JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Molecular Mechanism of Processive 3′ to 5′ RNA Translocation in the Active Subunit of the RNA Exosome Complex

Lela Vuković,[*,†,‡,§,∥] Christophe Chipot,[†,§,⊥] Debora L. Makino,[#,○] Elena Conti,[#] and Klaus Schulten[*,†,‡,§]

[†]Department of Physics, [‡]Center for the Physics of Living Cells, and [§]Beckman Institute, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States

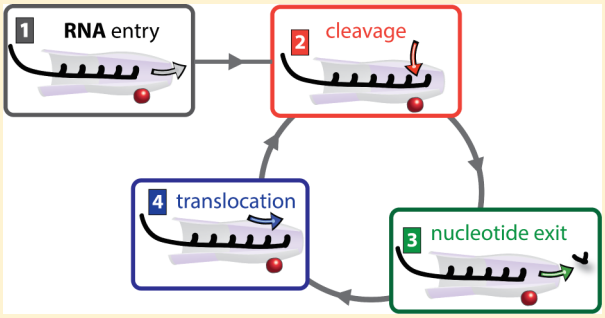
[∥]Department of Chemistry, University of Texas at El Paso, El Paso, Texas 79968, United States

[⊥]Laboratoire International Associé CNRS-University of Illinois, Université de Lorraine, Vandoeuvre-lès-Nancy 54000, France

[#]Department of Structural Cell Biology, Max Planck Institute of Biochemistry, Martinsried 82152, Germany

Supporting Information

ABSTRACT: Recent experimental studies revealed structural details of 3′ to 5′ degradation of RNA molecules, performed by the exosome complex. ssRNA is channeled through its multisubunit ring-like core into the active site tunnel of its key exonuclease subunit Rrp44, which acts both as an enzyme and a motor. Even in isolation, Rrp44 can pull and sequentially cleave RNA nucleotides, one at a time, without any external energy input and release a final 3−5 nucleotide long product. Using molecular dynamics simulations, we identify the main factors that control these processes. Our free energy calculations reveal that RNA transfer from solution into the active site of Rrp44 is highly favorable, but dependent on the length of the RNA strand. While RNA strands formed by 5 nucleotides or more correspond to a decreasing free energy along the translocation coordinate toward the cleavage site, a 4-nucleotide RNA experiences a free energy barrier along the same direction, potentially leading to incomplete cleavage of ssRNA and the release of short (3−5) nucleotide products. We provide new insight into how Rrp44 catalyzes a localized enzymatic reaction and performs an action distributed over several RNA nucleotides, leading eventually to the translocation of whole RNA segments into the position suitable for cleavage.



1. INTRODUCTION

In eukaryotes, a major fraction of DNA is copied into RNA molecules that fulfill then manifold functions, namely protein translation and control of protein translation.[1] Yet, the operation of the cell has to be tightly controlled, so that synthesized molecules contain no errors and exist for appropriate amounts of time at appropriate concentrations.[2,3] To achieve a tight regulation of the RNA content in the cell, proteins and protein complexes act as molecular machines and recognize and degrade RNA transcripts whenever needed.[2,4] The RNA exosome complex is a major RNA degradation complex in eukaryotes,[5] responsible for most of 3′ to 5′ RNA degradation,[6] and is involved in the degradation of multiple RNA types, including rRNAs, small nuclear and nucleolar RNAs, mRNAs, as well as various noncoding and defective RNAs.[7−9]

The exosome complex can recognize and degrade RNA on its own or with the help of cofactors, some of which are powered by ATP.[10−12] However, the exosome core can degrade substrates *in vitro* even without external energy input. A key and only subunit of the exosome that has the ability to processively degrade RNA, both within the exosome and by itself,[13,14] is the protein subunit Rrp44, a 3′ exonuclease.[15] Besides cleaving ssRNA substrates, Rrp44 is also able to unwind and degrade duplex and hairpin structures, as long as 3′ single-stranded RNA (ssRNA) extensions are present.[14]

Recent crystal structures and cryo-electron microscopy maps have revealed the architecture of the exosome complex and its binding mode to RNA during the degradation process,[16−19] shown schematically in Figure 1a. Three main regions, including the cap, the core, and the Rrp44 subunit, assemble into the exosome complex, while forming a long internal channel within it. ssRNA is then channeled within the exosome from cap proteins at the channel entry to the Rrp44 exonuclease active site at the channel exit,[16,18] as shown in Figure 1a. The length of the exosome channel spans approximately 30 nucleotides (nt).[16]

The cap and core subunits of the exosome initially recruit RNA substrates that have ssRNA extensions.[16] When ssRNA extensions are longer than 25 nucleotides, the 3′-end of RNA can reach the part of the exosome's interior channel formed by

Received: November 17, 2015
Published: March 1, 2016

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

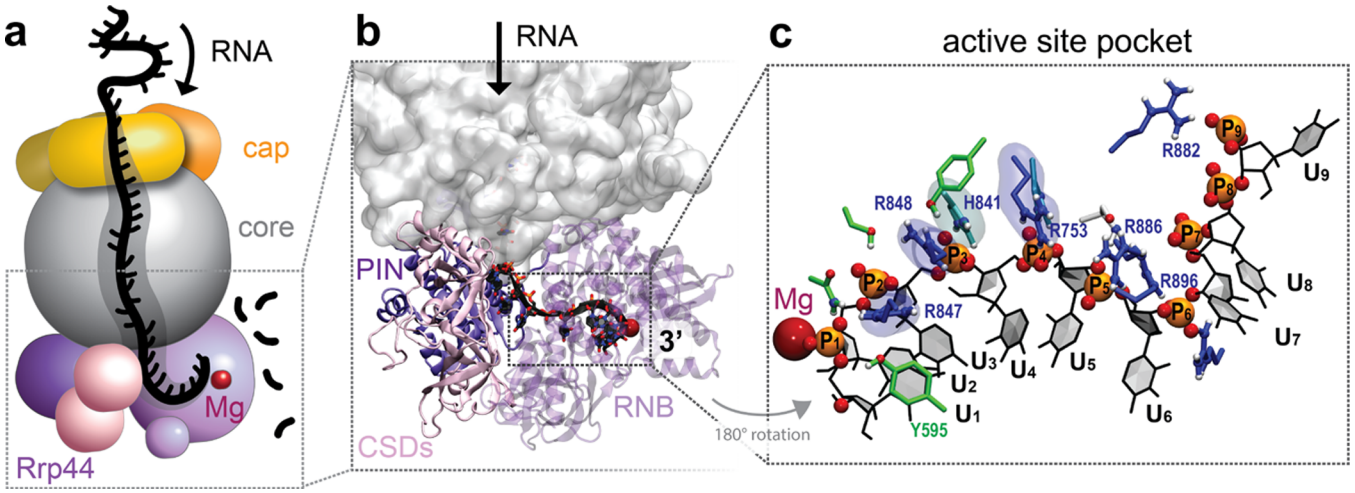


Figure 1. Exosome complex bound to ssRNA. (a) Schematic view of the exosome complex and bound ssRNA, the latter getting cleaved one nucleotide at-a-time. (b) Magnified structure of the exosome barrel-like core, shown as a silver density, is attached to the Rrp44 enzyme, whose PIN (dark purple), CSD (pink), and RNB and S_1 domains (transparent purple) are shown in ribbon representation. RNA is bound to the exonuclease binding site in the RNB domain of Rrp44. (c) Structure of the exonuclease active site bound to ssRNA. Key protein residues that bind to ssRNA are shown in licorice representation, and positively charged residues that coordinate phosphate groups of ssRNA tightly are shown as transparent densities. Positively charged residues are shown in blue color, polar residues in green, and nonpolar residues in white. Heavy atoms of RNA are shown in black, except for the atoms of the phosphate group (oxygen in red, phosphorus in orange).

the catalytic active site tunnel of Rrp44.[16] Then, the Rrp44 subunit can act both as a motor, by unwinding and pulling RNA substrates toward the cleavage site at the end of the interior channel, and as an enzyme, by hydrolytically cleaving RNA substrates into single 5′ monophosphate nucleotides.[15,20]

Upon entry of RNA into the active site of Rrp44 (step 1), the following steps of the RNA degradation cycle, which involve enzyme and motor actions of Rrp44, are shown schematically in Figure 2: (step 2) The 3′-end RNA nucleotide is hydrolytically cleaved at the Rrp44 exonuclease active site; (step 3) the cleaved nucleotide leaves the active site; and (step 4) RNA translocates forward by one "step," moving its 3′-end again into the position required for cleavage. Step 4 in the cycle can sometimes require that RNA at the entry of the exosome channel be unfolded, since only single stranded RNA can move into the channel.[16,18] The final products of RNA degradation by Rrp44 are 3−5 nt RNA pieces, which are released into the solution; the 4-nt final product is the most abundant.[15,16,18]

While the available structural data provide a snapshot of RNA within the internal channel of the exosome, they do not offer any explanation as to how the Rrp44 subunit achieves processivity in RNA degradation. What energy drives the repetition of the exosome's functional cycle (Figure 2) and how does the associated driving force become actually strong enough to enable unfolding of folded/assembled RNA substrates? To answer these intriguing questions we note first that the exosome cleavage cycle does not necessarily require external energy input. Nevertheless, substantial energy is released in the hydrolytic cleavage of RNA.[8,9] However, just like for ATP hydrolysis, the released energy can drive other processes in the cell, such as translocation or unfolding of RNA in the present case, albeit only if there is a coupling mechanism that channels the released energy into a mechanical force,[21−23] here directed toward pulling the 3′ end of RNA toward the Rrp44 active site, as it is being cleaved.

In the present study, we examine the energetics of the RNA translocation reaction in the key subunit of the exosome, the Rrp44 exonuclease, to determine how the energy released in the RNA hydrolysis reaction is required to be funneled to power RNA translocation. We also examine how the inner design of the exosome and Rrp44 exonuclease contributes to the processive degradation of RNA. The above issues are addressed by means of atomistic molecular dynamics (MD) simulations and free energy calculations, based on the recently solved structure of the exosome.[16]

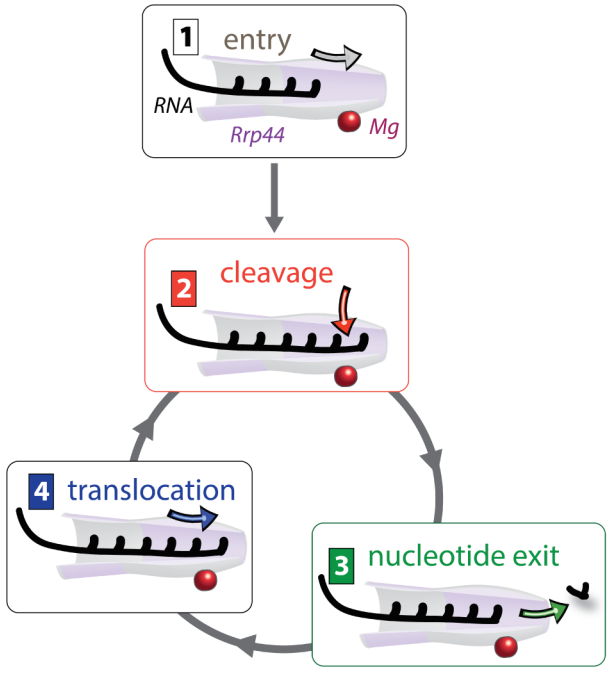


Figure 2. Functional cycle of RNA exosome. With its 3′-end, ssRNA enters the exosome and the active site tunnel of Rrp44 (step 1), followed by a repeated cycle of a 3′ nucleotide cleavage, cleaved nucleotide exit and further translocation of a whole ssRNA (steps 2−4).

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

2. RESULTS

The present study examines how the exosome powers processive RNA translocation and cleavage reactions by means of its subunit Rrp44. An interesting aspect of these reactions is that they occur in a continuous cycle and couple a highly localized reaction, the cleavage of the 3′-end RNA nucleotide, and an action that is distributed over several RNA nucleotides, namely, the translocation of a whole RNA segment toward the vacant position for cleavage. The fact that the mechanism of the cycle summarized in Figure 2 is not localized, but distributed along several RNA nucleotides (and eventually the whole RNA molecule that is to be degraded), makes the computational description conceptually and methodologically challenging.

To examine the mechanism of RNA translocation, the model of a complete exosome complex bound to RNA was first relaxed in MD simulations. The relaxed key part of the exosome, Rrp44 bound to RNA substrate and attached to a part of the exosome core, was examined in the subsequent simulations. First, the relative binding affinities of individual RNA nucleotides to the active site tunnel of Rrp44 were determined. Then, the binding and translocation dynamics of a representative RNA segment in the active site tunnel of Rrp44 were characterized. The simulations identified that the active site tunnel of Rrp44 possesses a characteristic binding pattern to RNA, with two very strong RNA binding sites separated from each other by a portion of space capable of accommodating three nucleotides, and a significantly weaker binding site in between. Since only RNA molecules that contain five or more nucleotides can simultaneously bind to both strong RNA binding sites, the active site tunnel of Rrp44 is well attuned to translocation of RNA segments five or more nucleotides long.

2.1. Molecular Dynamics Simulations of the Exosome. A model of the complete yeast exosome complex bound to RNA, based on the crystal structure reported in,[16] was used as a precursor of the reduced system including the Rrp44 subunit, which was studied in all subsequent simulations. To relax Rrp44 in the state that it occupies in the crystal structure reported in,[16] the whole exosome model was simulated for 24 ns, while the parts of the exosome core proteins were held restrained (α helices and β sheets). The RMSD of the Rrp44 subunit within the exosome complex is displayed in Figure 3. In relaxation simulations, the RMSD of Rrp44 (parts of defined secondary structure) plateaus at ≈2.3 Å. The relaxed Rrp44 bound to the RNA substrate and attached to a part of the exosome core was used as the starting structure for the ensuing simulations.

2.2. Exonuclease Active Site Tunnel of Rrp44 Contains Multiple Strong Binding Sites for ssRNA. To investigate the factors that contribute to processive translocation and cleavage of RNA in Rrp44[13] and the exosome,[16] the present study focuses on the interaction of RNA with Rrp44 exonuclease, the bottom subunit of the exosome highlighted in Figure 1a,b. The sequence of RNA (poly(U)) examined in the simulations performed here is based on the sequence of RNA in the crystal structure.[16] We first quantify binding of the exosome's internal channel to individual terminal RNA nucleotides covalently bound to longer ssRNA strands, overall 5−9 nucleotides long (as detailed in Table S1). Binding is characterized in terms of transfer free energies, which were determined in free energy perturbation (FEP) calculations as described in Methods, according to the thermodynamic cycle shown in Figure S1. The procedure employed ensures that the binding free energies of single nucleotides take into account the presence of longer covalently bound ssRNA fragments in the active site tunnel of Rrp44, i.e., a realistic environment of single nucleotides is studied. The results of individual FEP calculations are listed in detail in Table S1.

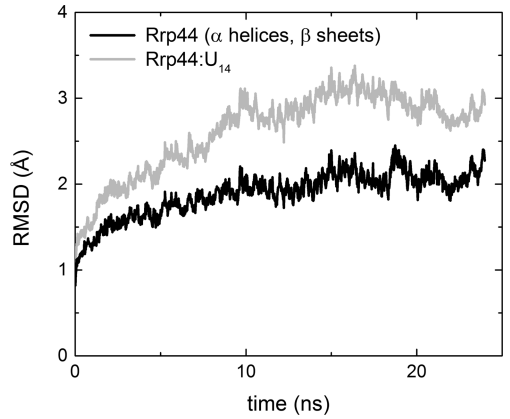


Figure 3. RMSD of Rrp44 in equilibration MD simulations of the complete exosome complex. RMSD of the whole Rrp44 and its ssRNA substrate is shown in gray, while the RMSD of the structurally defined parts of Rrp44 is shown in black.

Figure 4 shows transfer free energies of single uracil nucleotides from water to selected positions within the exosome, labeled in Figure 1c. The values reported take into account the presence of longer ssRNA strands to which single nucleotides are added or deleted in the course of the free

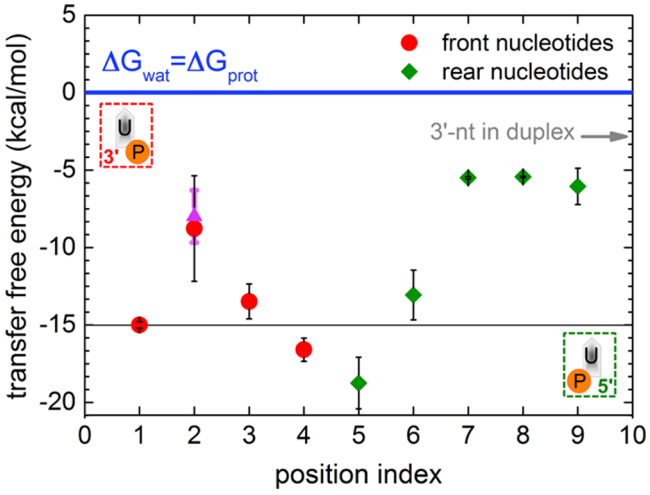


Figure 4. Interaction of RNA with the exonuclease active site of Rrp44. Shown are transfer free energies of single RNA nucleotides, covalently bound to longer ssRNA strands, from water into the active site tunnel of Rrp44. Transfer free energies are obtained for single terminal RNA nucleotides, covalently bound to longer ssRNA strands, overall 5−9 nucleotides long, as shown in Figure S1 and Table S1. All the data from FEP calculations, which were used to determine the plotted values, are reported in Table S1. Transfer free energies at position P_2 are obtained in two independent simulations, marked by red and purple points at position index 2. Inset: A scheme of annihilated nucleotide pieces from the 3′ end (red) and from the 5′ end (green); the 5′ end terminates with the 5′-OH group (based on parametrization of RNA nucleotides in the Amber force field). The position indices correspond to those shown in Figure 1c.

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

energy calculations (Table S1). Transfer of nucleotides from water to the protein environment is favorable at all examined positions of the extended RNA binding pathway; in fact, the transfer free energies range from −5 to −19 kcal/mol. The exonuclease active site tunnel contains several sites binding RNA very strongly, including positions P_1, P_4, and P_5, where the transfer free energies are as large as −14 kcal/mol, akin to the energy gained through ATP hydrolysis. The weakest binding sites for RNA are at positions P_2, P_7, P_8, and P_9. Interestingly, position P_1, the position where 3′ nucleotide gets cleaved, does not exhibit the tightest binding to RNA nucleotides. Positions P_4 and P_5 have more favorable binding than position P_1, by −2 and −4 kcal/mol, respectively.

Most of the errors in Figure 4 are small or moderate (<1.7 kcal/mol). The largest error of 3.4 kcal/mol is obtained for the binding free energy of a 3′-end nucleotide at position P_2 of the Rrp44 active site tunnel. This large error is most likely related to variations in coordination of Mg^{2+} in different λ-windows of FEP calculations, which should have a significant effect on the 3′-end nucleotide in position P_2: the 3′-end hydroxyl group of this nucleotide is very close to Mg^{2+}. The binding free energies at all other positions are accompanied by small errors, since Mg^{2+} coordination is either well determined with respect to RNA (when the 3′-nucleotide is at position P_1, as it occurs in pdbID: 4IFD), or not as influential (in all other positions, 3′-end nucleotides do not directly interact with Mg^{2+} and its coordination shell). To ascertain that the binding free energy at position P_2 is reproducible despite the magnitude of the error, an independent FEP calculation at this position was performed starting from a different initial state. The second simulation confirmed reproducibility of the binding free energy at position P_2, as shown in Figure 4.

Previous experiments showed that structured (duplex) RNA can be unwound and degraded by the exosome, as long as the RNA has a 3′ ssRNA overhang that can enter the interior channel of the exosome.[16] In order to pull the duplex RNA in and cleave it, the exosome can unwind the duplex. The stabilization conferred to the RNA through duplex formation and through binding to the interior channel of the exosome can be easily obtained. The absolute free energy for creation/annihilation of the 3′-end nucleotide in a duplex is determined in an additional FEP calculation in which a terminal nucleotide of a solvated duplex is transformed into a terminal hydrogen and vice versa. The relative free energy of stabilization of a 3′-end nucleotide in a duplex structure is then determined by a similar procedure as reported in Figure S1 and Table S1, by subtracting the absolute free energy of the reference state, in which a 3′-end nucleotide is tethered to ssRNA in solvent. Figure 4 reveals that a single terminal nucleotide in a duplex is more stable by approximately −3 kcal/mol than when in ssRNA. Therefore, stabilization of a terminal nucleotide in a duplex is modest compared to the stabilization provided through binding to the exonuclease pathway, which provides, as already stated above, between −5 and −19 kcal/mol stabilization at positions P_1 to P_9 within the exosome (as labeled in Figure 1c).

2.3. Translocation of Long RNAs into Active Site Tunnel of Rrp44 Is Thermodynamically Favored. Exosome function involves continuous translocation of the whole RNA toward the cleavage site. To examine the energetics of translocation of the whole RNA, the approach shown for a representative example in Figure 5a was used. In this approach, free energy differences between sets of pre-translocation and post-translocation states of the system are obtained by assuming that translocation of an RNA segment of a fixed length is equivalent to the process of "deleting" the 5′-end RNA nucleotide while "creating" the 3′-end RNA nucleotide. Accordingly, for the poly(U_5) system shown schematically in Figure 5a, the relative free energy of the state labeled as (P_2,U_5), i.e., a state where a 3′-end nucleotide of U_5 RNA is in position P_2, can be obtained as

$$\Delta\Delta G(P_2, U_5) = \Delta\Delta G(P_1, U_5) + \Delta G_{transfer}(P_6) - \Delta G_{transfer}(P_1) \quad (1)$$

where $\Delta\Delta G$ (P_2,U_5) is the relative free energy of the system in state (P_2,U_5), defined with respect to a reference, $\Delta\Delta G(P_1,U_5)$, the free energy of the system in state (P_1,U_5), which is arbitrarily set to zero. In eq 1, $\Delta G_{transfer}(P_1)$ and $\Delta G_{transfer}(P_6)$ are transfer free energies of single RNA nucleotides from water into the active site tunnel positions P_1 and P_6, respectively.

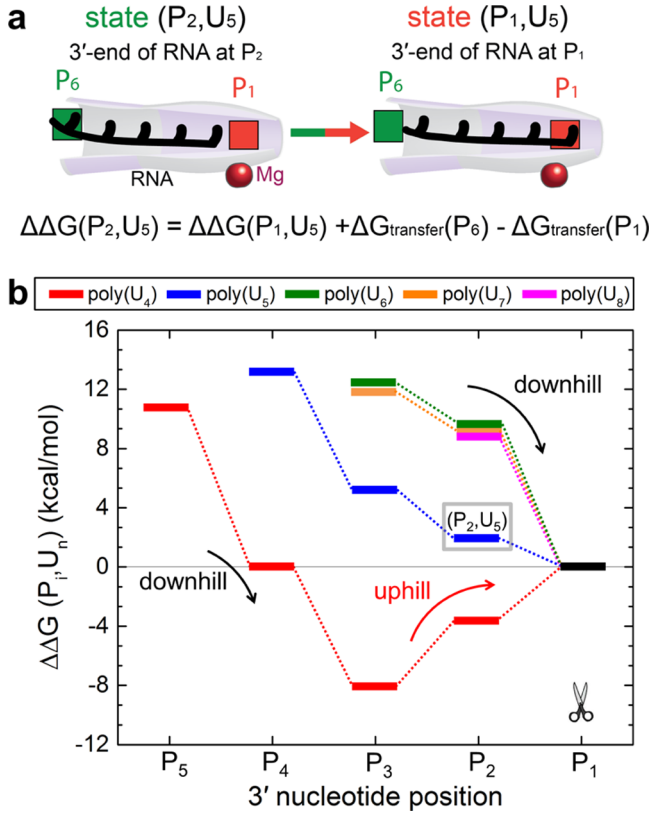


Figure 5. Translocation of RNA segments in the exonuclease active site tunnel of the exosome. (a) Scheme of the simplified left-to-right translocation process and the strategy to obtain free energy profiles for RNA translocation. The scheme shows a representative example for translocation of U_5 RNA, in which its 3′-end moves from position P_2 to position P_1. Pre- and post-translocation states are labeled as states (P_i, U_n), where i is position of the 3′-end of RNA, and n is the number of uracil nucleotides in RNA examined. Relative free energy of the system on the left, $\Delta\Delta G$ (P_2,U5), is calculated with respect to relative free energy of the system on the right, $\Delta\Delta G$ (P_1,U_5), according to eq 1. (b) Free energy profiles for translocation of RNA segments in the exonuclease active site tunnel of Rrp44. The RNA segment is a poly-U strand, 4−8 nucleotides in length. $\Delta\Delta G$ (P_i,U_n) profiles are derived from data in Figure 4, as defined in eq 2. Values of $\Delta\Delta G$ (P_1,U_n) are arbitrarily set to zero for all n examined (n = 4−8). The value of $\Delta\Delta G$ (P_2,U_5), a representative example in panel a, is framed in a gray rectangle.

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

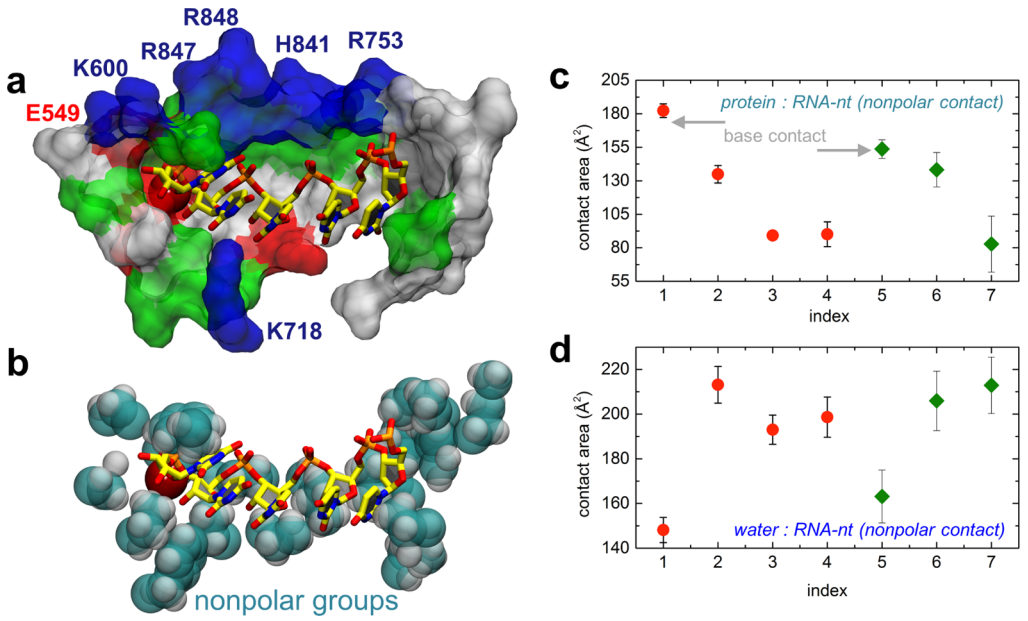


Figure 6. Interactions of RNA with its protein and water environment. (a) The exonuclease active site tunnel in Rrp44, colored according to residue type: nonpolar, white; polar, green; basic, blue; acidic, red. RNA atoms are shown in yellow (C), blue (N), red (O), and orange (P). (b) Nonpolar protein groups (side chains and backbone) of residues within 4 Å of the nonpolar groups of RNA (colored the same as in panel a). Mg^{2+} ions are in dark red. (c) Contact areas between single terminal RNA nucleotides (nonpolar and base atoms) and the nonpolar atoms of the protein environment. (d) Contact areas between single terminal RNA nucleotides (nonpolar and base atoms) and water. The values shown in panels (c,d) are computed as described in Methods and are averaged over 15 ns of trajectories. Red data points mark 3′-end nucleotides, and green data points mark 5′-end nucleotides. Nonpolar atoms of nucleotides are defined in Figure S6.

Transfer free energies of single nucleotides take into account the presence of tethered longer ssRNA strands as shown in the scheme of Figure 5a. These transfer free energies as well as others reported in Figure 4 provide the corresponding free energy changes for "deletion" and "creation" of single nucleotides at the two ends of an ssRNA strand in the Rrp44 active site.

In general, relative free energies for all system states, $\Delta\Delta G(P_i,U_n)$, are obtained by an approach analogous to the approach in eq 1:

$$\Delta\Delta G(P_i, U_n) = \Delta\Delta G(P_{i-1}, U_n) + \Delta G_{transfer}(P_{i-1+n}) - \Delta G_{transfer}(P_{i-1}); \Delta\Delta G(P_1, U_n) = 0 \quad (2)$$

where i is the position of the 3′-end RNA nucleotide within the active site tunnel, n is the number of uracil nucleotides in the RNA segment, and $\Delta G_{transfer}$ a transfer free energy of single RNA nucleotides from water to different positions within Rrp44 exonuclease, reported in Figure 4. All the relative free energies are defined with respect to $\Delta\Delta G(P_1,U_n)$ values, which are arbitrarily set to zero for all RNA strands examined (n = 4–8).

Free energy profiles $\Delta\Delta G(P_i,U_n)$ for the translocation of the studied RNA segments are gathered in Figure 5b as a function of the position of the 3′-end RNA nucleotide. As the free energy profiles are all shown with respect to arbitrary reference points ($\Delta\Delta G(P_1,U_n)$ = 0 kcal/mol), only the free energy differences, rather than absolute values, represent meaningful quantities in the subsequent analysis. The plot in Figure 5b shows that initial entries of U_5, U_6, and U_7 RNAs into the active site tunnel of Rrp44, prior to any cleavage reactions, i.e., for (P_2,U_n), (P_3,U_n), (P_4,U_n), or (P_5,U_n) where n = 5, 6, 7, are accompanied by downhill free energy profiles. Therefore, initial entries of these RNAs are energetically favorable. Furthermore, 5–8-nt RNA segments always have last translocation steps that are accompanied by downhill free energy profiles (step 4 of the cleavage cycle shown in Figure 2). These steps are thus also energetically favorable. However, the extent of the free energy gain in the last translocation step is length dependent, i.e., while translocation of the longer U_6, U_7, and U_8 RNA segments is strongly favorable (≈ −9 to −10 kcal/mol), translocation of the U_5 segment is only modestly so (by approximately −2 kcal/mol).

Interestingly, the free energy profile for translocation of the shortest piece of RNA examined in Figure 5b, U_4, exhibits a different trend. A downhill free energy profile is observed for U_4 entering the active site tunnel to the position where its 3′-end is at position P_3. After that point, the profile changes slope, and an uphill free energy profile is observed for translocation of U_4 toward the position of cleavage, estimated at ≈8 kcal/mol. This uphill free energy profile indicates that translocation of U_4 RNA into position for cleavage is not energetically favorable.

2.4. Design of the Exonuclease Active Site Tunnel.

The strength of the interaction of the active site tunnel with individual RNA nucleotides, quantified through single nucleotide transfer free energies in Figure 4, decreases in the following order: $P_5 > P_4 > P_1 > P_3 \approx P_6 > P_2$. To identify the factors that contribute to large differences in these interactions (and transfer free energies), we next examine the local environment of the active site.

Figure 6a shows that the design of the active site tunnel of Rrp44 is complementary to that of the RNA strand. Indeed, the protein tunnel surface that coordinates phosphate groups of RNA is made of a patch of positively charged residues. Furthermore, a distinct patch of hydrophobic residues lines the tunnel surface that faces the bases and nonpolar (alkyl) groups of RNA. In addition, the nonpolar groups of other polar

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069–4078

residues also contribute to the presence of the extended nonpolar patch in the protein tunnel (Figure 6b).

Electrostatics plays a key role in RNA-protein interactions. Therefore, this contribution was further examined. Figure 7b depicts the electrostatic potential in the active site pocket of Rrp44, projected onto the RNA surface, which includes contributions of protein residues, Mg^{2+}, and RNA (calculated in implicit solvent). Figure 7c shows the electrostatic potential that Rrp44 and Mg^{2+} create in the active site tunnel where RNA is located; the potential is projected on the RNA surface. Rrp44 and Mg^{2+} create a positive electrostatic potential at the position occupied by the phosphate groups of nucleotides U_1, U_3, and U_4 (Figure 7c). This positive potential is still visible in Figure 7b, even though negative potential of phosphate groups counterbalances the positive Rrp44 and Mg^{2+} contributions. Interestingly, the protein also creates a negative electrostatic potential at the phosphate group of nucleotide U_2 as well as at the 3′-end of nucleotide U_1. Therefore, trends in the electrostatic potential that RNA experiences in the Rrp44 active site tunnel are well correlated with trends in transfer free energies reported in Figure 4. Namely, the negatively charged phosphate groups of RNA are electrostatically stabilized at two locations in the active site tunnel, corresponding to very favorable transfer free energies, which are separated by a site of a weaker stabilization, P_2, corresponding to a more modest transfer free energy.

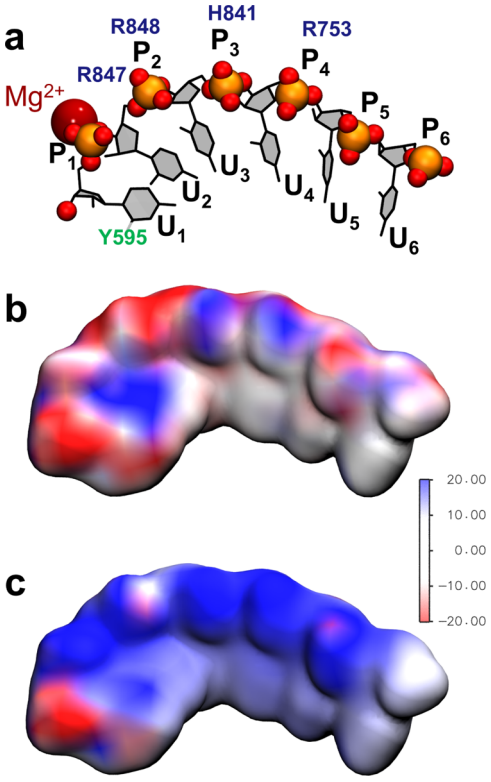


Figure 7. Electrostatic potential within the active site tunnel of Rrp44. (a) Conformation of the RNA segment in the active site tunnel of Rrp44. The color scheme of RNA and Mg^{2+} is as in Figure 1 (c). Labels of key amino acids are placed at their approximate positions with respect to RNA. (b) Electrostatic potential within the active site pocket of Rrp44, projected on the surface of the RNA molecule; the potential shown includes the contributions of protein residues, Mg^{2+}, and RNA (calculated in implicit solvent). (c) Electrostatic potential within the active site pocket of Rrp44, projected onto the surface of the RNA molecule; the potential shown here includes the contributions of only protein residues and Mg^{2+} (calculated in implicit solvent). The electrostatic potential was evaluated by means of adaptive Poisson−Boltzmann solver (APBS) implicit solvent calculations within VMD.[24] The orientation of RNA in panels (b,c) is identical to that in panel (a). The electrostatic potential scale bar is shown in units of k_BT/e, where k_B is Boltzmann constant, T is temperature, and e is the charge of an electron.

Transfer free energies of RNA nucleotides should also be influenced by hydrophobic interactions and solvation effects. For example, mutual prevention of water exposure of the hydrophobic groups of RNA nucleotides and the protein surface should have a favorable contribution to the transfer free energies. Figure 6c plots contact areas formed between nonpolar groups of the protein and the RNA nucleotides. Of all the nucleotides, those at positions P_1 and P_5 cover the largest hydrophobic areas of the protein surface; both of these nucleotides have bases that are in direct contact with hydrophobic groups on the protein surface. Figure 6d shows the average contact area between nonpolar groups of RNA and water. The nucleotides at positions P_1 (3′-end) and P_5 (5′-end) also have nonpolar groups that are least exposed to water. On the other hand, the 3′-end nucleotide at position P_2 has a significant exposure to water.

2.5. Dynamics of RNA Translocation in Rrp44 on a Representative Pathway. The free energy profiles for RNA translocation in Figure 5b confirm that ssRNA segments longer than 4-nt favor translocation toward the cleavage site, whereas 4-nt ssRNA segments are unlikely to translocate toward the cleavage site. However, this plot does not provide the complete free energy landscape that underlies translocation, lacking an estimate of a barrier for the translocation process. In addition, the simulations that provided the results in Figure 5b examined only the states of the system that represent local minima of the free energy surface and did not directly probe the translocation process.

To examine the process of ssRNA translocation, we simulated a U_6 segment at different positions along the pathway in the Rrp44 active site tunnel, so that its 3′-end spans the space between positions P_3 and P_2. This process corresponds to one translocation step during the initial entry of RNA into the active site pocket, rather than to the final translocation step prior to cleavage; the simulation performed was designed to avoid difficulties in sampling Mg^{2+}-RNA interactions that accompany the final translocation step prior to cleavage. The initial and final states of the simulated translocation reaction are displayed in Figure 8b,c. In both states, Mg^{2+} is coordinated by Rrp44 residues D552, D543 as well as by the solvent. Translocation of the U_6 segment by one position involves a relatively small displacement along a curved cylindrical tunnel (≈4 Å between center of mass coordinates of ssRNA molecules in the initial and final states of the system), nevertheless breaking and establishing of several strong adhesion contacts between RNA phosphate groups and basic residues of the protein.

Our initial simulations showed that RNA translocation cannot be properly described with a simple one-dimensional reaction coordinate model, such as a distance or a projected distance; the sampling of strong and slowly exchanging contacts is limited in such simulations. However, the transition path sampling approach can be used to describe complex transition paths,[23,25−27] such as the RNA translocation step. Given the small net displacement of RNA during translocation, we

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

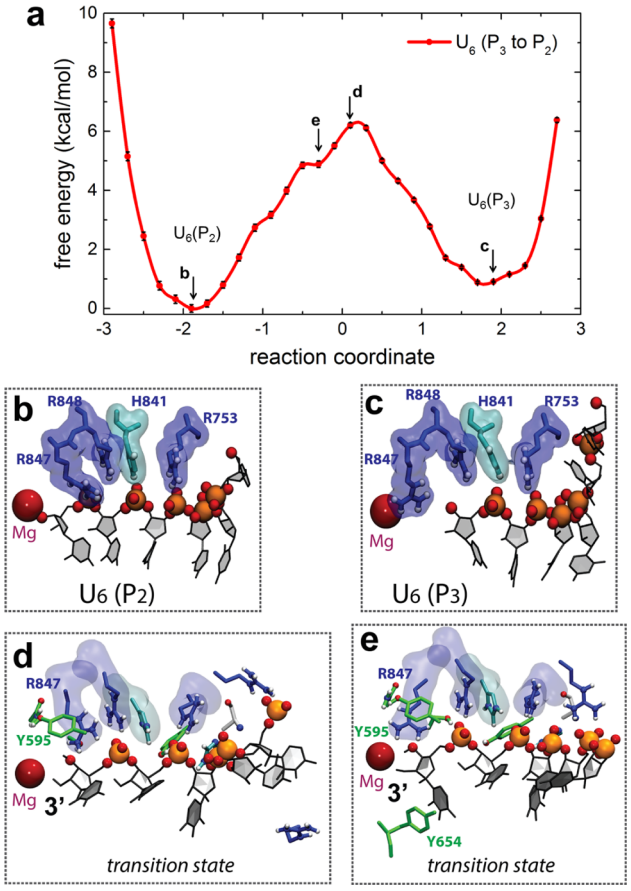


Figure 8. Translocation of poly(U_6) ssRNA segment in the active site tunnel of Rrp44. (a) Free energy profile of poly(U_6) ssRNA segment translocating, with the 3′-end translocating between positions P_3 and P_2. (b) Configuration of RNA in the active site tunnel in the U_6 (P_2) state. (c) Configuration of RNA in the active site tunnel in the U_6 (P_3) state. (d,e) System configurations observed during translocation.

adopted a simplified version of transition path sampling. For this purpose, we defined the reaction coordinate model $\chi = r_2 - r_1$, where r_1 is the RMSD between the current state of the system and the initial state (ssRNA with its 3′-end at P_3), and r_2 is the RMSD between the current state of the system and the final state (ssRNA with its 3′-end at P_2), as further described in Methods. The selected reaction coordinate model based on RMSDs should be a reasonable choice for the studied reaction, as it involves a small net displacement of RNA during translocation. Indeed, such RMSD-based transition coordinates have been used successfully in studies of other complex systems.[28,29]

The free energy profile along the selected reaction coordinate is obtained by replica-exchange MD coupled to umbrella sampling (REMD-US) windows, as described in Methods. REMD-US trajectories describe the transition of the simulated system between the defined initial and final states, as indicated in Figure S4. To confirm that this transition actually accounts for the translocation reaction, the behavior of REMD-US trajectories was examined when projected onto a different RMSD-based collective variable, defined in Figure S5a. This RMSD collective variable includes atom coordinates of the whole RNA and its protein environment and offers a better representation of the translocation reaction. Figure S5 confirms that the transition in REMD-US trajectories indeed coincides with the translocation of the U_6 RNA segment in the active site tunnel of Rrp44, as reflected in the overlap of consecutive histograms in the pairwise RMSD plot shown in Figure S5.

The free energy profile for the system evolving along the aforementioned RMSD-based reaction coordinate is shown in Figure 8a. This profile features two local minima, corresponding to the initial and final states of the transition, which are separated by a free energy barrier of ≈6 kcal/mol. The difference in free energies between the initial and final states is ≈1 kcal/mol, which is less than that determined in Figure 5 (≈ 4 kcal/mol). However, the difference observed is within error bounds; an error of 3.4 kcal/mol is reported for the binding free energy of a 3′-end nucleotide at position P_2 (Figure 4 and Table S1). In the initial state of the system, the two phosphate groups that are on the 3′-end side of the ssRNA segment are coordinated by H841 and R753, as seen in Figure 8c. On the other hand, in the final state of the system, these phosphate groups are coordinated by R848 and H841, and R753 coordinates another phosphate group that shifted to occupy position P_4 in the Rrp44 active site tunnel (Figure 8b).

In the states characterized by the barrier estimated in Figure 8a, the system exchanges contacts of RNA phosphate groups with basic protein residues. This exchange of contacts is gradual. As the system moves away from the initial state, R848 starts interacting with the phosphate group of the 3′-end RNA nucleotide and thus displaces H841 that previously coordinated the same phosphate group. As a result, H841 loses its partner phosphate and is found between the R848 and R753 residues, both of which chelate two neighboring phosphate groups. Eventually, in the states where RNA is further shifted toward the final state, H841 starts coordinating the phosphate group that is second in line from the 3′-end of RNA, while R753 begins to coordinate a new (incoming) phosphate group of the shifted RNA.

In Figure 8d,e, two characteristic transition-state configurations are depicted. These configurations indicate that residue Y654 can form characteristic stacking interactions with the 3′-end residue of translocating RNA and potentially stabilize the translocation process. However, such a role could not have been anticipated from the structure of the precleavage state, in which RNA occupies the full active site tunnel. In the precleavage state, the hydroxyl group of Y654 forms a hydrogen bond with the sugar backbone of RNA and stabilizes the RNA binding pose.

3. DISCUSSION AND CONCLUSION

In the present study, we have examined in atomic detail the processive RNA translocation in the exosome complex. MD simulations and free energy calculations were performed to examine the driving forces responsible for the translocation process, which involves exchange of tightly bound contacts in a curved internal channel of the exosome. To overcome the difficulties arising in the simulation of the complex translocation process, we used a combination of computational strategies, including FEP calculations and replica-exchange umbrella sampling simulations.[30,31]

The performed simulations quantify the driving forces at play for the entry of RNA into the active site tunnel of Rrp44 exonuclease within the exosome, where the interaction of the nucleotides with their environment is particularly strong. For example, transfer of single RNA nucleotides from aqueous solution to the entrance of the active site tunnel is favored by −7 kcal/mol (position P_7), −14 kcal/mol (P_6), or up to −19

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

kcal/mol (P_5) (Figure 4). Previous experimental studies pointed to the relevance of these positions of strong interaction, as RNA substrates are able to enter and get degraded by the exosome, only if they have 3′-end ssRNA overhangs that can reach positions P_6 and P_7 in the internal channel of the exosome.[16]

As hydrolytic cleavage of RNA is a highly exothermic process, we were intrigued by the possibility that the energy released in the cleavage reaction powers the functional cycle of the exosome, which involves processive translocation of RNA (Figure 2). Our simulations show that such scenario may not be necessarily true. The design of the active site tunnel of the exosome is such that translocation of RNA molecules longer than 4-nt into position for cleavage is thermodynamically favorable. On the other hand, translocation of 4-nt-long RNA molecules into the position for cleavage is not thermodynamically favorable, in agreement with the experimental results.[15,16,18]

The present simulations characterized the design of the exonuclease active site tunnel and the consequences of this design on RNA translocation. The active site tunnel has RNA binding sites of variable binding strengths, including two strong binding sites (P_1 and P_4−P_5), two significantly weaker binding sites (P_2 and P_7−P_8), and binding sites of intermediate strength (P_3 and P_6).

Two strong RNA binding sites contribute to the active site tunnel having a strong grip on RNA. These sites also lead to discrimination in length of RNA substrates. The strong RNA binding sites are spatially separated, at positions P_1 (the catalytic center) and at positions P_4 and P_5. RNA substrates that are 4-nt long cannot span both strong binding sites and therefore cannot simultaneously bind to both sites. In fact, simulations show that 4-nt long RNA thermodynamically favors binding to P_4 and P_5. The existence of a strong RNA binding site that is distinct from the cleavage site (P_1) was previously hypothesized for a processive exonuclease RNase R,[32] a close relative of the exosome's Rrp44 exonuclease. Our results in Figures 4 and 5 confirm the existence of a strong binding site at positions P_4 and P_5 in Rrp44.

The finding that position P_2 is the least favorable binding spot for RNA within the active site tunnel has two important consequences. The first consequence is that once the 3′-end nucleotide of an RNA segment is already at position P_2, it has a large local thermodynamic incentive to move forward into the much more favorable position P_1, potentially providing a local pull on RNA. The second consequence is that short (4-nt) ssRNA molecules thermodynamically favor binding to the active site tunnel two translocation steps away from the position for cleavage (with the 3′-end located at P_3). We would like to note that the simulated structure contains only one Mg^{2+} ion in the catalytic center (P_1), whereas two Mg^{2+} ions are present in the actual catalytic center of Rrp44. While our results for RNA binding strength at P_1 could change when two Mg^{2+} ions are at P_1, RNA at position P_2 is not coordinated by Mg^{2+}, and P_2 is therefore likely to remain the weakest RNA binding site.

Our simulations also examined the driving forces behind the ability of the exosome to unwind structured RNAs,[13,16] once the RNA substrate initially enters the exosome. As the RNA spans the internal channel of the exosome down to position P_2, unwinding and translocation of RNA toward the cleavage position P_1 should be favored as long as the 3′-end nucleotide at P_1 is in a thermodynamically more favorable state than the incoming RNA nucleotide that is about to enter the exosome channel. The simulations confirm that RNA binding to P_1 is significantly more favored thermodynamically over RNA staying in a typical structured RNA form (duplex), the effect of which can provide the driving force for RNA duplex unwinding by the exosome, in agreement with the experiments.[13,16]

A good correlation between the strength of Coulombic interactions and binding free energies of RNA to protein is observed for the simulated system. At the same time, solvation effects also seem to influence the RNA-protein binding strength. In the future, it will be of interest to examine if Coulombic interactions are usually the strongest determinants of RNA binding and RNA processing for other classes of nucleic acid-interacting proteins.

4. METHODS

In the present paper, we examined the interaction of RNA with the key part of the exosome complex, namely the Rrp44 exonuclease, which acts as an enzyme for RNA cleavage and a motor that powers RNA translocation and RNA unwinding. Equilibrium MD simulations were performed to relax the exosome complex and obtain a structure of the reduced system investigated in detail in further simulations. To understand RNA interaction with the active site tunnel of Rrp44, binding of individual RNA nucleotides to different positions along the tunnel was characterized. Behavior of longer RNA segments was examined to describe translocation of RNA toward the cleavage position at the end of the active site tunnel. Simulating translocation of a whole RNA segment was found to be challenging due to the curved geometry of the tunnel and the presence of multiple strong interactions of RNA with the amino acids that line the interior of the channel. To solve this challenging problem, we employed umbrella sampling coupled to replica exchange MD.

4.1. Atomic models. To examine the factors that contribute to RNA translocation and cleavage within the exosome complex, simulations were carried out on extracted portions of the crystal structure of an RNA-bound 11-subunit exosome complex of *Saccharomyces cerevisiae* (yeast) (pdb id: 4IFD), partially shown in Figure 1. MolProbity[33] assessment of this structure indicates an excellent geometry with 95.8% in favored and 99.9% in allowed regions of Ramachandran space, as shown in Figure S7. Furthermore, Clashcore and MolProbity scores are in the 100th percentile. Missing residues of the complex were added as flexible loops with MODELER.[34] All systems prepared were neutralized with the VMD plugin *cionize*.[24] The resulting structures were solvated in TIP3P water and ionized in 0.15 M NaCl with the *solvate* and *ionize* VMD plugins, respectively.[24] In most of the performed simulations, the systems contained a complete Rrp44 enzyme and parts of the exosome core proteins bound to Rrp44, with poly(U) ssRNA pieces placed at different positions along the RNA pathway inside the exosome. An independent duplicate calculation of binding free energy of a 3′-end nucleotide at position P_2 of Rrp44 was performed by simulating a smaller system containing only ssRNA and RNB and S1 domains of Rrp44.

Molecular Dynamics Simulations. The MD simulations were performed with NAMD2.9,[35] where the systems were described with the Amber force field with SB[36] and BSC0[37] corrections, a suitable choice for describing nucleic acids.[38,39] The particle mesh Ewald (PME) method[40] was used for the evaluation of long-range Coulomb interactions. The time step was set to 2.0 fs; all bonds involving hydrogen were constrained with SHAKE.[41] All simulations were performed in the isobaric-isothermal ensemble, at a constant temperature of 310 K, with a friction constant of 1.0 ps^{-1}, and at a constant pressure of 1 bar. The equations of motion were integrated with the r-RESPA multiple time-step propagator with an effective time step of 2 and 4 fs for short- and long-range interactions. The MD simulations involved typically altogether (protein, RNA, solvent) ~629,000 atoms, for the complete model of the exosome complex,

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

~197,200 atoms, for the first type of the reduced system, or ~50,900 atoms, for the second type of the reduced system; two types of reduced systems are described in the paragraph below.

All systems prepared were minimized for 2000 steps. Then, ions and water molecules were equilibrated for 2 ns around each of the prepared protein-RNA systems, which were restrained using harmonic forces with a spring constant of 1 kcal/(molÅ^2). The complete model of the exosome complex was simulated for 20 ns, while relaxing the added missing parts, keeping the protein-RNA structure restrained. After relaxation of the added missing parts, the complete model of the exosome was simulated for 24 ns, while keeping α-helices and β-sheets of core proteins restrained. The relaxed model of the exosome complex was used to extract two types of reduced systems investigated in further simulations. The first type contained the complete Rrp44 bound to ssRNA and attached to a part of the core proteins of the exosome; the second type included only RNB and S_1 domains of Rrp44 bound to ssRNA. After minimization followed by ion and water equilibration, the reduced systems were further relaxed for several nanoseconds prior to subsequent free energy calculations, as described below.

4.2. Free Energy Perturbation Calculations. FEP calculations were performed to obtain the transfer free energies of single RNA nucleotides from aqueous solution to eight different binding sites of the exosome complex. The structures of the exosome bound to ssRNA pieces in different positions were based on relaxed initial systems and prepared by removing RNA nucleotides from unwanted positions, along with a corresponding number of counterions. The transfer free energies were obtained for 3′-end or 5′-end single RNA nucleotides (covalently bound to longer pieces of ssRNA) in eight different positions, according to the thermodynamic cycle shown in Figure S1. The alchemical transformations for all single nucleotides were carried out bidirectionally (by changing to and from nothing) and used 80 λ-windows of even width, with 250,000 steps of equilibration followed by 500,000 steps of data collection. In total, each single nucleotide simulation (backward and forward) lasted 240 ns. In the FEP calculations, the single nucleotide was unrestrained, whereas the position of the P atoms of the RNA phosphate groups of the remaining nucleotides was restrained with the force constant k = 1 kcal/(mol A^2). The probability distribution functions for the forward and backward transformations exhibited a good overlap, as shown in a representative example in Figure S2, suggestive of an appropriate convergence of the free energy calculations.

For comparison and in order to obtain transfer free energies of nucleotides from water to protein environments, absolute solvation free energies of single 3′-end and 5′-end RNA nucleotides in water were obtained in separate simulations. The ParseFEP plugin in VMD[42] was used to calculate the free energy changes from the FEP simulations by means of the Bennett acceptance ratio (BAR) algorithm. As all the FEP simulations involved a creation or annihilation of charged nucleotides, the computed free energies were corrected,[43] as detailed in Table S1.

The results of all individual FEP simulations performed, together with the errors associated with the BAR free energies and hystereses for the backward and forward FEP transformations, are reported in Table S1.

4.3. Potential of Mean Force Calculations. To validate the obtained translocation free energies and gain insight into the translocation dynamics, we performed a test simulation of 6-nt ssRNA translocating through the exonuclease active site of Rrp44. The potential of mean force (PMF) was obtained by REMD-US simulations (replica exchange molecular dynamics-umbrella sampling) to increase the sampling efficiency;[30,31,44] REMD methods have been essential in studies of RNA systems.[45,46] The single translocation step was partitioned into 60 windows, where the reaction coordinate was defined as $\chi = r_2 - r_1$, where r_1 is the RMSD between the current state and the initial state, and r_2 is the RMSD between the current state and the final state. Confinement potentials were introduced in the form of harmonic restraints with a force constant k = 50 kcal/(mol Å^2). Exchanges between neighboring windows were attempted every 200 time steps and were accepted or rejected according to a Metropolis energy criterion. Each window was run for 20 ns, which amounts to a total time of 1200 ns. The underlying simulations involved altogether (protein, RNA, solvent) ~142,800 atoms. After sorting of the resulting trajectories into specified χ windows, the weighted histogram analysis method (WHAM) was used to reconstruct the PMF.[47,48] Monte Carlo bootstrap error analysis was performed also with the WHAM algorithm (with num_MC_trials set to 10). The histograms of the US windows used to reconstruct the PMF overlap appropriately, as can be seen in Figure S3.

4.4. Calculation of Contact Areas. To analyze contact areas of single RNA nucleotides with their protein (Figure 6c) or water environment (Figure 6d), we computed and averaged time-dependent contact area, defined as

$$a_{\mathrm{con}}(t) = \frac{a_{\mathrm{RNA\text{-}nt}}(t) + a_{\mathrm{env}}(t) - a_{\mathrm{RNA\text{-}nt:env}}(t)}{2} \quad (3)$$

where $a_{\mathrm{RNA\text{-}nt}}(t)$ is the solvent accessible surface area (SASA) of nonpolar atoms of a single RNA nucleotide (defined for uracil in Figure S6), $a_{\mathrm{env}}(t)$ is the SASA of the nucleotide environment (either nonpolar atoms of the Rrp44 protein or water), and $a_{\mathrm{RNA\text{-}nt:env}}(t)$ is the SASA of combined nonpolar atoms of the single RNA nucleotide and the environment (as above, either nonpolar atoms of protein or water). The computation was performed for systems containing poly(U) RNA segments in the active site tunnel of Rrp44. Values reported for single nucleotides 1−4 correspond to 3′-end nucleotides of a poly(U_6) ssRNA, and values reported for nucleotides 5−7 correspond to 5′-end nucleotides of a poly(U_5) (for index 5) or poly(U_6) ssRNAs. The evaluation was performed with the built-in SASA VMD plugin,[24] where a van der Waals radius of 1.4 Å was assigned to atoms to identify the points on a sphere that are accessible to the solvent.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b12065.

A representative thermodynamic cycle used to determine transfer free energies of single RNA nucleotides; probability distribution plots and free energy change as a function of λ-window value in a representative FEP calculation; histogram distributions from REMD-US simulations; REMD-US trajectories projected onto RMSD-based collective variables forming the reaction coordinate; REMD-US trajectories projected onto extended RMSD-based collective variables, including atom coordinates of the whole RNA and its protein environment; definition of nonpolar atoms of uracil nucleotide in SASA calculations; MolProbity Ramachandran analysis of the RNA exosome structure simulated; and a table with detailed results of FEP calculations (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*kschulte@ks.uiuc.edu

*Lvukovic@utep.edu

Present Address

○CRELUX GmbH, Martinsried 82152, Germany

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Wei Han for useful discussions and suggestions. This work was supported by National Science

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078

Foundation (NSF) grants PHY0822613 and PHY1430124 as well as by National Institutes of Health (NIH) grant 9P41GM104601. The authors gratefully acknowledge computer time provided by the NSF-funded Extreme Science and Engineering Discovery Environment (XSEDE) MCA93S028, Rechenzentrum Garching, and the Computational Science and Engineering Program at the University of Illinois (Taub computing cluster). L.V. acknowledges support as a Center for the Physics of Living Cells (CPLC) Postdoctoral Fellow. C.C. gratefully acknowledges the support of the Beckman Institute for Advanced Science and Technology.

REFERENCES

(1) Mattick, J. S. *BioEssays* **2003**, *25*, 930−939.
(2) Doma, M. K.; Parker, R. *Cell* **2007**, *131*, 660−668.
(3) Chen, B.; Retzlaff, M.; Roos, R.; Frydman, J. *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a004374.
(4) Parker, R.; Song, H. *Nat. Struct. Mol. Biol.* **2004**, *11*, 121−127.
(5) Mitchell, P.; Petfalski, E.; Shevchenko, A.; Mann, M.; Tollervey, D. *Cell* **1997**, *91*, 457−466.
(6) Houseley, J.; LaCava, J.; Tollervey, D. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 529−539.
(7) Houseley, J.; Tollervey, D. *Cell* **2009**, *136*, 763−776.
(8) Schneider, C.; Tollervey, D. *Trends Biochem. Sci.* **2013**, *38*, 485−493.
(9) Mitchell, P.; Tollervey, D. *Nat. Struct. Biol.* **2000**, *7*, 843−846.
(10) Makino, D. L.; Halbach, F.; Conti, E. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 654−660.
(11) Schaeffer, D.; Clark, A.; Klauer, A.; Tsanova, B.; van Hoof, A. In *Advances in Experimental Medicine and Biology*; Jensen, T., Ed.; Springer: New York, 2010; Vol. *702*, pp 79−90.
(12) Schuch, B.; Feigenbutz, M.; Makino, D. L.; Falk, S.; Basquin, C.; Mitchell, P.; Conti, E. *EMBO J.* **2014**, *33*, 2829−2846.
(13) Lee, G.; Bratkowski, M. A.; Ding, F.; Ke, A.; Ha, T. *Science* **2012**, *336*, 1726−1729.
(14) Wasmuth, E.; Lima, C. *Mol. Cell* **2012**, *48*, 133−144.
(15) Dziembowski, A.; Lorentzen, E.; Conti, E.; Seraphin, B. *Nat. Struct. Mol. Biol.* **2007**, *14*, 15−22.
(16) Makino, D. L.; Baumgartner, M.; Conti, E. *Nature* **2013**, *495*, 70−75.
(17) Makino, D. L.; Schuch, B.; Stegmann, E.; Baumgartner, M.; Basquin, C.; Conti, E. *Nature* **2015**, *524*, 54−58.
(18) Bonneau, F.; Basquin, J.; Ebert, J.; Lorentzen, E.; Conti, E. *Cell* **2009**, *139*, 547−559.
(19) Wasmuth, E. V.; Januszyk, K.; Lima, C. D. *Nature* **2014**, *511*, 435−439.
(20) Liu, Q.; Greimann, J. C.; Lima, C. D. *Cell* **2006**, *127*, 1223−1237.
(21) Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton*; Sinauer Associates: Sunderland, MA, 2001.
(22) Vale, R. D.; Milligan, R. A. *Science* **2000**, *288*, 88−95.
(23) Ma, W.; Schulten, K. *J. Am. Chem. Soc.* **2015**, *137*, 3031−3040.
(24) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(25) Pan, A. C.; Sezer, D.; Roux, B. *J. Phys. Chem. B* **2008**, *112*, 3432−3440.
(26) Shukla, D.; Meng, Y.; Roux, B.; Pande, V. S. *Nat. Commun.* **2014**, *5*, 1−11.
(27) Radhakrishnan, R.; Schlick, T. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5970−5975.
(28) Ito, Y.; Oroguchi, T.; Ikeguchi, M. *J. Am. Chem. Soc.* **2011**, *133*, 3372−3380.
(29) Noy, A.; Perez, A.; Laughton, C. A.; Orozco, M. *Nucleic Acids Res.* **2007**, *35*, 3330−3338.
(30) Jiang, W.; Phillips, J. C.; Huang, L.; Fajer, M.; Meng, Y.; Gumbart, J. C.; Luo, Y.; Schulten, K.; Roux, B. *Comput. Phys. Commun.* **2014**, *185*, 908−916.
(31) Gumbart, J. C.; Roux, B.; Chipot, C. *J. Chem. Theory Comput.* **2013**, *9*, 794−802.
(32) Vincent, H. A.; Deutscher, M. P. *J. Biol. Chem.* **2006**, *281*, 29769−29775.
(33) Chen, V. B.; Arendall, W. B., III; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.
(34) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779−815.
(35) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(36) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.
(37) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(38) Besseova, I.; Otyepka, M.; Reblova, K.; Sponer, J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10701−10711.
(39) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T.; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Sponer, J. *Nucleic Acids Res.* **2010**, *38*, 299−313.
(40) Darden, T.; York, D.; Pedersen, Y. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(41) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(42) Liu, P.; Dehez, F.; Cai, W.; Chipot, C. *J. Chem. Theory Comput.* **2012**, *8*, 2606−2616.
(43) Gumbart, J.; Roux, B. *Biophys. J.* **2012**, *102*, 795−801.
(44) Hazel, A.; Chipot, C.; Gumbart, J. C. *J. Chem. Theory Comput.* **2014**, *10*, 2836−2844.
(45) Garcia, A. E.; Paschek, D. *J. Am. Chem. Soc.* **2008**, *130*, 815−817.
(46) Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Swails, J. M.; Roitberg, A. E.; Cheatham, T. E. *J. Chem. Theory Comput.* **2014**, *10*, 492−499.
(47) Grossfield, A. *WHAM: the weighted histogram analysis method*; University of Rochester Medical Center: Rochester, NY, 2012, accessed February 2016.
(48) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

DOI: 10.1021/jacs.5b12065
J. Am. Chem. Soc. 2016, 138, 4069−4078