Identification of dynamical hinge points of the L1 ligase molecular switch

George M. Giambasu, Tai-Sung Lee, Carlos P. Sosa, et al.

Supplemental Material	http://rnajournal.cshlp.org/content/suppl/2010/02/01/rna.1897810.DC1.html
References	This article cites 50 articles, 20 of which can be accessed free at: http://rnajournal.cshlp.org/content/16/4/769.full.html#ref-list-1
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

To subscribe to *RNA* go to:
http://rnajournal.cshlp.org/subscriptions

Identification of dynamical hinge points of the L1 ligase molecular switch

GEORGE M. GIAMBAŞU,[1,2] TAI-SUNG LEE,[1,2] CARLOS P. SOSA,[3] MICHAEL P. ROBERTSON,[4,5] WILLIAM G. SCOTT,[4,5] and DARRIN M. YORK[2]

[1]Biomedical Informatics and Computational Biology, University of Minnesota, Minneapolis, Minnesota 55455, USA
[2]Department of Chemistry, University of Minnesota at Minneapolis, Minneapolis, Minnesota 55455, USA
[3]IBM and Biomedical Informatics and Computational Biology, University of Minnesota, Rochester, Minnesota 55901, USA
[4]Department of Chemistry and Biochemistry, University of California at Santa Cruz, Santa Cruz, California 95064, USA
[5]Center for the Molecular Biology of RNA, Sinsheimer Laboratories, University of California at Santa Cruz, Santa Cruz, California 95064, USA

ABSTRACT

The L1 ligase is an in vitro selected ribozyme that uses a noncanonically base-paired ligation site to catalyze regioselectively and regiospecifically the 5′ to 3′ phosphodiester bond ligation, a reaction relevant to origin of life hypotheses that invoke an RNA world scenario. The L1 ligase crystal structure revealed two different conformational states that were proposed to represent the active and inactive forms. It remains an open question as to what degree these two conformers persist as stable conformational intermediates in solution, and along what pathway are they able to interconvert. To explore these questions, we have performed a series of molecular dynamics simulations in explicit solvent of the inactive–active conformational switch in L1 ligase. Four simulations were performed departing from both conformers in both the reactant and product states, in addition to a simulation where local unfolding in the active state was induced. From these simulations, along with crystallographic data, a set of four virtual torsion angles that span two evolutionarily conserved and restricted regions were identified as dynamical hinge points in the conformational switch transition. The ligation site visits three distinct states characterized by hydrogen bond patterns that are correlated with the formation of specific contacts that may promote catalysis. The insights gained from these simulations contribute to a more detailed understanding of the coupled catalytic/conformational switch mechanism of L1 ligase that may facilitate the design and engineering of new catalytic riboswitches.

Keywords: RNA world hypothesis; riboswitch; conformational transition

INTRODUCTION

The RNA world hypothesis proposes that life originated from catalytic RNA molecules capable of replication via template-dependent assembly of RNA fragments of nucleotide monomers. To do so, RNA replicases would be required to catalyze the regiospecific condensation of a 3′ ribose hydroxyl upon the 5′ α phosphate of an RNA or nucleotide 5′ triphosphate, to yield the ubiquitous 3′ to 5′ phosphodiester linkage that is characteristic of all extant genomic nucleic acids. However, no known naturally occurring ribozyme catalyzes this phosphodiester assembly reaction. The concern that RNA might be inherently incapable of catalyzing this reaction was put to rest in 1993 with

the first in vitro evolution of a ribozyme ligase (Bartel and Szostak 1993). Subsequently, several other ribozyme ligases have been produced using in vitro selection techniques (Ekland et al. 1995; Jaeger et al. 1999; Robertson and Ellington 1999; Rogers and Joyce 1999; McGinness and Joyce 2003; Ikawa et al. 2004; Bagby et al. 2009; Shechner et al. 2009), including a small subset that specifically catalyzes regiospecific 3′ to 5′ phosphodiester linkages characteristic of all extant RNA and DNA polymerases. The L1 ligase ribozyme (Robertson and Ellington 1999) is one such example, and is unusual in that it uses an intrinsically flexible noncanonically base-paired ligation site (Ekland et al. 1995; Landweber and Pokrovskaya 1999). The crystal structure of the L1 ligase product has been solved recently (Robertson and Scott 2007), providing "a glimpse of biology's first enzyme" (Joyce 2007).

In addition to its potential relevance to the origin of life, the L1 ligase ribozyme, presumably as a fortuitous consequence of in vitro selection, is an allosteric ribozyme

molecular switch (Ellington and Szostak 1990; Marshall and Ellington 1999; Chen et al. 2007). It is possible to take advantage of this property to engineer new L1 ligase ribozymes by further selections that enable derivatives to be controlled by small molecules, peptides, or even proteins; in effect, creating artificial allosteric ribozymes (Robertson and Ellington 1999, 2000, 2001; Robertson et al. 2004).

The intrinsic flexibility of the L1 ligase was verified from the crystallized ligation product in which two crystallographically independent conformations, both products of self-ligation, were resolved in the same asymmetric unit. These conformers differed in the orientation of one of the stems (stem C) by a movement of the stem tip by ~80 Å (Robertson and Scott 2007). Based on the presence/absence of specific contacts between the ligation site and evolutionarily conserved regions of the U_{38} loop of stem C, it was proposed that the conformers represented catalytically active "on" and inactive "off" states (Robertson and Scott 2007).

Here, we address how changes in the L1 ligase conformation affect its function and catalytic activity. Generally, RNA switching and folding pathways occur along a rugged free energy landscape, characterized by multiple minima, low transition barriers, and kinetic traps (Treiber and Williamson 1999; Williamson 2000; Leulliot and Varani 2001; Bokinsky and Zhuang 2005; Xia 2008). The ground states of RNA molecules are thus best characterized by an ensemble of conformers. The factors that can shift the equilibrium between conformers include formation/disruption of tertiary contacts, binding of ligands, metal ions (typically Mg^{2+}), or proteins by induced fit or tertiary structure capture (Treiber and Williamson 1999; Williamson 2000; Leulliot and Varani 2001; Bokinsky and Zhuang 2005; Xia 2008).

The minimal set of eight torsion angles per nucleotide affords flexibility to the polynucleotide backbone, and has been used as the basis for the classification of different RNA folds (Murray et al. 2003; Sims and Kim 2003; Schneider et al. 2004; Richardson et al. 2008). An alternative two-dimensional representation of the RNA backbone conformations can be achieved using a reduced set of so-called *virtual torsion angles* formed along the virtual bonds between $C_{4'}$–P atoms that have been shown to be able to discriminate between major RNA folds (Duarte and Pyle 1998; Wadley et al. 2007).

In the current study, we identify dynamical hinge points of the L1 ligase ribozyme using large scale molecular dynamics simulations. We depart from an analysis of the two crystallized conformers, and show—using over 600 ns of molecular dynamics simulations—that the transition between on and off conformational states can be almost entirely described by changes in only four virtual torsion angles. The insights gained from these simulations are a first step toward a detailed understanding of the coupled catalytic/conformational riboswitch mechanism of L1 ligase.

RESULTS

Four virtual torsions can be used to distinguish between the two crystallized conformers

Restricted parts of the L1 ligase junction differ between the docked and undocked conformers

Figure 1A illustrates the crystallographic presumably active "docked" and inactive "undocked" conformers. Superposition of the three individual stems (A, B, and C), as found in the crystal structure, show small deviations (RMSDs of 1.71, 0.71, and 2.06 Å, respectively) between the docked and undocked conformational states (Table 1). Moreover, the base-pair hydrogen bonding patterns are almost completely conserved between conformers (Table 2); the main differences being located at the terminal GAAA tetraloops at the ends of the stems. Most of the structural variation is observed in the junction region connecting the stems where the RMSD between crystal conformers is 8.30 Å. This variation in the junction propagates to a large scale swing of stem C by ~80 Å. When stems A and B of the docked and undocked conformers are superimposed together with root-mean-square (RMS) fitting, the combined RMSD for stems A and B is 4.9 Å, whereas the RMSD for stem C (based on the same superposition) is 48.4 Å. This indicates that the relative orientations of stems A and B are similar between the two crystallographic conformers, whereas the orientation of stem C relative to stems A and B is vastly different.

Reorientation of stem C is affected mainly by changes in four virtual torsion angles: θ_{18}, θ_{37}, θ_{44}, and η_{38}

The two crystallographic conformers can be distinguished by changes in their virtual torsion angles (Fig. 1A; Duarte and Pyle 1998; Wadley et al. 2007). A survey of the 142 virtual torsion angles in the L1 ligase (two torsions per each of the 71 nucleotides) reveals that only four show a significant change (>45°): θ_{18}, θ_{37}, θ_{44}, and η_{38} (Fig. 1B). These torsions span regions that contain the evolutionarily conserved residues: U_{37}, U_{38}, and A_{39} for θ_{37} and η_{38} and the five-base motif ($C_{39} = G_{18}$ G_{37} A_{38} $C_{17}/U_{39} = A_{18}$ U_{37} G_{38} U_{17}), and the neighboring U_{19} for θ_{18} and θ_{44} (for a detailed statistical analysis of the conserved residues, see Robertson and Ellington 2001). Both U_{38} and U_{19} are of particular interest in the current work: U_{38} contributes to the docking of stem C into stem A in the active conformation; mutation data demonstrate that it is critical for catalysis; and U_{19} is evolutionarily conserved (97%), although its role in catalysis remains unclear (Robertson and Ellington 1999, 2000, 2001; Robertson et al. 2004).

It is of interest to note that the shortest pathway (in the space of the virtual torsion angles) to affect a transition between the docked and undocked conformers would involve changes in θ_{18} and θ_{44} by arcs of 82.4° and 128.2°, respectively. This transition, however, would *not* be possible
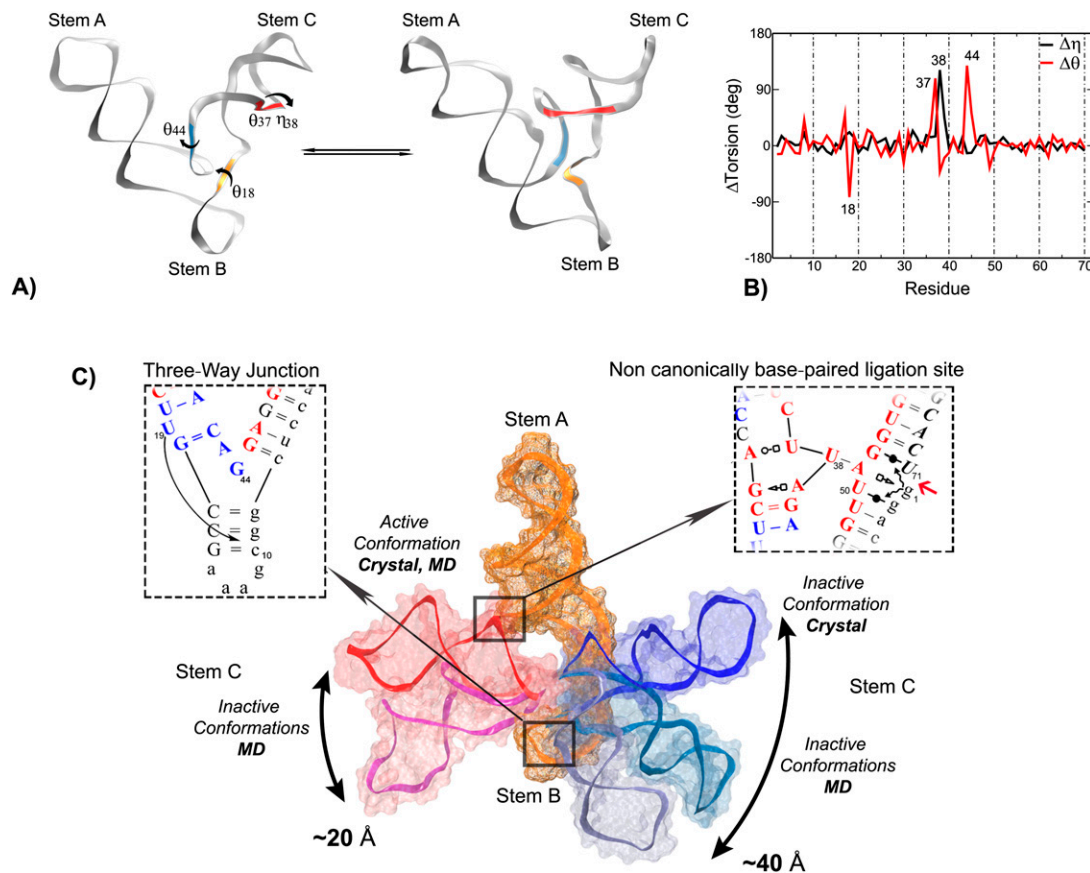
FIGURE 1. (*A*) Ribbon representations of the crystallographic presumed inactive undocked (*left*) and active docked (*right*) conformations of the L1 ligase. Changes in four virtual torsions (labeled and colored) distinguish between the conformers. The regions that do not change significantly are shown in gray. (*B*) Variations of the virtual torsions (ΔTorsion, where Torsion indicates either η or θ) between the docked (active) and undocked (inactive) conformers as found in the crystal structure. The $\Delta\eta$ (black) and $\Delta\theta$ (red) are shown. Only four virtual torsions, θ_{19}, θ_{37}, η_{38}, and η_{44}, show significant deviations (80.0° or more), whereas all other virtual torsions show relatively minor changes (typically, less than 25.0°). (*C*) Representative snapshots from the simulations illustrating the coupled on–off conformational switch–catalytic pathway starting from the active and inactive conformations, both in precursor and product states. (*Middle* panel) Representative snapshots from the simulations are shown with stems A and B (in yellow wire-frame surface) aligned and stem C in ribbons with transparent surfaces. (*Left* panel) Contacts important for stabilization of intermediate states that involve the conserved U_{19} and stem B. (*Right* panel) Interaction patterns for three states observed in the simulations between G_1/GTP_1 and the noncanonically base-paired ligation site. The states that do not appear in the crystal structure are indicated with curly arrows.

since the path would require stem C to pass through the region occupied by stem A. Rather, the sterically allowed pathway would require the transition to proceed along the longer complementary arcs of 277.6° and 231.8° (Fig. 1A).

Design and overall structure and dynamics of the simulations

The molecular dynamics simulations that have been carried out in this study are summarized in Table 3. These include: product in the docked conformation (Prod-D); product in the undocked conformation (Prod-D-U); product departing from the docked conformation with forced unfolding (Prod-D-UF); reactant precursor in the undocked conformation with extended triphosphate (Prec-D-XTP); and reactant precursor in the undocked conformation with Mg^{2+}-coordinated polymerase-like triphosphate (Prec-D-MgTP).

Simulations of the docked and undocked products indicate a large reorientation of stem C

A useful index to aid in the quantification of the positional deviation of one structure relative to another is to use an RMS fitting procedure to superimpose the structures, and measure the resulting average RMSD. In order to isolate the precise nature of the positional deviations between the two structures, it is useful to define separately a fitting set (FS) of atoms used to define the superposition, and a possibly different measured set (MS) of atoms from which the RMSD is averaged. We designate this type of generalized RMSD as RMSD(FS,MS). Note, the RMSD(FS,MS) corresponds to the standard RMSD when the FS and MS are the same. Table 1 lists the RMSD(FS,MS) for various fitting sets and measuring sets in order to help quantify the structural deviation between crystallographic structures and structures from the molecular dynamics simulations.

TABLE 1. RMSD(FS,MS) values comparing crystallographic and simulation structures

(MS,FS)/simulation	Prod-D	Prec-D-MgTP	Prec-D-XTP	Prod-D-UF	Prod-U	Crystal (reference U)
(A,A)	3.0 (0.7)	2.7 (0.6)	2.9 (0.6)	3.3 (0.6)	2.5 (0.6)	1.7
(B,B)	1.8 (0.6)	1.2 (0.3)	1.2 (0.2)	1.6 (0.4)	3.3 (0.6)	0.7
(C,C)	2.3 (0.4)	2.6 (0.3)	2.5 (0.4)	3.2 (0.5)	2.9 (0.4)	2.1
(AB,C)	7.2 (1.6)	4.9 (1.3)	5.8 (1.6)	13.5 (4.0)	19.1 (4.8)	48.4
(AC,B)	5.4 (1.9)	4.2 (1.3)	6.2 (2.2)	5.5 (1.6)	11.3 (2.4)	8.5
(BC,A)	8.2 (1.9)	5.0 (1.5)	4.9 (1.4)	16.7 (5.9)	19.3 (6.3)	44.0
(AB,AB)	3.7 (0.8)	2.8 (0.6)	3.1 (0.7)	3.5 (0.7)	3.4 (0.6)	4.9
(AC,AC)	4.1 (0.7)	3.1 (0.5)	3.4 (0.6)	7.0 (1.9)	7.6 (1.8)	15.1
(BC,BC)	3.0 (0.6)	3.0 (0.5)	3.7 (0.8)	3.8 (0.7)	5.8 (1.0)	7.8
(ABC,ABC)	4.2 (0.7)	3.1 (0.5)	3.5 (0.6)	6.7 (1.8)	8.0 (1.7)	14.3
(J,J)	2.1 (0.3)	2.0 (0.3)	2.3 (0.3)	2.3 (0.4)	4.1 (0.5)	8.3

The RMSD(FS,MS) is defined by superimposing two structures so as to minimize the RMSD based on a "fitting set" of atoms (FS), and then measuring the average RMSD of a "measuring set" of atoms (MS). In this way, one can use the RMSD(FS,MS) to quantify the internal structural deviations of each of the L1 ligase stems A, B, and C, and junction J, and also assess the relative orientation of stems. For example, if RMSD(AB,AB) and RMSD(C,C) are small, but RMSD(AB,C) is large, this indicates that the internal structure of all the stems, as well as the relative orientation of stems A and B, are generally unchanged, whereas there is a large reorientation of stem C relative to stems A and B. The reference coordinates used to superimpose the simulation structures were the crystallographically derived structures also used as the starting points of the simulations (see Materials and Methods). Standard deviations are shown in parentheses. The RMSD(FS,MS) for the crystallographic docked and undocked conformers are also compared (Crystal). The FS and MS sets are defined as follows: (A) stem A, residues 1–7 and 45–71; (B) stem B, residues 8–17; (C) stem C, residues 18–44; and (J) the junction, residues 6–10, 15–22, and 39–47. Note that the definitions of stems A, B, and C involve atom sets that do not intersect, and the union of these sets constitutes the entire L1 ligase. The definition of the junction, on the other hand, is a subset of atoms that intersects with parts of all of the stems. Hence, the set "ABC" implies the entire L1 ligase. The average values and standard deviations from the simulations are calculated from the entire trajectory, excluding the first 10 ns, with data collected every picosecond. All values in the table are in angstroms.

In the simulation of the docked product (Prod-D), the overall heavy-atom RMSD with respect to the initial structure (based on the docked crystal structure) was 4.2 Å, and that of the individual stems ranged from ~1.8 to 3.0 Å (Table 1). On the contrary, the undocked product simulations (Prod-U) showed a much higher overall heavy-atom RMSD of ~7.9 Å relative to the starting structure (based on the undocked crystal structure), despite having RMSD values for the individual stems (2.5–3.3 Å) similar to that of the docked simulation. This indicates that the structure in the simulation exhibits significant variation. The relatively low value of the RMSD(AB,AB) of 3.4 Å indicates that the relative orientation of stems A and B remain similar to the starting structure, whereas the RMSD(AB,C) value of 19.1 Å indicates that it is the orientation of stem C relative to stems A and B that exhibits large variations. Whereas the Prod-D simulation exhibits fairly small

structural variations in the junction region as indicated by the RMSD(J,J) value of 2.1 Å, the Prod-U simulation exhibits much larger variations having a RMSD(J,J) value of 4.1 Å. Together, these results indicate that restricted regions of the junction in the Prod-U simulation exhibit large local conformational variations, and as will be characterized in further detail below, act as hinge points for the reorientation of stem C.

Simulations of unfolding of the docked conformer

The unfolding procedure of the docked conformer is described in detail in the Materials and Methods section. The Prod-D-UF simulation shows a larger overall RMSD relative to the crystallographically derived initial structure of 6.7 Å, whereas the corresponding RMSD for the Prod-D simulation is 4.2 Å. The RMSD of the individual stems, on the other hand, range from ~1.6 to 3.3 Å in the Prod-D-UF simulation. Whereas the orientation of stem B relative to stems A and C is fairly small [RMSD(AC,B) value of 5.5 Å], and similar to the value of 5.4 Å in the Prod-D simulation, the orientations of stem A relative to stem C are markedly larger as expected [RMSD(AB,C) and RMSD(BC,A) values of 13.5 and 16.7 Å, respectively], and are intermediate between those of the Prod-D and Prod-U simulations. These movements of the stems are accompanied by a modest increase in the average RMSD of the junction from 2.1 Å in the Prod-D simulation to 2.3 Å in the Prod-D-UF simulation,

TABLE 2. Hydrogen bonds specific to the docked and undocked crystallographic conformers

Chain P (undocked conformer)		Chain Q (docked conformer)	
HB donor–acceptor	Location	HB donor–acceptor	Location
G_1:O2'–A_{51}:N6	$A_{51}\square\rightarrow G_1$	G_{11}:N2–A_{14}:O2P	GAAA (stem B)
G_{22}:O2'–A_{39}:N6	$G_{22}\leftarrow\square A_{39}$	G_{11}:O2'–A_{13}:N7	GAAA (stem B)
A_{23}:N6–U_{38}:O2'	—	A_{13}:N6–G_{11}:O2'	GAAA (stem B)
G_{33}:N1–C_{28}:N3	$C_{28} = G_{33}$	G_{22}:N1–U_{37}:O2'	—
G_{33}:N2–C_{28}:O2	$C_{28} = G_{33}$	U_{37}:N3–A_{23}:N7	$A_{23}\square- \bigcirc U_{38}$
G_{29}:N2–A_{32}:O2P	GAAA (stem C)	U_{38}:N3–A_{51}:N1	U_{38}–A_{51}
G_{29}:N2–A_{32}:N7	GAAA (stem C)	A_{51}:N6–U_{38}:O2	U_{38}–A_{51}
A_{31}:N6–G_{29}:O2'	GAAA (stem C)	G_{60}:N2–A_{63}:N7	GAAA (stem C)
G_{60}:O2'–A_{62}:N7	GAAA (stem A)	G_{60}:N2–A_{63}:O2P	GAAA (stem C)
		A_{62}:N6–G_{60}:O2'	GAAA (stem C)

The main differences occur at the ends of the stems in the GAAA tetraloops, consistent with conserved internal base pairing within the interior of the stems; notation for canonical and noncanonical base-pairing schemes follow Leontis et al. (2002).

TABLE 3. Summary of simulations presented in the current work

Label	Sequence/structure	Ions	Length (ns)
Prod-D	Chain Q (PDBID:2oiu)	Six Mg^{2+} from crystal structure	100
Prod-U	Chain P (PDBID:2oiu)	No Mg^{2+}	370
Prec-D-XTP	Chain Q (PDBID:2oiu) Precursor, TP in extended conformation	Six Mg^{2+} from crystal structure	80
Prec-D-MgTP	Chain Q (PDBID:2oiu) Precursor, TP in two Mg^{2+}-bound conformation[a]	Seven Mg^{2+}, six from crystal structure	100
Prod-D-UF	Chain P (PDBID:2oiu) Forced unfolding	Five Mg^{2+} from crystal structure, deleted one Mg^{2+} to induce unfolding	60

[a]As in the DNA polymerase β structure (PDBID:1BPY).

and a more notable increase in the fluctuations (e.g., standard deviation of the RMSD in Table 1).

It is apparent from the Prod-U and Prod-D-UF simulations that when stems A and C are not docked (stabilized by tertiary interactions between stems), L1 ligase exhibits a higher degree of conformational flexibility. The origin of this flexibility can be traced to increased fluctuations within the junction region with individual stems keeping their internal structure close to that found in crystal.

Design of two arrangements of the ligation site in the reactant state

It has been proposed that involvement of two Mg^{2+} ions enhances recognition or specificity of the ligation reaction (Yang et al. 2006). L1 ligase is a Mg^{2+}-dependent ribozyme. Generally, the ligation reaction consists of the formation of a phosphodiester bond between the $O_{3'}$ of the terminal nucleotide of the nucleic acid strand and the incoming P_α atom of a nucleotide triphosphate. The reaction proceeds through a pentacovalent transition state/intermediate with inversion of the P_α center stereochemistry, and results in the release of pyrophosphate.

To explore the structure and dynamics of the reactant precursor state of the L1 ligase, we designed two possible models based on the L1 ligase product structure in the docked conformation. These models differed only in the conformation and metal ion coordination of the extra GTP_1 pyrophosphate moiety (G_1 in the product state). The first precursor simulation (Prec-D-XTP) was based on an extended starting conformation of the pyrophosphate. The second precursor simulation (Prec-D-MgTP) was based on a homology model derived from the crystal structure of DNA polymerase β as follows. The motivating premise for the Prec-D-MgTP model was that, in order to accommodate highly charged substrates such as GTP, catalytic sites require two Mg^{2+} ions (Yang et al. (2006). We observed a

high degree of similarity of the disposition of the charged groups and Mg^{2+} between the L1 ligase product structure and crystal structures of several protein polymerases. We have included an additional Mg^{2+} ion and used homology modeling to replicate the conformation of the pyrophosphate based on the crystal structure of human DNA polymerase β (PDB ID 1BPY) (Sawaya et al. 1997). The similarity between the mechanism of the transition state stabilization of the RNA and protein polymerases has been pointed out recently based on the analysis of the crystallized catalytic core of an in vitro optimized class I ligase ribozyme (Bagby et al. 2009; Shechner et al. 2009). We have followed the naming convention for the two Mg^{2+} ions following Yang et al. (2006), with the ion closest to the reaction site labeled A and the other labeled B. Data presented in Table 1 indicate that the overall structure and dynamics of these two simulations followed a similar trend as the Prod-D simulation.

Intrinsic flexibility of the U_{38} loop

The observation that U_{38}, a conserved residue critical for catalytic activity in the L1 ligase family, is docked into the ligation site and makes a canonical base pair with a constituent of the ligation site (A_{51}) in the docked-active conformation, whereas in the undocked-inactive conformation it is positioned ~40 Å away from the site, has led to the postulate that the former is more likely representative of a catalytically active state (Robertson and Scott 2007). The simulation results presented here provide insight into the transition between docked and undocked conformations observed in the crystal and their implications for catalysis. Of particular focus in this section is the conformational change of the U_{38}-containing loop (henceforth designated as the "U_{38} loop").

The conformational change of the U_{38} loop can be tracked by the θ_{37} and η_{38} virtual torsions

In the crystal structures, the conformation of the U_{38} loop varies considerably between the docked and undocked conformers, showing two different conformations. Two virtual torsions that span the loop, θ_{37} and η_{38}, show a large variation (~120.0°) between the two crystal conformers, whereas all the other virtual torsions in the immediate base-paired vicinity show a much smaller variation (≤20.0°) (Fig. 1B). This is suggestive that during the undocked ⇌ docked conformational transition, the U_{38} loop is highly flexible along the θ_{37} and η_{38} virtual torsions. In particular, the loop has to transition from a state where U_{38}

is buried into stem C and makes a sugar-base hydrogen bond with A_{23} (Table 2) into an open conformation that will expose U_{38} in solution, giving it the possibility of docking into stem A. The docking consists in the formation of a canonical base $U_{38} = A_{52}$ and several phosphate–Mg^{2+} contacts, supported by A_{39} and G_{40} of stem C on one side and G_1 of stem A on the other.

The simulations span the full range of the L1 ligase conformational transition involving the U_{38} loop

In the Prod-D, Prec-D-XTP, and Prec-D-MgTP simulations, U_{38} remains base-paired with A_{51}, and the overall behavior of the θ_{37} and η_{38} virtual torsions is similar. For the Prod-D simulation, the average values for the θ_{37} and η_{38} torsions (220.3° and 27.0°, respectively) stay within 43° of the values found in the docked crystal conformer (243.8° and −15.5°, respectively). The θ_{37} and η_{38} torsions exhibit larger fluctuations with respect to other nearby virtual torsion angles (e.g., the set of four torsions upstream of η_{38} and four torsions downstream from θ_{37}). In the Prod-D simulation, the standard deviations for θ_{37} and η_{38} virtual torsions were 36.4° and 60.0°, respectively, whereas the corresponding values for the nearby virtual torsions remained within 16.7°. Similar trends were observed in the Prec-D-XTP and Prec-D-MgTP simulations (data not shown).

In the Prod-U simulation, the θ_{37} and η_{38} virtual torsions follow a similar trend with respect to their undocked conformer values. Averages were 192.1° and 128.7°, respectively, compared with 135.9° and 105.9° in the undocked crystal structure. The standard deviations (32.6° and 28.2°, respectively) were lower than in the Prod-D simulation. For these sets of simulations (Prod-D, Prec-D-XTP, Prec-D-MgTP, and Prod-U), U_{38} fluctuates relatively close to its starting geometry in two approximately separate states, indicated by different colors in Figure 2B. In Figure 2A, these two regions are labeled D (solvent exposed/docked) and U (buried), respectively, and two representative structures are depicted. As shown by the correlation plots of θ_{37} and η_{38} virtual torsions (Fig. 2B) in the Prod-D-UF simulation, the U_{38} loop is able to span the conformational space that covers almost exactly the same area that includes the reunion of the regions sampled during the Prod-D and Prod-U simulations, encompassing the conformations found in crystal.

Three-way junction flexibility: U_{19} makes specific contacts with stem B

Three- and four-way junctions constitute extremely important scaffolds for RNA riboswitches or aptamers (Lescoute and Westhof 2006). The flexibility in the junction plays a key role in the catalytic activity of the L1 ligase family (Robertson et al. 2001). It was found that there are only two pentuple nucleotide structural motifs ($C_{39} = G_{18} \, G_{37} \, A_{38} \, C_{17}$ or $U_{39} = A_{18} \, U_{37} \, G_{38} \, U_{17}$) located in the junction region that supports catalysis (Robertson et al. 2001). These motifs are

preserved during the evolutionary optimization of L1 ligase constructs, but it is yet unclear how they can affect the dynamics of the ligase in order to promote catalysis.

Additionally, U_{19} is a conserved residue in all of the L1 ligase constructs. However, in both crystallographic conformers, U_{19} is oriented toward the exterior of the L1 ligase body without making any contacts. The lack of clear structural basis that explains the conserved character of U_{19} in either the docked or undocked crystal structures raises the question as to whether this residue might be important in facilitating the undocked $\rightleftharpoons$ docked conformational switch. This section explores the possible role of U_{19} in stabilizing intermediate states in this transition.

Relatively larger fluctuations in the three-way junction suggest the presence of the hinge points

In the Prod-D and Prod-U simulations, the virtual torsions associated with the residues that form the junction region (J) are characterized by residue-average standard deviations (RASDs) of 16.1° and 18.5°, respectively. The values of the RASD for the entire set of 142 virtual torsions were 14.2° and 15.4°, respectively. As a second reference, the RASD for a base-paired structure, like stem A, is 12.5 Å for the Prod-U simulation. These values suggest that the junction possesses a higher degree of intrinsic flexibility relative to other regions of the L1 ligase. The major contributors (SD $\geq$30.0) to these large values of RASD are θ_{44} (SD = 48.7°) and η_{45} (SD = 71.0°), for the Prod-D simulation; and θ_{18} (SD = 56.8°), η_{19} (SD = 30.7°), θ_{19} (SD = 37.8°), η_{20} (SD = 53.6°), and θ_{20} (SD = 30.8°), for the Prod-U simulation. In the case of the Prod-D-UF simulation, θ_{44} and η_{45} virtual torsions have the largest standard deviations in the junction region ($\sim$30.0°).

A question that emerges from these data is: What is the connection between these relatively large fluctuations of some of the virtual torsions that span the junction region and the transition from the inactive to the active conformation? During Prod-U simulation, θ_{18} is observed to transition through a series of two relatively distant clusters (see Supplemental Material) that are disposed exactly on the same path that would allow the undocked $\rightleftharpoons$ docked conformational transformation to occur via changes in the backbone virtual torsions apparent from the crystal structure. The transition along θ_{18} is not unique in the junction region during the Prod-U simulation; the other angles that show large SDs during Prod-U simulations populated more than one state and were partially correlated with θ_{18}.

The conserved U_{19} interacts specifically with stem B, and accounts for its possible role in the stabilization of intermediate states on the conformational transition pathway. For the duration of the Prod-D, Prod-D-UF, and precursors simulations, U_{19} remains exposed to the solvent, with no specific contacts with any part of the L1 ligase. However, during the Prod-U simulations, a series of interaction patterns involving U_{19} and stem B, which correlate with the transition
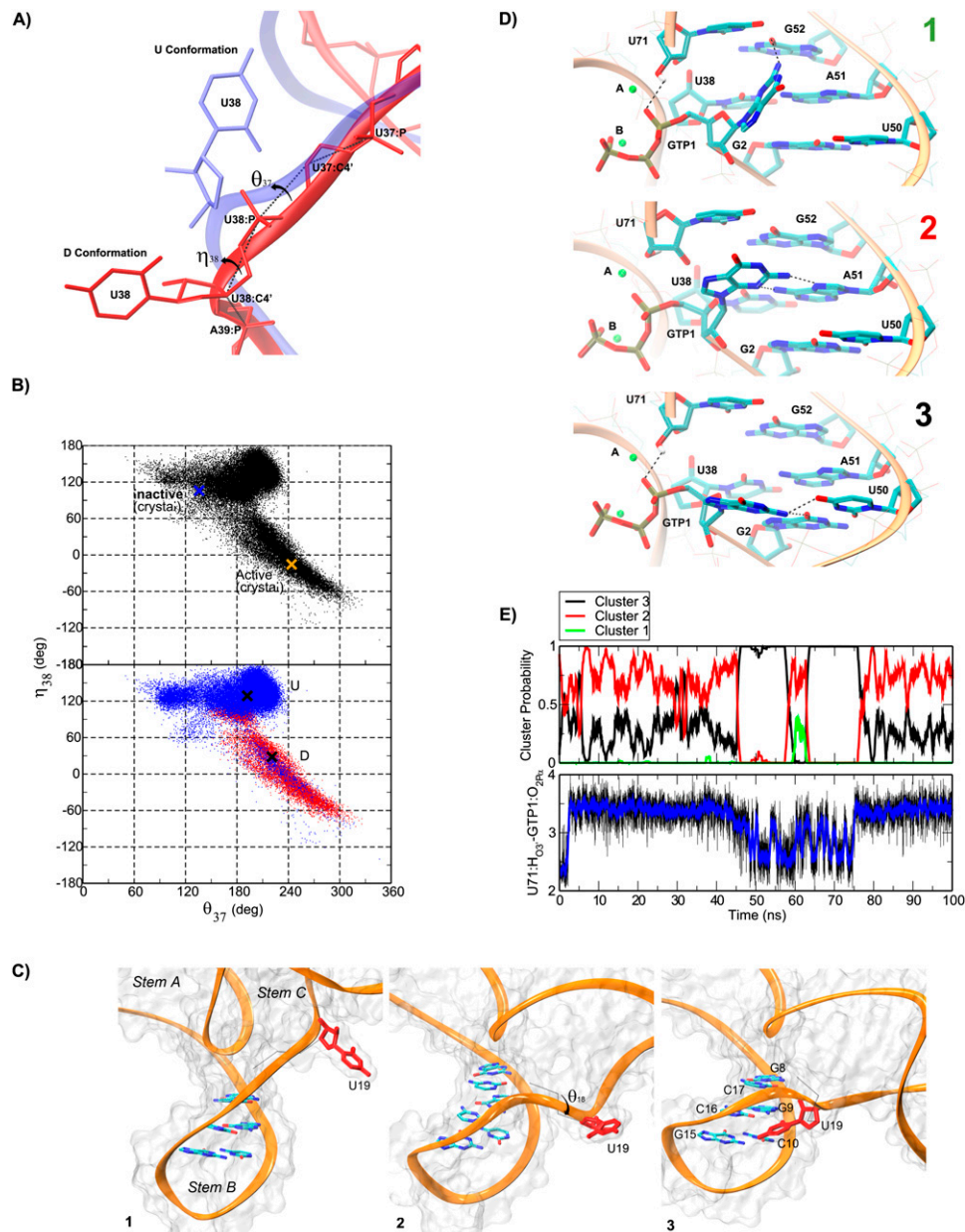
FIGURE 2. (A,B) The U_{38} loop is responsible for allosteric control of the catalytic step by transitioning from a "closed" conformation, labeled "U conformation" (specific, as shown by our simulations, to the inactive/undocked conformers), to an "open" conformation, labeled "D conformation" (specific to the active/docked conformations). (A) Representative conformations for the U and D conformations. (B) *Upper* panel: During the Prod-D-UF simulation, where unfolding of the docked state was induced by removal of key interstem interactions, the complete transition between the two states was observed by monitoring θ_{37} and η_{38} virtual torsions. The values of the θ_{37} and η_{38} found in the two crystallized conformers are shown with blue and orange ×'s. (*Lower* panel) The θ_{37}/η_{38} space sampled in the Prod-D-UF overlaps with regions sampled during the Prod-D (red) and Prod-U (blue) simulations. The center of each distribution is marked with black ×'s. For additional details on the distribution of the two virtual torsions during the Prod-D and Prod-U simulations relative to the Prod-D-UF simulation, see Supplemental Material. (C) Exploring connections between the flexibility of the three-way junction and the L1 ligase conformational switch. Representative structures for the conformational transition along θ_{18} toward the active conformation that brings U_{19} in the close proximity of stem B. Structure 1 is similar to the undocked crystal conformation ($\theta_{18} \approx -110°$). Structure 2 is the transition structure between structures 1 and 3 ($\theta_{18} \approx 170°$). Structure 3 shows a set of specific tertiary interactions between the conserved U_{19} and stem B ($\theta_{18} \approx 90°$). The time evolution of the hydrogen bonding contacts as well as the overall distribution of the virtual torsions can be found in Supplemental Material. (D) The noncanonically base-paired ligation site exhibits a high degree of conformational variability, passing through a series of three states (clusters 1, 2, and 3) characterized by specific hydrogen bond patterns between GTP_1/G_1 and the ligation site. Shown is the arrangement of the ligation site in the case of the Prec-MgTP simulations. (E) The time evolution of probabilities of the three cluster states, and their correlation with the formation of specific $U_{71}:H_{O3'}-GTP_1:O_{2P\alpha}$ interactions, are shown for the Prec-D-MgTP simulation. For the corresponding Prec-D-XTP simulation results, see Supplemental Material.

along the θ_{18} virtual torsion, were observed. Their succession is presented from different perspectives in Figure 2C and in Supplemental Material. We identify two major structural arrangements, labeled 1 and 3 (Fig. 2C). A transition structure, labeled 2, was chosen to be approximately equidistant from A and C in the θ_{18} virtual torsion space (the probability density function is shown in Supplemental Material). Once θ_{18} passes through the transition structure B (Supplemental Fig. 4, vertical dashed line at ~75 ns) after only ~10 ns, U_{19} starts to make contacts with stem B through a series of hydrogen bonding patterns that interchange during the course of the simulation. These interaction patterns include base–base ($U_{19}:N_3–C_{16}:O_2$, $U_{19}:O_2–G_9:N_2$, $U_{19}:N_3–C_{10}:O_2$), base–backbone ($U_{19}:O_{2'}–C_{16}:O_2$), and backbone–backbone ($U_{19}:O_{2'}–C_{17}:O_{2'}$, $C_{17}:O_{2'}–U_{20}:O_{1P}$) contacts. Given the conserved nature of the interacting partners, it is possible that these base–base interactions play an important role in the stabilization of the intermediate states during transitions between the two conformations, and contribute in part to the discrimination in the evolutionary optimization process between active and inactive constructs.

High variability of the noncanonically based-paired ligation site has implications on catalysis

The anatomy of the ligation site in the two crystal conformers is different

L1 ligase catalyzes the phosphodiester bond formation between $GTP_1:P_\alpha$ and $U_{71}:O_{3'}$. Generally, this type of nucleotidyl transfer reaction proceeds through a pentacovalent transition state or metastable intermediate that causes the inversion of the P_α stereochemistry, and results in the release of pyrophosphate (Steitz 1993; Steitz and Steitz 1993). The noncanonically base-paired ligation site on stem A is built from three base pairs: $U_{50}•G_2$ (*cis*-WC/WC), $A_{51} \square \rightarrow G_1$ (*trans*-Hoogsten/sugar edge), and $G_{52}•U_{71}$ (*cis*-WC/WC) (Fig. 1C, right inset). In the crystallographic docked product conformation A_{51} is also involved in a canonical base pair with U_{38}, giving rise to base triple $U_{38} — A_{51} \square \rightarrow G_1$. This interaction, along with the interactions between G_1, A_{39}, and G_{40} phosphates with a Mg^{2+} ion, are responsible for mediating the docking of stem C into stem A. The formation of the base triple $U_{38} — A_{51} \square \rightarrow G_1$ disrupts the hydrogen bond between $G_1:O_{2'}$ and $A_{51}:N_6$. This hydrogen bond is present in the undocked conformer crystal structure and is typical for a *trans*-Hoogsten/sugar edge base pair (Leontis et al. 2002). The distance between $G_1:O_{2'}$ and $A_{51}:N_6$ is 3.2 Å in the undocked conformer and 4.8 Å in the docked conformer (Table 2).

Dynamics of the ligation site in solution simulations

During the Prod-D, Prec-XTP, and Prec-MgTP simulations, the ligation site shows a strong variability, spanning a series of three conformational clusters characterized by the formation of specific hydrogen bond interactions between G_1/GTP_1 and the rest of the ligation site. G_1/GTP_1 oscillates between formation of a hydrogen bond with G_{52}, $GTP_1/G_1:N_2–G_{52}:O_6$ (Fig. 2D, cluster 1); a base triple with A_{51} and U_{38} that is identical to that found in crystallized docked conformer (Fig. 2D, cluster 2); and a base triple with U_{50} and G_2 (Fig. 2D, cluster 3).

The average values and standard deviations of the hydrogen bond distances for each cluster are shown in Table 4; we also indicate which contact yields a hydrogen bond in each simulation. The hydrogen bond pattern in the Prod-U simulation follows a different trend with respect to the Prod-D, Prec-XTP, and Prec-MgTP simulations, occupying only one of the clusters (cluster 2), identical to that found in the undocked crystal conformer that corresponds to a typical *trans*-Hoogsten/sugar edge base pair between A_{51} and G_1.

One of the structural features that confers an increased flexibility for G_1/GTP_1 and its ability to interact with different parts of the ligation site for the Prod-D, Prec-XTP, and Prec-MgTP simulations resides in the way $GTP_1/G_1:O_{2'}$ interacts with the ligation site and the solvent. During Prod-D, Prec-XTP, and Prec-MgTP simulations $GTP_1/G_1:O_{2'}$ shows no tendency to recover the $G_1:O_{2'}–A_{51}:N_6$ hydrogen bond present in the undocked crystal and Prod-U simulation (Table 4), but instead interacts alternately with bulk water molecules or $G_2:O_{2P}$. Docking of stem C into stem A causes the rupture of the $G_1:O_{2'}–A_{51}:N_6$ hydrogen bond in the active state, and has the effect of inducing a conformational variability in the ligation site of the product and precursor/reactant state.

During the simulations, the nucleophile $U_{71}:H_{O3'}$ is involved in contacts with $U_{71}:O_{2'}$ and $GTP_1:O_{2P\alpha}$, of which $GTP_1:O_{2P\alpha}$ appears to be a strong candidate for a general

TABLE 4. Important contacts in L1 ligase in the noncanonically base-paired ligation site: Precursor and product simulations

Contact/(cluster)	Prec-D-MgTP	Prec-D-XTP	Prod-D	Prod-U
GTP1/G1:N2–G52:O6 (1)	5.5 (1.9)*	6.3 (0.9)	6.0 (1.1)*	4.0 (0.5)
GTP1/G1:N3–A51:N6 (2)	5.4 (0.8)*	4.1 (1.1)*	6.2 (1.2)*	3.1 (0.2)*
GTP1/G1:N2–A51:N7 (2)	3.8 (1.3)*	5.6 (1.0)*	5.2 (1.3)*	3.0 (0.1)*
GTP1/G1:O2'–A51:N6 (2)	6.3 (0.9)*	7.4 (0.8)*	7.4 (1.1)*	3.0 (0.2)*
GTP1/G1:N2–G2:O6 (3)	3.4 (0.8)*	3.0 (0.3)*	3.3 (0.9)*	3.6 (0.5)
GTP1/G1:N2–U50:O4 (3)	3.7 (1.0)*	3.5 (1.1)*	3.7 (1.5)*	5.2 (1.0)

Simulations derived from the docked conformer show a higher degree of variability by populating three distinct states than those derived from the undocked conformer that exhibit a unimodal distribution. These three states correspond to the three clusters shown in Figure 2 and are characterized by specific hydrogen bonding interactions. Contacts that yield a hydrogen bond during a specific simulation are labeled with an asterisk (*). Interatomic distance averages and standard deviations (in parentheses) are listed in angstroms.

base. These hydrogen bonds are mainly formed from structures in clusters 1 and 3, with negligible probability for cluster 2, as shown in Figure 2E. The variability of the ligation site (i.e., the capacity to visit several conformational clusters) makes it possible to induce the formation of specific contacts that might promote the chemical step in catalysis (i.e., the contact between $U_{71}{:}H_{O3'}$ and $GTP_1{:}O_{2P\alpha}$). This is due to the noncanonical base pairing of the ligation site.

CONCLUSION

RNA has a diverse ensemble of conformations that interchange on time scales ranging from femtoseconds to microseconds (Crothers 2001). This conformational variability allows RNA molecules to be allosterically controlled, such as in the case of aptamers and aptazymes (Bunka and Stockley 2006), or riboswitches (Schwalbe et al. 2007). However, due to RNA's rugged conformational landscape, structural biology methods such as X-ray crystallography and NMR face challenges when it comes to capturing an accurate, complete picture of the range of conformations that might play important roles in function (Xia 2008). On the other hand, molecular simulation methods provide a wealth of detail concerning both structure and dynamics, and offer powerful tools to complement structural biology, biochemical and biophysical experiments. We report here the first insight obtained from molecular simulations on the conformational transition of the synthetic L1 ligase ribozyme.

We have identified a reduced set of four virtual torsions out of a total of 142 that can be used to distinguish between the active and inactive conformations found in crystal. These virtual torsions span two evolutionarily conserved and restricted regions located in the three-way junction and a U_{38} loop, the conserved residue that is postulated to be responsible for the allosteric control of the catalytic step (Robertson and Ellington 1999, 2000, 2001; Robertson and Scott 2007; Robertson et al. 2004), and they have been used in the present work as an index to follow the evolution of the conformational rearrangement and monitor important transitions. On the time scale of our simulations we were able to observe approximately two-thirds of the complete stem C's 80-Å conformational switch. Since this transition might correspond to the rate-controlling step in L1 ligase catalysis, it is of interest to identify the dynamic hinge points and key interactions that provide stabilization of conformational intermediates that allow the transition to occur. The origin of the large-scale reorientation was traced to variations of a restricted region of the junction. Progress along the transition coordinate was accompanied by the formation of new contacts, not apparent in the crystal structures, between two evolutionarily conserved portions of the L1 ligase: the CG stack of stem B and U_{19} of stem C. These contacts might play an important role in stabilization of intermediate states along the transition pathway.

Simulations were performed from two different initial arrangements of the reactant state differing in the conformation of the GTP_1 triphosphate and its Mg^{2+} coordination. We observed that the noncanonically base-paired ligation site shows a high degree of variability, visiting three distinct conformational states characterized by specific hydrogen bonding patterns. The $U_{71}{:}O_{3'}$ nucleophile was observed to make direct hydrogen bond interactions with the $O_{2\alpha}$ of GTP_1, in support of the role of this residue as a potential general base. These interactions arise as a direct result of the noncanonically base-paired nature of the ligation site and its intrinsic flexibility.

The present simulation results have identified important dynamical hinge points in the conformational transition from inactive to active forms of the L1 ligase, and have characterized interactions that stabilize intermediates along the transition pathway. The insights gained from this work advance our knowledge of the coupled catalytic/conformational riboswitch mechanism of L1 ligase and may have broader implications for prebiotic RNA enzymes.

MATERIALS AND METHODS

General system setup

Initial structures used in the simulations were based on a 2.6-Å resolution crystal structure of two co-crystallized L1 ligase ligation product conformers (PDBID:2OIU) (Robertson and Scott 2007). A series of five simulations were carried out using as starting structures the two separate conformers (chains P and Q in the PDB file), corresponding to the reputed active state with stem C docked with stem A, and in an alternate presumably inactive undocked state where stem C is positioned ∼80 Å away in the crystal.

Since the crystal structure contained the ligated product of L1 ligase, for the reactant state simulations, an additional pyrophosphate moiety not present in the crystal structure was added in the specified conformation (Table 3). The ribozyme was immersed in an orthogonal cell ($a = 90.0$ Å, $b = 90.0$ Å, $c = 90.0$ Å) of pre-equilibrated TIP3P (Jorgensen et al. 1983) waters so as to ensure a 15-Å buffer between any RNA atom and the boundary of the cell.

The ribozyme molecule was rotated so that the longest molecular axis was aligned with the longest water cell axis, it was centered about the site, and pruned such that any water molecule within 2.8 Å from the solute was removed. All water molecules and Mg^{2+} ions resolved in the crystal structure and within 7 Å of each conformer were kept in the starting structures. The ionic atmosphere consisted of Na^+ and Cl^- ions that were added at random positions at least 5.0 Å away from any solute to neutralize the system and reach the physiologic concentration of 0.14 M.

Simulation protocol

Simulations were performed with the NAMD simulation package (version 2.6) (Phillips et al. 2005) using the all-atom Cornell et al. (1995) force field (parm99), and generated with the AMBER 10 package (Pearlman et al. 1995; Case et al. 2002, 2005) and the TIP3P water model (Jorgensen et al. 1983). Periodic boundary

conditions were used along with the isothermal–isobaric ensemble (NPT) at 1 atm and 300 K using a Nosé–Hoover–Langevin pressure piston control (Martyna et al. 1994; Feller et al. 1995), with a decay period of 100.0 fs and a damping time scale of 50 fs, and the Langevin thermostat with a damping coefficient of 0.1 ps^{-1}. The smooth particle mesh Ewald (PME) method (Essmann et al. 1995; Sagui and Darden 1999) was employed with a B-spline interpolation order of 6 and the default κ value used in NAMD. The FFT grid points used for the lattice directions were chosen using ~ 1.0 Å spacing. Nonbonded interactions were treated using an atom-based cutoff of 12.0 Å with switching of nonbond potential beginning at 10.0 Å. Numerical integration was performed using the leap-frog Verlet algorithm with a 1 fs time step (Allen and Tildesley 1987). Covalent bond lengths involving water hydrogens were constrained using the SHAKE algorithm (Ryckaert et al. 1977).

Equilibration procedure

The positions of the solute atoms, including the crystallographic waters and Mg^{2+} ions (if any), were kept fixed, letting the water and counter-ion and co-ion positions to equilibrate for 15 ns, after 5000 steps of initial minimization. The solute atom positions were energy optimized and then were allowed to move under harmonic restraints over 5 ns. The initial harmonic force constant was 5.0 kcal/mol^{-1} and exponentially released over 5 ns. The whole system was energy optimized and unconstrained dynamics simulation began from 30 K under a constant pressure of 1 atm. The temperature was increased to 300 K in steps of 5 K, each step taking 10 ps. The same equilibration process was applied for each simulation. The motions and relaxations of solvent and counterions are notoriously slow to converge in nucleic acid simulations (Ponomarev et al. 2004), and careful equilibration is critical. For each simulation, a total of 20 ns of equilibration (15 ns of water/ion relaxation and 5 ns of solvent and structure relaxation) was carried out before production of the trajectory used for analysis (Ponomarev et al. 2004).

Analysis

Analysis of the trajectories was carried out using VMD (version 1.8.6) (Humphrey et al. 1996). L1 Ligase Stem A (A) was defined to comprise residues 1–7 and 45–71, stem B (B) residues 8–17, stem C (C) residues 18–44, and the junction (J) residues 6–10, 15–22, and 39–47.

For structural comparison, we do not include the three GAAA tetraloops since they were introduced into the L1 ligase structure for crystallization purposes (Robertson and Scott 2007). The notation used to denote base pairing follows that of Leontis et al. (2002). The definitions of virtual torsions used in the present work follows that of Duarte and Pyle (1998), and are designated as follows: η_n is the virtual torsion angle defined by the atoms C4$'_{n-1}$, P$_n$, C4$'_n$, and P$_{n+1}$; and θ_n is the virtual torsion angle defined by the atoms P$_n$, C4$'_n$, P$_{n+1}$, and C4$'_{n+1}$. The coefficient n stands for the position in the sequence (residue number) of the nucleotide that contains that specific atom.

Clustering of structures from the trajectories was performed based on the hydrogen bond pattern. For this purpose, it was useful to define a smooth hydrogen bond probability function to use as an index to aid in the definition of the cluster. The cluster probability, $P_i(t)$, in which the simulation structure at time t belongs to cluster i, is defined as

$$P_i(t) = \frac{\dfrac{1}{N_i}\sum_{j=1}^{N_i}\mathrm{HB}P_{ij}(t)}{\sum_{k=1}^{\mathrm{clusters}}\dfrac{1}{N_k}\sum_{j=1}^{N_k}\mathrm{HB}P_{kj}(t)}, \tag{1}$$

where N_i is the total number of hydrogen bonds associated with cluster i; and HB$P_{ij}(t)$ is the probability of formation of the hydrogen bond j in cluster i at time t, and is defined by

$$\mathrm{HB}P_{ij}(t) = \left\{1 + \exp\left[c\Delta E_{ij}(t)\right]\right\}^{-1}, \tag{2}$$

where $\Delta E_{ij}(t) = E[r_{ij}(t), \theta_{ij}(t)] - E[r_0, \theta_0]$, and $E[r, \theta]$ is defined as

$$E[r, \theta] = -\frac{\cos\theta}{r^2}. \tag{3}$$

For hydrogen bonds of the form D$-$H $\cdots$ A, where D and A are the hydrogen bond donor and acceptor, respectively; r is the H $\cdots$ A distance (Å); and θ is the D$-$H $\cdots$ A angle (deg). The constants r_0, θ_0, and c were chosen empirically to be 3.2 Å, $\pi/3$ rad, and 100 Å^2, respectively.

Unfolding the docked conformer

To explore the conformational switch pathway in the vicinity of the docked conformation we have conducted a simulation in which we induced the unfolding of the docked product by removing key interactions responsible for docking stem C into stem A. These interactions include a Mg^{2+} ion that forms an interstem coordination, and two hydrogen bonds between the canonical U$_{38}$ and A$_{51}$ base pair. Departing from the docked crystallographic structure, the Prod-D-UF simulation was prepared analogously to the Prod-D simulation except without the interstem Mg^{2+} ion. Once the solvent was equilibrated, a harmonic force constant of 30.0 kcal/mol/Å^2 was gradually applied over 5.0 ns to force the distance between U$_{38}$:N$_3$ and A$_{51}$:N$_1$ from 3.5 to 8.0 Å in order to gently break the interaction between stems.

SUPPLEMENTAL MATERIAL

Supplemental material can be found at http://www.rnajournal.org.

REFERENCES

Allen M, Tildesley D. 1987. *Computer simulation of liquids*. Oxford University Press, Oxford, UK.

Bagby SC, Bergman NH, Shechner DM, Yen C, Bartel DP. 2009. A class I ligase ribozyme with reduced Mg^{2+} dependence: Selection, sequence analysis, and identification of functional tertiary interactions. *RNA* **15:** 2129–2146.

Bartel DP, Szostak JW. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science* **261:** 1411–1418.

Bokinsky G, Zhuang X. 2005. Single-molecule RNA folding. *Acc Chem Res* **38:** 566–573.

Bunka DHJ, Stockley PG. 2006. Aptamers come of age—at last. *Nat Rev Microbiol* **4:** 588–596.

Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Crowley M, Walker RC, Zhang W, et al. 2002. *AMBER 10*. University of California at San Francisco, San Francisco.

Case DA, Cheatham TE III, Darden TA, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling CL, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. *J Comput Chem* **26:** 1668–1688.

Chen X, Li N, Ellington AD. 2007. Ribozyme catalysis of metabolism in the RNA world. *Chem Biodivers* **4:** 633–655.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* **117:** 5179–5197.

Crothers D. 2001. RNA conformational dynamics. In *RNA* (ed. D Stoll et al.), pp 61–71. Elsevier, Oxford, UK.

Duarte CM, Pyle AM. 1998. Stepping through an RNA structure: A novel approach to conformational analysis. *J Mol Biol* **284:** 1465–1478.

Ekland EH, Szostak JW, Bartel DP. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **269:** 364–369.

Ellington AD, Szostak JW. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346:** 818–822.

Essmann U, Perera L, Berkowitz ML, Darden T, Hsing L, Pedersen LG. 1995. A smooth particle mesh Ewald method. *J Chem Phys* **103:** 8577–8593.

Feller S, Zhang Y, Pastor R, Brooks B. 1995. Constant pressure molecular dynamics simulation: The Langevin piston method. *J Chem Phys* **103:** 4613–4621.

Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. *J Mol Grapics* **14:** 33–38.

Ikawa Y, Tsuda K, Matsumura S, Inoue T. 2004. De novo synthesis and development of an RNA enzyme. *Proc Natl Acad Sci* **101:** 13750–13755.

Jaeger L, Wright MC, Joyce GF. 1999. A complex ligase ribozyme evolved in vitro from a group I ribozyme domain. *Proc Natl Acad Sci* **96:** 14712–14717.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79:** 926–935.

Joyce GF. 2007. A glimpse of biology's first enzyme. *Science* **315:** 1507–1508.

Landweber LF, Pokrovskaya ID. 1999. Emergence of a dual-catalytic RNA with metal-specific cleavage and ligase activities: The spandrels of RNA evolution. *Proc Natl Acad Sci* **96:** 173–178.

Leontis NB, Stombaugh J, Westhof E. 2002. The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res* **30:** 3497–3531.

Lescoute A, Westhof E. 2006. Topology of three-way junctions in folded RNAs. *RNA* **12:** 83–93.

Leulliot N, Varani G. 2001. Current topics in RNA–protein recognition: Control of specificity and biological function through induced fit and conformational capture. *Biochemistry* **40:** 7947–7956.

Marshall KA, Ellington AD. 1999. Training ribozymes to switch. *Nat Struct Biol* **6:** 992–994.

Martyna GJ, Tobias DJ, Klein ML. 1994. Constant pressure molecular dynamics algorithms. *J Chem Phys* **101:** 4177–4189.

McGinness KE, Joyce GF. 2003. In search of an RNA replicase ribozyme. *Chem Biol* **10:** 5–14.

Murray LJW, Arendall WB, Richardson DC, Richardson JS. 2003. RNA backbone is rotameric. *Proc Natl Acad Sci* **100:** 13904–13909.

Pearlman DA, Case DA, Caldwell JW, Ross WR, Cheatham TE III, DeBolt S, Ferguson D, Seibel G, Kollman P. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics, and free energy calculations to simulate the structure and energetic properties of molecules. *Comput Phys Commun* **91:** 1–41.

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kaleé L, Schulten K. 2005. Scalable molecular dynamics with NAMD. *J Comput Chem* **26:** 1781–1802.

Ponomarev SY, Thayer KM, Beveridge DL. 2004. Ion motions in molecular dynamics simulations on DNA. *Proc Natl Acad Sci* **101:** 14771–14775.

Richardson JS, Schneider B, Murray LW, Kapral GJ, Immormino RM, Headd JJ, Richardson DC, Ham D, Hershkovits E, Williams LD, et al. 2008. RNA backbone: Consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). *RNA* **14:** 465–481.

Robertson MP, Ellington AD. 1999. In vitro selection of an allosteric ribozyme that transduces analytes to amplicons. *Nat Biotechnol* **17:** 62–66.

Robertson MP, Ellington AD. 2000. Design and optimization of effector-activated ribozyme ligases. *Nucleic Acids Res* **28:** 1751–1759.

Robertson MP, Ellington AD. 2001. In vitro selection of nucleoprotein enzymes. *Nat Biotechnol* **19:** 650–655.

Robertson MP, Scott WG. 2007. The structural basis of ribozyme-catalyzed RNA assembly. *Science* **315:** 1549–1550.

Robertson MP, Hesselberth JR, Ellington AD. 2001. Optimization and optimality of a short ribozyme ligase that joins non-Watson–Crick base pairings. *RNA* **7:** 513–523.

Robertson MP, Knudsen SM, Ellington AD. 2004. In vitro selection of ribozymes dependent on peptides for activity. *RNA* **10:** 114–127.

Rogers J, Joyce GF. 1999. A ribozyme that lacks cytidine. *Nature* **402:** 323–325.

Ryckaert JP, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J Comput Phys* **23:** 327–341.

Sagui C, Darden TA. 1999. Molecular dynamics simulations of biomolecules: Long-range electrostatic effects. *Annu Rev Biophys Biomol Struct* **28:** 155–179.

Sawaya MR, Prasad R, Wilson S, Kraut J, Pelletier H. 1997. Crystal structures of human DNA polymerase β complexed with gapped and nicked DNA: Evidence for an induced fit mechanism. *Biochemistry* **36:** 11205–11215.

Schneider B, Morávek Z Berman HM. 2004. RNA conformational classes. *Nucleic Acids Res* **32:** 1666–1677.

Schwalbe H, Buck J. Fürtig B, Noeske J, Wöhnert J. 2007. Structures of RNA switches: Insight into molecular recognition and tertiary structure. *Angew Chem Int Ed* **46:** 1212–1219.

Shechner DM, Grant RA, Bagby SC, Koldobskaya Y, Piccirilli JA, Bartel DP. 2009. Crystal structure of the catalytic core of an RNA–polymerase ribozyme. *Science* **326:** 1271–1275.

Sims GE, Kim SH. 2003. Global mapping of nucleic acid conformational space: Dinucleoside monophosphate conformations and transition pathways among conformational classes. *Nucleic Acids Res* **31:** 5607–5616.

Steitz TA. 1993. DNA- and RNA-dependent DNA polymerases. *Curr Opin Struct Biol* **3:** 31–38.

Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci* **90:** 6498–6502.

Treiber DK, Williamson JR. 1999. Exposing the kinetic traps in RNA folding. *Curr Opin Struct Biol* **9:** 339–345.

Wadley LM, Keating KS, Duarte CM, Pyle AM. 2007. Evaluating and learning from RNA pseudotorsional space: Quantitative validation

of a reduced representation for RNA structure. *J Mol Biol* **372:** 942–957.

Williamson JR. 2000. Induced fit in RNA–protein recognition. *Nat Struct Biol* **7:** 834–837.

Xia T. 2008. Taking femtosecond snapshots of RNA conformational dynamics and complexity. *Curr Opin Chem Biol* **2:** 1–8.

Yang W, Lee JY, Nowotny M. 2006. Making and breaking nucleic acids: Two-Mg^{2+}-ion catalysis and substrate specificity. *Mol Cell* **22:** 5–13.