Supporting Text

Single-Molecule Fluorescence Resonance Energy Transfer (FRET) Analysis. Structural details of the RNA modifications introduced are shown in Fig. 7. The durations of their docked and undocked events are obtained from single-molecule FRET time trajectories, examples of which are given in Fig. 2. The distributions of docked and undocked times are plotted as histograms (Fig. 4), each bar representing the number of dwell times that fall into its particular time range. Integrating the histograms generates cumulative plots where each point represents the number of events counted that have a dwell time less than or equal to the specified time (Fig. 4). Fitting these cumulative plots with Eq. **S1** yields the rate constants for docking and undocking, respectively:

$$N(t) = \sum_{i=1}^{M} A_i \left(1 - e^{k_{i,\text{observed}} t}\right) \qquad \textbf{[S1]}$$

The number of exponentials M indicates the number of rate constants detected from dwell times obtained with one particular time resolution. However, most ribozyme variants exhibited more rate constants than could be accurately determined from experiments at only one time resolution. Therefore, dwell times obtained from two time resolutions were separately fit with multiexponentials with the minimum number of rate constants necessary to produce a low-residuals fit. Rate constants that were extracted from both sets of dwell times corresponded well. For example, dwell times of the docked states of the wild-type ribozyme in our 2-s time-resolution data were fit with a triple exponential to yield $k_{\text{undock},1\text{-}3}$. The 0.1-s time-resolution data were fit with a triple exponential to yield $k_{\text{undock},2\text{-}4}$. The two middle rate constants, $k_{\text{undock},2}$ and $k_{\text{undock},3}$, extracted from these two separate data sets agree within experimental error (1). The undocking rate constants of the dC12, dA38, and C39S3 variants were extracted in a similar fashion, except that fewer rate constants were obtained (Table 1).

In the case of the RzAS3 and RzAS3/C39S3 variants, the 0.1-s time-resolution data were fit with a triple exponential to yield $k_{\mathrm{undock},2\text{-}4}$. The 1-s time-resolution data exhibit all four undocking rate constants, including the fastest one corresponding to $k_{\mathrm{undock},4}$. This is due to the rapid docking rate constant of these variants and the memory effect, that is, the observation that one trajectory is characterized by only a single undocking rate constant. To avoid having too many variable fitting parameters, we fit the dwell time histogram using a quadruple exponential decay, but with the two fastest undocking rate constants, $k_{\mathrm{undock},3,4}$, fixed to the values determined from the 0.1-s time-resolution data. This procedure allowed us to determined $k_{\mathrm{undock},1,\,2}$ from the fit. It is worth noting that even when leaving $k_{\mathrm{undock},3,4}$ as free fitting parameters, the derived values for $k_{\mathrm{undock},3,4}$ were similar.

The histograms of the dwell times of the undocked states of all variants were fit with single- or double-exponential decays to yield $k_{\mathrm{dock},1(2)}$. The determination of docking rate constants is thus completely uncoupled from the determination of the docking rate constants. Errors in our rate constants (Table 1) were estimated by independently analyzing subsets of our data and comparing the resulting rate constants.

The use of finite observation windows to measure the dwell times introduces a systematic bias in the observed rate constants. This was corrected for by measuring the rate at which either Cy3 or Cy5 is inactivated by photobleaching and subtracting this rate $k_{\mathrm{photobleach}}$ from the observed rate k_{observed} along with the reciprocal of the observation window t_w (2):

$$k_{i,\mathrm{observed}} = k_{i,\mathrm{actual}} + k_{photobleach} + \frac{1}{t_w} \qquad \textbf{[S2]}$$

Because the time resolutions used in the experiments may not be sufficient to resolve all docked events that undock rapidly, the amplitude A_i for the relatively fast undocking populations could be underestimated with Eq. **S1**. This was corrected with Eq. **S3**, which assumes that all dwell times shorter than τ are not observed:

$$A_{\text{corrected}} = \frac{A_{\text{observed}}}{e^{-k_i \tau}} \quad \textbf{[S3]}$$

For trajectories taken with 2-s time resolution (i. e., the laser was turned on for 200 ms every 2 s), $\tau = 1$ s; for trajectories taken with 1-s time resolution (i. e., the laser was turned on for 200 ms every 1 s), $\tau = 0.5$ s; for trajectories taken at 100-ms time resolution (i. e., the laser was kept on), $\tau = 50$ ms. In all cases, the camera refreshing rate was 100 ms.

Furthermore, Cy5 dyes occasionally "blink." As a result, the effective FRET value goes to zero. These events can thus be distinguished from the undocked states with FRET values of ≈0.2 and are not counted as undocking events. [The most challenging variants in this regard are RzAS3 and RzAS3/C39S3, because they both exhibit very fast docking rate constants. The resulting short dwell times in the undocked state together with our data noise make it difficult to distinguish "blinking" events from undocked states in our 1-s time-resolution experiments, yet our 100-ms time-resolution experiments again yield a sufficient number of FRET data points to clearly distinguished FRET = 0.0 ("blinking") from FRET = 0.2 (undocked state). We found that at most 22% of the low FRET states are "blinking" events.]

Next, we determined the fraction of molecules f_i that exhibit the undocking rate constants $k_{\text{undock},i}$. First, FRET time traces were divided into two fractions: those with detectable docked states ($f_{\text{docked,detectable}}$); and those that do not dock ($f_{\text{undetectable}}$). The fractions f_i (i = 1, M) were determined from the amplitudes A_i (i = 1, M) obtained from multiexponential fits (Eq. **S1**) using Eq. **S4**:

$$f_i = \left(\frac{A_i \left(\frac{1}{k_{dock}} + \frac{1}{k_{undock,i}} \right)}{\sum_{k=1}^{M} A_k \left(\frac{1}{k_{dock}} + \frac{1}{k_{undock,k}} \right)} \right) f_{\text{docked,detectable}} \quad \textbf{[S4]}$$

For the dA38 and dC12 variants, we observed two undocking rate constants, both of which are much faster than those of the wild-type ribozyme. Therefore, the comparably large $f_{\text{docked,undetectable}}$ fractions determined for both variants are likely in part due to undocking rate(s) too fast to be detected at even our highest time resolution (100 ms).

For the C39S3 variant, we observed three undocking rate constants. Two time resolutions were used to determine the fractions f_1, f_2, f_3, and $f_{\text{undetectable}}$. Trajectories with 1-s time resolution give fractions $f_1^{'}$, $f_2^{'}$, and $f_{un\det ectable}^{'}$ using Eqs. **S1**-**S4**, whereas the 100-ms resolution trajectories give fractions $f_2^{''}$, $f_3^{''}$, and $f_{undetectable}^{''}$. The rate constant $k_{\text{undock,2}}$ determined from trajectories with 1-s time resolution was comparable with that determined from trajectories with 100-ms time resolution. Because the 1-s time resolution was sufficient to resolve the rate constants $k_{\text{undock,1}}$ and $k_{\text{undock,2}}$, $f_1 = f_1^{'}$ and $f_2 = f_2^{'}$. f_3 was calculated from $\left(\frac{f_3^{''}}{f_2^{''}}\right)f_2^{'}$, whereas $f_{\text{undetectable}}$ was determined as $f_{undetectable}^{'} - f_3$, because traces exhibiting $k_{\text{undock,3}}$ are expected to be counted in the low-resolution data as part of $f_{undetectable}^{'}$.

For the wild-type hairpin construct, we observed four undocking rate constants. Two time resolutions were used to determine the fractions f_1, f_2, f_3, f_4, and $f_{\text{undetectable}}$. Trajectories with 2-s time resolution gave fractions $f_1^{'}$, $f_2^{'}$, $f_3^{'}$, and $f_{undetectable}^{'}$. Trajectories with 100-ms time resolution yielded fractions $f_2^{''}$, $f_3^{''}$, $f_4^{''}$ and $f_{undetectable}^{''}$. The rate constants $k_{\text{undock,2}}$ and $k_{\text{undock,3}}$ determined from trajectories with 2-s time resolution were comparable to those determined from trajectories with 100-ms time resolution. We then determined f_1, f_2, f_3, f_4, and $f_{\text{undetectable}}$ with $f_1 = f_1^{'}$, $f_2 = f_2^{'}$, $f_3 = f_3^{'}$, $f_4 \approx \left(\frac{f_4^{''}}{f_3^{''}}\right)f_3$, and $f_{\text{undetectable}} = f_{undetectable}^{'} - f_4$.

Because the RzAS3 and RzAS3/C39S3 variants exhibit docking rate heterogeneity in addition to undocking rate heterogeneity, a more sophisticated algorithm was used to determine what fraction of the trajectories exhibits a given pair of docking and undocking rate constants, denoted $p_{dock,i;undock,j}$. We first computed the conditional likelihood for each possible pair of docking and undocking rate constants to give rise to the dwell times observed in a given single-molecule trace, denoted $P(trace_k \mid k_{dock,i}; k_{undock,j})$. Bayes' rule then allows us to express the conditional probability that a trace belongs to a given subpopulation, $P(k_{dock,i}; k_{undock,j} \mid trace_k)$, in terms of $P(trace_k \mid k_{dock,i}; k_{undock,j})$ and a knowledge of $p_{\mathrm{dock,i;undock,j}}$:

$$P(k_{dock,i}; k_{undock,j} \mid trace_k) = \frac{P(trace_k \mid k_{dock,i}; k_{undock,j})\, p_{dock,i;undock,j}}{\sum_{n,m} P(trace_k \mid k_{dock,n}; k_{undock,m})\, p_{dock,n;undock,m}} \qquad \textbf{[S5]}$$

Because we do not know the $p_{dock,i;undock,j}$ *a priori*, we begin with an initial guess for these occupancies. After computing the above conditional probabilities, we revise our estimate of the occupancy of each species to:

$$p_{dock,i;undock,j} \rightarrow \frac{1}{N} \sum_k P(k_{dock,i}; k_{undock,j} \mid trace_k) \qquad \textbf{[S6]}$$

where N is the total number of traces. This process is then iterated until all $p_{dock,i;undock,j}$ converge to a self-consistent set of occupancies for the subpopulations exhibiting each pair of docking and undocking rate constants.

For the RzAS3 variant, two time resolutions were used to measure the rate constants and to determine the fractions $f_{\mathrm{dock,i;undock,j}}$, which are denoted as $f_{i,j}$ hereafter. First, the fractions $f'_{docked,detectable}$ and $f'_{undetectable}$ were determined by using 1-s time-resolution trajectories. The algorithm described above was used to determine the set of occupancies $p'_{1,1}$, $p'_{1,2}$, $p'_{1,3}$ $p'_{2,1}$, $p'_{2,2}$, and $p'_{2,3}$ from the 1-s time-resolution trajectories and $p''_{1,3}$, $p''_{1,4}$, $p''_{2,3}$ and $p''_{2,4}$ from the 100-ms time-resolution trajectories. Here we use $p_{i,j}$ to denote $p_{dock,i;undock,j}$. We determine $f_{i,j}$ using $f_{1,1}$ = $p'_{1,1} f'_{docked,detectable}$, $f_{1,2} = p'_{1,2} f'_{docked,detectable}$, $f_{2,1} = p'_{2,1} f'_{docked,detectable}$, and $f_{2,2} = p'_{2,2} f'_{docked,detectable}$. Because of the faster docking rate, the number of docked events per traces in this variant is relatively high compared to that of the wild-type hairpin ribozyme. Thus, a significant number of molecules with undocking rate constant $ki_{\text{undock},4}$ show detectable docked events in the 1-s time-resolution trajectories. From the observed docking and undocking rate constants, we predict that nearly all molecules with $k_{\text{dock},1}$ and $k_{\text{undock},4}$ and 63% of molecules with $k_{\text{dock},2}$ and $k_{\text{undock},4}$ will show detectable docked events in the 1-s time-resolution trajectories, whereas the remaining 37% of molecules with $k_{\text{dock},2}$ and $k_{\text{undock},4}$ will not. These molecules are grouped into $p'_{1,3}$ and $p'_{2,3}$. We use $p'_{1,3}$, $p'_{2,3}$, $p''_{1,3}$, $p''_{1,4}$, $p''_{2,3}$, and $p''_{2,4}$ to determine the actual $f_{1,3}$, $f_{2,3}$, $f_{1,4}$, and $f_{2,4}$:

$$f_{1,3} = \frac{p''_{1,3}\left(p'_{1,3} + p'_{2,3}\right)f'_{\text{docked,detectable}}}{p''_{1,3} + p''_{2,3} + p''_{1,4} + 0.63p''_{2,4}},\ f_{2,3} = \frac{p''_{2,3}\left(p'_{1,3} + p'_{2,3}\right)f'_{\text{docked,detectable}}}{p''_{1,3} + p''_{2,3} + p''_{1,4} + 0.63p''_{2,4}},$$

$$f_{1,4} = \frac{p''_{1,4}\left(p'_{1,3} + p'_{2,3}\right)f'_{\text{docked,detectable}}}{p''_{1,3} + p''_{2,3} + p''_{1,4} + 0.63p''_{2,4}} \quad \text{and} \quad f_{2,4} = \frac{p''_{2,4}\left(p'_{1,3} + p'_{2,3}\right)f'_{\text{docked,detectable}}}{p''_{1,3} + p''_{2,3} + p''_{1,4} + 0.63p''_{2,4}}.$$

The fraction of molecules that are either docking inactive or undock too rapidly for their docked states to be detected with the available time resolutions is: $f_{\text{undetectable}} = f'_{\text{undetectable}} - 0.37f_{2,4}$. The populations $f_{i,j}$ of the RzAS3/C39S3 variant were determined similarly.

Cleavage and Ligation Assays. 5'-^{32}P-labeled substrate was prepared by phosphorylation with T4 polynucleotide kinase and [γ-^{32}P]-ATP. All cleavage reactions were conducted under single turnover (presteady-state) conditions in standard buffer (50 mM Tris•HCl, pH 7.5/12 mM Mg^{2+}) at 25ºC. Ribozyme (final concentrations: 100 nM strand RzA and 200 nM strand RzB) and radiolabeled substrate (<1 nM) were preannealed separately in standard buffer by heating to 70°C for 2 min and slow cooling over 5 min to room temperature. After preincubation for 15 min at 25°C, an equal volume of standard buffer containing labeled substrate was added to the ribozyme to initiate the reaction. Two-microliter reaction aliquots were taken at appropriate time intervals and quenched with 13-µl stop solution (80% formamide/0.025% xylene-cyanol/0.025% bromophenol blue/50 mM EDTA). The 5' cleavage product was separated from uncleaved substrate by denaturing 20% polyacrylamide, 8 M urea, gel electrophoresis, and it was quantified and normalized to the sum of the substrate and product bands by using a PhosphorImager Storm 840 with IMAGEQUANT software (Molecular Dynamics). Error bars are calculated from at least two independent cleavage assays. Time traces of product formation were fit to the double-exponential first-order rate equation $y(t) = y_0 + A_f\left(1 - e^{(-k_{cleav,obs,f}\,t)}\right) + A_s\left(1 - e^{(-k_{cleav,obs,s}\,t)}\right)$, using Marquardt-Levenberg nonlinear least-squares regression, where $A_f + A_s$ is the final extent of cleavage and the two $k_{\mathrm{cleav,obs}}$s are the first-order rate constants of the fast and slow phases.

5'-^{32}P-labeled 5' product with a 2',3'-cyclic phosphate (5'P$_c$) was prepared by cleaving the radio-labeled substrate using wild-type hairpin ribozyme for 4 h in standard buffer (50 mM Tris•HCl, pH 7.5/12 mM Mg^{2+}) at 25ºC. 5'P$_c$ was purified from uncleaved substrate by denaturing 20% polyacrylamide, 8 M urea, gel electrophoresis, elution into water at 65ºC for 1 h, and desalting using a CentriSpin-10 column (Princeton Separations, Adelphia, NJ). Ribozyme [final saturating concentrations: 8 µM strand RzA/12 µM strand RzB/radiolabeled 5'P$_c$ (<1 nM)] and 12 µM 3'P were preannealed separately in standard buffer as described for the cleavage reaction, preincubated for 15 min, then mixed to initiate ligation. After 4 h at 25°C, the reactions were quenched with 13-µl stop solution as described above. Ligated substrate was separated from 5'P$_c$ by denaturing 20% polyacrylamide, 8 M urea, gel electrophoresis, and it was quantified and normalized to the sum of the substrate and 5'P$_c$ bands as described above, yielding the fraction ligated, f_{lig} (Fig. 5). The internal equilibrium of the chemical step was determined as $k_{\mathrm{lig}}/k_{\mathrm{cleav}} = f_{\mathrm{lig}}/(1-f_{\mathrm{lig}})\times(1+k'_{\mathrm{undock}}/k'_{\mathrm{dock}})/(1+k_{\mathrm{undock}}/k_{\mathrm{dock}})$. [This considers that the binding equilibrium constants for the substrate and products become very large at the high concentration of ribozyme used in the experiments. Note that the (un)docking rate constants of the major subpopulation were used as they contribute predominantly to the chemistry equilibrium due to the longer time that the products stay bound to the major subpopulation.] Error bars are calculated from at least two independent ligation assays. Forward cleavage assays (with 8 μM unlabeled and traces of radiolabeled substrate, 10 μM RzA, and 12 μM RzB) yielded very similar final product/substrate distributions, showing that the internal and overall equilibrium are indeed reached in both the cleavage and ligation experiments.

Steady-State FRET Kinetic Assays. Steady-state FRET measurements of the hairpin ribozyme doubly labeled with Cy3 and Cy5 were performed on an Aminco-Bowman Series 2 (AB2) spectrofluorometer (Thermo Spectronic, Woburn, MA), similar to previously described experiments (3). Annealed ribozyme (final concentration 100 nM; with a 4-fold, i.e., saturating, excess of unlabeled strand RzB) was incubated at 25°C for at least 15 min in standard buffer (50 mM Tris•HCl, pH 7.5/12 mM Mg^{2+}) supplemented with 25 mM dithiothreitol (DTT) as oxygen scavenger. In a 150-μl cuvette, a saturating 8-fold (in the case of the dC12 variant, 20-fold) excess of noncleavable substrate analog was manually added to form the ribozyme-substrate complex and initiate docking. Cy3 was excited at 540 nm (4-nm bandwidth), and for kinetic experiments, the fluorescence emission F was recorded simultaneously at the Cy3 (560 nm, 8-nm bandwidth) and Cy5 (665 nm, 8-nm bandwidth) maximum emission wavelengths. The FRET ratio (=$F_{665}/(F_{665}+F_{560})$) was calculated and normalized with its initial value (Fig. 6). The resulting time traces were fit to a single exponential increase function of the form $y(t) = y_0 + A\left(1 - e^{(-k_{dock,obs}t)}\right)$, where A and $k_{\mathrm{dock,obs}}$ are the extent and rate constant of the FRET increase, respectively (Table 2).

Kinetic Simulations of Steady-State FRET Assays. The (un)docking rate constants of each hairpin ribozyme variant, $k_{\mathrm{dock},i}$ and $k_{\mathrm{undock},i}$, and their subpopulation fractions, f_i, measured by single-molecule FRET, were used to generate a kinetic simulation that was compared to a typical experimental steady-state FRET time course determined as described above (Fig. 6). The kinetic simulation was generated by using Eq. **S7**:

$$F(t) = F_{\mathbf{max}}\left[\sum_i\left(f_i\,\frac{k_{dock,i}}{k_{dock,i}+k_{undock,i}}\left(1-e^{-\left(k_{dock,i}+k_{undock,i}\right)t}\right)\right)\right] \qquad \textbf{[S7]}$$

where i sums over all subpopulations, and $F_{\max}$, the maximum FRET ratio reached at equilibrium, was fit to match the total amplitude observed experimentally, using Marquardt-Levenberg nonlinear least-squares regression. In the case of RzAS3 and RzAS3/C39S3, the sum in Eq. **S7** is a double sum that applies over both sets of docking rate constants, and f_i becomes $f_{i,j}$, as described above.

Kinetic Simulations of Cleavage Reactions. The reaction pathway of the hairpin ribozyme (Fig. 1*B*) can be approximated by four sequential unimolecular reactions with three reversible steps (substrate complex docking, cleavage, and product complex undocking) and one irreversible step (product dissociation), because $k_{\mathrm{off}} \approx 0$ and $k'_{\mathrm{on}} \approx 0$. The dynamics of the resulting set of coupled rate equations are dictated by the following master equation (4, 5):

$$\dot{\mathbf{q}}_i = \mathbf{K}_i \mathbf{q}_i \qquad \textbf{[S8]}$$

where $\mathbf{q}$ is the population matrix, $\mathbf{K}_i$ is the rate constant matrix, the point denotes derivation with respect to time, and the index i indicates each subpopulation distinguishable in their docking and/or undocking rate constants. For the reaction pathway of the hairpin ribozyme $\mathbf{K}_i$ is as follows:

$$\mathbf{K}_i = \begin{pmatrix} -k_{dock,i} & k_{undock,i} & 0 & 0 & 0 \\ k_{dock,i} & -k_{undock,i} - k_{cleav} & k_{lig} & 0 & 0 \\ 0 & k_{cleav} & -k_{lig} - k'_{undock,i} & k'_{dock,i} & 0 \\ 0 & 0 & k'_{undock,i} & -k'_{dock,i} - k'_{off} & k'_{on} \\ 0 & 0 & 0 & k'_{off} & -k'_{on} \end{pmatrix} \qquad \textbf{[S9]}$$

where the individual rate constants are defined as in Fig. 1*B* and Table 1. Here we assume that k_{cleav} and k_{lig} are the same for all molecular subpopulations (see further discussion of this point below). Because the matrix elements of $\mathbf{K}_i$, i.e., the rate constants of the reaction pathway, are not time dependent, the general solution for the master equation is:

$$\mathbf{q}_i(t, t_0) = \mathbf{exp}(\mathbf{K}_i(t - t_0))\mathbf{q}_i(t_0) \qquad \textbf{[S10]}$$

The rate matrix, $\mathbf{K}_i$, is real, quasisymmetric, and tridiagonal, with values strictly negative in the main diagonal, strictly positive in the diagonals above and below, and zero elsewhere. Thus, $\mathbf{K}_i$ can be transformed into a symmetric matrix and diagonalized. The eigenvalues of $\mathbf{K}_i$ are then strictly negative or zero. The analytical solution for the general rate equation for each reactant is given by:

$$\begin{cases} c_{S_{U,i}}(t) = \sum\limits_{k=1\to 5} A_{k,i}\exp(\lambda_{k,i}t) \\ c_{S_{D,i}}(t) = \sum\limits_{k=1\to 5} B_{k,i}\exp(\lambda_{k,i}t) \\ c_{P_{D,i}}(t) = \sum\limits_{k=1\to 5} C_{k,i}\exp(\lambda_{k,i}t) \\ c_{P_{U,i}}(t) = \sum\limits_{k=1\to 5} D_{k,i}\exp(\lambda_{k,i}t) \\ c_{P_i}(t) = \sum\limits_{k=1\to 5} E_{k,i}\exp(\lambda_{k,i}t) \end{cases} \qquad \textbf{[S11]}$$

where the $c_{S_{U,i}}(t)$, etc., are the reaction intermediate concentrations, $\lambda_{k,i}$ are the eigenvalues of $\mathbf{K}_i$, and $A_{k,i}$ to $E_{k,i}$ are obtained by normalizing the eigenvectors of $\mathbf{K}_i$ using the initial and final boundary conditions at $t = 0$ and $t = \infty$.

Based on the above matrix algebra, each molecule subpopulation yielded its own analytical solution for the overall cleavage time course, which was weighted by their fractions and summed up. The ratio $r = k_{\text{lig}}/k_{\text{cleav}}$ was experimentally determined as the internal chemistry equilibrium (see above and Fig. 5) and kept constant throughout the simulation. The cleavage and ligation rate constants were then obtained by a single-variable fit of the simulated cleavage time courses to the experimental ones using Marquardt-Levenberg least-squares regression (MATLAB, MathWorks) and experimentally measured docking/undocking rate constants, the subpopulation fractions and the internal equilibrium constant *r*. Errors on k_{cleav} (dk_{cleav}) arise from the 95% confidence intervals of the least-squares fit calculated from the fit residuals and the Jacobian matrix (MATLAB). Errors on k_{lig} (dk_{lig}) were calculated from the error on k_{cleav} and the experimental error on *r* (*dr*, Table 2) as follows:

$$dk_{lig} = rdk_{cleav} + k_{cleav}dr \qquad \textbf{[S12]}$$

Using this analysis, we find that modifications of essential residues distant from the site of catalysis alter not only the rate constants of docking and/or undocking but also those of catalytic chemistry. This is illustrated in the free energy reaction profile of the dominant subpopulation of each modified ribozyme in comparison to the wild-type in Fig. 9, calculated from the rate constants obtained in this study.

In our calculations, we have assumed that k_{cleav} and k_{lig} are shared among all molecular subpopulations with different docking and undocking kinetics. This is necessary to keep the number of fitting parameters reasonably low. Given the substantial differences between the docking/undocking and cleavage/ligation rate constants of the various site-specifically modified variants, it is possible that whatever causes the molecular heterogeneity observed in form of the subpopulations of each variant also affects their chemistry rate constants. From the high reaction extent in both cleavage and ligation reactions, it is apparent that most, if not all, of the noninterchanging subpopulations must be active; yet, we cannot formally prove that they share the same k_{cleav} and k_{lig}. Our analysis therefore yields k_{cleav} and k_{lig} values that are not necessarily specific for any of the molecular subpopulations but are certain averages over all active subpopulations of a given ribozyme variant. Fig. 3 illustrates that the shape of the overall cleavage time course for all variants is determined by the fact that the fast cleavage phase is composed mainly of contributions from subpopulations I, II, and in part III, whereas the slow cleavage phase is dominated by contributions from subpopulation IV and in part III. This general behavior is similar for all variants, suggesting that our relative comparisons of the averaged k_{leav} and k_{lig} values between them are likely valid, even if some subpopulations may have chemistry rate constants that differ from those of the major subpopulation.

To further evaluate this notion, we performed analytical fits in which we varied the ratio ρ of the chemistry rate constants of the minor and major subpopulations:

$$\rho = k_{\text{cleav,II-IV}}/k_{\text{cleav,I}}$$

Table 3 shows the least-square minimized values for $k_{\text{cleav,I}}$ and $k_{\text{cleav,II-IV}}$ of all variants at several different ρ values, together with the squared residuals to evaluate the relative quality of the fits, the fraction of the major subpopulation (taken from Table 1), and the average chemistry rate constant $k_{\text{cleav,ave}}$ over all subpopulations that results from these values. For all variants, the fit quality improves somewhat with $\rho > 1$ (and the optimal ρ, with values between 3 and 10, is similar for all variants), which suggests that the higher subpopulations behave somewhat similarly in all variants and perhaps are characterized by a faster reaction chemistry than the corresponding major subpopulation. However, $k_{\text{cleav,ave}}$ does not change significantly and even with an optimized ρ is within 2-fold of the reported values in Fig. 3 (derived for $\rho = 1$). Likewise, when ρ is chosen to be significantly <1 (e.g., $\rho = 0.33$), which results in significantly worse fits for all variants (Table 3), the average $k_{\text{cleav,ave}}$ values change by at most a factor of only 2-3 from the reported values in Fig. 3. Thus, even if the catalytic behavior of the different subpopulations were more complex than assumed for the analyses presented in Fig. 3, our comparisons of the cleavage and ligation rate constants of ribozyme variants are likely valid to within 2- to 3-fold. Considering that the k_{cleav} values of RzAS3 and RzAS3/C39S3 are both ≈20-fold smaller than that of the wild type, and that the k_{lig}/k_{cleav} values of dC12 and dA38, directly measured in experiments, are also markedly different from that of the wild type, we conclude that modifications remote from the active site affect the chemistry rate constants.

1. Zhuang, X., Kim, H., Pereira, M. J., Babcock, H. P., Walter, N. G. & Chu, S. (2002) *Science* **296,** 1473-1476.
2. Bartley, L. E., Zhuang, X., Das, R., Chu, S. & Herschlag, D. (2003) *J. Mol. Biol*. **328,** 1011-1026.
3. Walter, N. G., Hampel, K. J., Brown, K. M. & Burke, J. M. (1998) *EMBO J*. **17,** 2378-2391.
4. Gutfreund, H. (1995) *Kinetics for the Life Sciences: Receptors, Transmitters and Catalysts* (Cambridge Univ. Press, Cambridge, U. K.).
5. Steinfeld, J. I., Francisco, J. S. & Hase, W. L. (1989) *Chemical Kinetics and Dynamics* (Prentice-Hall, Englewood Cliffs, NJ).