Supporting Information

All-atom Simulations Decrypt the Molecular Terms of RNA Catalysis in the Exon-ligation Step of the Spliceosome

Jure Borišek[‡] and Alessandra Magistrato[*,†]

[‡]National Institute of Chemistry, Slovenia, Hajdrihova 19, 1001 Ljubljana, Slovenia

[†]CNR-IOM-Democritos national Simulation Center c/o SISSA, via Bonomea 265, 34136 Trieste, Italy

Corresponding Author*:

*E-mail: alessandra.magistrato@sissa.it

Table of Contents

Supporting Results

Exploration of an alternative reaction coordinate

In order to investigate a potential alternative reaction pathway we selected difference between breaking (P@U1 – $O^{3'}$@G2) and forming (O_{nuc}@G-1 – P@U1) bonds for the reaction coordinate RC4. In distinct studies of DNA/RNA processing enzymes[1-4], this reaction coordinate was able, at the same level of theory, capture both the reaction of phosphodiester bond cleavage and formation, and indirectly the associated proton transfer events. Starting from the value of RC4 = -1.5 Å after 17 windows the reaction never reached the proper transition state, and at the end value of RC = 1.5 Å the nucleophile was still protonated (Figure S3). In comparison to the direct proton transfer mechanism (see Manuscript), M1 from RC = 0 on does not coordinate O_{nuc}, and simultaneously the distance between M2 and O^{Rp} starts to increase in agreement with a previous study.[1] However, the proton transfer from the nucleophile to nearby water never occurred, at variance with other similar studies.[3] For this reaction pathway, where products are not fully formed and transition state is not properly reached, the ΔF is 16.2 ± 1.1 kcal/mol.

Supporting Figures

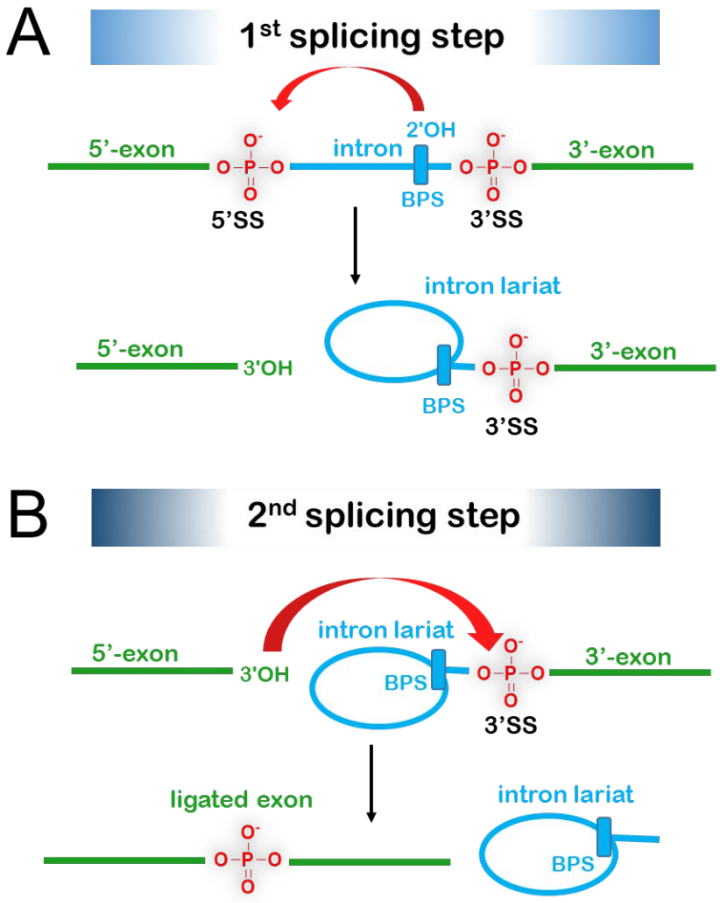


Figure S1. Schematic presentation of splicing reactions. (A) In the 1st splicing step the 2'OH of adenine nucleotide at the branch point site (BPS) performs a nucleophilic attack on the phosphate at the 5'-end of the 5'-splice site (5'SS), forming an intron lariat-3'-exon intermediate. (B) In the 2nd splicing step the 3'OH at the 3'-end of the cleaved 5'-exon attacks as a nucleophile the phosphate at the 5'-end of the 3'-exon, yielding the ligated exon and the intron lariat as products.

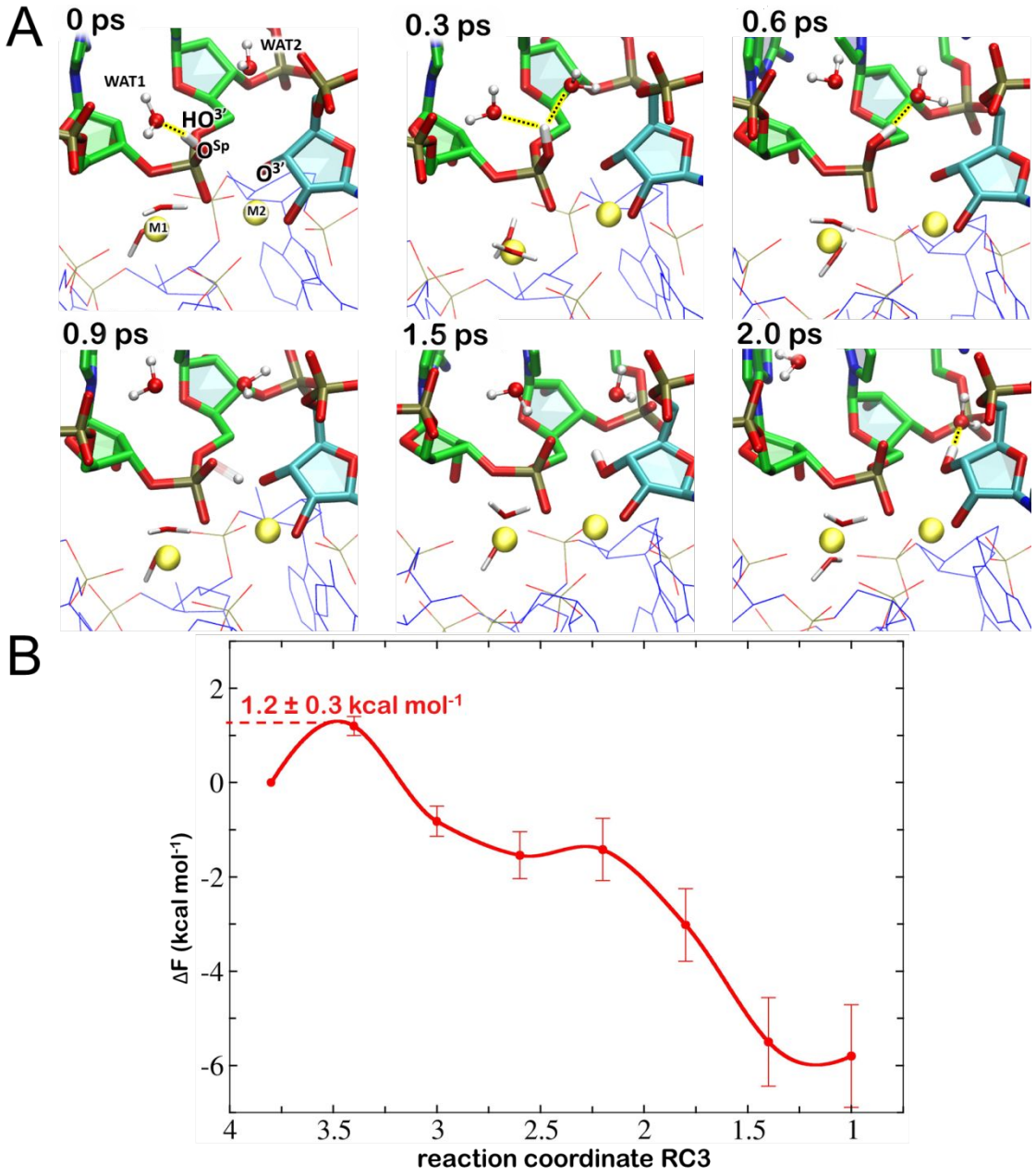


Figure S2. (A) Snapshots from the constrained QM/MM MD simulation featuring the second proton transfer event at fixed reaction coordinate window RC2 = 17. In the time frame of 2 ps the proton ($HO^{3'}$) is transferred from the scissile phosphate oxygen (O^{Sp}) to the leaving group ($O^{3'}$) facilitated by two water molecules. Hydrogen bonds between transferred proton and nearby water molecules are shown in dash lines and highlighted in yellow. (B) Helmholtz free energy profile (ΔF [kcal/mol]) for the second proton transfer from scissile phosphate oxygen (O^{Sp}) to the leaving group ($O^{3'}$) monitored by scanning RC3 (the distance of the $HO^{3'}$ to $O^{3'}$) at a fixed RC2 window (RC2=17) (see Figure 2).

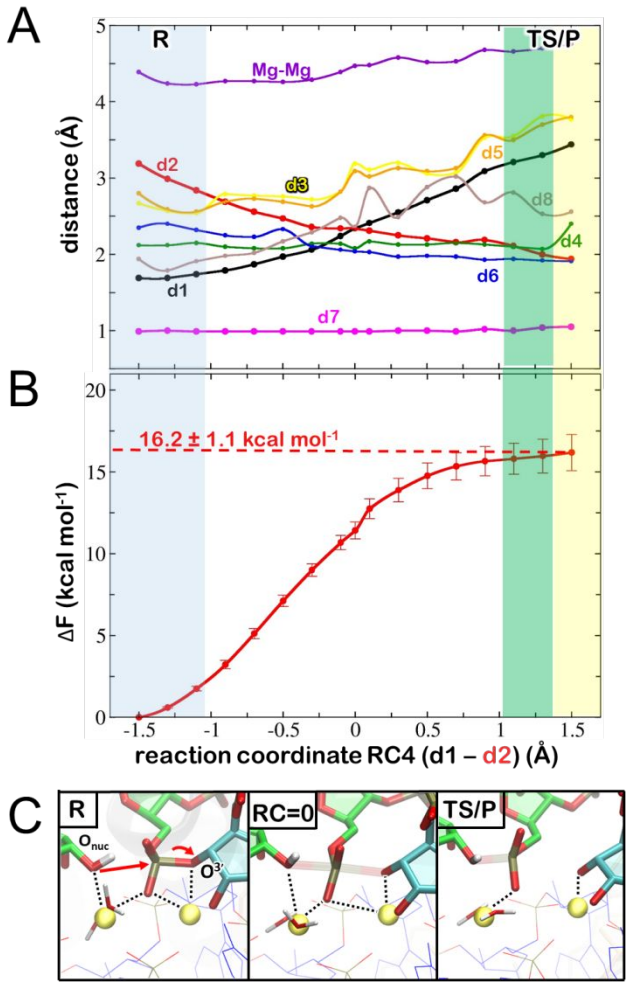


Figure S3. Structural and energetic properties for an alternative reaction pathway. Highlighted are the areas for reactants (R), first proton and transition state/products (TS/P). (A) Monitored bond distances along the reaction pathway. (B) Free energy profile (ΔF [kcal/mol]) calculated by thermodynamic integration at the DFT-BLYP level of theory with a double zeta plain wave basis sets and Amber ff14SB/ff99+bsc0+χOL3 level for the QM and MM part, respectively. (C) Representative snapshots of the reactants (R), reaction coordinate at value 0 (RC=0), and not fully reached products (P). Exon and 3'-exon-intron carbon atoms are depicted in green and cyan, respectively, and Mg^{2+}-ions in yellow spheres; the rest of the catalytic site is presented as lines. Dashed lines depict coordination of metal ions with exon and intron-exon residues. Non-Mg^{2+}-ions coordinating waters are omitted for clarity.

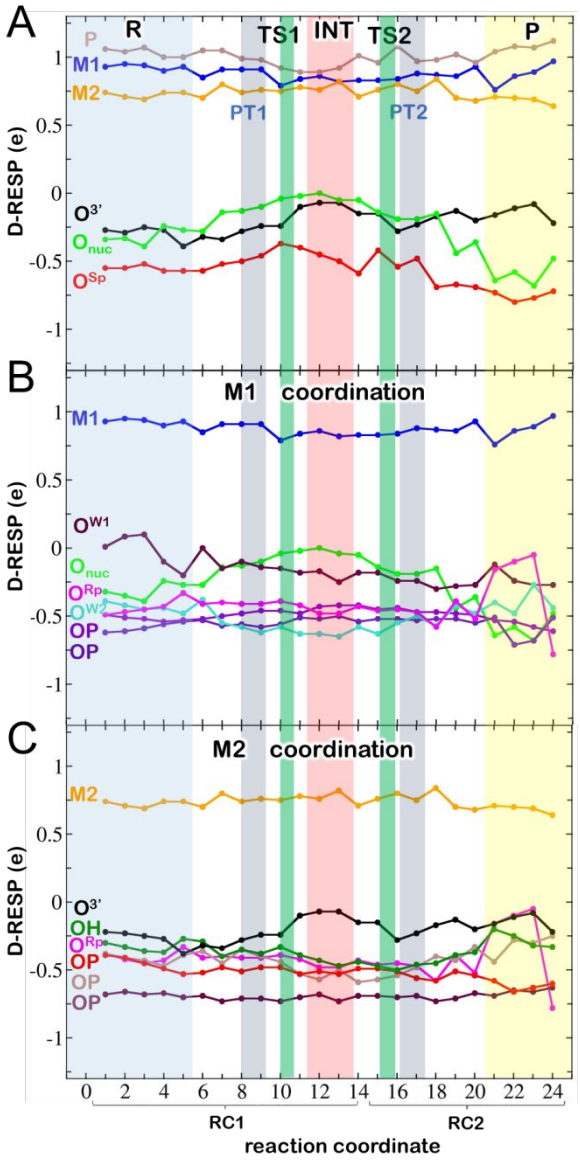


Figure S4. Dynamical RESP (D-RESP) charges with highlighted areas corresponding to reactant (R), first proton transfer (PT1), transition state 1 (TS1), intermediate (INT), transition state 2 (TS2), second proton transfer (PT2) and product (P) states. Monitored are D-RESP for (A) atoms directly involved in reaction (P, M1, M2, $O^{3'}$, O_{nuc} and O^{Sp}). (B) M1 coordinating atoms (M1, O^{W1}, On_{uc}, O^{Rp}, O^{W2} and two OP atoms), and (C) M2 coordinating atoms (M2, $O^{3'}$, OH, O^{Rp} and three OP atoms). We remark that the sudden drop of O^{Rp} charge between windows 23 and 24 is most probably due to the reorientation of the transferred proton of $O^{3'}$ oxygen away from the scissile phosphate oxygen O^{Rp}, which, in these windows, starts hydrogen bonding with water.

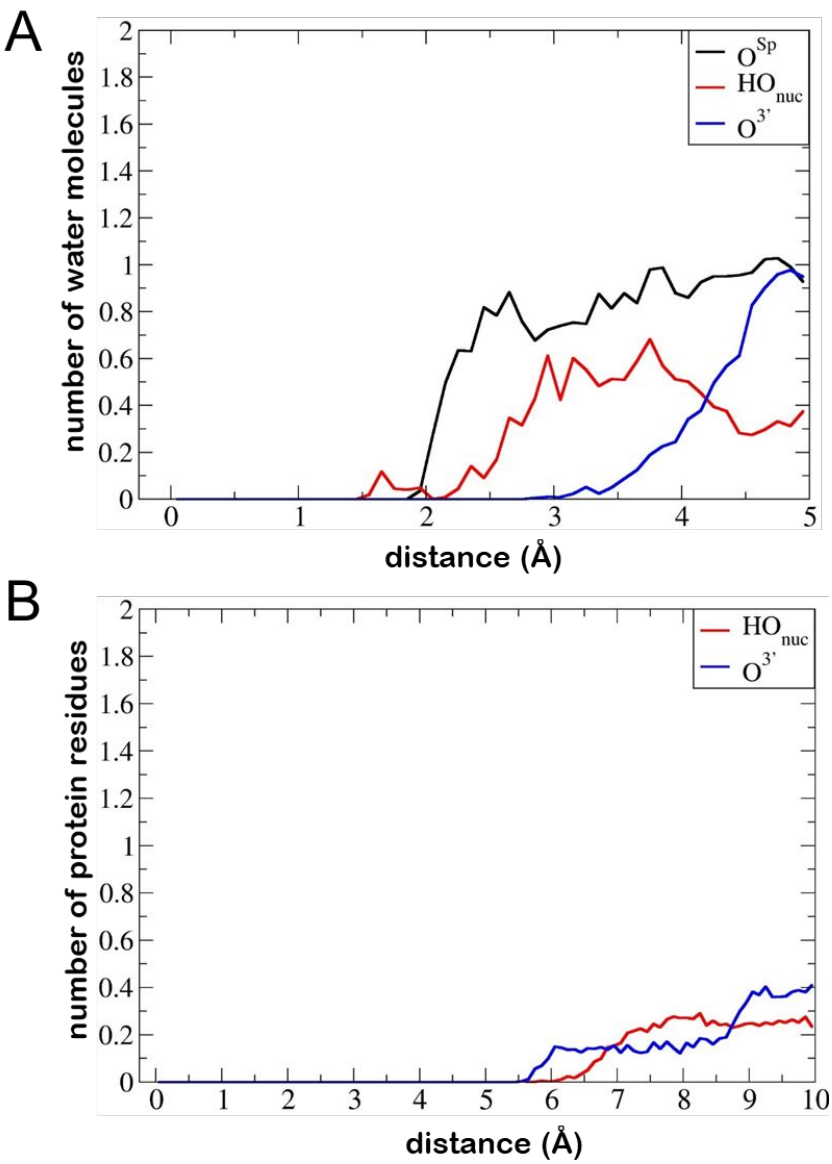


Figure S5. Hydration of catalytic site and investigation of protein environment by radial distribution function to explore possible proton transfer events. (A) The distances of water molecules are considered with respect to hydrogen of nucleophilic group HO_{nuc}@G-1 ($HO^{3'}$), scissile phosphate oxygen O^{Sp}@U1 and leaving group oxygen $O^{3'}$@G2. (B) The distances of the protein residues are presented relative to catalytic site hydrogen of nucleophilic group HO_{nuc}@G-1 and leaving group oxygen $O^{3'}$@G2.

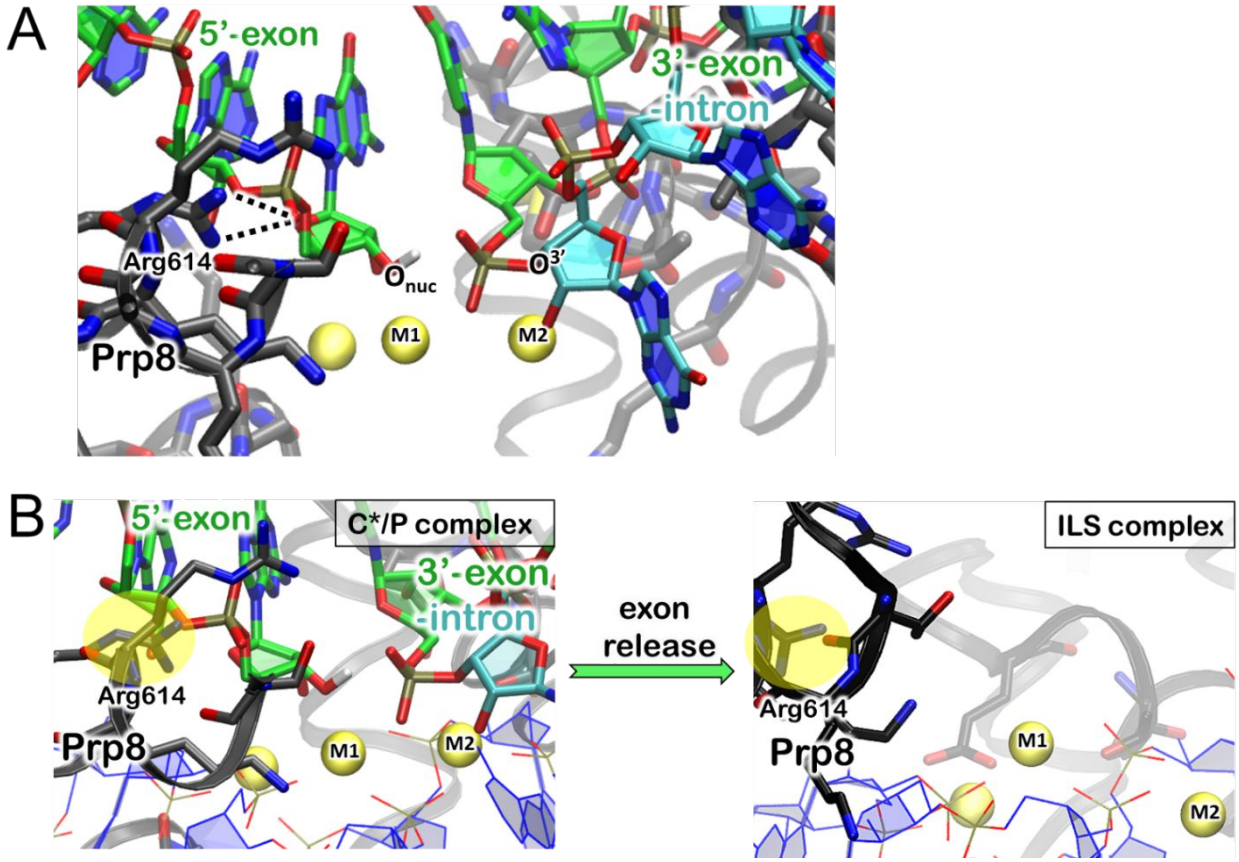


Figure S6. Representation of the Arg614 position and the catalytic site consisting of intron (cyan), exon (green), U6 snRNA (blue), metal ions (yellow sphere), and surrounding Prp8 protein residues (black). (A) Protein residues within 14 Å of nucleophile (O_{nuc}) and leaving group ($O^{3'}$) are shown. (B) Comparison of Prp8 protein residues between C*/P[5-6] and ILS complex[7] showing the switch of Prp8's Arg614 residue (highlighted in yellow circle) upon exon release. Unlabeled metal ion is one of the three magnesium ions that play only a structural role in stabilizing the active site (the other two Mg^{2+} ions are not shown here).

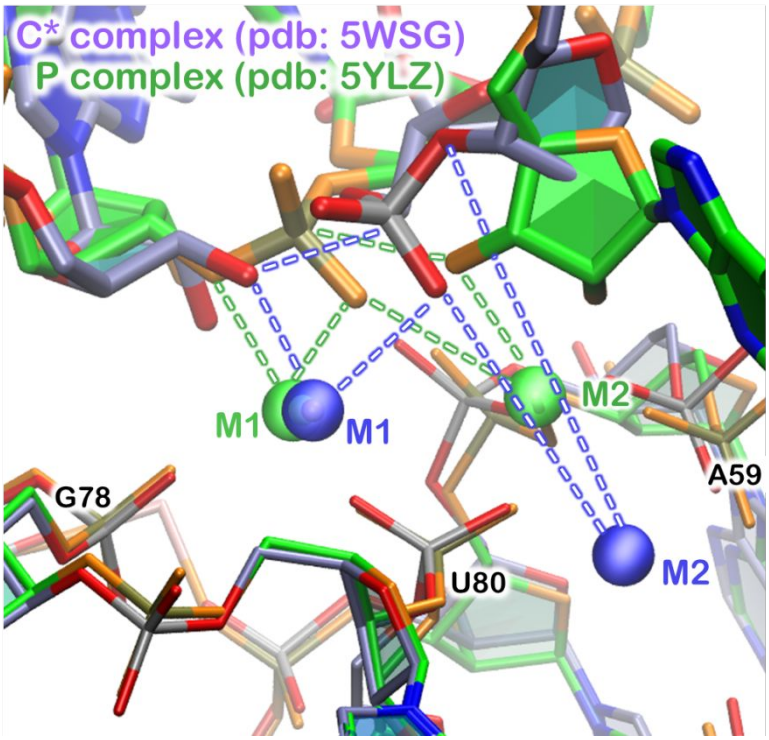


Figure S7. Superimposition of spliceosome C* (pdb id: 5WSG) and P (pdb id: 5YLZ) complexes depicted in blue and green color for the carbon atoms, and red and orange color for oxygens, respectively, showing a more suitable configuration of P complex to undergo and study catalysis. M1 and M2 denote Mg^{2+} ions and are shown as blue and green spheres for the C* and P complex, respectively.

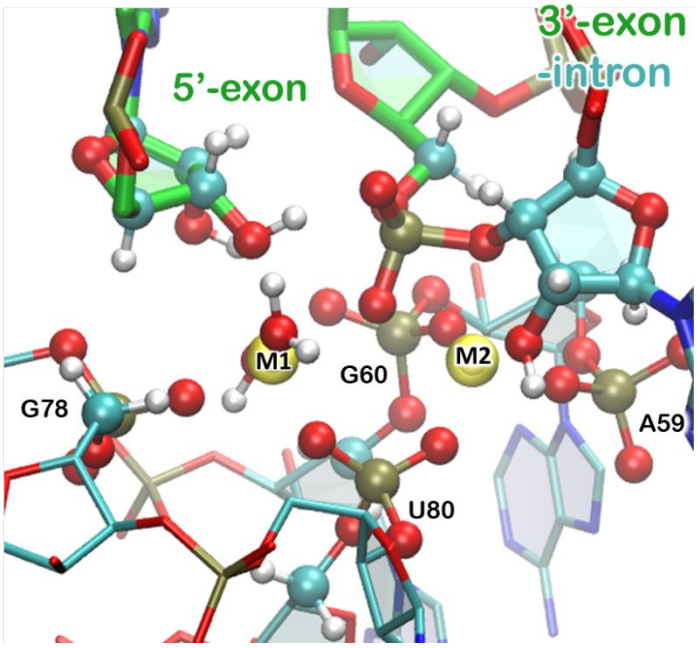


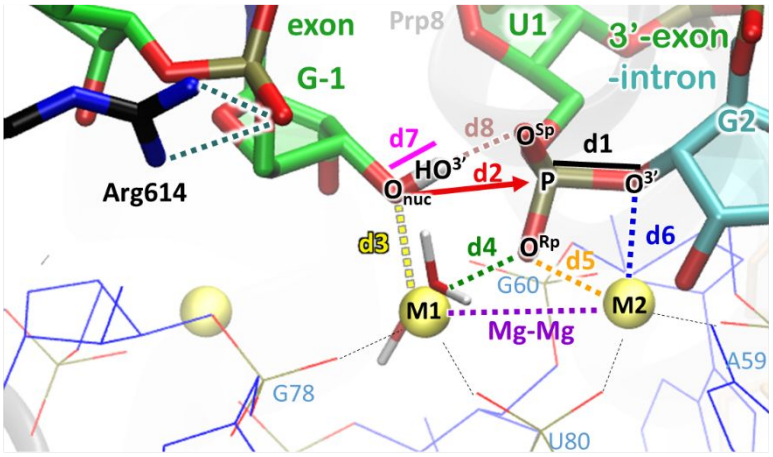
Figure S8. Selected atoms of the QM region shown in balls and sticks employed for most of QM/MM MD simulations. M1 and M2 denote Mg^{2+} ions. The carbon atoms of MM region are shown in green and cyan for exon and intron/U6 RNA, respectively. QM region carbon, phosphorous, oxygen, hydrogen and Mg^{2+} atoms are shown in cyan, brown, red, white and yellow, respectively.

Supporting Tables

Table S1. Details of the proteins included in the spliceosome P model. Columns CHAIN, MOLECULE and RESOLUTION refer to original names and reported resolution from the original PDB structure (pdb id: 5ylz). Each considered molecule is highlighted with the same color as presented throughout the manuscript.

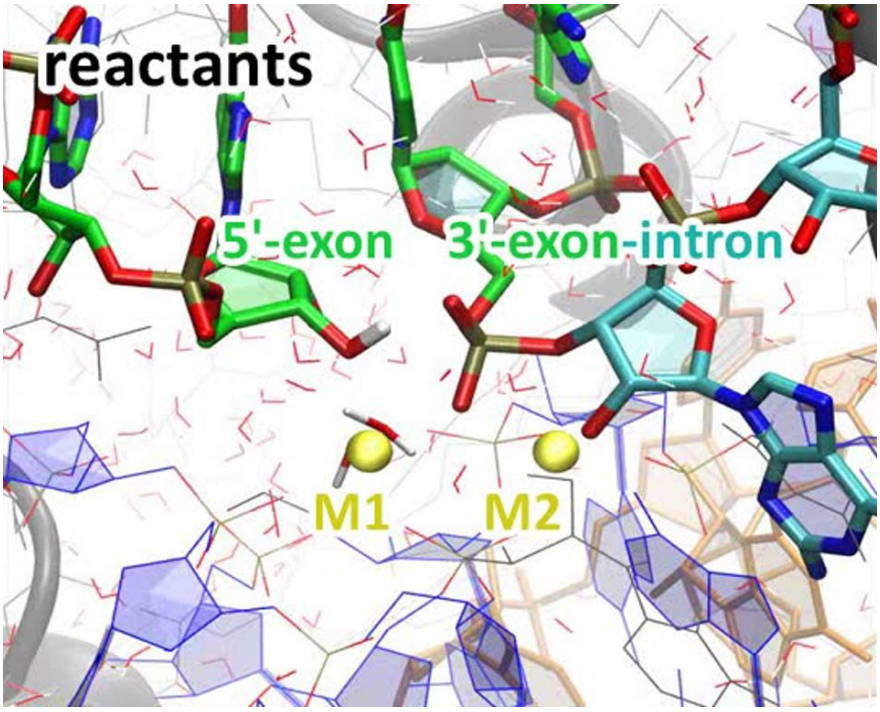
Spliceosome P model			
Cryo-EM 3.6 Å (5YLZ)			
Total number of atoms (water included) = 369767			
CHAIN	MOLECULE	CONSIDERED	RESOLUTION (Å)
A	Prp8	563 - 2085	3.0 ~ 4.5
B	U5 snRNA	85 - 109	3.0 ~ 4.0
D	U6 snRNA	45 - 94	3.0 ~ 4.5
E	mRNA: 5'-exon 3'-intron-exon Intron lariat	 -1 - 10 345 - 355 1 - 8, 295 - 304	3.0 ~ 6.0
F	U2 snRNA	11 - 40	3.0 ~ 6.0
O	Prp46	111 - 447	3.0 ~ 4.5
P	Cwc15	4 - 41	3.0 ~ 4.5
-	Mg+	#5	
-	Na+	#102	
-	Wat	#111357	

Table S2. Monitored distances along the reaction pathway and Mg^{2+}-ions coordination for the wild type model and Arg614>Ala isoform during 10 ps of unrestrained QM/MM MD simulation to assess the effect of mutation on overall stability of the catalytic site.



Model	d1 (Å)	d2 (Å)	d3 (Å)	d4 (Å)	d5 (Å)	d6 (Å)	d7 (Å)	d8 (Å)	Mg-Mg (Å)
wild type	1.68 ± 0.05	3.20 ± 0.13	2.59 ± 0.27	2.16 ± 0.14	2.61 ± 0.30	2.33 ± 0.15	0.99 ± 0.02	2.00 ± 0.23	4.27 ± 0.19
R614>A isoform	1.68 ± 0.05	3.26 ± 0.15	2.52 ± 0.19	2.29 ± 0.21	2.57 ± 0.35	2.46 ± 0.26	1.00 ± 0.02	2.00 ± 0.36	4.29 ± 0.24

Supporting Movie Caption



Movie S1. Second splicing step reaction comprising reactant, first proton transfer, intermediate, second proton transfer and product states. Intron (cyan) and exon (green) are pictured in sticks, Prp8 protein is shown as grey cartoon, and U2 (orange) and U6 (blue) RNAs as lines. M1 and M2 Mg^{2+}-ions are represented as yellow spheres.

Supporting references

(1) Casalino, L.; Palermo, G.; Rothlisberger, U.; Magistrato, A., Who Activates the Nucleophile in Ribozyme Catalysis? An Answer from the Splicing Mechanism of Group II Introns. *J. Am. Chem. Soc.* **2016,** 138, 10374-10377.

(2) Casalino, L.; Jinek, M.; Palermo, G., Two-Metal Ion Mechanism of DNA Cleavage in CRISPR-Cas9, 2019. *Preprint* **2019,** https://doi.org/10.26434/chemrxiv.9784082.v1.

(3) De Vivo, M.; Dal Peraro, M.; Klein, M. L., Phosphodiester Cleavage in Ribonuclease H Occurs Via an Associative Two-Metal-Aided Catalytic Mechanism. *J. Am. Chem. Soc.* **2008,** 130, 10955-10962.

(4) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G., Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc.* **2011,** 133, 8934-8941.

(5) Bai, R.; Yan, C.; Wan, R. X.; Lei, J. L.; Shi, Y. G., Structure of the Post-Catalytic Spliceosome from Saccharomyces Cerevisiae. *Cell* **2017,** 171, 1589-1598.

(6) Yan, C. Y.; Wan, R. X.; Bai, R.; Huang, G. X. Y.; Shi, Y. G., Structure of a Yeast Step II Catalytically Activated Spliceosome. *Science* **2017,** 355, 149-155.

(7) Wan, R. X.; Yan, C. Y.; Bai, R.; Lei, J. L.; Shi, Y. G., Structure of an Intron Lariat Spliceosome from Saccharomyces Cerevisiae. *Cell* **2017,** 171, 120-132.