D)$, was monitored in real time for each individual molecule. The standard reaction buffer contained 50 mM Tris-HCl, pH 7.5 and 12 mM MgCl₂. For buffer lacking Mg²⁺, 12 mM MgCl₂ was replaced with 50 mM NaCl and 1 mM (or 10 mM as specified in the text) EDTA. An oxygen scavenging system [10% (wt/vol) glucose, 1% (vol/vol) β-mercaptoethanol, 300 μg/ml glucose oxidase, and 40 μg/ml catalase] was added to all imaging buffers to reduce photobleaching. Measurements were performed at 25°C.

We thank Elio Abbondanzieri for critically reading the manuscript and Taekjip Ha for making the HMM analysis software publicly available. This work is supported in part by the National Science Foundation, a Packard Science and Engineering Fellowship (to X.Z.), and the National Institutes of Health (N.G.W.). X.Z. is a Howard Hughes Medical Institute Investigator.

- Weiss S (2000) *Nat Struct Biol* 7:724–729.
- Ha T (2001) *Methods* 25:78–86.
- Zhuang X (2005) *Annu Rev Biophys Biomol Struct* 34:399–414.
- Walter NG, Burke JM (1998) *Curr Opin Chem Biol* 2:24–30.
- Fedor MJ (2000) *J Mol Biol* 297:269–291.
- Hegg LA, Fedor MJ (1995) *Biochemistry* 34:15813–15828.
- Esteban JA, Banerjee AR, Burke JM (1997) *J Biol Chem* 272:13629–13639.
- Nesbitt SM, Erlacher HA, Fedor MJ (1999) *J Mol Biol* 286:1009–1024.
- Walter NG, Hampel KJ, Brown KM, Burke JM (1998) *EMBO J* 17:2378–2391.
- Murchie AI, Thomson JB, Walter F, Lilley DM (1998) *Mol Cell* 1:873–881.
- Walter NG, Burke JM, Millar DP (1999) *Nat Struct Biol* 6:544–549.
- Pljevaljcic G, Klostermeier D, Millar DP (2005) *Biochemistry* 44:4870–4876.
- Rupert PB, Ferre-D'Amare AR (2001) *Nature* 410:780–786.
- Rupert PB, Massey AP, Sigurdsson ST, Ferre-D'Amare AR (2002) *Science* 298:1421–1424.
- Salter J, Krucinska J, Alam S, Grum-Tokars V, Wedekind JE (2006) *Biochemistry* 45:685–700.
- Zhuang X, Kim H, Pereira MJ, Babcock HP, Walter NG, Chu S (2002) *Science* 296:1473–1476.
- Bokinsky G, Rueda D, Misra VK, Rhodes MM, Gordus A, Babcock HP, Walter NG, Zhuang X (2003) *Proc Natl Acad Sci USA* 100:9302–9307.
- Tan E, Wilson TJ, Nahas MK, Clegg RM, Lilley DM, Ha T (2003) *Proc Natl Acad Sci USA* 100:9308–9313.
- Rueda D, Bokinsky G, Rhodes MM, Rust MJ, Zhuang X, Walter NG (2004) *Proc Natl Acad Sci USA* 101:10066–10071.
- Nahas MK, Wilson TJ, Hohng S, Jarvie K, Lilley DM, Ha T (2004) *Nat Struct Mol Biol* 11:1107–1113.
- Pljevaljcic G, Millar DP, Deniz AA (2004) *Biophys J* 87:457–467.
- Fersht A (1999) *Structure and Mechanism in Protein Science* (Freeman, New York).
- Mckinney SA, Joo C, Ha T (2006) *Biophys J* 91:1941–1951.
- Rhodes MM, Reblova K, Sponer J, Walter NG (2006) *Proc Natl Acad Sci USA* 103:13381–13385.
- Bokinsky G, Nivon LG, Liu S, Chai G, Hong M, Weeks KM, Zhuang X (2006) *J Mol Biol* 361:771–784.

Supporting Information for

Dissecting the multi-step reaction pathway of an RNA enzyme by single-molecule kinetic “fingerprinting”

Shixin Liu, Gregory Bokinsky, Nils G. Walter and Xiaowei Zhuang

Supporting Figures:

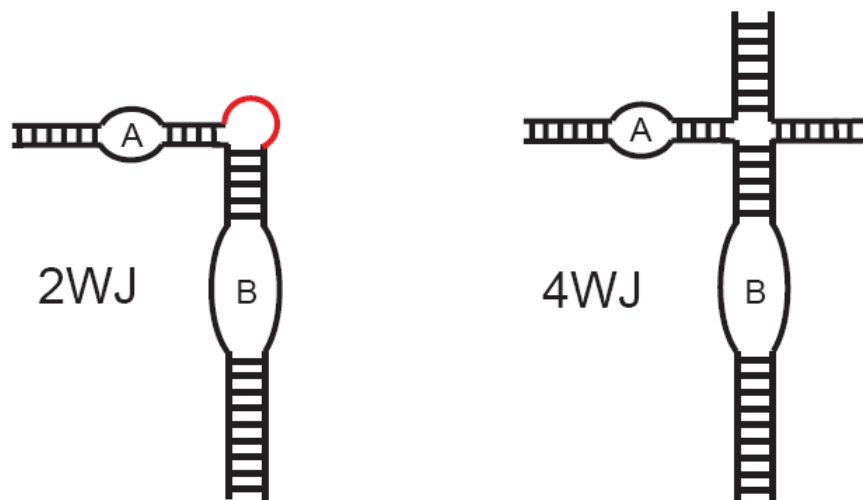


Fig. 7. Schematic secondary structures of the two-way junction (2WJ) and four-way junction (4WJ) hairpin ribozymes. The hairpin ribozyme consists of two helix-loop-helix domains. The two domains may be connected by a two-way helical junction in the minimal form or by a four-way helical junction in its natural form. The AC₅ loop connecting the two-way junction shown here as the red circular loop may be removed (as in the nicked form of the 2WJ ribozyme) without substantially affecting the conformational dynamics and catalytic activity of the enzyme (Bokinsky et al. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 9302-9307).

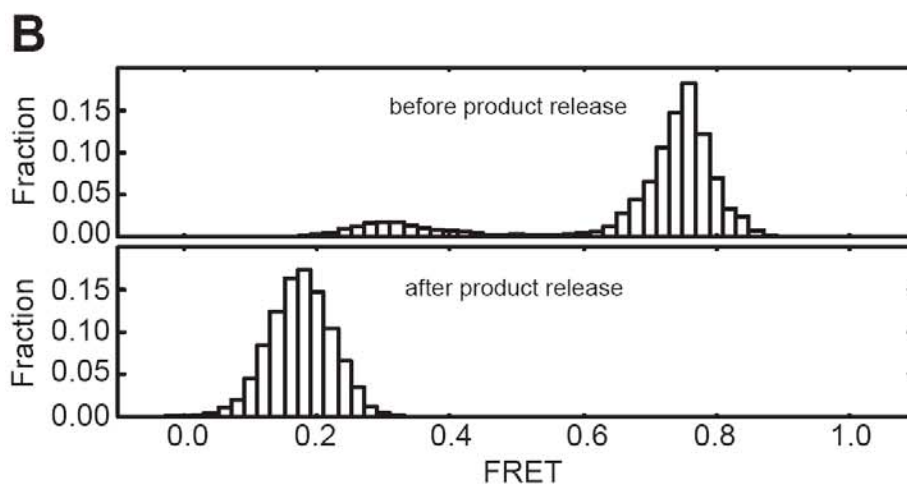
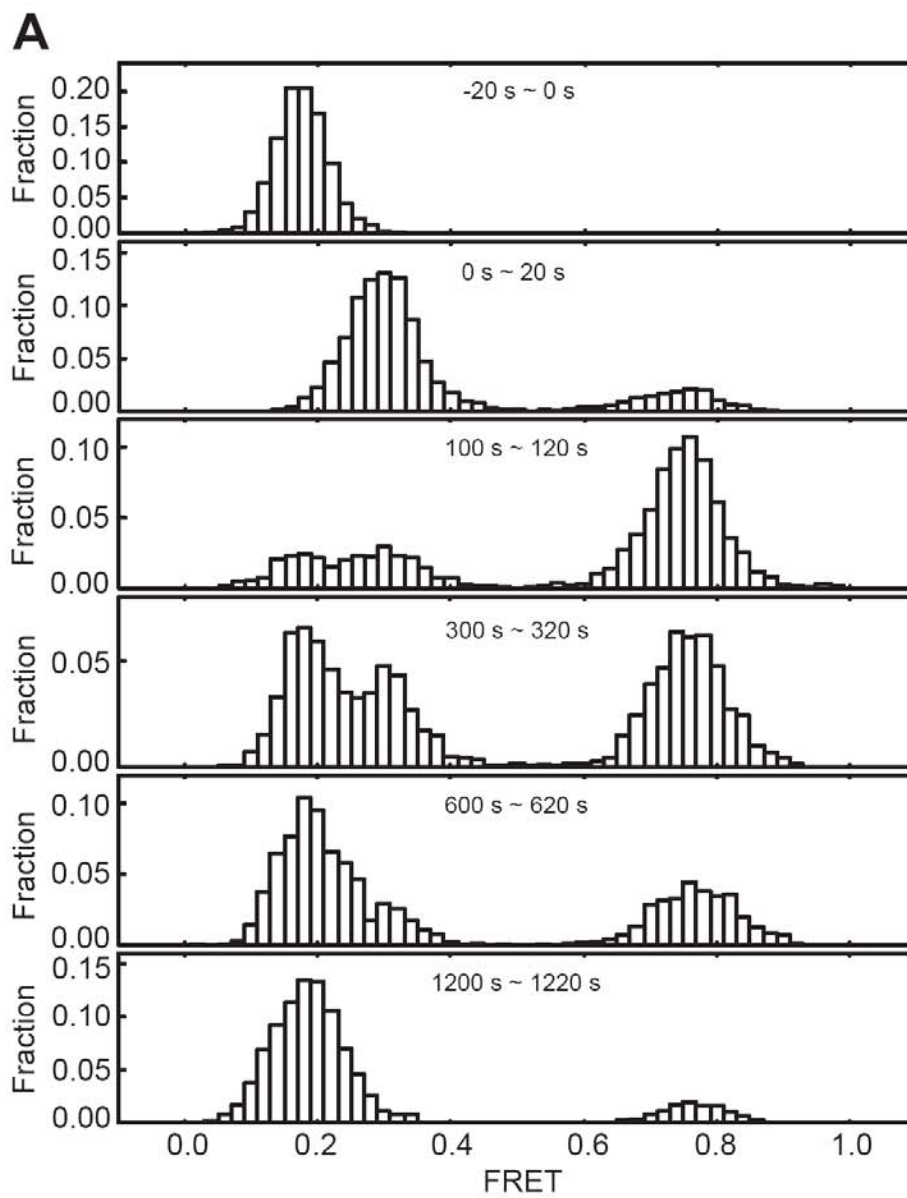


Fig. 8. FRET histograms of the 2WJ hairpin ribozyme. (A) FRET histograms constructed from different 20-second time intervals in the FRET time trace such as the one shown in Fig. 1C, with 0 s as the time point when Mg^{2+} was added. From top to bottom: -20 s to 0 s, 0 s to 20 s, 100 s to 120 s, 300 s to 320 s, 600 s to 620 s, and 1200 s to 1220 s. The evolution of FRET histograms clearly links the FRET levels to different physical states: before addition of Mg^{2+} , the ribozyme exhibits a FRET level averaged at 0.18; after addition of Mg^{2+} , two new FRET levels appear, with the 0.30 FRET level representing the undocked states (U_L and U_C) and the 0.76 FRET level representing the docked states (D_L and D_C); as time further increases, a FRET = 0.18 peak surfaces again, but this time due to cleavage and product release (the P_R state). As expected, when a high concentration of 3'P was added to the buffer to prevent product dissociation or when a ribozyme with a non-cleavable substrate was probed, the FRET level of 0.18 does not appear in the presence of Mg^{2+} (Fig. 1E). (B) FRET histograms before and after release of 3'P. For each trace, a single transition event from 0.76 FRET to 0.18 FRET was identified as the 3'P release event. In the before-product-release histogram, two peaks at FRET = 0.30 and FRET = 0.76 exist, representing the undocked (U_L and U_C) and docked (D_L and D_C) states. In the after-product-release histogram, only a single peak at FRET = 0.18 exists, representing the product-released state (P_R). The transitions from 0.76 FRET to 0.18 FRET were identified by visual inspection or by Hidden Markov Modeling (HMM) analysis (See Fig. 9 for details of this analysis method), both methods yielding the same answer.

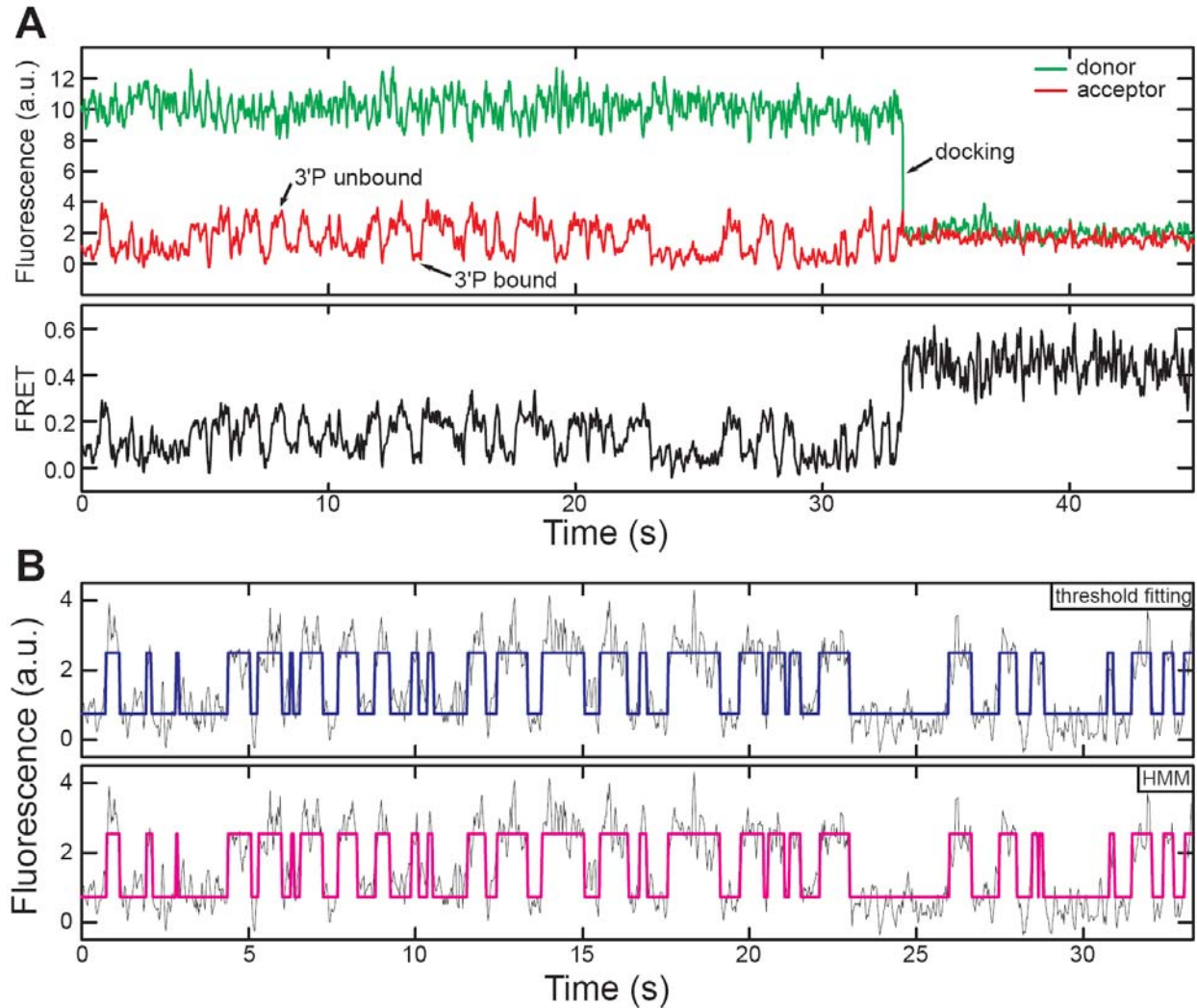


Fig. 9. Analysis of the binding and dissociation kinetics of the cleavage product 3'P. 3'P with fluorescence quencher dabcyI (3'P-D) was added to the immobilized and cleaved ribozyme in the presence of Mg^{2+} . (A) A representative time trace of Cy3 and Cy5 fluorescence and of FRET for a single ribozyme molecule in the presence of 500 nM 3'P-D. The trace was recorded at the frame rate of 33 fps. In the undocked state, 3'P-D stochastically bound to and dissociated from the ribozyme, leading to repetitive quenching and dequenching of the Cy5 signal, whereas the signal from the more distant Cy3 stayed large. The FRET trace fluctuates between a FRET = 0.2 level, as expected for the P_R state, and a FRET ≈ 0 level due to quenching of Cy5 in the 3'P-D bound state. In the docked state, 3'P-D remained bound, quenching signals from Cy3 and Cy5, now both in proximity to the quencher. (B) The upper panel shows analysis of the Cy5 fluorescence trace (the undocked portion, black) by a threshold fitting method. We assigned quenched and unquenched states by setting two threshold levels: a transition from quenched to

unquenched state is identified as the Cy5 fluorescence increases to larger than threshold #1 and a transition from unquenched to quenched state is identified as the Cy5 fluorescence decreases to smaller than threshold #2. The transition assignment is indicated by the blue line. The lower panel shows analysis of the same Cy5 fluorescence trace by the HMM method (McKinney et al. (2006) *Biophys. J.* **91**, 1941-1951). Based on probability only, this method determines the most likely FRET values of underlying states and the most likely time sequence of those states, which may be originally hidden in the FRET trajectories because of noise. The transition assignment is indicated by the pink line. The two analysis methods yield nearly identical transitions assignments. From these two methods we got nearly identical product binding and dissociation rate constants. For 500 nM 3'P, the threshold fitting method reported $k_{\text{off}} = 2.6 \text{ s}^{-1}$ and $k_{\text{on}(\text{obs})} = 2.8 \text{ s}^{-1}$; the HMM method reported $k_{\text{off}} = 2.7 \text{ s}^{-1}$ and $k_{\text{on}(\text{obs})} = 2.9 \text{ s}^{-1}$.

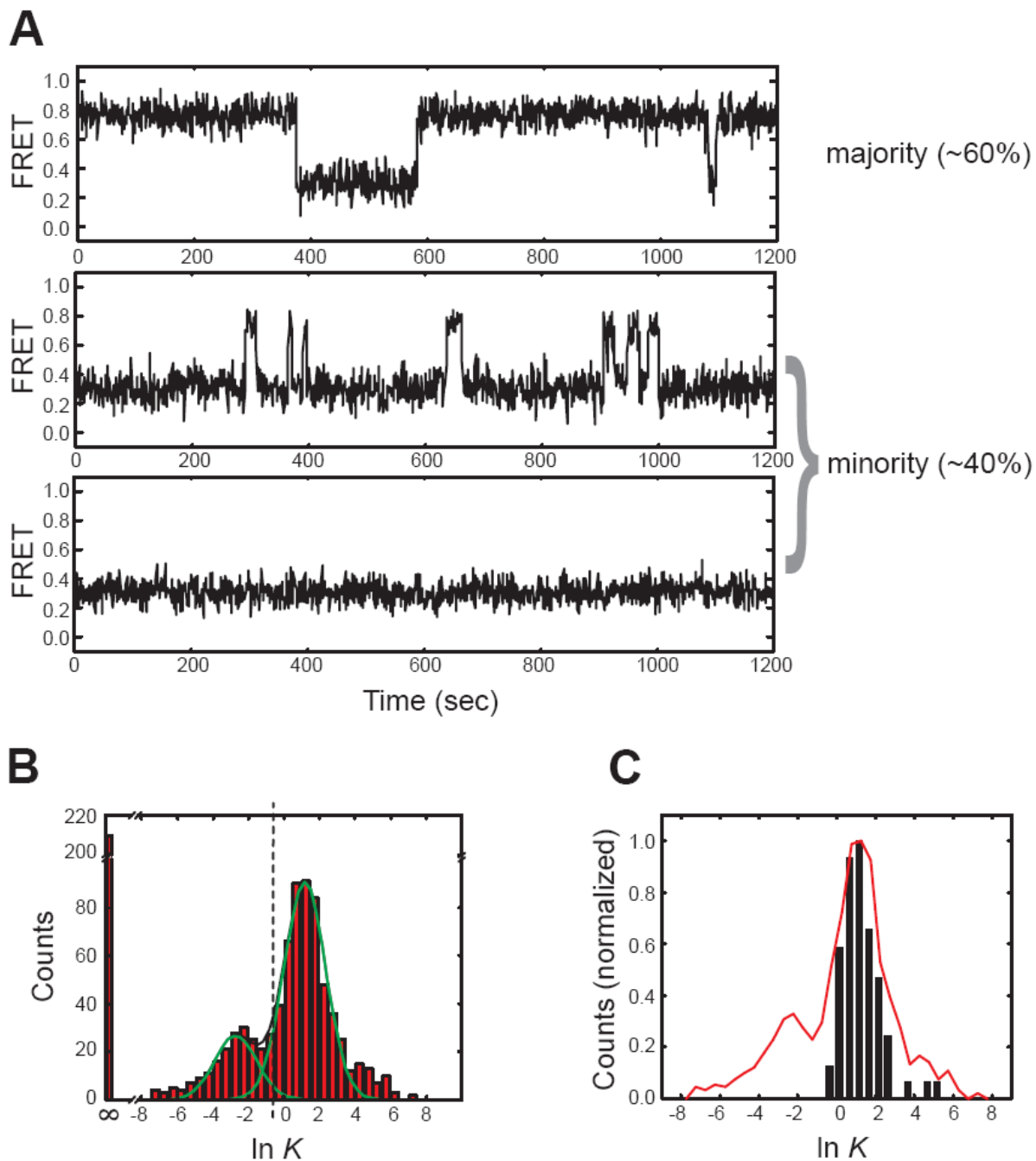


Fig. 10. Heterogeneous undocking kinetics of the hairpin ribozyme. (A) Representative FRET time traces (1 fps) of three ribozyme molecules at equilibrium. The majority of molecules (60%) display slow undocking kinetics as exemplified by the data in the upper panel, while the other molecules display significantly faster undocking kinetics or undetectable docking, as shown in the lower two panels. This is consistent with our previous observation of heterogeneous

undocking kinetics, in which a major population (61%) docks most stably and a few minor populations show one or more orders of magnitude faster undocking rates (Zhuang et al. (2002) *Science* **296**, 1473-1476). (B) Histogram of docking equilibrium constant (K) constructed from many individual molecules at equilibrium. The value K is defined as the ratio between total docked time and total undocked time for each trace (i.e. each molecule). A total of 989 single molecule FRET traces that lasted at least 600 s before photobleaching were used to construct the histogram, which is plotted in a semi-log scale since $\ln K$ scales with the free energy difference between the docked and undocked states. Three peaks are shown in the histogram: the peak centered at 1.2 corresponds to the major population; the peaks centered at -2.5 and at infinity (arising from the traces without detectable docking event) correspond to the minor populations. To isolate the major population, we only picked the FRET traces with K larger than 0.5 ($\ln K > -0.7$). From the multi-Gaussian fit of the histogram (green line), we calculated the misidentification error of this selection criterion to be $\sim 5\%$, which includes contributions from both missing traces in the major population and misidentifying traces in the minor populations as the major population due to stochasticity. Based on this criterion, we identified 595 FRET traces as belonging to the major population. The overall equilibrium constant between the docked and undocked states is 3.2, very close to the center of the major peak in this histogram ($\ln 3.2 = 1.16$). (C) After all the rate constants of the major population were determined (Fig. 6A), we simulated 200 FRET traces (600 s long) at equilibrium (with 30 μM 3'P) and constructed a histogram of $\ln K$ (black bars). The histogram of $\ln K$ for all the molecules obtained experimentally is superimposed as red line. The simulated histogram agrees well with the major peak of the experimental histogram, indicating that our selection criterion correctly identifies the major population.

Supporting Methods:

Analysis of the last docking event. Here we denote the probability that the dwell time of the last docked (FRET = 0.8) state adopts a value between t and $t + \Delta t$ as $p_{last}(t)\Delta t$. This probability is related to the normalized population of molecules that enter state U_C (or equivalently P_R) between t and $t + \Delta t$, i.e., $p_{last}(t)\Delta t = \Delta P_{U_C}(t) = k_{undock}^C P_{D_C}(t)\Delta t$, if we assume that all molecules were in state D_L at $t = 0$. The equations governing the kinetics of these last FRET = 0.8 events are:

$$\frac{dP_{D_L}(t)}{dt} = -(k_{undock}^L + k_{cleav})P_{D_L}(t) + k_{lig}P_{D_C}(t),$$

$$\frac{dP_{D_C}(t)}{dt} = -(k_{undock}^C + k_{lig})P_{D_C}(t) + k_{cleav}P_{D_L}(t),$$

$$\frac{dP_{U_C}(t)}{dt} = k_{undock}^C P_{D_C}(t).$$

Here, P_X denotes the normalized population of molecules in state X and the initial conditions for solving the above equations are $P_{D_L}(0) = 1$, $P_{D_C}(0) = 0$, and $P_{U_C}(0) = 0$. Solving the kinetic equations above yields

$$p_{last}(t) = \frac{mn}{n-m}(e^{-mt} - e^{-nt}),$$

where

$$m = \frac{(k_{undock}^L + k_{cleav} + k_{lig} + k_{undock}^C) - \sqrt{(k_{undock}^L + k_{cleav} + k_{lig} + k_{undock}^C)^2 - 4(k_{undock}^L k_{lig} + k_{undock}^C k_{cleav} + k_{undock}^L k_{undock}^C)}}{2},$$

$$n = \frac{(k_{undock}^L + k_{cleav} + k_{lig} + k_{undock}^C) + \sqrt{(k_{undock}^L + k_{cleav} + k_{lig} + k_{undock}^C)^2 - 4(k_{undock}^L k_{lig} + k_{undock}^C k_{cleav} + k_{undock}^L k_{undock}^C)}}{2}.$$

The cumulative distribution is then

$$P(t) = \int_0^t p_{last}(\tau) d\tau = \left(\frac{mn}{n-m}\right) \left(\frac{1-e^{-mt}}{m} - \frac{1-e^{-nt}}{n}\right).$$

This is a double-exponential distribution. From the already determined values of $k_{undock}^L = 0.00045 \text{ s}^{-1}$, $k_{undock}^C = 0.043 \text{ s}^{-1}$, and $k_{lig} = 13k_{cleav}$, we estimated m to be at least 20-fold smaller than n . Therefore the cumulative distribution can be approximated as a single-exponential function $P(t) \approx 1 - e^{-mt}$: this approximation introduces less than 5% error.

From 831 cleavage trajectories, we derived the cumulative dwell time distribution of the last FRET = 0.8 event. Indeed, the distribution can be quantitatively fit to a single exponential function with $m = 0.0032 \pm 0.0001 \text{ s}^{-1}$ (Fig. 5B). From the above expression of m and the already determined values of k_{undock}^L , k_{undock}^C and k_{lig}/k_{cleav} , we obtained $k_{cleav} = 0.022 \pm 0.006 \text{ s}^{-1}$ and $k_{lig} = 0.28 \pm 0.10 \text{ s}^{-1}$. From these values, n was calculated to be 0.34, 100-fold larger than m , consistent with our above estimate that e^{-mt} is by far the dominating exponential component.