Supporting Information

Elucidating the Catalytic Reaction Mechanism of Orotate Phosphoribosyltransferase by means of X-ray Crystallography and Computational Simulations

Maite Roca,[*,†] Sergio Navas-Yuste,[‡] Kirill Zinovjev,[#] Miguel López-Estepa,[‡] Sara Gómez,[‡] Francisco J. Fernández,[‡,¶] M. Cristina Vega,[*,‡] Iñaki Tuñón[*,§]

[†]Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain.

[‡] Center for Biological Research (CIB-CSIC), Structural and Chemical Biology, 28040 Madrid, Spain

[#] School of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom.

[§]Departament de Química Física, Universitat de València, 46100 Burjassot, Spain.

[¶]Present address: Abvance Biotech srl, 28003 Madrid, Spain.

*To whom correspondence should be addressed.

M. Roca; mroca@uji.es. Phone: (+34) 964 728069. Fax: (+34) 964 728066

I. Tuñón; Ignacio.Tunon@uv.es. Phone (+34) 963544880. Fax (+34) 963544564

M. C. Vega; cvega@cib.csic.es, cristina.vega@strubicib.org. Phone (+34) 911098070. Fax (+34) 915360432

1. Computational Details

In the *E. coli* structure with OA substrate and sulfate ions, the active site of subunit A contains more sulfate ions interacting with residues of the same subunit and with residues of the catalytic loop of the adjacent subunit (residues 99-109), simulating the PRPP binding mode. Both structures, those of *E. coli* and *S. typhimurium*, were superimposed using PyMOL software[1] and PRPP and Mg^{2+} were added into the active site of chain A of the *E. coli* structure to model the reaction. The PROPKA3.0 program[2-5] was used to estimate the p*K*a values of the titratable protein residues to verify their protonation states at a pH of 7. According to the results, all the lysine and arginine residues are positively charged and all the aspartate and glutamate residues are negatively charged, while the rest of the residues are neutral. Histidines are found in their neutral form.

All protein force field parameters are taken from the AMBER ff14SB force field[6] while substrates OA (in its amide and imidic acid forms) and PRPP were parametrized using the Antechamber program[7] from the AmberTools18 package,[6] based on the general amber force field.[8] The systems with PRPP, Mg^{2+} ion and OA in its amide and imidic acid forms were solvated by a box of TIP3P water molecules[9] with a buffer region of at least 13 Å from any protein or substrate atom to the limits of the box and was also neutralized adding 17 Na^{+} counterions by tleap tool[10] from AmberTools18. Firstly, the systems were minimized using 5000 steps of steepest descent method followed by 5000 steps of conjugate gradient minimization for a maximum of 10,000 minimization steps. After heating the systems to 298 K during a 500 ps molecular dynamic (MD) simulation, 1 ns of NPT MD simulation at 298 K and 1 bar was performed. Finally, a 100 ns MD production was performed using the NVT ensemble. A time step of 2 fs was employed and the SHAKE algorithm[11,12] was used to constraint all the bonds involving hydrogens. Long-range electrostatic interactions were treated by the particle mesh Ewald (PME) method[12-14] and short-range non-bonded interactions were truncated using a cutoff of 8 Å in periodic boundary conditions. Langevin thermostat[15] and Berendsen barostat[16] were used to control temperature and pressure, respectively. Classical MD simulations were performed using the single GPU version of PMEMD named pmemd.cuda from Amber18.[17,18]

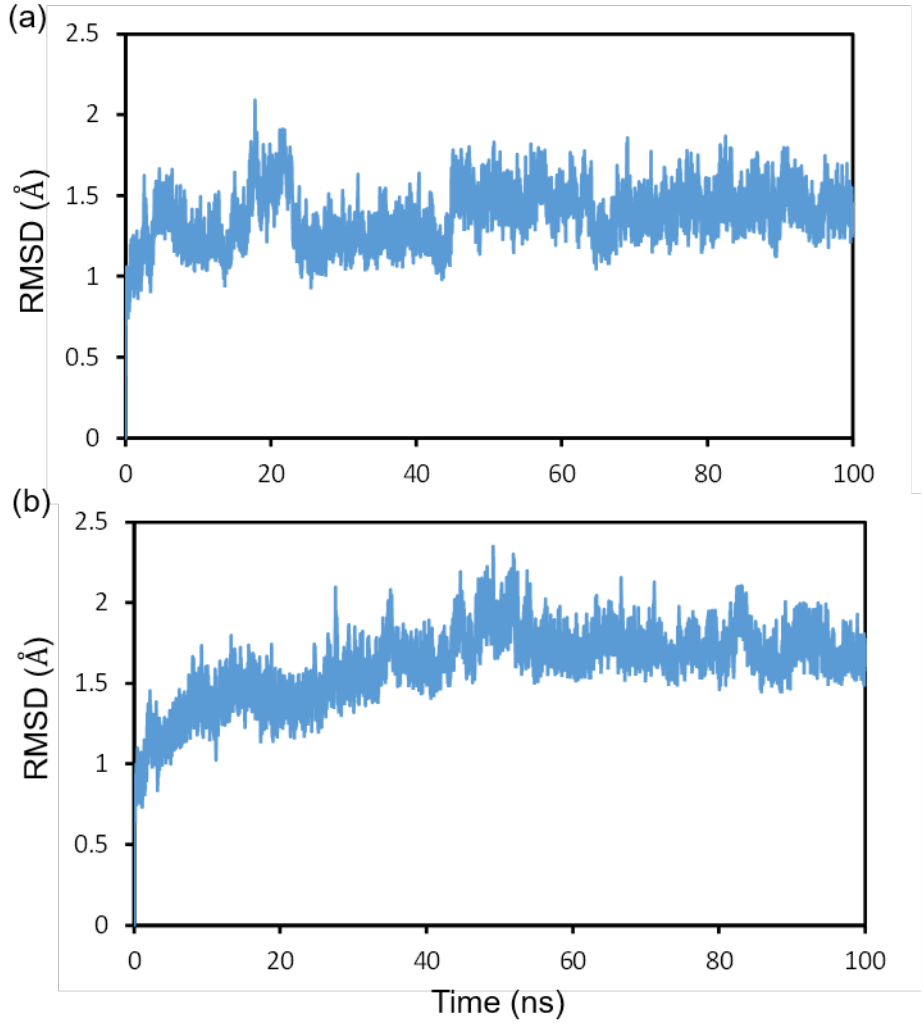

Figure S1. Root mean square deviation (RMSD) for (a) the amide form of OA and (b) the imidic acid form of OA along 100 ns MD simulation. X-axis represents the simulation time (in ns) and Y-axis the RMSD of the atoms belonging to the protein backbone with respect to the first frame (in Å).

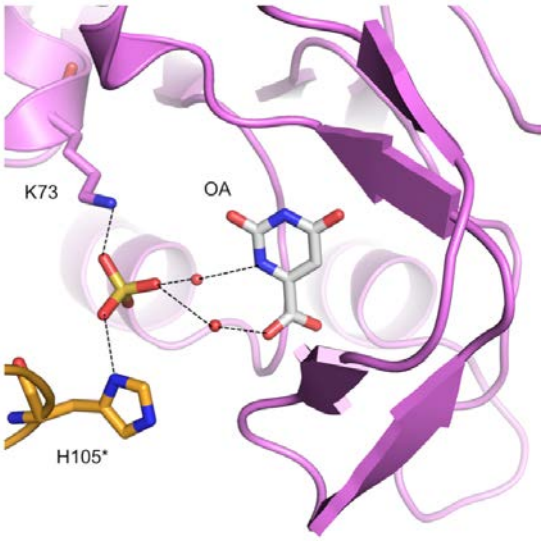

Figure S2. Close-up of the water-mediated interactions observed between the sulfate ion at the entry gate to the active site and the substrate OA.

2. Analysis of distances along the molecular dynamics simulations

The evolution of key distances along the MD simulation of the system with the amide form of OA are provided in Figure S3, S4, S5, S6 and S7; see Figure 3 for the atom and residue naming).

The interaction distances between OA and the neighboring residues are shown in Figure S3. The distance between the main-chain nitrogen atom (N) of Lys26 residue (Lys26N) and O71 atom of the carboxylic group of OA is larger than in the X-ray structures reported in this work (2.95 Å in the OA/SO_4^{2-} complex) and other structures.[19-21] The larger Lys26N-O71 distance can be explained because in the MD structure there is no direct hydrogen bond interaction between these two atoms but a water-mediated hydrogen bond interaction (Figure 3 of the manuscript). The other oxygen atom of the carboxylate (O72) of OA establishes a hydrogen bond interaction to the OG1 of Thr128 residue (Figure S3 and Figure 3 of the manuscript) This interaction was observed in the X-ray structure from *S. typhimurium*.[19] The Phe35 residue makes hydrogen bond interaction between its main-chain nitrogen atom (N) and the O4 atom of OA, and between the main-chain oxygen atom (O) and the N3 atom of OA. Moreover, the O4 atom of OA hydrogen bonds to the guanidinium of Arg156 (NH2). These interactions remain roughly constant along the MD simulation and are present in the X-ray structures of *Ec*OPRTase/OA (Figure 5 of the manuscript) and *Ec*OPRTase/OA/SO_4^{2-} and in the X-ray structures from *S. typhimurium*.[19,20] In the meantime, the distance between the O2 atom of OA and the side-chain nitrogen atom (NZ) of Lys73 (Lys73NZ) is sometimes larger along the MD simulation because, as the interaction with Lys26N, there is a water-mediated hydrogen bond interaction between these two atoms. In the X-ray structures of *Ec*OPRTase/OA/SO_4^{2-}, this distance is also a bit larger to form a hydrogen bond interaction (Figure S2) and in the X-ray structure from *S. typhimurium*, where there is a water-mediated hydrogen bond interaction.[19] In Figure S3, the donor-acceptor distance of the phosphoribosyl transfer reaction (distance between N1 atom of OA and C1 atom of PRPP) is shown to fluctuate in the range of 3-4 Å. Finally, Figure S3 also plots the distance between the N1 atom of OA and the oxygen atom of a water molecule that could be the proton acceptor of the previous step. This distance fluctuates but, on average, the water molecule is kept close to the proton

atom to be transferred. This critical interaction is also observed in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} (Figure S2).

The distances between the oxygen atoms of the 5'-phosphate moiety of PRPP and the surrounding residues along the MD simulation are shown in Figure S4. It is important to note that the three oxygen atoms of the 5'-phosphate form a network of hydrogen bonds with the main-chain nitrogen atoms (N) of Thr128, Ala129, Gly130, Thr131, Ala132 and to the Lys26NZ atom. Thus, the 5'-phosphate group is located within a well conserved PRPP-binding region that anchors the group. The preservation of this interaction network, as observed in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} and in the X-ray structures from *S. typhimurium*,[19,20] during the MD simulation indicates that the active-site pocket provides a very strong environment for PRPP binding (Figure 6 of the manuscript).

The distances between the oxygen atoms of the pyrophosphoryl group of PRPP and the neighboring residues of the same subunit of the active site are plotted in Figure S5 as well as the distances between the ribose hydroxyl groups of PRPP and the two conserved aspartate residues, Asp124 and Asp125. The O3 hydroxyl group forms a hydrogen bond with Asp124 and the O2 hydroxyl group with Asp125 and both of them remained stable along the MD simulation. These two interactions were also observed at the X-ray structures from *S. typhimurium*.[19,20] The Tyr72N and Lys73N atoms establish interactions with the O2B and O3B of the β-phosphate from the pyrophosphoryl group of PRPP, respectively, and the Lys100NZ atom with the O2B atom. These interactions are also present in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} between these residues and a sulfate ion (Figure 6) and the X-ray structures from *S. typhimurium*.[19,20] However, the latter interaction is not kept along all the time of the MD simulation. The same behavior is observed in the distances between Lys100NZ atom with the α-phosphate oxygen atoms (O1A and O3A) of PRPP, and the Lys26NZ atom and the α-phosphate O1A atom of PRPP. These interactions were not observed in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} since the sulfate ions were a bit farther from each other and then, the pyrophosphoryl group of PRPP was not present, but were observed in X-ray structure from *S. typhimurium*.[19]

The distances between the residues of the adjacent chain with PRPP substrate are represented in Figure S6. The hydrogen bond interactions between the guanidinium

nitrogen atoms (NH1 and NH2) of Arg99* and the β-phosphate oxygen atoms (O1B and O2B) of PRPP and between the side-chain nitrogen atom of Lys103* (Lys103*NZ) and O2A and O3A atoms of α-phosphate oxygen atoms PRPP are most of the time maintained along the MD simulation, as observed in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} with the sulfate ions (Figure 6 of the manuscript). In contrast, the hydrogen bond interaction distances between the Lys103*NZ atom and O1B atom of PRPP fluctuates much more along the MD simulation. Figure S6 also displays the distance corresponding to the hydrogen bond interaction between the water molecule, which could be the base to abstract the H1 atom of OA, and the O2A atom of PRPP. This interaction is also kept along the MD simulation and is also reported in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} with a sulfate ion (see the water molecule between N1 atom of OA and an oxygen atom of a sulfate ion in Figure S2).

Some of the interactions observed in the X-ray structure of *Ec*OPRTase/OA/SO_4^{2-} with the sulfate ion that mimics the α-phosphate of PRPP (Figure 6 of the manuscript) are not conserved during MD simulations. The strong interaction made by the positively charged lysine side chain of Lys103* (Lys103*NZ atom) and the oxygen atoms of α-phosphate of PRPP (O2A and O3A atoms) is maintained as observed in Figure S6. Thus, Lys103* residue, which belongs to the adjacent subunit, may play an important role in keeping the flexible loop in a closed or open conformation. However, the hydrogen bond interaction with His105* is lost due to the presence of more positively charged lysine residues such as Lys26 and Lys100 that interact with the oxygen atoms of α-phosphate moiety of PRPP (Figure 3 of the manuscript and Figure S5). Furthermore, the hydrogen bond interactions between Lys73 and the oxygen atoms of α-phosphate of PRPP, which are displayed in Figure S7, became weaker because Lys73 is closer to the O2 atom of OA (Figure S3).

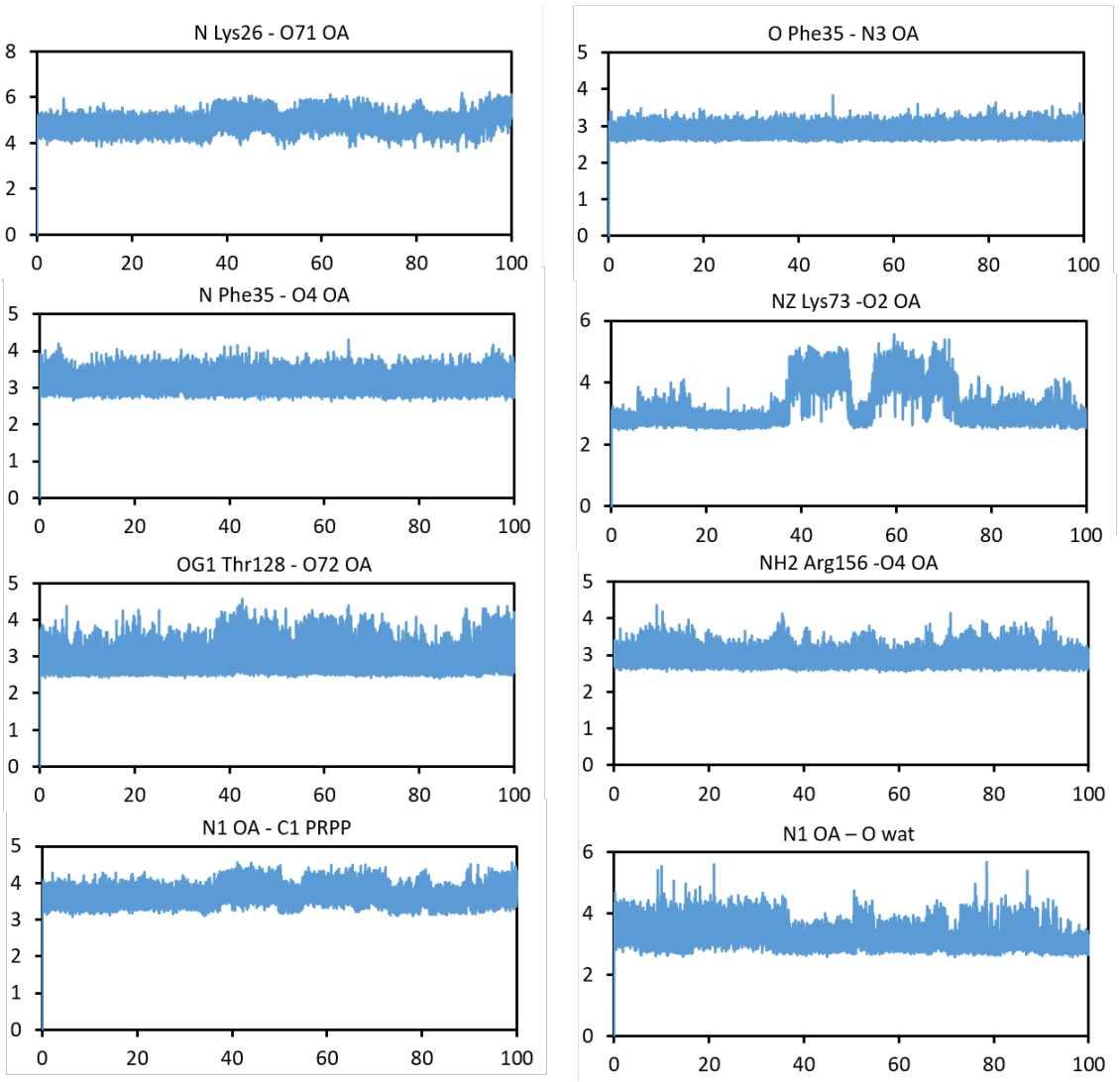

Figure S3. Time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between OA and the surrounding residues, the donor and acceptor distance of the phosphoribosyl transferase reaction between OA and PRPP (N1 OA - C1 PRPP) and the hydrogen bond interaction between a water molecule and N1 atom of OA along 100 ns of MD simulations. X-axis represents the simulation time (in ns) and Y-axis the distances (in Å). The title of each graph shows the distance plotted.

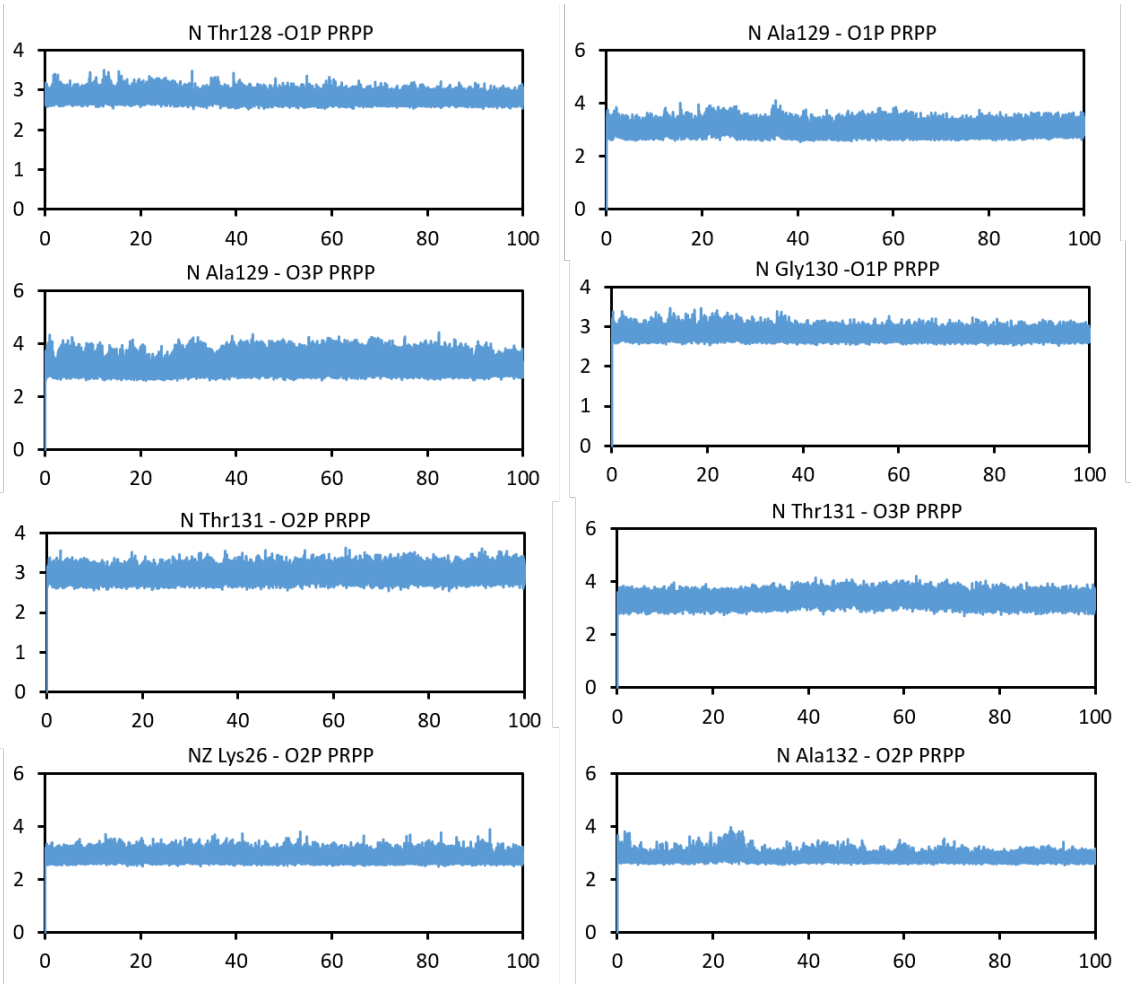

Figure S4. Time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the 5'-phosphate group of PRPP ligand and the surrounding residues along 100 ns of MD simulations. X-axis represents the simulation time (in ns) and Y-axis the distances (in Å). The title of each graph shows the distance plotted.

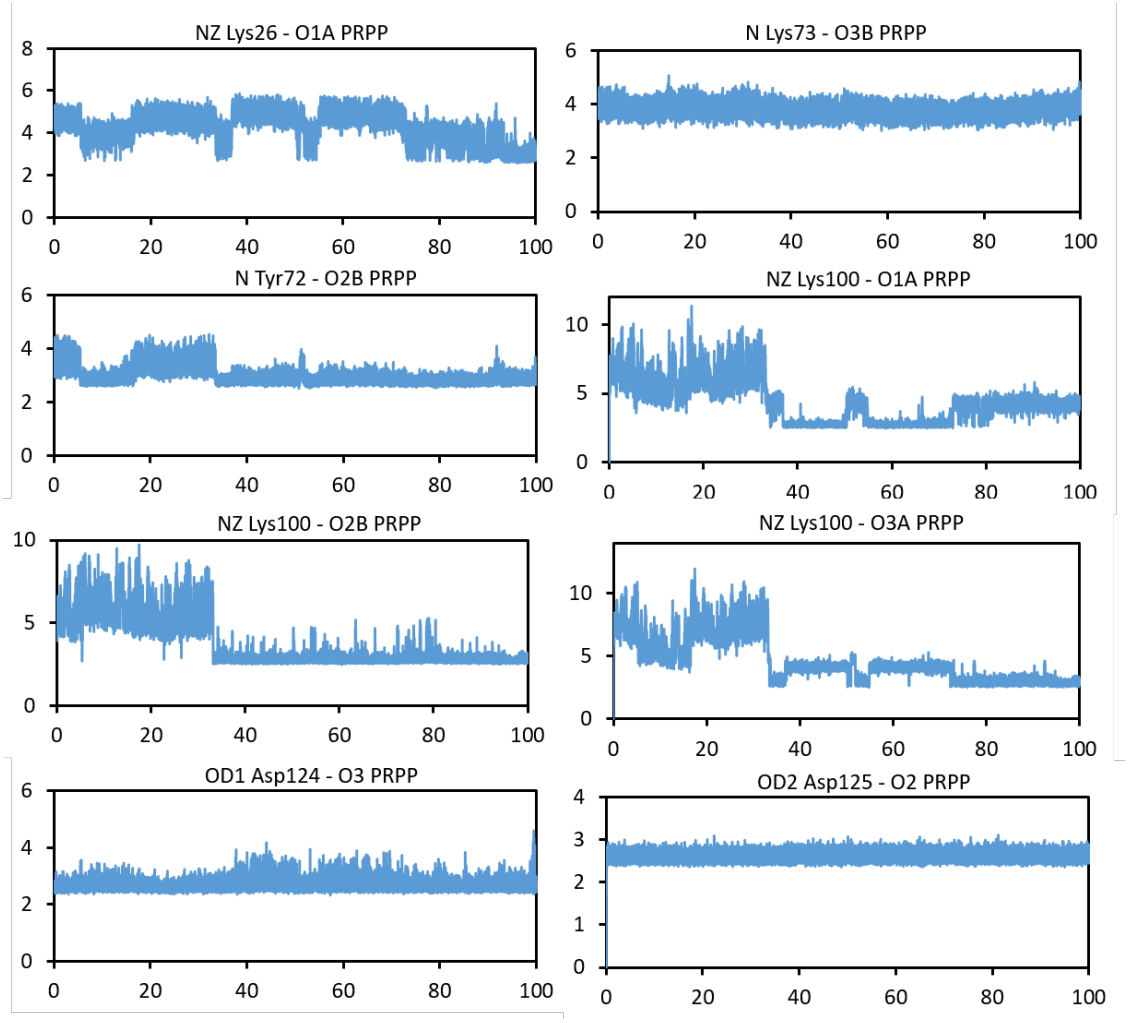

Figure S5. Time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the oxygen atoms of pyrophosphoryl group and the ribose hydroxyl groups of PRPP and the surrounding residues along 100 ns of MD simulations. X-axis represents the simulation time (in ns) and Y-axis the distances (in Å). The title of each graph shows the distance plotted.

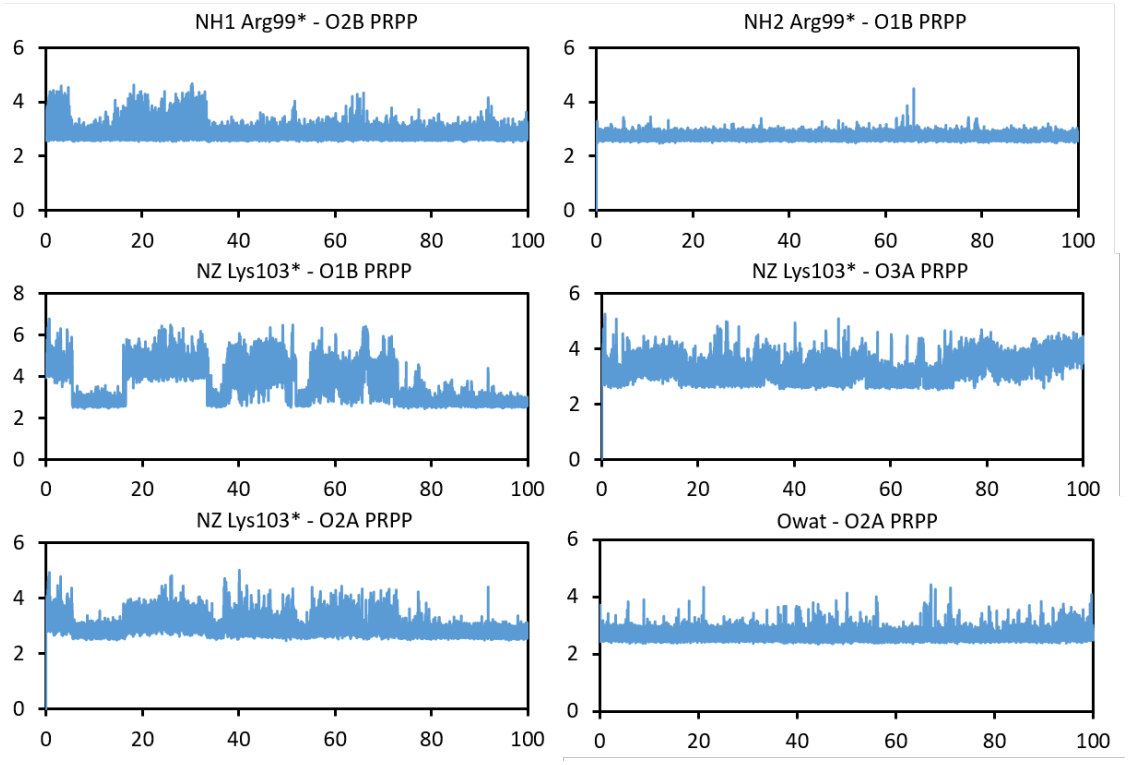

Figure S6. Time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the oxygen atoms of pyrophosphoryl group of PRPP and the surrounding residues of the adjacent subunit (denoted as *) and the hydrogen bond interaction between a water molecule and the O2A atom of PRPP along 100 ns of MD simulations. X-axis represents the simulation time (in ns) and Y-axis the distances (in Å). The title of each graph shows the distance plotted.

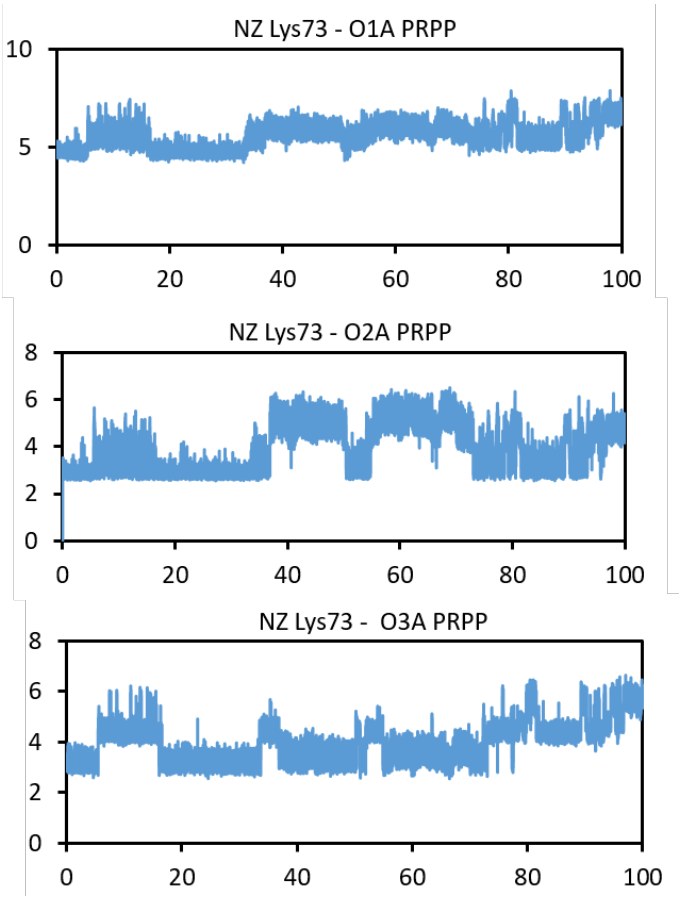

Figure S7. Time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the oxygen atoms of α-phosphate of PRPP and side-chain nitrogen atom of Lys73 residue along 100 ns of MD simulations. X-axis represents the simulation time (in ns) and Y-axis the distances (in Å). The title of each graph shows the distance plotted.

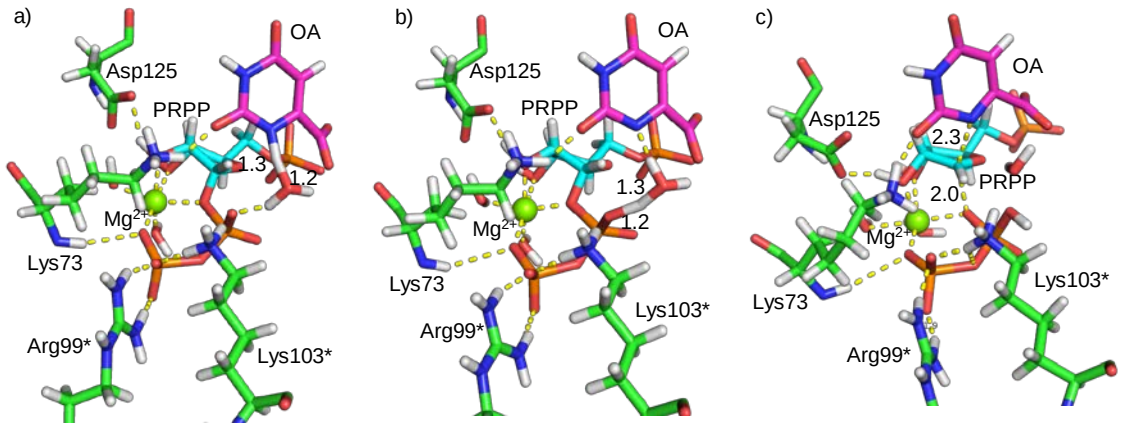

Figure S8. Localized transition state (TS) structures for the three steps of the most favorable mechanism a) for the first step, b) for the second step and c) for the last step, in the active site of OPRT enzyme by means of PM6/MM energy optimizations. The carbon atoms of PRPP and OA are shown in light blue and pink, respectively. The bond breaking and the bond forming distances are shown in Å. The octahedral coordination of the Mg^{2+} ion and the hydrogen bond interactions of key residues to stabilize the TS structures are depicted as dashed lines.

Table S1. Average values of some key distances and the angle that defines the chemical reaction of the nucleophilic attack (with their standard deviations) at the reactant state (R), reactant state previous to the rate-limiting step (R') and transition state (TS) along the MD simulation of the PMF. The distances are in Å and the angle in degrees.

	Mechanism 1			**Mechanism 2**			**Mechanism3**		
Distances	**R**	**R'**	**TS**	**R**	**R'**	**TS**	**R**	**R'**	**TS**
d(N1 OA, C1 PRPP)	3.38 ± 0.18	2.8 ± 0.16	2.34 ± 0.10	3.72 ± 0.23	3.05 ± 0.15	2.22 ± 0.09	3.66 ± 0.19	3.17 ± 0.16	2.53 ± 0.10
d(C1 PRPP, O1 PRPP)	1.43 ± 0.03	1.47 ± 0.03	2.04 ± 0.12	1.40 ± 0.04	1.48 ± 0.03	2.06 ± 0.10	1.46 ± 0.02	1.48 ± 0.04	2.10 ± 0.12
ang(N1, C1, O1)	153 ± 6	151 ±5	166 ± 5	125 ± 9	130 ± 6	153 ± 5	131 ± 7	133 ± 6	149 ± 7
d(O1 PRPP, Mg^{2+})	2.22 ± 0.10	2.20 ± 0.10	2.02 ± 0.07	2.09 ± 0.15	2.05 ± 0.12	2.15 ± 0.11	2.36 ± 0.13	2.30 ± 0.12	2.20 ± 0.10
d(O2 PRPP, OD2 Asp125)	3.05 ± 0.12	2.8 ± 0.10	2.73 ± 0.11	2.92 ± 0.09	2.80 ± 0.13	2.85 ± 0.10	2.83 ± 0.12	2.82 ± 0.11	2.80 ± 0.12
d(O3B PRPP, N Lys73)	3.6 ± 0.20	3.7 ± 0.10	3.43 ± 0.20	3.34 ± 0.23	3.50 ± 0.21	3.70 ± 0.22	3.43 ± 0.19	3.41 ± 0.16	3.44 ± 0.17
d(O2B PRPP, N Tyr72)	3.02 ± 0.11	3.11 ± 0.10	2.93 ± 0.10	3.21± 0.13	3.43 ± 0.12	3.12 ± 0.12	2.91 ± 0.09	2.95 ± 0.10	2.94 ± 0.08
d(O1B PRPP, NH2 Arg99*)	2.98 ± 0.10	2.82 ± 0.10	2.79 ± 0.10	2.80 ± 0.12	2.94 ± 0.13	2.83± 0.12	2.89 ± 0.11	2.87 ± 0.12	2.88 ± 0.12
d(O2B PRPP, NH1 Arg99*)	2.95 ± 0.11	2.90 ± 0.10	2.81 ± 0.10	2.88 ± 0.12	2.93 ± 0.12	3.00 ± 0.11	2.93 ± 0.10	2.94 ± 0.11	2.89 ± 0.11
d(O1B PRPP NZ Lys103*)	2.85 ± 0.10	2.73 ± 0.10	2.70 ± 0.09	2.90 ± 0.12	2.82 ± 0.12	2.92 ± 0.10	2.94 ± 0.11	2.94 ± 0.12	2.91 ± 0.12
d(O3A PRPP NZ Lys103*)	3.50 ± 0.22	3.61 ± 0.10	2.86 ± 0.16	3.10 ± 0.14	3.63 ± 0.15	3.10 ± 0.12	3.30 ± 0.13	3.29 ± 0.12	3.22 ± 0.12

References

(1) https://pymol.org/2/ **2019**.

(2) Li, H.; Robertson, A. D.; Jensen, J. H. Very Fast Empirical Prediction and Rationalization of Protein pK(a) Values. *Proteins.* **2005**, *61*, 704-721.

(3) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very Fast Prediction and Rationalization of pK(a) Values for Protein-Ligand Complexes. *Proteins.* **2008**, *73*, 765-783.

(4) Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525-537.

(5) Sondergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK(a) Values. *J. Chem. Theory Comput.* **2011**, *7*, 2284-2295.

(6) Case, D. A.; Ben-Shalom, I. Y.; S.R. Brozell, S. R.; Cerutti, D. S.; Cheatham, I., T.E.;; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; M.K.Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 18, University of California, San Francisco, **2018**.

(7) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247-260.

(8) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157-1174.

(9) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(10) Schafmeister, C. E. A. F.; Ross, W. S.; Romanovski, V. Universtity of California, LEAP, San Francisco, **1995**.

(11) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341.

(12) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - An N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

(13) York, D. M.; Darden, T. A.; Pedersen, L. G. The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - A Comparison of the Ewald and Truncated List Methods. *J. Chem. Phys.* **1993**, *99*, 8345-8348.

(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

(15) Grest, G. S.; Kremer, K. Molecular-Dynamics Simulation for Polymers in the Presence of a Heat Bath. *Phys. Rev. A.* **1986**, *33*, 3628-3631.

(16) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.

(17) Le Grand, S.; Goetz, A. W.; Walker, R. C. SPFP: Speed without Compromise-A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374-380.

(18) Salomon-Ferrer, R.; Goetz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878-3888.

(19) Grubmeyer, C.; Hansen, M. R.; Fedorov, A. A.; Almo, S. C. Structure of Salmonella Typhimurium OMP Synthase in a Complete Substrate Complex. *Biochemistry.* **2012**, *51*, 4397-4405.

(20) Scapin, G.; Ozturk, D. H.; Grubmeyer, C.; Sacchettini, J. C. The Crystal-Structure of the Orotate Phosphoribosyltransferase Complexed with Orotate and Alpha-D-5-Phosphoribosyl-1-Pyrophosphate. *Biochemistry.* **1995**, *34*, 10744-10754.

(21) Gonzalez-Segura, L.; Witte, J. F.; McClard, R. W.; Hurley, T. D. Ternary Complex Formation and Induced Asymmetry in Orotate Phosphoribosyltransferase. *Biochemistry.* **2007**, *46*, 14075-14086.