Mechanism of substrate translocation by a ring-shaped ATPase motor at millisecond resolution

Wen Ma and Klaus Schulten*

E-mail: kschulte@ks.uiuc.edu

Supporting Information

*To whom correspondence should be addressed

The large scale conformational transition I $\rightarrow$ F in Rho proceeds on a millisecond timescale, which is still beyond the reach of conventional molecular dynamics simulations. In order to describe the process nevertheless by molecular dynamics simulation, pathway sampling techniques can be used,[1,2] in particular the string method.[3–5] Free energy values and kinetics along the transition pathway can be determined by means of milestoning analysis.[6–10] In Section 1, we first introduce a criterion to select the key atoms involved in the reaction coordinate for the transition pathway calculation and we describe then the details of the string method. In Section 2 we describe the milestoning method.

In Section 3, 4 and 5 we describe three detailed analyses carried out in the present study to characterize the I $\rightarrow$ F transition pathway resulting from our simulations. The first analysis reveals the energetic coupling between Rho and RNA by calculating interaction energies between individual RNA residues and individual Rho subunits during RNA translocation. The second analysis scans interaction energy contributions from all residues of Rho. The third analysis depicts the conformational changes of key residues responsible for ATP binding transitions, namely $T^- \rightarrow T$ and $T \rightarrow T^*$ during one rotary reaction step.

Movie S1 (Section 6) shows the motion of the six K326 side chains that translocate RNA in six sequential rotary reaction steps. Movie S2 (Section 7) shows the relative motion of Rho subunits during six sequential rotary reaction steps.

Figure S1, just below, shows schematically dwell and motor-action phases of Rho dynamics. The motor-action phase is responsible for the RNA translocation process.

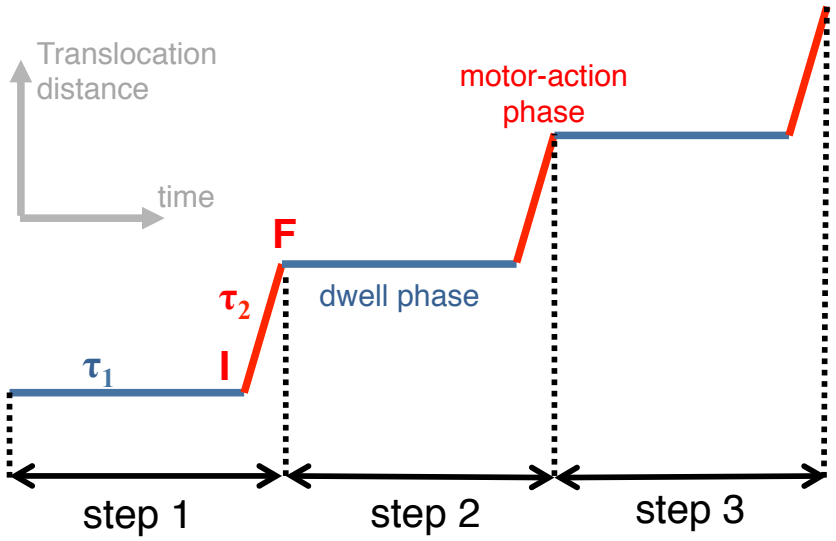

Figure S1: Schematic illustration of dwell and motor-action (RNA-translocation) phases of Rho. During the dwell phase a new ATP molecule is bound to the empty interface, ADP and Pi are released and a bound ATP molecule in a T^* state becomes hydrolyzed. The release of hydrolysis product (ADP + Pi) is rate limiting ($\tau_1 \sim 10$ ms).[11,12] The completion of all three dwell phase processes is found in the present simulations to trigger RNA translocation. In fact, the motor-action (RNA-translocation) phase ($\tau_2 \sim 0.1$ ms) corresponds to the I→F transition that is described in detail in the present study. Steps 1, 2, 3, ... correspond to the 60°, 120°, 180°, ... conformational rotations that induce together a series of RNA translocation steps as also described in Table S4.

1. String method protocol.

In this section, we describe the computational procedure employed to perform the "string method with swarms-of-trajectories".[4] In (1.1) key atoms are chosen for the collective variables describing the I $\rightarrow$ F transition; in (1.2) the detailed protocol for carrying out the string method calculation is outlined; in (1.3) the choice of the initial pathway is specified.

1.1 Choice of collective variables. Due to the high-dimensional conformational space of Rho (altogether $\sim 40,000$ atoms), one needs to select a smaller, yet representative set of coordinates that is suitable to search for and describe the most probable transition path (MPTP);[4] this representative set of coordinates is referred to in the present study as the collective variables $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$.[3] By analyzing the difference of the structural features among different interface states, we propose below that the collective variable space can be reduced to positions of key residues at the six subunit-subunit interfaces $m_1/m_2, m_2/m_3, \ldots, m_6/m_1$, which contribute to relative motion of subunits, but not to internal motion of the individual subunits.

In order to identify the participation of subunit degrees of freedom in the rotatory reaction we first compare the internal coordinate difference between the R and F states separately for Rho subunits $m_j, j = 1, 2, \ldots 6$. We found that the root mean square deviation (RMSD) for an individual subunit is only around 1 Å to 1.5 Å (Table S1). C_α coordinates of the C-terminal domain (residue 136 to residue 405) were included to calculate the RMSD value, $\mathrm{RMSD}_j(\mathrm{R},\mathrm{F})$, for subunits $j = 1, 2, \ldots 6$ according to

$$\mathrm{RMSD}_j(\mathrm{F},\mathrm{R}) \;=\; \left\{ \frac{1}{N_c} \sum_{i=1}^{N_c} \left[T^{(j)}(\mathrm{F}\rightarrow\mathrm{R}) \mathbf{r}_i^{(j)}(\mathrm{F}) \;-\; \mathbf{r}_i^{(j)}(\mathrm{R}) \right]^2 \right\}^{1/2} \quad . \tag{S1}$$

Here $\mathbf{r}_i^{(j)}(\mathrm{R})$ and $\mathbf{r}_i^{(j)}(\mathrm{F})$ are the C_α coordinate vectors of the ith residue of subunit j in state R and state F, respectively; $T^{(j)}(\mathrm{F}\rightarrow\mathrm{R})$ is the three-dimensional coordinate transformation that subtracts any overall translation and rotation of subunit j of state F compared to state R. N_c is the number of C_α coordinates of the C terminal domain ($N_c = 270$). The small

RMSD values indicate that the participation of the subunits in the rotary reaction step involves mainly relative motions of subunits, but not internal motion of individual subunits. This assumption is supported by a hydrogen/deuterium exchange experiment[13] that showed little exchange signal in the β-sheet core of the Rho C-terminal domain, indicating that Rho conformational transitions in response to ligand binding mainly involve monolithic subunit movements. This interpretation is strongly supported also by the results of the extensive string method calculations carried out for the present study, which showed that, for the internal motions, the maximum RMSD from the prior R state along the transition pathway is limited per subunit to 1.6 Å (Table S1). Here the maximum RMSD_j(Y,R) in Table S1 is calculated for protein conformation Y, which was taken from the discretized images 0, 1, ..., M of the transition pathway.

Table S1: Subunit internal RMSD changes

j	1	2	3	4	5	6
RMSD_j(F,R) (Å)	1.42	1.33	1.21	1.06	1.21	1.27
$\max_{0\le Y\le M} \mathrm{RMSD}_j$(Y,R) (Å)	1.55	1.39	1.42	1.45	1.37	1.44

While there is little motion of residues in subunit interiors, the residues at the interfaces between neighboring subunits exhibit a crucial involvement in the rotary reaction step as illustrated in Figure S2b. In this figure the surface residues are colored according to the deviation of interaction energy among the six $m_1/m_2, m_2/m_3, \ldots, m_6/m_1$ interfaces, $\Delta {E_s}^{\mathrm{prior}}$, experienced by residue s in the prior state R (counted are for every residue s on a particular subunit the interaction with all residues of neighboring subunits). The $\Delta {E_s}^{\mathrm{prior}}$ values provide a good criterium for the selection of collective variables involved in the rotary reaction step as a larger interaction energy deviation of residue s in six interfaces indicates a more important contribution of residue s to the transition between interface states. Here s denotes the number of a particular residue in one of the six Rho subunits, say m_1, but it

is understood that s represents actually at this point the residue with number s also in the other five subunits $m_2, m_3, \ldots, m_6$, not just the residue for a specific subunit. ${\Delta E_s}^{\rm prior}$ has been calculated through the expressions

$$
\begin{aligned}
{\Delta E_s}^{\rm prior} &= \left[\frac{1}{6}\sum_{j=1}^{6}(E_{s,j} - \bar{E}_s)^2\right]^{1/2} && \text{(S2)}\\
E_{s,j} &= \sum_{s'\in S(j')}\sum_{j'\in N(j)} E_{s,j}(s',j') && \text{(S3)}\\
\bar{E}_s &= \frac{1}{6}\sum_{j=1}^{6} E_{s,j} \quad . && \text{(S4)}
\end{aligned}
$$

Here $N(j) = \{j \bmod 6 + 1, (j-2) \bmod 6 + 1\}$ is the set of two subunits that are neighbors of subunit j. $S(j')$ is the set of residues of subunit j'. $E_{s,j}(s',j')$ is the interaction energy between residue s on subunit j and residue s' on subunit j' in state R; $E_{s,j}$ is the interaction energy between residue s on subunit j and all the residues on its neighboring subunits in state R; $\bar{E}_s$ is the average interaction energy among six subunits between residue s and neighboring subunits in state R. Generalized Born/solvent-accessible surface area implicit solvent (GB/SA) calculations[14,15] were employed to determine the residue-residue interaction energy $E_{s,j}(s',j')$ in state R.

We selected for inclusion in the collective variables residues s with ${\Delta E_s}^{\rm prior} > 1$ kcal/mol. The collective variables $\mathbf{z}$ include Cartesian coordinates of the C_α atoms of the identified residues; residues Lys (NZ), Arg (CZ), Glu (CD), Gln (CD), Asp (CG), Asn (CG) and Tyr (CZ), which exhibit long side chains, were represented in the collective variables through a second atom (atom names given in parentheses after the respective amino acids). The list of the identified residues' indices is: 140 173 180 181 184 210 211 212 214 215 217 218 221 233 265 269 272 275 283 284 286 287 290 298 299 305 308 323 326 333 334 336 337 339 340 342 347 351 352 353 361 362 364 366 367 368 384 385 388. These residues are located at the subunit-subunit interfaces, spanning over loops and four helices namely the ones along the quadruple of segments (213 - 222, 267 - 277, 326 - 337, 347 - 353). The C_α atoms of the

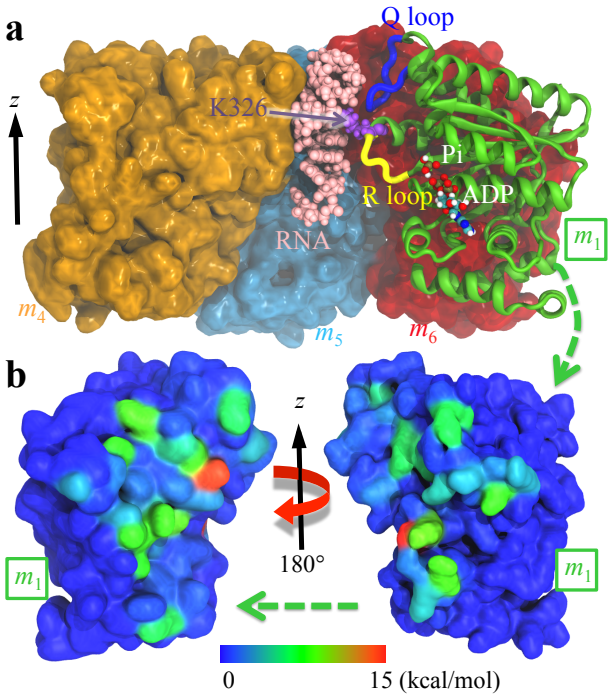

Figure S2: Key structural elements and surface interactions of the C-terminal domain in the initial state. (a) Front view of Rho's structural components in contact with RNA. For clarity, m_2 and m_3 are not shown. m_1, m_4, m_5 and m_6 are shown in green, orange, light blue and red, respectively; ADP and Pi are shown in van der Waals representation colored by atom name; Q loop and R loop are highlighted in blue and in yellow, respectively. (b) Contributions of individual residues to the interaction energy deviation, $\Delta E_s^{\text{prior}}$, among six interfaces (projected onto the surface of subunit m_1 surface). (Right) m_1 is in the same view as in a. (Left) m_1 is rotated 180° around the z-axis.

four helices were also added to the collective variable set, if not included yet in the original list. The total number of selected atoms is 660, constituting a 660×3-dimensional collective variable space. Below we will employ a similar criterion to identify key residues involved in Rho's conformational dynamics along the simulated transition pathway.

An important treatment involved in the following string method calculation was that we discounted the rigid-body translation and rotation of the entire system from the collective variables **z**. To simplify the procedure we actually discounted overall translation and rotation of just the C_α atoms in the beta-sheet core of subunit m_1 by fixing the respective center of mass and orientation.

1.2 String method with swarms of trajectories. In our study we discretized the I

$\rightarrow$ F pathway into $M+1=51$ points, the points being referred to as images. As the path is iteratively refined, one needs to start from an initial choice, often making the trivial choice of a linear path connecting I with F. The details of how we build the initial path is introduced further below. Subsequently, one refines the images iteratively to obtain an optimal pathway as outlined for example in Refs. 3,4,16. In the "string method with swarms-of-trajectories"[4] the images are refined at iterations $j = 1, 2, \ldots$ following swarms of molecular dynamics trajectories. The trajectory swarms for each image starts from an ensemble of initial conditions close to image point $\mathbf{z}^{(\alpha)}(j)$, but the trajectories quickly diverge due to random initial thermalized positions and velocities. Each swarm involves S trajectories, each trajectory lasting time τ; in our case we chose $S = 30$ and $\tau = 4$ ps. The image moves $\mathbf{z}^{(\alpha)}(j)$ to $\tilde{\mathbf{z}}^{(\alpha)}(j+1) = \mathbf{z}^{(\alpha)}(j) + \Delta_\alpha(j)$ where $\Delta_\alpha(j)$ is the average displacement of the S trajectories. At this point one obtains 51 images $\tilde{\mathbf{z}}^{(\alpha)}(j+1)$ that are, in general, not smoothed or equally spaced along the pathway. The pathway smoothing and equal image spacing was achieved through an algorithm described next. In explaining this algorithm acting on $\tilde{\mathbf{z}}^{(\alpha)}(j+1)$ we replace the latter for the sake of simpler notation by $\tilde{\mathbf{z}}^{(\alpha)}$.

As pointed out in Results, the large numbers of degrees of freedom according to well established principles of statistics brings about stronger noise than do smaller numbers of degrees of freedom. To avoid problems due to the strong noise and resulting kinks in the transition pathway we followed Refs. 17,18 and smoothed the transition pathway (string) by means of multidimensional curve fitting. For this purpose one introduces a low-pass filter through a truncated Fourier series

$$\mathbf{z}^*(t) = \tilde{\mathbf{z}}^{(0)} + (\tilde{\mathbf{z}}^{(M)} - \tilde{\mathbf{z}}^{(0)})t + \sum_{i=1}^{n}\sum_{j=1}^{P} \beta_{ij}\mathrm{sin(j\pi t)}\mathbf{e}_\mathrm{i} \quad . \tag{S5}$$

Here $\mathbf{z}^*(t)$ ($t \in [0,1]$ is the progress variable along the transition pathway I $\rightarrow$ F) is an analytical function connecting the initial image $\tilde{\mathbf{z}}^{(0)}$ to the final image $\tilde{\mathbf{z}}^{(M)}$. $n = 660 \times 3$ is the dimension of the collective variable space spanned by the basis vectors $\mathbf{e}_i$. P is the

number of Fourier components kept; for the present study we use $P = 8$. $\{\beta_{ij}\}$ and $\{t_\alpha\}$ are optimized by minimizing the error χ, where

$$\chi^2 = \sum_{\alpha=0}^{M} |\mathbf{z}^*(t_\alpha) - \tilde{\mathbf{z}}^{(\alpha)}|^2 \quad . \tag{S6}$$

After the above multidimensional curve fitting one obtains $\mathbf{z}^*(t_\alpha)$ ($\alpha = 0, 1, ..., M$) which are, however, not separated by equal arc lengths along the new string. In order to make the arc lengths between each two adjacent images the same one needs to enforce the following condition and determine thereby new curve parameters t^*_α ($\alpha = 0, 1, ..., M$)

$$\int_{t^*_0}^{t^*_\alpha} ||\dot{\mathbf{z}^*}(t)||\, dt = \frac{\alpha}{M} L \quad . \tag{S7}$$

Here $t^*_0 = 0$, $t^*_M = 1$, and L is the total arc length of the string, namely $L = \int_0^1 ||\dot{\mathbf{z}^*}(t)||\, dt$. The corresponding redistributed images are $\mathbf{z}^{(\alpha)}(j+1)$.

In the last part of the current iteration step one provides an opportunity to all degrees of freedom of Rho to adjust themselves through relaxation to the new selected collective variables $\mathbf{z}^{(\alpha)}(j+1)$. For this purpose one carries out a 200 ps MD simulation in which all atomic coordinates are free to move except the atomic coordinates belonging to the collective variable set which were restrained around $\mathbf{z}^{(\alpha)}(j+1)$ with a harmonic potential whose spring constant is 2 kcal/(mol Å^2). From the last 60 ps of the trajectory, 30 snapshots are selected every 2 ps to serve as an ensemble of initial conditions for each swarm in the next iteration, namely $j + 2$. We also saved the averaged protein conformations $\overline{\mathbf{y}}(j)$ of the selected 30 snapshots at every iteration for the analysis proposed below.

The procedure above was carried out iteratively until one reached convergence. During the iteration process, the average RMSD for $\mathbf{z}^{(\alpha)}(j)$ of the 51 images in the string compared with the initial string $\mathbf{z}^{(\alpha)}(0)$ was monitored and it reached a plateau after convergence. For the current study, the string converged after 200 iterations. The average RMSD for the 51 images with reference to the initial configuration as a function of the iteration number

j is plotted in Figure S3. The expected values and error bars of the structural properties in Figure 3a, Figure S7 and Figure S9 were calculated over the last 70 iterations using the averaged protein conformations $\overline{\mathbf{y}}(j)$ (here j goes from 180 to 250).

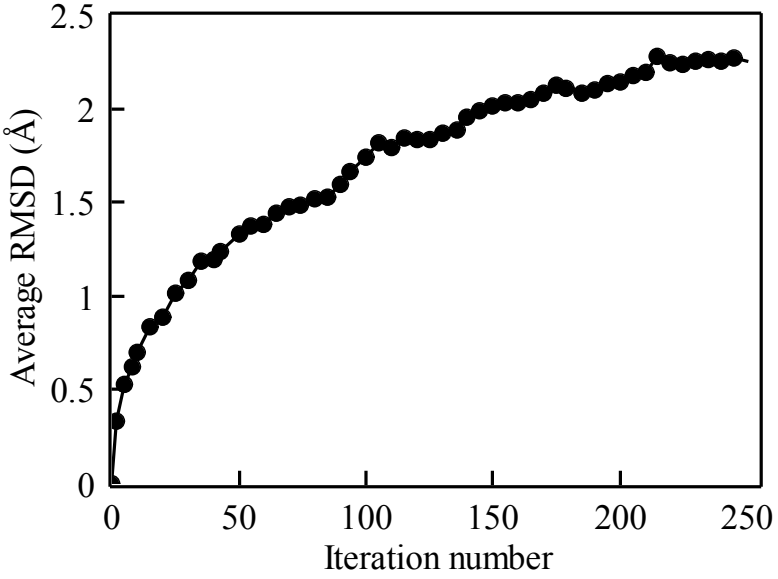

Figure S3: Average RMSD for the 51 images from the initial string as a function of iteration number. Given the small changes of the RMSD value after 200 iterations we consider the optimal path calculation converged.

1.3 Choice of the initial pathway. The initial transition pathway, also presented through 51 images, was generated through a 10 ns targeted MD simulation.[19] The targeted MD simulation applied a harmonic restraint to the I $\rightarrow$ F RMSD value of the selected Cartesian coordinates (660 atoms) driving the system straight from the I conformation to the F conformation. The center of the restraint was moved in 5 ns from I to F, and then kept restrained to F for another 5 ns of simulation. The images for the initial choice of transition pathway $\mathbf{z}^{(\beta)}$, $\beta = 1, 2, \ldots 50$ were then chosen as snapshots every 200 ps from the 10 ns targeted MD trajectory. With this choice one can proceed with the "string method with swarms-of-trajectories" computation.

2. Calculations of free energy and kinetics with milestoning.

Based on the converged transition pathway (images $\mathbf{z}^{\{i\}}, i = 0, 1, \ldots, M$) obtained through the string method one can explore the corresponding free energy and kinetics of the transition. In the following we demonstrate how Voronoi tessellation[9,10] along with sampling trajectories within the Voronoi cells allows one to obtain the free energy profile along the transition pathway. The mean first passage times (MFPTs) can be calculated through the milestoning analysis[7,9,10] by extracting statistical information of the Voronoi cell surface hitting events.

2.1 Free energy calculation. To obtain the free energy profile that hexameric helicase encounters along the I $\rightarrow$ F transition pathway, we employed the enhanced sampling method with Voronoi tessellation suggested in Ref. 9,10. The tessellation is obtained by making the images $\mathbf{z}^{\{i\}}$ along the transition pathway the centroids of associated Voronoi cells B_i, the latter defined through

$$B_i = \left\{ \mathbf{z} \in \mathbb{R}^n \mid ||\mathbf{z} - \mathbf{z}^{\{i\}}|| < ||\mathbf{z} - \mathbf{z}^{\{j\}}|| \ \forall j \neq i \right\} \tag{S8}$$

where $i, j = 0, 1, \ldots, M$ and $n = 660 \times 3$. The interfaces between neighboring Voronoi cells B_i and B_{i+1} ($B_i \cap B_{i+1}$) approximate the isocommittor surfaces of the reaction;[7] however interfaces between non-neighboring cells may also be non-empty, allowing then transitions between the corresponding non-adjacent cells.

In order to derive information on free energy and kinetics we launched twenty 0.8 ns-long MD simulations restrained in each Voronoi cell B_i with random initial velocities. The initial all-atom Cartesian coordinates $\mathbf{x}^{\{i\}}$ were taken from the snapshots near the end of the 200 ps restrained simulations during the 200th iteration in the string method calculation. The restrained behavior was implemented by adding a harmonic potential to conformations $\mathbf{z}$ which reached the outside of the B_i boundary, with force gradients normal to the planar surfaces and directed to the inside of the cell.[10] The parts of trajectories outside B_i were

dropped in the following calculations.

During the stated MD simulation in cell B_i we monitored how often the trajectories hit a common boundary with Voronoi cell B_j, i.e., one monitors how often a trajectory tried to move from cell i to cell j, defining thereby the numbers $N(i \to j)$. From the numbers $N(i \to j)$ an escape rate constant can be calculated through $k_{i,j} = N(i \to j)/T_i$ (here T_i is the total simulation time inside cell B_i). From $k_{i,j}$ a rate matrix $\mathbf{K}$ is obtained, namely

$$\mathbf{K}_{ij} = \left\{ \begin{array}{ll} k_{j,i} & \text{if } i \neq j \\ -\sum\limits_{\substack{l=0 \\ l \neq i}}^{M} k_{i,l} & \text{if } i = j \end{array} \right\} . \tag{S9}$$

One needs to check at this point if the resulting matrix $\mathbf{K}$ describes a random process of 1st order kinetics along the transition pathway that is ergodic such that there is a finite positive probability to pass from any cell B_i to any other cell B_j. If the system is not ergodic a free energy profile cannot be determined and one needs to increase the simulation time T_i in each cell until one obtains indeed a rate matrix $\mathbf{K}$ corresponding to an ergodic process. In the present case the sampling of $\mathbf{K}$ over trajectories required a total simulation time of 1.6 μs for the two free energy profiles in Figures 2b, c (0.8 μs each). The calculation for the mean first passage times (MFPTs) of the transition in the next step will be based on the same data set.

In case ergodicity has been reached the kernel of $\mathbf{K}$ is 1-dimensional and one can determine the single equilibrium probability vector $\boldsymbol{\pi}$ through

$$\mathbf{K} \cdot \boldsymbol{\pi} = 0 \tag{S10}$$

where the normalization $\sum_{i=0}^{M} \pi_i = 1$ is enforced. The equilibrium probability vector will be used again for the estimation of MFPTs. The free energy as a function of the reaction

coordinate, i.e., for cells B_i, $i = 0, 1, \ldots M$, can be calculated by means of

$$G_i = -k_{\mathrm{B}}\mathrm{T}\ \log(\pi_i) \tag{S11}$$

where k_{B} is the Boltzmann constant and T is the temperature.

The 1-dimensional free energy profile can be then projected onto a 2-dimensional x-y space through

$$F(x, y) = -k_{\mathrm{B}}\mathrm{T}\ \log[\sum_{i=0}^{M} \pi_i p_i(x, y)]\ . \tag{S12}$$

Here $p_i(x, y)$ is the normalized probability distribution for coordinates x and y projected from the trajectories associated with Voronoi cell B_i. Eq. S12 was used for plotting the 2-D free energy landscapes in the cases of Figures 3c, 4a, 5, and S5.

2.2 Calculation of mean first passage times. Milestoning, introduced by Faradjian and Elber,[6] estimates the long time kinetics of the transition from split trajectories crossing the milestones along the transition pathway. Recent developments of the milestoning technique introduced Markovian milestoning with Voronoi tessellations (MMVT)[9,10] and directional milestoning (DiM).[8,20] Here we employed MMVT which uses the interfaces shared by two Voronoi cells (non-empty $B_m \cap B_n$, $m \neq n$ and $m, n = 0, 1, \ldots, M$) as milestones labeled by indices $0, 1, \ldots, N_T$. The analysis takes transitions between non-neighboring cells into account so that N_T does not necessarily equal M. One has to first determine rates q_{ij} from milestone i to milestone j from the above mentioned trajectories inside the Voronoi cells. It has been proven that for the most likely rate matrix, given the observed trajectories, q_{ij} satisfies the condition:[9,10]

$$q_{ij} = \left\{ \begin{array}{cc} \frac{\sum_{l=0}^{M} \pi_l (N_{ij}^l / T_l)}{\sum_{l=0}^{M} \pi_l (R_i^l / T_l)} & \text{if } i \neq j \\ -\sum\limits_{\substack{l=0 \\ l \neq i}}^{N_T} q_{il} & \text{if } i = j \end{array} \right\}\ . \tag{S13}$$

Here π_l ($l = 0, 1, \ldots, M$) is the equilibrium probability for cell B_l determined by solving Eq. S10. N_{ij}^l, R_i^l and T_l are determined from the trajectory in cell B_l: N_{ij}^l is the number of transitions from milestone i to milestone j, R_i^l is the total time during which i has been the most recent milestone crossed by the system, and T_l is the total time of the trajectory. The numerator $\sum_{l=0}^{M} \pi_l (N_{ij}^l / T_l)$ and denominator $\sum_{l=0}^{M} \pi_l (R_i^l / T_l)$ inside the fraction represent averages of N_{ij}^l / T_l and R_i^l / T_l with weight π_l, respectively. After obtaining q_{ij}, one solves the following system of linear equations to determine the mean first passage times (MFPTs)[7,9,10,21] from milestone m to a target milestone n ($\tau_{m,n}$, $m \neq n$),

$$\sum_{\substack{m=0 \\ m \neq n}}^{N_T} q_{km} \tau_{m,n} = -1, \quad k \neq n. \tag{S14}$$

3. Analysis of interaction between Rho and RNA.

In this section we describe the analysis of the interaction between Rho and RNA leading to RNA translocation. The first part of the analysis carried out in the present study focussed on the residue level. We analyzed the interactions between individual RNA residues (one RNA residue j includes a phosphate group P_j and a nucleobase U_j, $j = 3, 4, ..., 6$) and individual Rho subunits $m_1, m_2, \ldots, m_6$ in four key intermediate states I, IM1, IM2 and F. The second part of the analysis determined the free energy of the Rho-RNA complex along Rho's conformational transition path and along the RNA translocation distance.

During the RNA translocation process, the interactions between RNA residues and Rho subunits $m_1, m_2, \ldots, m_6$ undergo significant changes. Figure 4 in Results illustrates the key role of K326 side chains which switch their interactions with RNA backbone phosphates to aid RNA translocation. In Figure S4 the interaction energies between Rho subunits $m_1, m_2, \ldots, m_6$ and RNA residues $3, 4, \ldots, 6$ in the states I, IM1, IM2, and F are characterized. Subunits m_2, m_3, m_5 and m_6 do not change significantly their interactions with RNA during the I $\rightarrow$ F transition; however subunits m_1 and m_4 alter their interactions with RNA greatly. During the transition I $\rightarrow$ IM1, the interaction between m_1 and RNA residue 3 is greatly weakened; during the transition IM2 $\rightarrow$ F, the interaction between m_4 and RNA reside 6 is notably strengthened. The timings of these transitions are consistent with the conformational transitions of K326 and RNA phosphate groups shown in Figure 4. By decomposing the RNA - Rho interaction energy into individual residue contributions (see Figure S6a), we will demonstrate below that the Rho-RNA interaction change is mainly controlled by K326, along with a contribution from the Q loop (drawn in Figure S2a).

In Results (see Figure 5) it had been already demonstrated that in case of the situation X = E RNA translocation is an exergonic process coupled to Rho conformational transitions. For comparison, we show in Figure S5 the free energy along the two progress variables Rho conformation and RNA translocation distance for the situation X = D. Here Rho conformation is defined by $\Delta\mathrm{RMSD} = \mathrm{RMSD(Y, I)} - \mathrm{RMSD(Y, F)}$ given that RMSD(Y, Z) is the

conformational RMSD difference between two conformations Y and Z. The I $\rightarrow$ F transition is strongly endergonic (the state F is energetically very unfavorable); in contrast to the 2.4 Å RNA translocation in the case X = E, no clear RNA translocation is observed in the case X = D. This comparison shows clearly that coupling between Rho conformation transition and RNA translocation is controlled by the hydrolysis product release.

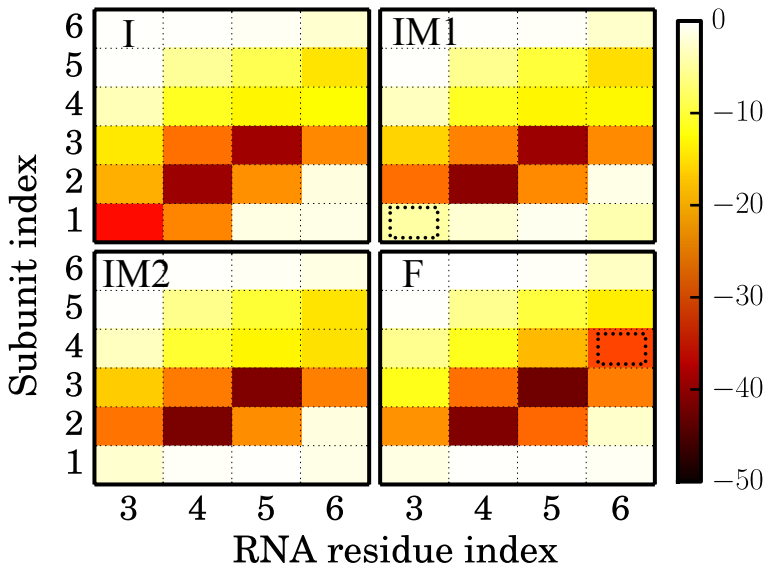

Figure S4: Interaction energy decomposition for six subunits and four RNA residues in states I, IM1, IM2 and F. The figure describes the I $\rightarrow$ F transition for the case X = E. The color in each block exhibits the interaction energy between one particular subunit and one particular RNA residue. The energy values are color-coded as shown in the color bar that uses the energy unit kcal/mol. The two dashed rectangles in IM1 and F indicate RNA residue release from m_1 (during the transition I $\rightarrow$ IM1) and RNA residue binding to m_4 (during the transition IM2 $\rightarrow$ F), respectively. The RNA residues not shown, namely 1, 2, 7, 8, 9, exhibit only negligible interaction strength with Rho.

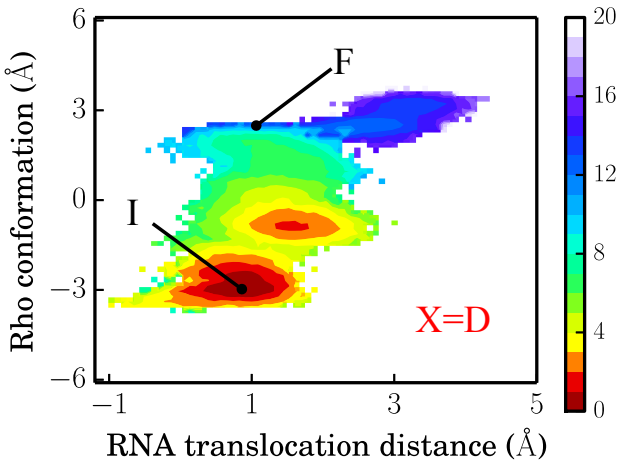

Figure S5: Free energy landscape along RNA translocation distance and Rho conformation for the X = D case. The Rho conformation is defined by ΔRMSD as introduced in the text; the RNA translocation distance is monitored by the center of mass distance between RNA backbone (residues 3 - 8) and Rho backbones (the C-terminal domain). States I and F (same as states I and F in Figure 2c) are labeled with black dots.

4. Residue-based interaction energy contribution in the I $\rightarrow$ F transition.

The conformational transition arising in the Rho-RNA complex involves switching of interactions at the Rho-RNA binding interface and at the subunit-subunit interfaces. In order to capture the interaction energy contribution from all residues, we follow into the footsteps of the definition of $\Delta E_s^{\text{prior}}$ stated in Eqs. (S2 – S4). $\Delta E_s^{\text{prior}}$ accounts for the deviation of interaction energy of individual residues among the six interfaces in the prior state R. In the following we seek to describe the root mean square deviation in the subunit-RNA and in the subunit-subunit interaction energy, $\Delta E_s^{\text{subunit-RNA}}$ and $\Delta E_s^{\text{subunit-subunit}}$, respectively, along the transition pathway, the latter represented as a string of $M+1 = 51$ discretized images. $\Delta E_s^{\text{subunit-RNA}}$ describes the interaction energy deviation along the pathway between residue s of Rho and all residues of RNA; $\Delta E_s^{\text{subunit-subunit}}$ describes the interaction energy deviation along the pathway between residue s of Rho and all residues of neigboring Rho subunits. The larger the value of $\Delta E_s^{\text{subunit-RNA}}$ or $\Delta E_s^{\text{subunit-subunit}}$, the more significant is the residue s for the coupling to RNA translocation and in the Rho conformational transitions. By s we denote again, as in case of $\Delta E_s^{\text{prior}}$, the respective residue in any of the six Rho subunits.

Before we specify the analysis in terms of $\Delta E_s^{\text{subunit-RNA}}$ and $\Delta E_s^{\text{subunit-subunit}}$ further, we summarize the overall result. The analysis reveals that residues T286, K326, R269, R272, K283, K298, R299, E333, E342 and R353 are significantly involved in the function of Rho. Indeed the key role of these residues had been singled out in earlier mutagenesis studies.[22–26] However, our analysis identifies residues D210, E215, D233, R305, E334, K367 and E368 to also be largely involved in Rho's function and recommends them for further investigation by means of mutagenesis studies. We note that the above mentioned residues fully cover the identified residues.

$\Delta E_s^{\text{subunit-RNA}}$ is defined through

$$\Delta E_s^{\text{subunit-RNA}} = \frac{1}{6}\sum_{j=1}^{6}\left[\frac{1}{(M+1)}\sum_{k=0}^{M}(E_{s,j}^{k} - \bar{E}_{s,j})^2\right]^{1/2} \quad \text{(S15)}$$

$$E_{s,j}^{k} = \sum_{r\in RNA} E_{s,j}^{k}(r) \quad \text{(S16)}$$

$$\bar{E}_{s,j} = \frac{1}{(M+1)}\sum_{k=0}^{M} E_{s,j}^{k} \quad . \quad \text{(S17)}$$

Here $M+1 = 51$ is the total number of images along the transition pathway. s, j, r and k are, respectively, subunit residue index (1 – 415), subunit index (1 – 6), RNA residue index (1 – 9) and image index (0 – 50). $E_{s,j}^{k}(r)$ is the interaction energy between residue s on subunit j and residue r on RNA for image k; $E_{s,j}^{k}(r)$ is averaged over the pathways of the last 70 iterations in the string method calculation for image k. $E_{s,j}^{k}$ is the interaction energy between residue s on subunit j and all residues of RNA for image k. $\bar{E}_{s,j}$ is the average interaction energy among all the images between residue s on subunit j and RNA. The first summation from $j = 1$ to 6 in Eq. S15 reveals $\Delta E_s^{\text{subunit-RNA}}$ as an average root mean square deviation among the six subunits $m_1, m_2, \ldots, m_6$. $\Delta E_s^{\text{subunit-RNA}}$ as a function of residue index s is shown in Figure S6a. Table S2 lists the important residues that exhibit a $\Delta E_s^{\text{subunit-RNA}}$ value of more than 2 kcal/mol along with information of related mutagenesis studies.

$\Delta E_s^{\text{subunit-subunit}}$ is defined through

$$\Delta E_s^{\text{subunit-subunit}} = \frac{1}{6}\sum_{j=1}^{6}\left[\frac{1}{(M+1)}\sum_{k=0}^{M}(E_{s,j}^{k} - \bar{E}_{s,j})^2\right]^{1/2} \quad \text{(S18)}$$

$$E_{s,j}^{k} = \sum_{s'\in S(j')}\sum_{j'\in N(j)} E_{s,j}^{k}(s', j') \quad \text{(S19)}$$

$$\bar{E}_{s,j} = \frac{1}{(M+1)}\sum_{k=0}^{M} E_{s,j}^{k} \quad . \quad \text{(S20)}$$

Here $M+1$, s, j and k are same as introduced above for Eqs. (S15 – S17). $N(j)$ and $S(j')$

are the set of neighbors of subunit j and the set of residues of subunit j' used in Eqs. (S2 – S4), respectively. $E^k_{s,j}(s', j')$ is the interaction energy between residue s on subunit j and residue s' on subunit j' for image k; $E^k_{s,j}(s', j')$ is averaged over the pathways of the last 70 iterations in the string method calculation for image k. $E^k_{s,j}$ and $\bar{E}_{s,j}$ are defined in similar ways as in Eqs. (S16 and S17). $\Delta E_s{}^{\text{subunit}-\text{subunit}}$ as a function of residue index s is shown in Figure S6b. Table S3 lists the important residues that exhibit a $\Delta E_s{}^{\text{subunit}-\text{subunit}}$ value of more than 5 kcal/mol along with information of related mutagenesis studies.

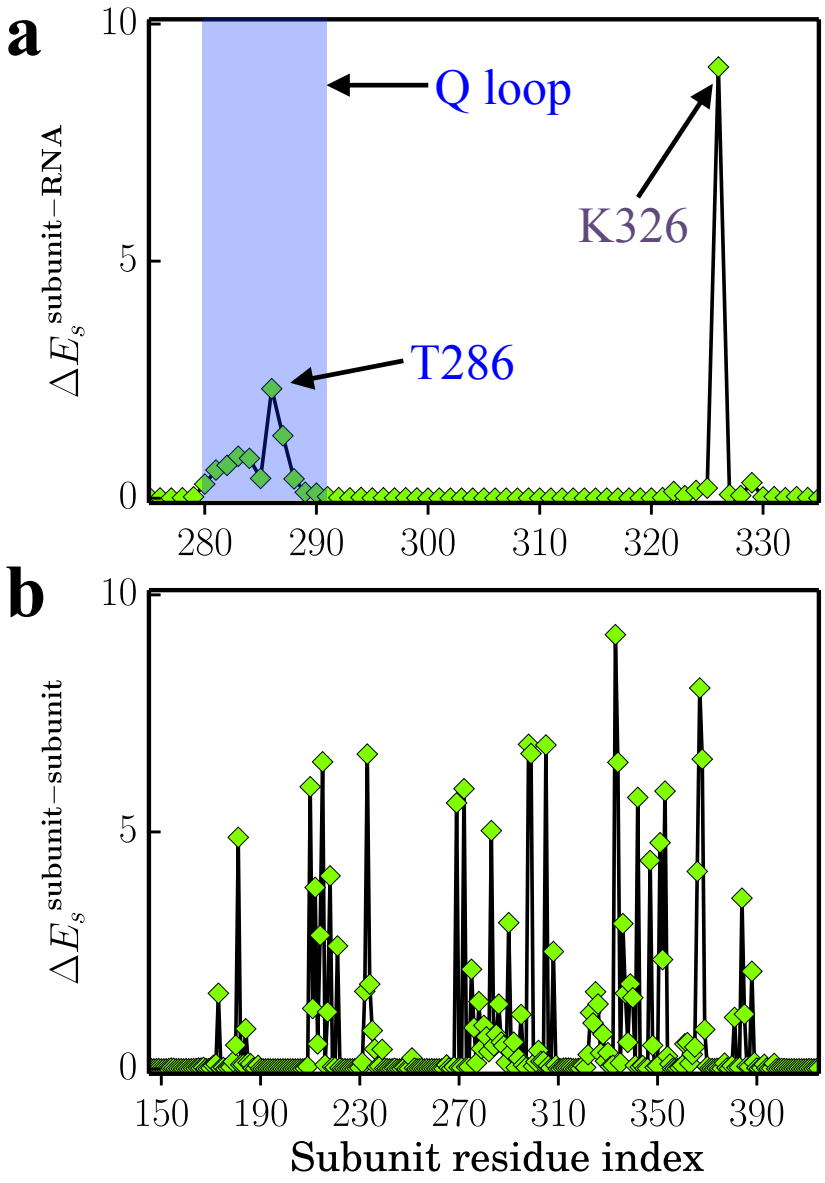

Figure S6: Residue-based contribution to the interaction energy deviation along the transition pathway of Rho. Residue-based $\Delta E_s^{\text{subunit-RNA}}$ and $\Delta E_s^{\text{subunit-subunit}}$ values are defined by Eqs. S15 and Eq. S18, respectively; they show the amplitude of interaction energy deviation in kcal/mol. (a) $\Delta E_s^{\text{subunit-RNA}}$ shows residue contributions to the interaction energy deviation between residue s of Rho and all residues of RNA. The contribution from K326 and T286 are highlighted by arrows. The blue shadow region highlights the contribution from residues in the Q loop (Figure S2a). Residues that contribute a value of more than 2 kcal/mol are listed in Table S2. (b) $\Delta E_s^{\text{subunit-subunit}}$ shows residue contributions to the interaction energy deviation between residue s of Rho and all residues of neigboring Rho subunits. Residues that contribute a value of more than 5 kcal/mol are listed in Table S3.

Table S2: Important residues contributing to subunit-RNA interaction predicted by the transition pathway

residue	$\Delta E_s^{\text{subunit-RNA}}$ [a]	mutation [b]	activity [c]
T286	2.3	T286A [22]	– –
K326	9.1	K326A [23]	– –

[a] The unit is kcal/mol;
[b] Citation of the corresponding report is given after the mutation;
[c] The *in-vitro* transcription termination activity of the mutant is indicated by the following symbols: – –, strong defects.

Table S3: Important residues contributing to subunit-subunit interaction predicted by the transition pathway

residue	$\Delta E_s^{\text{subunit}-\text{subunit}}$ [a]	mutation [b]	activity [c]
D210	6.0	NT	
E215	6.5	NT	
D233	6.6	NT	
R269	5.6	R269A [24]	see note [d]
R272	5.9	R272A [24]	see note [e]
K283	5.0	K283A [23]	– –
K298	6.8	K298A [25]	–
R299	6.6	R299A [25]	–
R305	6.8	NT	
E333	9.2	E333D [26]	– –
E334	6.5	NT	
E342	5.7	E342G [26]	– –
R353	5.9	R353A [23]	– –
K367	8.0	NT	
E368	6.5	NT	

[a] The unit is kcal/mol;
[b] Citation of the corresponding report is given after the mutation;
[c] The *in-vitro* transcription termination activity of the mutant is indicated by the following symbols: – –, strong defects; –, moderate loss of activity; NT, not tested.
[d] The transcription termination activity of this mutant was not tested before, however it resulted in strong defects in the ATPase activity.
[e] This mutant resulted in forming insoluble inclusion bodies.

5. ATP binding transitions before Rho becomes competent to hydrolyze ATP.

During the I $\rightarrow$ F transition the six domain-domain interfaces m_1/m_2, m_2/m_3, ..., m_6/m_1 adjust themselves to the ligand states changes of the R $\rightarrow$ I step, namely, in case of X = E, the changes T^*, T^*, T, T, E, D $\rightarrow$ D^*, T^*, T, T, T^-, E (see Figure 2a). The I $\rightarrow$ F transition involves two major ATP binding state changes, in particular, (i) the loosely bound state T^- to the moderate ATP binding state T at the m_5/m_6 interface and (ii) the moderate binding state T to the tight ATP binding state T^* at the m_3/m_4 interface. Right between the m_5/m_6 and m_3/m_4 interfaces, the m_4/m_5 interface also undergoes a major state change, which is required before the T $\rightarrow$ T^* transition could happen at the m_4/m_5 interface in the next rotary step.

The key changes involved in process (i) at the m_5/m_6 interface are depicted in Figure S7a, b, c. In state I the ATP molecule is bound loosely near the binding site, whereas in the final state F a close interaction between the ATP γ-phosphate and side chain of R366(m_6), the ATP binding motif commonly referred to as the arginine finger, develops. One can recognize in Figure S7b, which shows the distance R366(m_6)-P_γ along the reaction path, that R366(m_6) approaches the γ phosphate in state IM1, but that the interaction between R366(m_6) and P_γ exhibits still some fluctuations reflecting the fact that the tight-binding ATP state T^*, competent of actual hydrolysis, has not formed yet.

The key changes at the m_4/m_5 interface are illustrated in Figure S7d, e, f. The side chain of E211(m_4) is seen to disengage from the ATP-bound Mg^{2+} ion such that it can bind a water molecule and prepare a nucleophilic attack on the γ-phosphate that initiates hydrolysis. This shift of E211(m_4) arises during the step IM2 $\rightarrow$ F as can be clearly seen in Figure S7e, which shows the distance between E211(m_4) and Mg^{2+} along the reaction path; indeed the distance changes from a value of 2 Å to 4.5 Å in reaching F.

The key changes involved in process (ii) at the m_3/m_4 interface are illustrated in Figure S8a, b, c. The fluctuation of the E211(m_3) side chain is decreased so that it can stabilize

nearby water molecules to form water wires which facilitate proton transfer.[27,28] Here the degree of fluctuation of E211(m_3) is analyzed in Figure S8a, by calculating distributions of its position deviation from average positions, demonstrated for the four states I, IM1, IM2 and F. E211(m_3) at states IM1 and IM2 already becomes less flexible than state I; however, only at state F the distribution of deviation is centered at a small value 0.4 Å with a width of $\sim$ 0.4 Å. The dynamics of E211(m_3) is correlated with water distributions at the catalytic site. Water distributions in states I and F are then visualized around the γ-phosphate by drawing their positions which have ever been close to P_γ, presented in Figure S8b, c, respectively. The observed water occupancy of multiple coordination sites located between E211(m_3) and the γ-phosphate in state F is necessary for ATP hydrolysis, which is supposed to take place via a multi-center proton pathway as reported in QM/MM studies for F_1-ATPase[29,30] that is homologous to Rho. Several important water sites in Figure S8c (state F) are not observed in Figure S8b (state I).

In Figure S9 we show the interaction energy change between m_4 and its neighboring subunits m_3 and m_5 as an indication of the T $\rightarrow$ T* transition at the m_3/m_4 interface. The interaction becomes stronger, namely decreases in value, after state IM2 has formed, indicating tight binding between m_4 and its neighboring subunits.

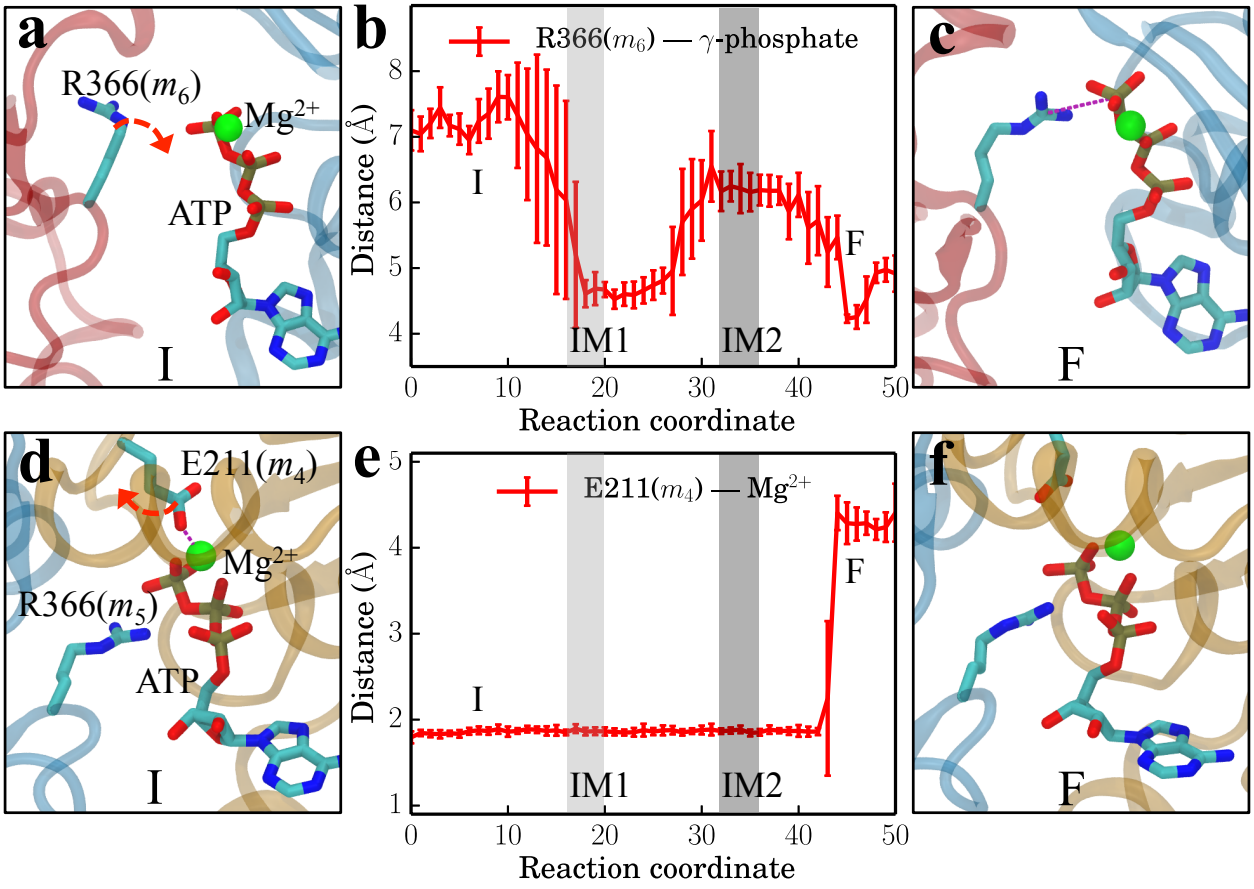

Figure S7: Participation of key residues in interface state changes during the I → F transition. (a-c) Side chain of R366(m_6), the so-called arginine finger, coordinates with the γ phosphate of ATP during the T$^-$ → T transition at the m_5/m_6 interface. The juxtaposition of R366(m_6) and P$_\gamma$ of ATP in states I and F is shown in (a) and (c), respectively. The distance between atom CZ of R366(m_6) and P$_\gamma$ of ATP along the transition pathway is presented in (b). (d-f) Side chain of E211(m_4), the so-called catalytic glutamate, disengages from the magnesium ion at the m_4/m_5 interface, as required before the T → T* transition could happen. The juxtaposition of E211(m_4), R366(m_5) and Mg^{2+} in states I and F is shown in (d) and (f), respectively. The time evolution of the distance between atom OE2 of E211(m_4) and Mg^{2+} in complex with ATP is shown in (e). Two intermediates IM1 and IM2 are highlighted by light grey and dark grey areas, respectively. Magnesium ions are shown as green spheres, protein residues and ATP molecules in licorice; adjacent subunits are displayed in transparent cartoon representation.

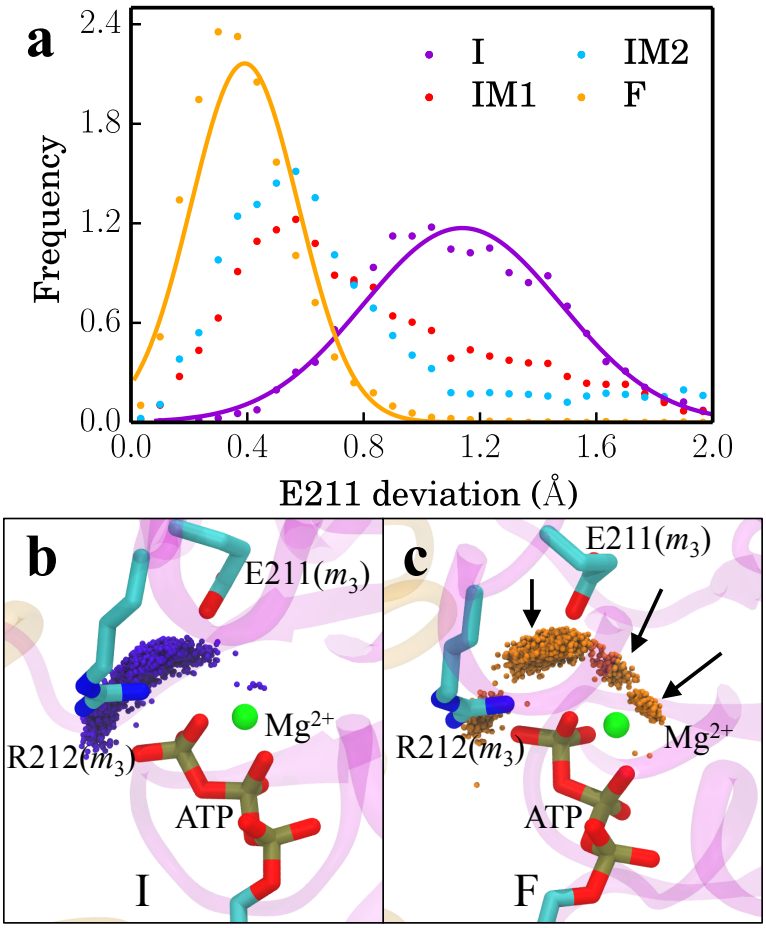

Figure S8: Key changes at the m_3/m_4 interface during the I $\rightarrow$ F transition. (a) The distributions of E211(m_3) position deviation (actually the distance between a snapshot and the averaged position of atom OE2 of E211(m_3)) for states I, IM1, IM2 and F are plotted as violet, red, cyan and orange dots, respectively. The distributions for states I and F are fitted to normal distributions (solid lines). (b, c) The water distributions at the catalytic site in states I and F are shown in (b) and (c), respectively. A water molecule is displayed if one of its hydrogen atoms has ever been within 2.6 Å of P_γ. Only oxygen atoms of water molecules are shown, namely as violet or orange spheres for state I or F, respectively. Black arrows indicate the coordination sites occupied by catalytic water molecules in (c). Magnesium ions are shown as green spheres, protein residues and ATP molecules in licorice; adjacent subunits are displayed in transparent cartoon representation.

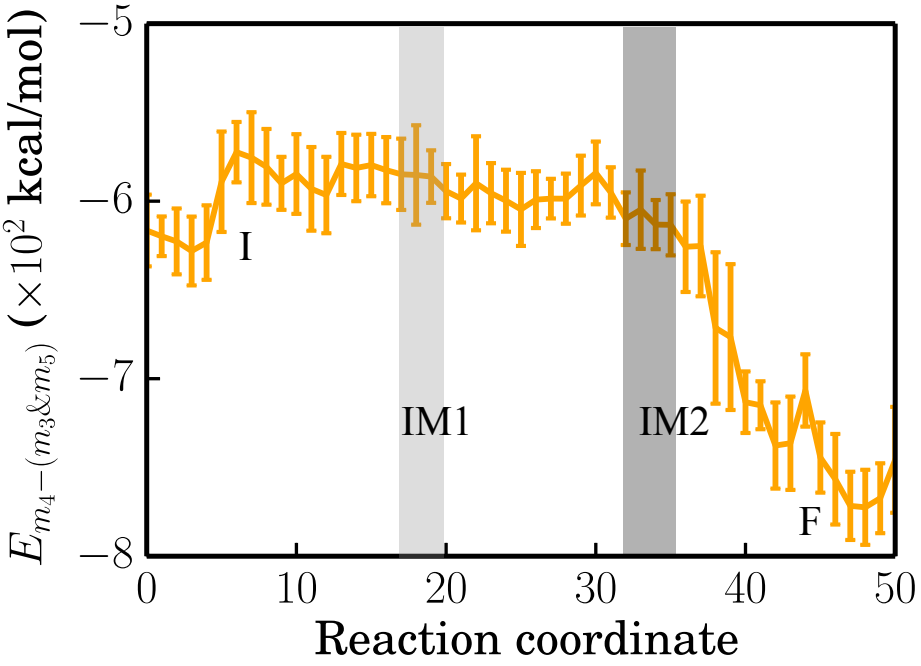

Figure S9: Interaction energy between subunit m_4 and its neighboring subunits m_3 and m_5 along the discrete reaction coordinate. The states I, IM1, IM2 and F are labeled, states IM1 and IM2 being highlighted by light grey and dark grey areas, respectively.

6. Movie S1

RNA translocation during six sequential Rho rotary reaction steps, namely steps 1-6 defined in Table S4, is shown in Movie S1. For the purpose of the movie the conformational transition of the RNA-Rho system during the R $\rightarrow$ F step were extended to five further rotary steps, the coordinates of $M + 1 = 51$ frames in the I $\rightarrow$ F transition pathway being repeatedly rotated by 60° around the z-axis to generate the 51 $\times$ 5 frames in the five further steps. The extended trajectory with a total of 51 $\times$ 6 frames has been smoothed by a gliding average over 3 frames. The first mean passage time (MFPT) in μs from the initial state to a state of a given frame is shown in the movie; the MFPT values show the time evolution of the Rho-RNA conformational transitions.

As the ligand state transformation proceeds from state R to state F, along with it proceed the interactions of the K326 side chains with RNA phosphate groups: in step 1 K326(m_1) disengages from phosphate P_3, while K326(m_4) engages with phosphate P_6; likewise in step 2 K326(m_2) disengages from phosphate P_4, while K326(m_5) engages with phosphate P_7, etc. The numbering of the RNA backbone phosphates involved in the disengagement and engagement increases during steps 1 to 7 from (P_3, P_6) to (P_9, P_{12}) = (P_{3+6}, P_{6+6}), and accordingly the RNA experiences during the first six steps a translocation by a length of 6 $\times$ 2.4 Å.

The P loops and Q loops of subunits m_1, m_2, m_3 and m_4 are shown in the movie in transparent cartoon representation in green, violet, magenta and orange respectively; for clarity the loops of subunits m_5 and m_6 are not shown. The side chains of the lysines K326 of subunits m_1, m_2, m_3 and m_4 are shown in opaque colors, namely green, violet, magenta and orange, respectively. The RNA is shown in pink cartoon representation. RNA backbone atoms are shown in transparent bonds and spheres colored by atom name; the phosphorus atom of P_7 is shown as an opaque tan sphere.

The movie highlights the important role for RNA translocation of the six K326 side chains accompanying the motion of the Rho subunits. The motion of the lysines is well adjusted

to the geometry of the bound RNA such that the RNA gets pulled in a linear direction through engagement with and disengagement from the backbone phosphate groups; the RNA adopting a helical geometry in its bound state does not rotate as it is translocated relative to Rho.

Table S4: Continuation of RNA translocation processes in Rho. Columns m(off) and m(on) state the subunits to which subunits K326 belong that disengage from or engage with, respectively, an RNA phosphate group in a particular step; columns P(off) and P(on) state the respective phosphate groups.

step	R	F	m(off)	P(off)	m(on)	P(on)
1	T*, T*, T, T, E, D	D, T*, T*, T, T, E	m_1	P_3	m_4	P_6
2	D, T*, T*, T, T, E	E, D, T*, T*, T, T	m_2	P_4	m_5	P_7
3	E, D, T*, T*, T, T	T, E, D, T*, T*, T	m_3	P_5	m_6	P_8
4	T, E, D, T*, T*, T	T, T, E, D, T*, T*	m_4	P_6	m_1	P_9
5	T, T, E, D, T*, T*	T*, T, T, E, D, T*	m_5	P_7	m_2	P_{10}
6	T*, T, T, E, D, T*	T*, T*, T, T, E, D	m_6	P_8	m_3	P_{11}
7	T*, T*, T, T, E, D	D, T*, T*, T, T, E	m_1	P_9	m_4	P_{12}
⋮	⋮	⋮	⋮	⋮	⋮	⋮

7. Movie S2

Movie S2 shows the relative motions of Rho subunits during six sequential Rho rotary reaction steps.

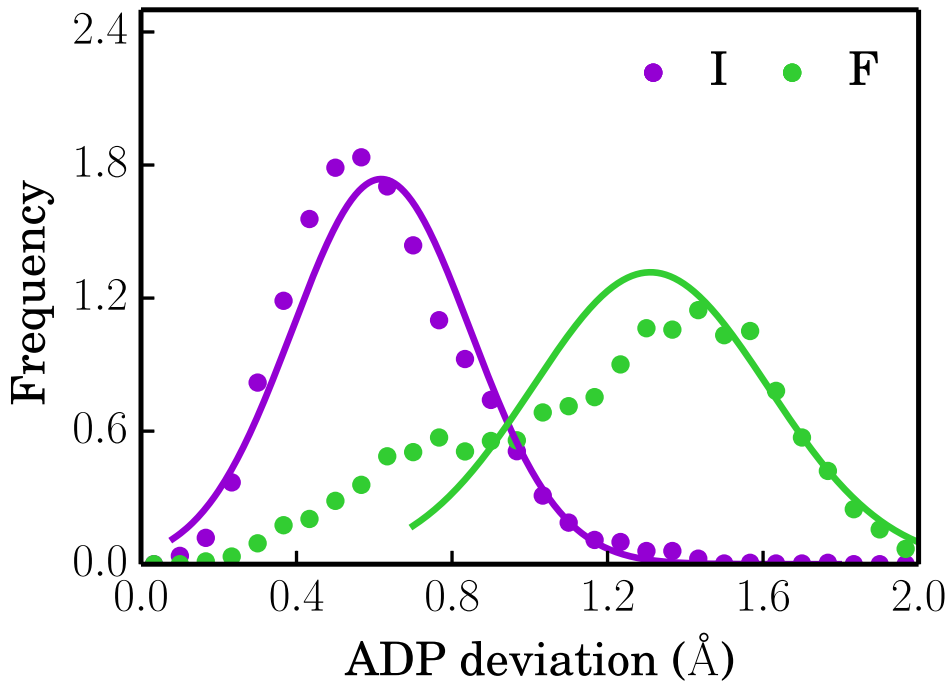

Figure S10: Firmness of ADP binding at the m_1/m_2 interface at the beginning and end of a conformational rotation I $\rightarrow$ F. The firmness is characterized through the average RMSD values of atoms C4′ and C6. Shown are the distributions of the RMSD values sampled from the trajectories simulated in Voronoi cells as described in Methods and Sec. 2 in Supporting Information. The distributions are fitted to normal distributions (solid lines).

Table S5: Subunit binding pocket[a] RMSDs between crystal structure (pdb: 3ICE) and state I[b]

subunit j	1	2	3	4	5	6
RMSD$_j$(3ICE,I) (Å)	0.78	1.20	0.82	0.87	0.94	1.00

[a] A binding pocket for subunit j is defined by the C$_\alpha$ atoms of residues 158, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 189, 190, 211, 212, 215, 265, 269, 318, 320, 336, 337, 338, 345, 355, 356, 357, 365, 366, 367, 368, 369.
[b] The state I conformation is taken from the optimal transition pathway.

References

(1) Elber, R. *Curr. Opin. Struct. Biol.* **2011**, *21*, 167 – 172.

(2) Vanden-Eijnden, E. In *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*; Bowman, G. R., Pande, V. S., Noé, F., Eds.; Springer Netherlands, 2014; Vol. 797; pp 91–100.

(3) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 024106.

(4) Pan, A.; Sezer, D.; Roux, B. *J. Phys. Chem. B* **2008**, *20*, 3432–3440.

(5) Johnson, M. E.; Hummer, G. *J. Phys. Chem. B* **2012**, *116*, 8573–8583.

(6) Faradjian, A. K.; Elber, R. *J. Chem. Phys.* **2004**, *120*, 10880–10889.

(7) Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R. *J. Chem. Phys.* **2008**, *129*, 174102.

(8) Májek, P.; Elber, R. *J. Chem. Theor. Comp.* **2010**, *6*, 1805–1817.

(9) Vanden-Eijnden, E.; Venturoli, M. *J. Chem. Phys.* **2009**, *130*, 194101 – 194101–13.

(10) Maragliano, L.; Vanden-Eijnden, E.; Roux, B. *J. Chem. Theor. Comp.* **2009**, *5*, 2589–2594.

(11) Adelman, J. L.; Jeong, Y.-J.; Liao, J.-C.; Patel, G.; Kim, D.-E.; Oster, G.; Patel, S. S. *Mol. Cell* **2006**, *22*, 611 – 621.

(12) Chen, X.; Stitt, B. L. *J. Biol. Chem.* **2009**, *284*, 33773–33780.

(13) Stitt, B. L.; Xiao, H. *J. Mol. Biol.* **2010**, *402*, 813 – 824.

(14) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins: Struct., Func., Bioinf.* **2004**, *55*, 383–394.

(15) Tanner, D. E.; Phillips, J. C.; Schulten, K. *J. Chem. Theor. Comp.* **2012**, *8*, 2521–2530.

(16) Maragliano, L.; Roux, B.; Vanden-Eijnden, E. *J. Chem. Theor. Comp.* **2014**, *10*, 524–533.

(17) Khavrutskii, I. V.; Arora, K.; Brooks III, C. L. *J. Chem. Phys.* **2006**, *125*, 174108.

(18) Zhu, F.; Hummer, G. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 19814–19819.

(19) Schlitter, J.; Engels, M.; Krüger, P. *J. Mol. Graph.* **1994**, *12*, 84–89.

(20) Kirmizialtin, S.; Elber, R. *J. Phys. Chem. A* **2011**, *115*, 6137–6148.

(21) Park, S.; Şener, M. K.; Lu, D.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 1313–1319.

(22) Wei, R. R.; Richardson, J. P. *J. Mol. Biol.* **2001**, *314*, 1007 – 1015.

(23) Rabhi, M.; Gocheva, V.; Jacquinot, F.; Lee, A.; Margeat, E.; Boudvillain, M. *J. Mol. Biol.* **2011**, *405*, 497 – 518.

(24) Xu, Q.; Baciou, L.; Sebban, P.; Gunner, M. R. *Biochemistry* **2002**, *41*, 10021–10025.

(25) Xu, Y.; Kohn, H.; Widger, W. R. *J. Biol. Chem.* **2002**, *277*, 30023–30030.

(26) Miwa, Y.; Horiguchi, T.; Shigesada, K. *J. Mol. Biol.* **1995**, *254*, 815 – 837.

(27) Colavin, A.; Hsin, J.; Huang, K. C. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 3585–3590.

(28) McCullagh, M.; Saunders, M. G.; Voth, G. A. *J. Am. Chem. Soc.* **2014**, *136*, 13053–13058.

(29) Dittrich, M.; Hayashi, S.; Schulten, K. *Biophys. J.* **2003**, *85*, 2253–2266.

(30) Hayashi, S.; Ueno, H.; Shaikh, A. R.; Umemura, M.; Kamiya, M.; Ito, Y.; Ikeguchi, M.; Komoriya, Y.; Iino, R.; Noji, H. *J. Am. Chem. Soc.* **2012**, *134*, 8447–8454.