

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-Esteva, Sara Gómez, Francisco J. Fernández, M. Cristina Vega,* and Iñaki Tuñón*



Cite This: *ACS Catal.* 2020, 10, 1871–1885



Read Online

ACCESS |



Metrics & More



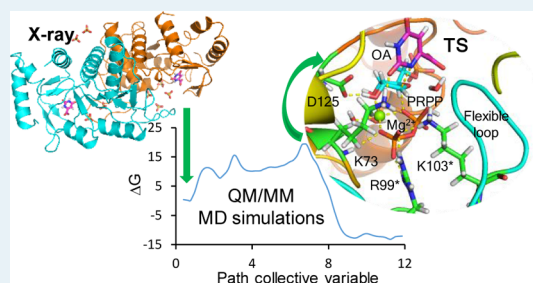
Article Recommendations



Supporting Information

ABSTRACT: Orotate phosphoribosyltransferase (OPRTase) catalyzes the reaction between the ribose donor α -D-5-phosphoribosyl-1-pyrophosphate (PRPP) and orotate (OA) in the presence of Mg^{2+} ion to obtain pyrophosphate and pyrimidine nucleotide orotidine 5'-monophosphate (OMP), a key precursor in de novo biosynthesis of pyrimidine nucleotides. In this work, several structures of the dimeric *Escherichia coli* OPRTase (EcOPRTase) have been determined at high resolution, and kinetic measurements have been carried out to obtain the catalytic rate and Michaelis constants. Molecular dynamics (MD) simulations have been carried out, and structural analysis from the X-ray and MD simulation structures reveals conformational changes related to the flexible catalytic loop that establishes hydrogen bond interactions with the pyrophosphoryl group of PRPP. It is proposed that the OA substrate can be in equilibrium in its tautomeric forms. Starting from the most stable tautomeric form, all the plausible mechanisms have been explored by means of quantum mechanics/molecular mechanics (QM/MM) MD simulations using the adaptive string method. The most feasible mechanism consists of the proton transfer from the N1 atom of OA to a water molecule and from the water molecule to the α -phosphate O2A atom of PRPP. After that, the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP proceeds to yield OMP and pyrophosphate. The free energy barrier obtained is in very good agreement with the experimental data reported. Analysis of some relevant distances between key residues and the substrates (OA and PRPP) at the reactant state and transition state (TS) of the rate-limiting step allows us to understand the role of some conserved residues (Lys73, Asp125, Lys103*, Arg99*, and Mg^{2+} ion) electrostatically stabilizing the TS and preserving the flexible catalytic loop in a closed conformation during the enzymatic reaction.

KEYWORDS: X-ray structure, orotate phosphoribosyltransferase, pyrimidine metabolism, MD simulations, QM/MM methods, reaction mechanism, enzyme catalysis



1. INTRODUCTION

Phosphoribosyltransferases (PRTases) are involved in the synthesis of pyrimidine nucleotides, the key precursors of DNA and RNA, and also in the synthesis of some amino acids such as histidine and tryptophan and the pyridine coenzymes nicotinamide adenine dinucleotide (NAD) and NAD phosphate.¹ In particular, orotate PRTase (OPRTase) is an enzyme that catalyzes the formation of the pyrimidine nucleotide orotidine 5'-monophosphate (OMP) and pyrophosphate from the ribose donor α -D-5-phosphoribosyl-1-pyrophosphate (PRPP) and orotate (OA) in the presence of Mg^{2+} ion (Scheme 1). The product OMP is converted by OMP decarboxylase to uridine 5'-monophosphate, the precursor of all pyrimidine nucleotides.

OPRTases are widely distributed among different organisms such as insects, bacteria, fungi, and mammals, *Plasmodium falciparum* (PfOPRTase),² *Mycobacterium tuberculosis* (MtOPRTase),³ *Salmonella typhimurium* (StOPRTase),⁴ *Sac-*

charomyces cerevisiae (ScOPRTase),⁵ and humans (*HsOPRTase*)⁶ to mention a few. For all these organisms, de novo pyrimidine biosynthesis is the major pathway of nucleotide production.

P. falciparum is the most virulent parasite that causes human malaria. The development of new antimalarial drugs is urgently required because of the increased resistance of *P. falciparum* parasite to current therapeutic treatments.⁷ *M. tuberculosis* is the bacterium that causes the severe human infectious tuberculosis disease, and the rise of drug resistance to tuberculosis represents an important threat for public health.⁸ Thus, structural investigations for the discovery of new drugs

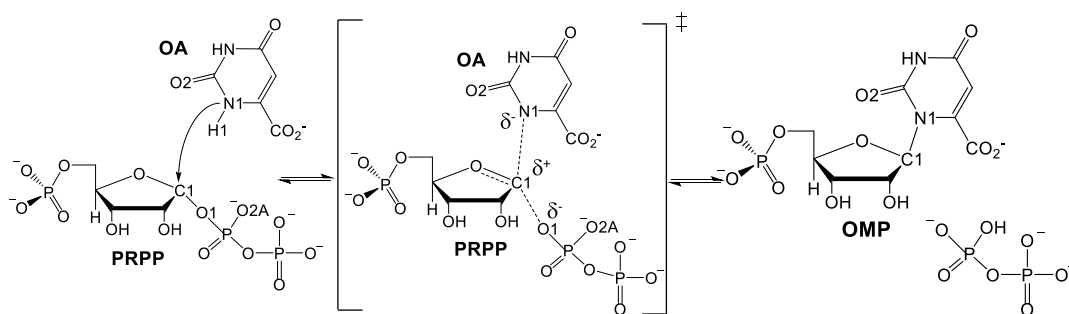
Received: December 8, 2019

Revised: December 30, 2019

Published: January 2, 2020



Scheme 1. Enzymatic Reaction between PRPP and OA To Yield OMP and Pyrophosphate Catalyzed by OPRTase in the Presence of Mg^{2+}



for tuberculosis treatment are a top priority.⁹ In humans, inherited defects in the nucleotide synthases lead to severe inborn disorders of metabolism, such as orotic aciduria, 2,8-dihydroxyadenine lithiasis, and Lesch–Nyhan syndrome causing human pathologies such as gout and mental retardation.¹⁰ Moreover, *HsOPRTase* plays a key role in quickly proliferating cells to satisfy the increased demand for nucleic acid synthesis through de novo pyrimidine biosynthesis. Therefore, new ways of interrupting nucleotide synthesis in rapidly proliferating human cancer cells may offer alternative therapeutic options for this disease. Indeed, therapies against pyrimidine production are already in use for the treatment of autoimmune diseases and malignant neoplastic disease such as cancer.^{11,12}

All in all, the diseases mentioned above can be treated by the inhibition of the OPRTase enzyme by means of blocking the OMP production. Thus, OPRTase is an attractive target for the rational design of antimalarial, antitubercular, and anticancer drugs. For that purpose, it is important to elucidate the reaction mechanism of OPRTase in a detailed molecular way to understand its enzymatic action and perform a thorough analysis of the transition state (TS) structure to enable the computational design of novel and potent drugs resembling the TS structure in order to control all these diseases. Computational optimizations of the TS structures in vacuo for *S. typhimurium*, *P. falciparum*, and human have been performed^{13–15} using a small model of the system, and TS analogue (TSA) inhibitors have been proposed.^{16,17} However, consideration of the complex effects of the protein environment, including its flexibility, has lagged behind even though it would allow a better characterization of the reaction mechanism, TS structures, and the role of protein residues in catalysis.

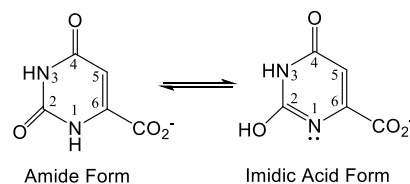
The vast majority of the OPRTase structures found in the literature are composed of two subunits with a flexible loop over each catalytic site with a well-conserved amino acid sequence and a well-defined PRPP binding domain.^{2,18–20} Nuclear magnetic resonance and proteolysis studies²¹ have characterized the dynamics of the flexible loop, which moves between its open and closed positions to interact with bound substrates and sequester them from bulk solvent.²² The main characteristic of this flexible loop (residues 99–109 in *S. typhimurium* and in *Escherichia coli*)^{18,20} is that it closes the active site of the adjacent subunit establishing hydrogen bond interactions with the PRPP substrate. This feature has been proposed to be essential for PRPP binding and catalysis.^{18,20,23} The OPRTase structures reveal that the enzyme undergoes different conformational changes depending on the complex formed (apoenzyme, primary complex with OA, binary

product complex with OMP, or ternary complex with OA, Mg^{2+} , and PRPP). While the apoenzyme is similar between two microorganisms (*S. cerevisiae* and *S. typhimurium*), large conformational changes take place in those structures that contain the substrates.¹⁹ Moreover, a structural asymmetry has been described for some OPRTase structures of the ternary complex, whereby the flexible catalytic loop of one active site is closed and prepared to catalyze the reaction, while the other loop remains open into the solvent.^{18,20} Then, the dimeric architecture of OPRTase plays an essential role in the catalysis because the flexible catalytic loop interacts with the substrates of the active site of the adjacent subunit, thus demonstrating the cooperativity of OPRTase.

The OPRTase reaction mechanism is under debate. In the literature, it has been described that the reaction proceeds with inversion of stereochemistry at the anomeric carbon C1 (Scheme 1), and a loosely oxocarbenium-like TS for phosphoribosyl transfer has been proposed⁴ to develop through an S_N1 -like mechanism.²⁴ However, the precise pathway for the proton transfer (H1, in Scheme 1) from the N1 atom of OA to the enzyme or PRPP is still unclear.

It has been proposed that OA could be in equilibrium in its tautomeric forms (the amide and imidic acid forms) and the imidic acid form might be the reactant state existing in the active site of OPRTase (Scheme 2).¹⁹ A viable chemical

Scheme 2. Equilibrium between the Tautomeric Forms of OA



mechanism for activation of the N1 atom is the stabilization of the imidic acid form of OA in the enzyme active site (Scheme 2). In this tautomeric form, the N1 atom has been activated through proton transfer to the O2 atom, now found in its hydroxyl form. An examination of the OPRTase ternary complex structures^{19,20} shows that there are two ordered water molecules that interact with the O2 atom of OA. These water molecules interact in turn with two lysine residues (Lys73 and Lys103 of the adjacent subunit), one aspartate residue (Asp125) and the other oxygen atom of the pyrophosphate product (residue numbering from *S. typhimurium* and *E. coli*). Thus, it is reasonable to assume that the acceptor of the proton

coming from either the N1 atom (amide form) or O2 atom (imidic acid form) is probably one of these water molecules.

With regard to the role of the residues in the enzymatic reaction, mutagenesis studies of the conserved lysine (Lys19, 26, 73, 100, and 103)²⁵ and aspartate residues (Asp124 and 125)²⁶ of *S. typhimurium* have been carried out. In the study of lysine residues, Lys73 and Lys103 mutations to alanine or glutamine residues produced a greater decrease in the catalytic rate constant, the Lys103 mutation causing the most significant reduction. In the case of the aspartate residues, mutation to asparagine residues resulted in an important loss of enzymatic activity. As a consequence, these residues must play remarkable roles in catalysis.

In order to unveil enzymatic reaction mechanisms, the associated free energy landscapes must be explored. However, exploration of a mechanistic proposal may require the consideration of a highly dimensional space. In addition, for the case under study here, several mechanistic routes are possible, and all of them must be analyzed. To make this study affordable, we took advantage of the fact that for each mechanism most reactive trajectories are confined within a narrow tube that goes along the minimum free energy path (MFEP) of the reaction.^{27,28} A methodological strategy that allow us to explore different possible mechanisms at an affordable computational cost is provided by the string method^{29,30} and, in particular, its combination with a path collective variable that defines the reaction coordinate (RC) as the advance of the system along the MFEP.³¹

In this work, we have combined X-ray crystallography with quantum mechanics/molecular mechanics (QM/MM) molecular dynamics (MD) simulations using the adaptive string method to provide novel insights into the OPRTase reaction mechanism. A more accurate understanding of this mechanism and an analysis of the TS structure should facilitate the discovery of new inhibitors capable of blocking the synthesis of OMP for the control of deadly infectious diseases such as malaria and tuberculosis as well as for developing novel compounds with antitumoral activity.

2. MATERIALS AND METHODS

2.1. Recombinant EcOPRTase Production. The gene encoding full-length *E. coli* OPRTase (EcOPRTase) was amplified by polymerase chain reaction from the genomic DNA of *E. coli* BL21 using forward and reverse primers incorporating NcoI and XhoI restriction sites, which was subsequently ligated into similarly digested pET21b. The reverse primer carried a stop codon so that the pET21b:EcOPRTase plasmid would drive the expression of untagged EcOPRTase. The correct clone was verified by sequencing. For expression, the recombinant plasmid was transformed into *E. coli* BL21(DE3) cells. One liter of Luria–Bertani broth supplemented with 100 $\mu\text{g}/\text{mL}$ of ampicillin was inoculated with 1 mL of an overnight starter culture and allowed to grow at 37 °C with shaking until an optical density at 600 nm of 0.8 was reached. Then, isopropyl β -D-1-thiogalactopyranoside was added to a final concentration of 1 mM to induce expression, which was continued for 8 h at 37 °C. Cells were harvested by centrifugation and stored at –80 °C until used. For lysis, the pellet was resuspended in buffer A (25 mM Tris-HCl, pH 7.6, 0.5 mM ethylenediaminetetraacetic acid, and 2 mM 2-mercaptoethanol) supplied with protease inhibitor (1 mM phenylmethylsulfonyl fluoride) and DNase (0.2 μL benzonase). The suspension was then lysed by sonication and

clarified by centrifugation (15,000g for 30 min at 4 °C) and filtration through a 0.45 μm membrane. EcOPRTase was purified in three steps. The first step consisted of a 10 min heat shock at 62 °C, followed by centrifugation and filtration (as before). In the second step, the clarified heat-shocked lysate was loaded onto a 5 mL HiTrap Q anion exchange chromatography column (GE Healthcare) pre-equilibrated with buffer A. Unbound impurities were removed by washing the column with 10 column volumes of buffer A. Protein elution was brought about by a linear NaCl gradient from 0 to 1 M over 20 column volumes. Peak fractions containing EcOPRTase were pooled, concentrated, and loaded onto a HiLoad 16/60 Superdex 200 size-exclusion chromatography column pre-equilibrated with buffer A. EcOPRTase eluted as a single symmetric peak. Fractions were again concentrated to 12 mg/mL, and either fractions were flash-frozen in 50- μL aliquots (for structural studies) or sterile glycerol was added to a final concentration of 55% (v/v) before freezing (for enzymatic studies) and stored away at –80 °C.

2.2. Enzymatic Activity Assay. A continuous spectrophotometric method was used to follow the forward reaction catalyzed by EcOPRTase at 25 °C, which measured the decreasing OA concentration as previously described.² The assay was performed in UV-compatible plastic cuvettes with a BioSpectrometer Kinetics (Eppendorf) spectrophotometer equipped with a temperature-controlled cuvette holder. The reaction mixture (150 μL) contained 20 mM Tris-HCl, pH 7.6, 2 mM MgCl_2 , 100 μM PRPP, 1 μg EcOPRTase (28 μM), and 5–100 μM OA.^{2,32} The enzyme was incubated for 1 min with all components except PRPP. The reaction was started by the addition of 100 μM PRPP, and the linear progress curve of absorbance change at 295 nm was followed over 2 min. Molar extinction coefficients at 295 nm were calculated for all substrates under experimental conditions; for OA, the molar extinction coefficient was 4.11 $\text{mM}^{-1} \text{cm}^{-1}$ (compare with 3.67 $\text{mM}^{-1} \text{cm}^{-1}$ in other publications). Initial velocity measurements were fitted to the Michaelis–Menten model to calculate the kinetic parameters with SigmaPlot v13 (Systat Software Inc.).

2.3. Crystallization and Data Collection. EcOPRTase was crystallized with an empty active site and in complex with OA. There are two crystals with OA depending on the presence or absence of sulfate ions, which mimic the second substrate (PRPP) and the reaction's product (OMP). Sitting-drop vapor-diffusion experiments were set up at 10 mg/mL by mixing 1 μL of protein and 1 μL of the crystallization solution. For the empty active-site crystals and the crystals with OA without sulfate anions, the crystallization condition was 0.1 M sodium acetate, pH 5.0, 0.2 M MgCl_2 , and 20% (w/v) PEG 6000. For crystals with both OA and sulfate anions, however, the crystallization condition was 0.3 M sodium citrate, pH 4.8, 2.1 M $(\text{NH}_4)_2\text{SO}_4$, and 0.2 M NaCl. In all cases, crystals of various morphologies (cubic, bipyramidal, and rhombohedral plates) appeared within 4–6 d. Crystals were cryoprotected in 20% (v/v) glycerol or CryoMix 3 (CryoProtX, Molecular Dimensions) immediately before flash-freezing in liquid nitrogen. Complete X-ray diffraction datasets were collected from all three crystal classes at the XALOC beamline (ALBA, Barcelona, Spain). Crystals of EcOPRTase with an empty active site belonged to the $P2_12_12$ (num. 18) orthorhombic space group with unit-cell dimensions $a = 75.5 \text{ \AA}$, $b = 131.8 \text{ \AA}$, $c = 39.5 \text{ \AA}$, and they diffracted up to 1.55 \AA resolution. The two crystal forms for EcOPRTase in complex with OA

belonged to a related space group ($P2_12_12_1$, num. 19) or a different space group ($P4_12_12$, num. 92). The former had unit-cell dimensions of $a = 53.7$ Å, $b = 69.5$ Å, $c = 104.4$ Å, and it diffracted to 1.25 Å, while the latter had unit-cell dimensions $a = b = 83.2$ Å, $c = 123.2$ Å, and it diffracted to 1.60 Å. Data were processed with XDS^{33,34} and scaled and merged with Aimless³⁵ from the CCP4 suite of programs.³⁶ Detailed data processing and merging statistics are shown in Table 1.

2.4. Structure Determination and Refinement. The crystal structures of EcOPRTase were solved by molecular replacement with PHASER.³⁷ The search model for the empty active-site structure, the OA complex structure without sulfate anions, and the OA complex structure with sulfate anions was PDB 1ORO after removing the solvent molecules.¹⁸ After initial placement, an omit map was calculated to ascertain the presence of substrates in the active site. Complete atomic models were built by iterative model building in Coot³⁸ and maximum likelihood refinement with phenix.refine.³⁹ The quality of the intermediate and final macromolecular models was gauged with MolProbity.⁴⁰ Crystallographic refinement and validation statistics are presented in Table 1.

The coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with accession codes 6TAI (EcOPRT), 6TAJ (EcOPRT/OA), and 6TAK (EcOPRT/OA/SO₄²⁻).

2.5. MD Simulations. The OA complexes from *E. coli* (OA with the presence or absence of sulfate ions) are dimers and contain the substrate OA in the two active sites of the two subunits, but the Mg²⁺ ion and PRPP are missing. The dimeric crystal structure of *S. typhimurium* OPRTase, which resembles the structure of *E. coli*, in complex with the Mg²⁺ ion, PRPP, and OA (PDB entry 1LH0)²⁰ was used as a template to build the starting structure for the Michaelis complex. Details of preparing the systems for the MM MD simulations are given in the Computational Details section in the Supporting Information.

Once the systems with the amide and imidic acid form of OA were built, minimized, heated, and pre-equilibrated, 100 ns of MM MD simulations was performed using Amber force field from Amber18 package.⁴¹ All the computational details are described in the Supporting Information. According to the time-dependent evolution of the root-mean-square deviation (rmsd) of the protein backbone (Figure S1 in the Supporting Information), the system can be considered equilibrated after 50 ns. The configurations of the active sites after the MM equilibration are shown in Figure 1. It is important to mention that the octahedral coordination of Mg²⁺ (with four oxygen atoms of PRPP and two water molecules) remained intact during the MD simulations.

From the analysis of the crystal structures and the MD simulations of the system with the amide form of OA (Figure 1a), we investigated which residues could act as a base to abstract the proton from the OA substrate, and three different mechanisms were proposed (Figure 2). In the first mechanism (mechanism 1 in Figure 2), as shown in Figure 1a, a water molecule found near OA makes hydrogen bonding interactions with the H1 atom of OA and the O2A atom of the α -phosphate group of the pyrophosphoryl fragment of PRPP. This activated water molecule could conceivably act as a base (specific base) to extract the H1 proton from the N1 atom in the amide form of OA, subsequently shuttling it to the O2A atom of PRPP. After this water-mediated proton transfer, the N1 atom can perform a nucleophilic attack on the C1 atom of

Table 1. X-ray Diffraction Data Collection and Refinement Statistics^a

	EcOPRTase (empty)	EcOPRTase (OA)	EcOPRTase (OA/SO ₄ ²⁻)
PDB code	6TAI	6TAJ	6TAK
Data Collection			
wavelength (Å)	1.03076	0.97926	0.97926
space group	$P2_12_12$	$P4_12_12$	$P2_12_12$
Cell Dimensions			
a, b, c (Å)	75.46, 131.75, 39.52	83.12, 83.12, 122.79	53.66, 69.47, 104.38
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
resolution range (Å)	49.63–1.55 (1.56–1.55)	49.38–1.60 (1.64–1.60)	47.72–1.25 (1.28–1.25)
total reflections	390,120 (2808)	720,144 (44,233)	452,457 (12,799)
unique reflections	57,472 (510)	57,672 (3947)	104,599 (5479)
completeness (%)	98.9 (89.5)	99.5 (93.6)	95.9 (68.8)
multiplicity	6.8 (5.5)	12.5 (11.2)	4.3 (2.3)
$\langle I \rangle / \sigma(I)$	20.8 (1.6) ^b	15.1 (0.8)	13.2 (0.6)
R_{meas}^c	0.047 (0.997)	0.079 (2.372)	0.054 (1.722)
$CC1/2^d$	0.999 (0.680)	0.999 (0.442)	0.999 (0.292)
Refinement and Validation			
resolution (Å)	49.63–1.55 (1.58–1.55)	49.38–1.60 (1.64–1.60)	47.72–1.25 (1.28–1.25)
reflections used	57,394 (2547)	57,358 (3540)	101,164 (4568)
reflections for R_{free}	2807 (485)	2000 (128)	1982 (91)
R_{work}^e	0.1768 (0.3100)	0.1949 (0.3601)	0.1811 (0.4621)
R_{free}^f	0.2036 (0.3400)	0.2234 (0.3541)	0.2059 (0.5036)
protein residues model/sequence (chains)	411/426 (2)	409/426 (2)	420/426 (2)
ligand atoms	10	22	22
solvent	196	193	375
rmsd			
bond lengths (Å)	0.004	0.009	0.012
bond angles (°)	0.69	0.85	1.21
B-Factors (Å²)			
average	40.44	45.88	26.23
protein	40.23	46.10	24.20
ligands	59.80	33.91	20.55
solvent	42.89	41.11	37.75
Ramachandran Plot			
favoured/allowed (%)	98.75/1.25	98.75/1.25	99.03/0.97
outliers (%)	0.0	0.0	0.0
rotamer outliers (%)	0.00	0.29	0.56
clashscore	2.78	2.12	6.14

^aValues for the highest resolution shell are given in parentheses. ^bThe maximum resolution of the dataset was chosen so that $CC1/2 > 0.25$.

^c $R_{\text{meas}} = \sum_i \left(\frac{n}{n-1} \right)^{1/2} \sum_i |I_i(h) - I(h)| / \sum_i \sum_i |I_i(h)|$, where n is the number of independent observations of $I(h)$. ^d $CC1/2$ is the Pearson correlation coefficient calculated between two random half datasets. ^e $R_{\text{work}} = \sum_i |F_o(h) - F_c(h)| / \sum_i F_o(h)$. ^f R_{free} is calculated as R_{work} but with 2–5% of randomly chosen reflections omitted from refinement.

PRPP to yield products. A second possible mechanism (mechanism 2 in Figure 2), also starting from the amide form of OA, could involve the direct proton transfer from the N1 atom of OA to the O2A atom of the α -phosphate group of PRPP, with the nucleophilic attack of the N1 atom of OA onto the C1 atom of PRPP taking place concomitantly. As a third

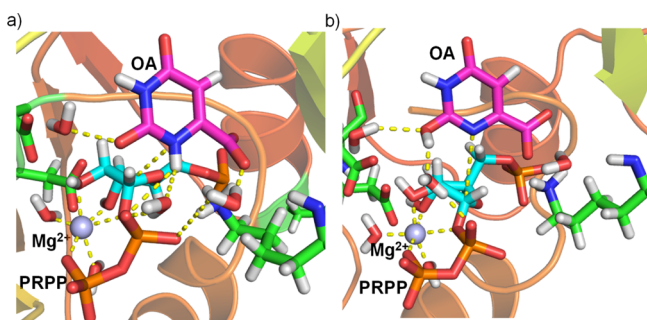


Figure 1. Active site of the OPRTase enzyme after 100 ns of MD simulation (a) with the amide form of OA and (b) with the imidic acid form of OA. The carbon atoms of PRPP and OA are shown in light blue and pink, respectively. Hydrogen bond interactions between the substrates (OA and PRPP) and solvent molecules are plotted as dashed lines. The bond-forming distance (N1 atom from OA and C1 atom from PRPP) and the octahedral coordination of the Mg^{2+} ion are also displayed as dashed lines.

alternative (mechanism 3 in Figure 2), we could examine a mechanism whereby an initial proton-transfer event would occur from the N1 atom to one of the carboxylate oxygens of OA before the proton is transferred to a water molecule and, consecutively, to the O2A atom of the α -phosphate group of PRPP. Then, the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP would proceed. It is important to point out that no other candidates to act as bases were observed in the X-ray structures or during the MD simulations.

In the case the reaction mechanism would start from the imidic acid form of OA, a water molecule involved in a hydrogen bond interaction with the H1 atom bonded with the O2 atom of OA could also be the proton acceptor (Figure 1b), and from this water molecule, the proton could also be transferred to the O2A atom from the α -phosphate group of

the pyrophosphoryl fragment of PRPP or to the Asp125 residue.

From the relaxed structures of both systems, the amide and imidic acid forms of OA, 200 ps of QM/MM simulations were performed with the sander.MPI module of AmberTools18.^{41,42} OA, PRPP, the Mg^{2+} ion, and three water molecules (a total of 54 atoms) were treated at a quantum mechanical level using the PM6 method,⁴³ while the rest of the protein and solvent water molecules were described by ff14SB⁴⁴ and TIP3P⁴⁵ force fields, respectively, as implemented in AmberTools18.⁴¹ The water molecules in the QM region are the two water molecules coordinated with the Mg^{2+} cation and one water molecule that could be a potential base to abstract the proton atom (H1) from OA (Figures 1–3). Figure 3 shows a schematic representation of the active site with the QM region and the hydrogen bond interactions between the substrates (OA, in its amide form, and PRPP) and enzymatic residues. In Figure 3, all the residues belong to the subunit selected for the study except Arg99* and Lys103*, which belong to the adjacent subunit.

2.6. Free Energy Perturbation. The amide form of OA is known to be in equilibrium with its tautomeric form, the imidic acid form. The latter form could conceivably start the reaction mechanism.¹⁹ Therefore, we decided to carry out free energy perturbation calculations to establish which tautomer is the most stable form in the active site of the enzyme and which is the free energy difference between them. If we consider the thermodynamic cycle shown in Scheme 3, the free energy of conversion of the amide form to the imidic acid form in the protein ($\Delta G_{\text{Protein}}$) is obtained as the difference of the free energy of interaction of the imidic acid form ($\Delta G_{\text{Imidic,p}}$) and the amide form ($\Delta G_{\text{Amide,p}}$) with the protein plus the free energy difference in the gas phase (ΔG_{Gas}).

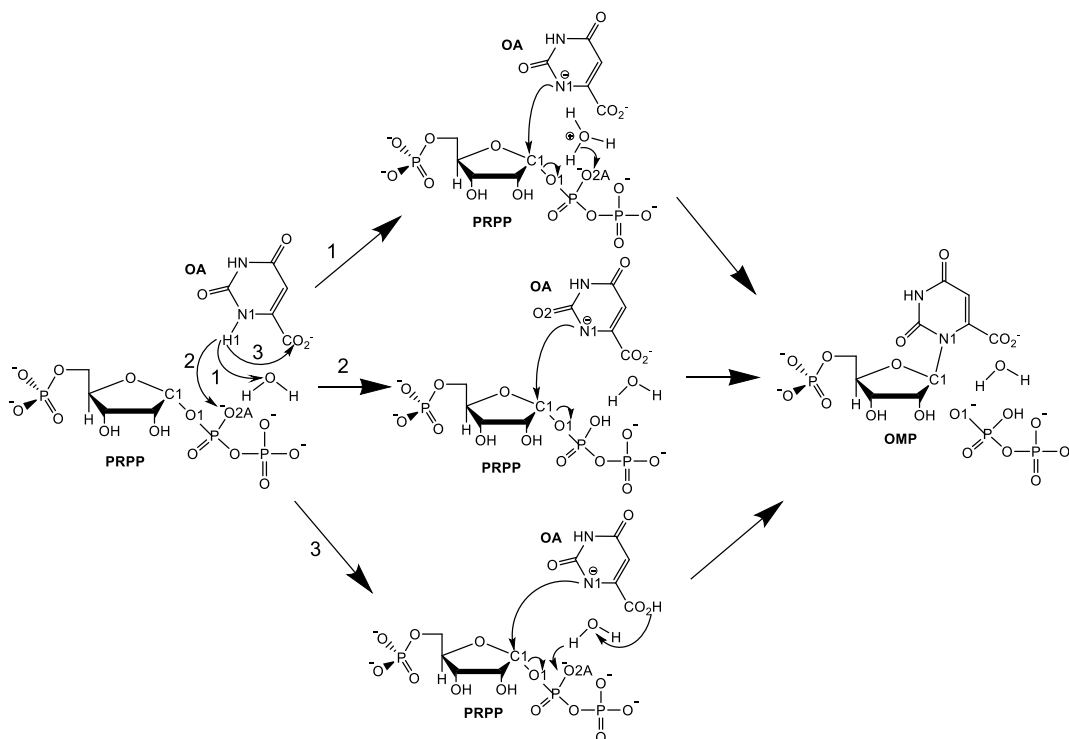


Figure 2. Proposed reaction mechanisms starting from the amide form of the OA substrate.

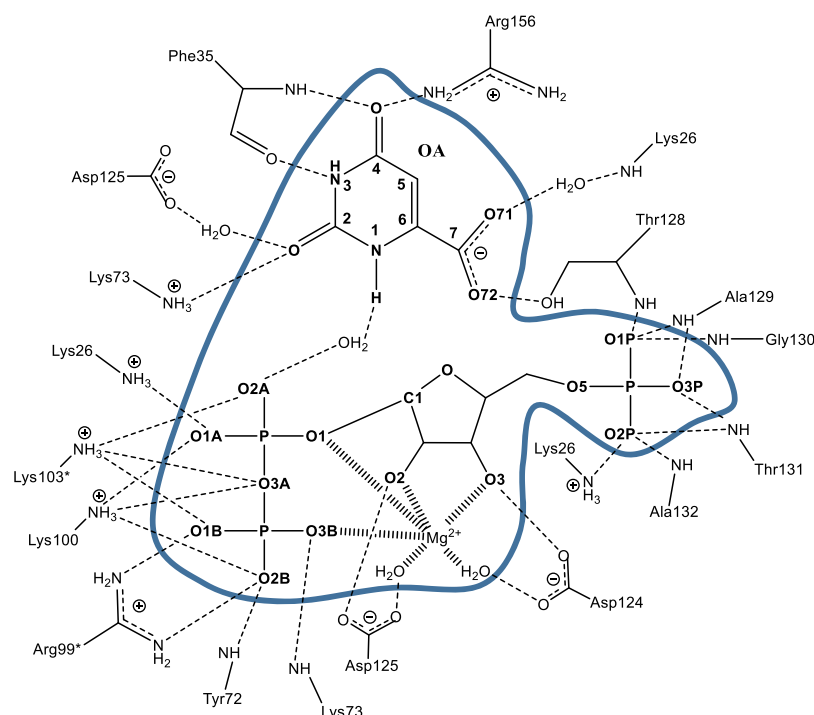
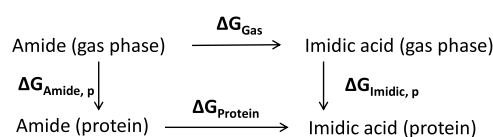


Figure 3. Schematic representation of the active site. The atoms inside the blue curve are described quantum mechanically. OA and PRPP are shown in bold. Mg^{2+} coordination bonds are shown as hashed lines. Hydrogen bond interactions are shown in dashed lines. Lys73 and Asp125 are shown twice and Lys26 three times. Residues Arg99* and Lys103* belong to the adjacent subunit.

Scheme 3. Thermodynamic Cycle To Calculate $\Delta G_{\text{Protein}}$



$$\Delta G_{\text{Protein}} = \Delta G_{\text{Gas}} + (\Delta G_{\text{Imidic,p}} - \Delta G_{\text{Amide,p}}) \quad (1)$$

On the one hand, the free energy difference in the gas phase was obtained from gas-phase optimizations of the amide form and the imidic acid form using the M06-2X density functional with the 6-311+G(2df,pd) basis set with the Gaussian09 program.⁴⁶ The free energy was obtained using the standard rigid-rotor and harmonic oscillator approximations. On the other hand, the difference of the interaction free energy of the imidic form and the amide form in the protein environment was computed by means of free energy perturbation calculations. For the transformation from the amide form to the imidic acid form, the chemical space was divided into 21 λ -values from 0.00 to 1.00 with a $\Delta\lambda = 0.05$, and minimization and MD simulations were performed. First, 2000 steps of steepest descent minimization were done in each window. After that, each window was heated from 50 to 298 K during 50 ps. Finally, 5 ns of NPT MD simulation was run in each of the 21 windows. The rest of conditions were as previously described.

The analysis of the data was done by a Python tool, alchemical-analysis.py, freely available on GitHub as part of the pymbar package (located at <https://github.com/MobleyLab/alchemical-analysis>).⁴⁷ For the analysis, the first 1 ns of MD production was discarded as equilibration data. In the results file, the free energy of the conversion of the amide form to the imidic acid form in the protein was calculated by means of

different methods. The chosen method to estimate the free energy difference is the Bennett acceptance ratio (BAR) method.⁴⁸

2.7. Minimum Free Energy Path. In order to elucidate the enzymatic mechanism catalyzed by EcOPRTase, all the possible paths mentioned above and displayed in Figure 2 were studied by means of the exploration of the potential energy profiles. The obtained reactant and product structures were used as the initial and final points, respectively, for the exploration of the MFEP. The MFEP for each possible path was computed with an implementation of the recently developed adaptive string method³¹ in the AmberTools18 package.⁴¹ A set of collective variables (CVs) involved in the OPRTase reaction and sufficient to follow the progress of the reaction (the bond-forming and bond-breaking distances and hybridization coordinate of the C1 atom of PRPP) were chosen. Details of the chosen CVs for each possible path are reported in the Results and Discussion section. An initial guess was defined along the CVs, and a set of equidistant points along the path, called the string nodes, were selected. The string consisted of 80 nodes. The initial structures for the string nodes were obtained from the potential energy profiles previously explored. For each node, a separate MD simulation was launched using a time step of 1 fs and for a maximum of 250 ps. The rest of conditions were as described above. From the MD trajectories, the approximate slope of the underlying free energy was obtained and each node was moved toward the lower free energy, while keeping them equidistant. To improve the sampling of the configurational space, replica exchange was employed, with exchange between neighboring nodes attempted every 50 steps. By repeating this procedure iteratively, the string eventually reached the MFEP. The convergence of the string was monitored by measuring the mean distance between the past and current positions of the

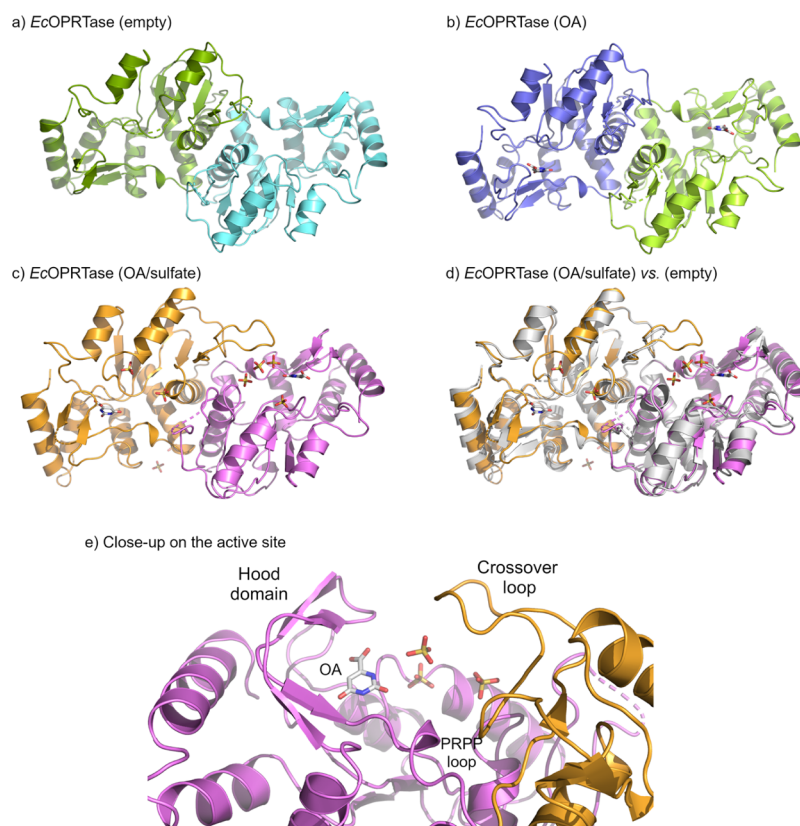


Figure 4. Crystal structures of EcOPRTase. Cartoon representations of the three EcOPRTase structures are shown in chain colors. OA and sulfate ions are shown in sticks and CPK colors where appropriate. (a) Empty active site. (b) OA-only complex. (c) OA/SO₄²⁻ complex. (d) Superposition of the OA/SO₄²⁻ complex (in chain colors) with the empty active site (white). (e) Close-up of the active site of the OA/SO₄²⁻ complex to show the ordered crossover loop, the cap or hood domain, and the PRPP binding loop.

string nodes. The converged string was used to obtain the potential of mean force (PMF). This PMF was performed along a path collective variable, the so-called *s* coordinate, a continuous mathematical function defined by a set of points along the path that measures the advance of the system along the MFEP.³¹ The PMF was obtained using umbrella sampling (US),⁴⁹ which allowed to calculate the 95% confidence intervals (CIs) analytically, assuming a normal distribution of the sampling along the RC in each US window. The different lengths for the production of the PMF aimed to reach a value of the 95% CI for the free energy barrier of ± 1 kcal·mol⁻¹. The positions and force constants of US windows were taken from the node parameters adjusted during the adaptive string optimization.³¹

Finally, corrections to the PM6/MM PMF using a higher level method to describe the QM region were performed. The high-level method used was M06-2X density functional with the 6-311+G(2df,pd) basis set using the Gaussian09 program.⁴⁶ From the structures of the PMF, a minimization process was performed in each node along the MFEP using 500 steps of the steepest descent method, followed by 500 steps of conjugate gradient minimization for a maximum of 1000 minimization steps. Single point energy calculations at the high-level method (M06-2X/6-311+G(2df,pd)) and the low-level method (PM6) of the minimized structures along the PMF were performed. Moreover, zero point energy (ZPE) correction to the activation free energy barrier was also computed by means of localization and characterization of reactants and TS structures. The localization of the stationary

structures was performed using an all-purpose suite for multiscale calculations named QM3.⁵⁰

3. RESULTS AND DISCUSSION

3.1. Crystallographic Structures of EcOPRTase Complexes. **3.1.1. Overall Fold Architecture.** In agreement with other PRTases, EcOPRTase is a constitutive homodimer both in solution and in crystalline form (Table 1). Each monomer is characterized by an $\alpha + \beta$ structure consisting of a central three-layer α/β (Rossmann) fold furnished with N- and C-terminal extensions (Figure 4). The N-terminal extreme, referred to as the “cap” or hood domain, consists of the first α -helix ($\alpha 1$) and a long and functionally relevant β -hairpin, characterized by a symmetric kink at conserved glycine and proline residues at opposite locations along the β -strands that break them into $\beta 1/\beta 1'$ -strands and $\beta 2'/\beta 2$ -strands (Figure 4). The C-terminal end, which packs against the N-terminal end, consists of two antiparallel α -helices that run perpendicular to helix $\alpha 1$ (Figure 4). In total, the EcOPRTase structure includes 8 α -helices and 9 β -strands.

A long and flexible catalytic loop closes in over the active-site cavity of the adjacent monomer with every catalytic turnover (Figure 4e). The substrate-binding regions are defined as the binding region for the OA part of OMP (residues 25–27), situated at the tip of the N-terminal β -hairpin motif between β -strands $\beta 1'$ and $\beta 2'$; the pyrophosphate binding region (residues 72–74), created by a β -strand-loop- α -helix motif between $\beta 3$ and $\alpha 3$; and the PRPP-binding region, ¹²⁰VMLVDDVITAGT¹³¹, a conserved sequence motif

covering a second β -strand-loop- α -helix motif between $\beta 5$ and $\alpha 4$. The presence of bound substrates triggers subtle structural rearrangements in the active site and neighboring regions besides the closing of the flexible loop.

Several dramatic conformational changes have been proposed to accompany the substrate and product binding and dissociation events and the catalysis. These changes involve domain rotations, particularly of the upper hood domain down onto the bound substrates and products, and a disorder-to-order transition of the conserved flexible loop in the substrate ternary complex.¹⁹ Because all the EcOPRTase structures presented here correspond to either the apoenzyme or the OA complex, which elicits only moderate conformational adjustments, they are structurally similar and thus can be superimposed with 0.55–1.07 Å rmsd.

3.1.2. Crossover Loop Is Fully Disordered in the Apoenzyme. We have solved the first crystallographic structure of EcOPRTase with an empty active site at 1.55 Å (Figure 4a). Suitable conditions had to be found to avoid the presence of sulfate or phosphate ions in the active site because they become easily locked into it. This was considered important because a previous structure of EcOPRTase without bound substrates contained sulfate ions which immobilized an interface loop in an inactive conformation (PDB 1ORO).¹⁸ In contrast to the latter structure of EcOPRTase, in the apoenzyme without bound sulfates or phosphates, the two crossover loops are fully disordered. A similar observation was made in the structure of *S. cerevisiae* OPRTase with an empty active site determined by X-ray crystallography at 2.35 Å (PDB 2PRY).¹⁹ Several as yet unpublished crystal structures are also available for comparisons, including *Francisella tularensis* OPRTase (PDB 3MJD),⁵¹ *Burkholderia cenocepacia* OPRTase (PDB 4OHC),⁵² *Bacillus anthracis* OPRTase (PDB 3M3H),⁵³ and *Vibrio cholerae* OPRTase (PDB 3N2L).⁵⁴ Except for the OPRTase from *B. cenocepacia*, all other OPRTases that were crystallized as apoenzymes lack interpretable electron density for the catalytic loop.

Crystal structures of OPRTase obtained in the presence of phosphate or sulfate anions (but without substrates or products) show that these anions are tightly bound into the active site. Several homologous structures of OPRTase contain either sulfate or phosphate anions in the active site. Those containing inorganic phosphate include, for instance, *Aeropyrum pernix* OPRTase (PDB 2YZK)⁵⁵ and *M. tuberculosis* OPRTase (PDB 5HKL),⁹ which have a single phosphate anion bound to the PRPP-binding loop in a location coincident with that of 5'-phosphate of bound PRPP. There are more crystal structures of homologous OPRTases with sulfate ions: in *Streptococcus mutans* OPRTase (PDB 3DEZ),⁵⁶ two fully occupied SO_4^{2-} sites were interpreted as corresponding to the 5'-phosphate and the pyrophosphate moieties of the PRPP substrate; in the closely related OPRTase from *Streptococcus pyogenes* (PDB 2AEE),⁵⁷ there is a single sulfate anion whose position coincides with that of the phosphate ions in the *A. pernix* and *M. tuberculosis* OPRTase apoenzyme structures. In *P. falciparum* OPRTase (PDB 4FYM),⁵⁸ with eight independent protein chains in the asymmetric unit, between 2–8 sulfate ions could be identified in the electron density; the most interesting binding sites again matched those of the 5'-phosphate and the pyrophosphate group of PRPP. About half of the sulfate or phosphate-containing OPRTase structures show ordered catalytic loops (at least in one of the subunits), in contrast to the apoenzyme structures, indicating that the

presence of an anion at the 5'-phosphate position of PRPP stabilizes a closed active-site configuration. Based on the available structures and our EcOPRTase structure crystallized from ammonium sulfate with OA (cf. below), it is possible to model the position that the PRPP substrate should occupy with considerable accuracy.

3.1.3. Orotic Acid Binding Site Is Performed in the Absence of PRPP. In this work, we have solved the crystal structures of two OA complexes of EcOPRTase. The first structure, solved to 1.59 Å, contains only OA under phosphate-/sulfate-free crystallization conditions (Figure 4b); the second structure of the OA complex, refined to 1.25 Å, was obtained under high sulfate concentration and therefore contains sulfate anions in the active site (Figure 4c). The high resolution at which both structures were solved ensures that the structural information derived from them is very reliable. As in these previous structures, the natural substrate OA in our EcOPRTase OA complex is held in place by residues from the hood domain together with the $\alpha 2$ helix (Figure 5). Key residues for OA binding are Lys26, Phe34,

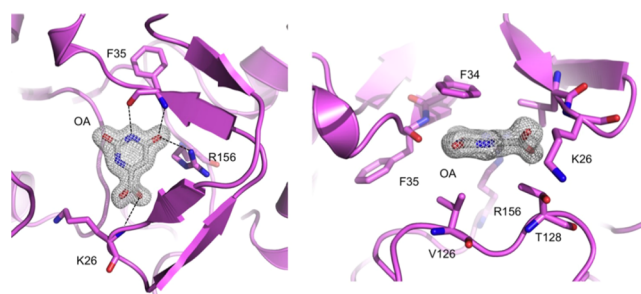


Figure 5. Active site of the EcOPRTase/OA complex. Close-up of the active site structure in the OA complex shown in cartoon representation. The σ_A -weighted $2mF_o - DF_c$ electron density map for OA is shown at the 1 rms contour level. Left, the side chains of the enzyme residues that establish hydrogen bonding interactions (dashed lines) with OA are shown in sticks. Right, the side chains of the residues involved in shaping the hydrophobic pocket for OA are shown in sticks.

Phe35, and Arg156. The main-chain amide N of Lys26 forms a hydrogen bond to the OA carboxylate that helps keep its plane at an angle of 41.7° with respect to the OA ring's plane. The side chain of Lys26 points out to the solvent and makes no interactions with any other protein residue. An extensive network of hydrogen bonds mediated by water molecules connects the OA carboxylate moiety with the surrounding area of the active site. The phenyl side chain of Phe34 plays an important role in OA binding by providing stacking interactions; the distance between the Phe34 side-chain phenyl and the OA ring is 3.5–4.2 Å. Finally, the main-chain conformation adopted by Phe35 allows its O and N atoms to form hydrogen bonds to the N3 and the carbonylic O4 atoms of OA. In our OA structure, the side chain of Arg156 also interacts with O4 while being held in place by interactions with Ser31, Tyr33, and Glu158. This constellation of interactions seems responsible for ensuring correct positioning and orientation for the reaction of OA with PRPP. Indeed, in the only two available structures of OPRTase with OA (PDB 1LHO²⁰ and 1OPR⁵⁹), from the *S. typhimurium* enzyme, which also contain PRPP, the interactions that stabilize OA are the same that we have observed for EcOPRTase.

In the second OA complex of EcOPRTase, which was crystallized in the presence of 2.1 M ammonium sulfate, the interactions between OA and EcOPRTase are essentially unchanged. In the OA/SO₄^{2−} cocomplex, up to four sulfate ions could be identified in electron density, three of which occupy functionally important positions in the active site (Figure 6). Two sulfate ions occupy conserved positions

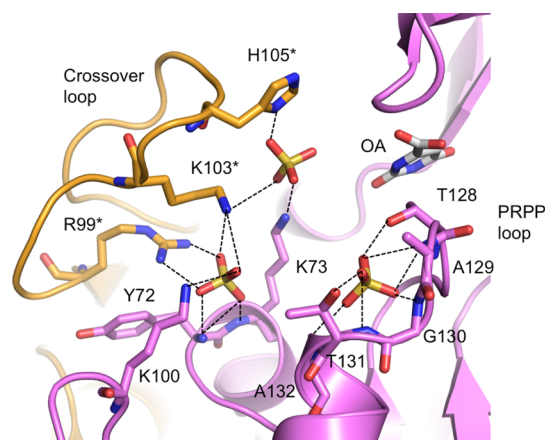


Figure 6. Active site of the EcOPRTase/OA/SO₄^{2−} complex. Close-up of the active site structure in the OA/SO₄^{2−} complex shown in cartoon representation. The side chains of the enzyme residues that establish hydrogen bonding interactions (dashed lines) with the SO₄^{2−} anions are shown in sticks. The flexible crossover loop, which comes from the adjacent subunit, is shown in orange.

corresponding to the binding sites for the PRPP phosphate groups, as previously noted for the subset of apoenzyme structures that have been cocrystallized with sulfate. One of these sulfate ions is located in the position of the 5′-phosphate group of the PRPP substrate and establishes hydrogen bond interactions with Thr128, Ala129, Gly130, Thr131, and Ala132 (the PRPP binding loop), as observed in the X-ray structures of *S. typhimurium* OPRTase.^{20,59} The second sulfate ion is located at the interface between the two subunits, where it engages residues from both protomers in a pseudo-symmetric arrangement: Tyr72, Lys73, and Lys100 from the same monomer and Arg99* and Lys103* from the adjacent monomer (Figure 6). These interactions are found in the X-ray structure of *S. typhimurium* OPRTase with OA and Mg²⁺-PRPP²⁰ between these residues and the β-phosphate moiety of PRPP.

The third sulfate ion was found at the entry gate to the substrate-binding pocket, where it was held in place by interactions with the side chains of Lys73 from one protomer and Lys103* and His105* from the opposite subunit (Figures 6 and S2 in the Supporting Information). Interestingly, the latter two interactions are found to stabilize the α-phosphate moiety of PRPP in the structure of *S. typhimurium* OPRTase,²⁰ where the crossover loop is deeply buried into the active site. The conservation of these interactions between the PRPP-bound state and those holding the sulfate ion at the entry gate of the active site as seen in our structure suggests that Lys73, Lys103*, and His105* may help to release the pyrophosphate product after the catalytic turnover.

3.1.4. Subunit Asymmetry and Active Site Cooperativity. Certain OPRTases including the yeast enzyme have been described to operate via a double Theorell–Chance (“hit-and-run”) mechanism whereby the timing of OA and PRPP

binding to one site is synchronized to the timing of OMP and pyrophosphate release from the opposite site, leading to a unique alternating site catalysis without the accumulation of the ternary complexes.⁶⁰ This mechanism requires the two subunits to act in a highly cooperative fashion that in the steady state is governed by an equation indistinguishable from that observed in classical ping-pong bi–bi kinetics. At the structural level, this asymmetric behavior manifests itself as local structural differences upon superposition and by the presence of half-site occupancy of catalytic sites for certain substrate and product combinations in preference over fully occupied ternary complexes. Thus, in OPRTase both active sites may have bound PRPP, while only one of the two subunits had OA. There is a correlation between substrate binding and the disorder-to-order transition of the active-site loop; hence, the active site asymmetry also becomes evident by observing to what degree this loop appears well-defined in the electron density as PRPP or OA sit in the active-site pocket.

In our structures of dimeric EcOPRTase, the two subunits adopt similar but distinct conformations. Interestingly, the disparity between the subunits is greater when the PRPP binding site is not stabilized by the presence of either its canonical substrate or a substrate analogue such as sulfate ions. Hence, the rmsd between the two subunits is 0.76 Å (empty active site), 0.75 Å (OA complex), or 0.55 Å (OA/SO₄^{2−}). Inspection of the structural superposition shows that the regions farther from the dimer interface (the N- and C-terminal domains and the β6 and β7 strands of the central α/β domain) are shifted by several angstroms, even though it is not obvious that crystal packing interactions played any role. In the apoenzyme and the OA complex, both subunits lack interpretable electron density between the catalytic loop residues 102–108. Last, in the OA/SO₄^{2−} complex, the crossover loop of chain B is fully folded with interpretable electron density, in essentially the same conformation observed for chain A of the EcOPRTase (empty) structure.

When the individual subunits are compared across the various complexes, it becomes apparent that there are two classes of subunits, where pairs of subunits in different complexes are more alike than between subunits of the same complex. This property makes the conformational asymmetry even more apparent. Thus, chain A of EcOPRTase (empty) has a lower rmsd (greater similarity) than chain B after superposition with either chain A or B of the OA complex, indicating that both chains in the OA complex adopt a conformation that is best mimicked by chain A of the EcOPRTase with an empty active site. However, when the comparison is made with the OA/SO₄^{2−} complex, the trend is reversed, and here chain B of the EcOPRTase (empty) structure is closest (Figure 4d). Finally, direct comparison of the two OA-containing complexes shows a strong difference between chains A (with an rmsd of 0.90 Å) despite a good agreement between chains B (0.59 Å) and chains A–B (0.65 and 0.77 Å). This observation can be explained by the greater similarity between the two chains in the OA/SO₄^{2−} complex (rmsd 0.55 Å) with respect to the OA complex (0.75 Å), and if the shared conformation in the OA/SO₄^{2−} complex corresponds to chain B in the OA complex. In contrast, chain A of the empty active-site structure is closest to chain A of the OA complex, thus establishing a three-way relationship between the three EcOPRTase structures presented here.

The functional consequences of such subunit asymmetries have been dealt with elsewhere.⁶⁰ For this work, accounting for

such complex cooperation events is of moderate importance because we focus our attention on the chemical steps during a productive catalytic cycle. However, it must be understood that the dimeric architecture of OPRTase plays a crucial role over the catalytic cycle beyond the fact that the catalytic loop of one subunit acts upon the substrates (OA and PRPP) bound in the adjacent subunit.

3.2. Enzyme Kinetics. To guide computational chemistry modeling, we determined the catalytic constant (k_{cat}) and the Michaelis constant (K_{M}) for the forward reaction catalyzed by EcOPRTase (OA + PRPP \rightarrow OMP + PPi). Initial velocity measurements at 25 °C were done in triplicate by following the decrease in absorbance at 295 nm as OA was consumed (cf. Experimental procedures). PRPP was kept fixed at 100 μM . The observed k_{cat} was $26.4 \pm 0.6 \text{ s}^{-1}$, and the K_{M} was $99 \pm 8 \mu\text{M}$ for OA, which yields a $k_{\text{cat}}/K_{\text{M}}$ value of $2.66 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$. This value is in accord with the values previously reported for related enzymes despite variable assay pH and temperature conditions.^{2,3,61}

3.3. Tautomeric Forms of the Substrate. The free energy difference between the imidic acid form and the amide form in the gas phase was calculated with Gaussian using M06-2X as the QM method and the 6-311+G(2df,pd) basis set, as described before. We obtained a value of $27.5 \text{ kcal}\cdot\text{mol}^{-1}$, consistent with the amide form of OA being significantly more stable than the imidic form. The contribution of the substrate–protein interactions to the relative stability of the two forms was estimated to be $-7.61 \pm 0.11 \text{ kcal}\cdot\text{mol}^{-1}$ using the BAR method. Thus, although the protein preferentially stabilizes the imidic acid form over the amide form, this stabilization is not enough to overcome the huge free energy difference observed in the gas phase. Therefore, the estimated free energy difference in the enzyme of $19.9 \text{ kcal}\cdot\text{mol}^{-1}$ between the imidic and the amide forms rendered the amide form the most energetically favorable chemical structure and the preferred form to study the reaction mechanism by QM/MM methods. Any reaction mechanism requiring the OA to acquire an imidic acid function was thus ruled out, as the energy penalty to reach this tautomer in the enzyme is too high. Consequently, all the subsequent steps, including the analysis of the MD simulations and the exploration of the reaction mechanisms, were performed starting exclusively from the amide form of OA.

3.4. Analysis of MD Simulation. A thorough analysis of the evolution of all interaction distances between OA in its amide form and PRPP with their surrounding residues is provided in the Supporting Information. From this comprehensive analysis, it became clear that the catalytic loops of both protomers, which interact with the oxygen atoms of α - and β -phosphates of the pyrophosphoryl group of PRPP, are more flexible than the binding regions for the 5'-phosphate of PRPP and OA and that this flexibility is necessary for catalysis. These catalytic loops remain in an open conformation for the system without bound PRPP and for the product release. In contrast, once the PRPP substrate is bound, the catalytic loops tend to adopt a closed conformation over the active site where some interactions are stable and others fluctuate during the MD simulation.

It is important to mention that the Mg^{2+} ion forms a complex with PRPP in an octahedral coordination, where two interactions involve the ribose hydroxyl groups of PRPP, two more are provided by the oxygen atoms of the pyrophosphoryl fragment and the remaining two interactions are contributed by two water molecules. All these interactions are conserved

along the MD simulations. This coordination is similar to that observed for the *S. typhimurium* OPRTase structure and to the first Mg^{2+} in the HGPRTase TS analogue structure.^{20,22}

As a summary, the average structure obtained from MD simulations agrees quite well with the X-ray structures reported in this work (EcOPRTase/OA and EcOPRTase/OA/ SO_4^{2-} complexes) and the X-ray structures from *S. typhimurium*,^{20,59} in particular regarding the interactions established between OA and the oxygen atoms of the 5'-phosphate group of PRPP with their neighboring residues. Most of the reported differences can be attributed to the presence of aqueous solution in the simulations.

3.5. Reaction Mechanisms and Free Energy Profiles.

Starting from the amide form of OA and from the analysis of the crystal structures and the MD simulations, three different mechanisms were proposed to be studied (Figure 2).

First, the simplest mechanism (mechanism 2 in Figure 2) was explored whereby the H1 proton from OA is transferred directly to the α -phosphate O2A atom of the pyrophosphoryl group of PRPP. Subsequently, the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP can take place, and the bond C1–O1 is broken to form OMP and the pyrophosphate leaving group. Four distances, the bond-forming and bond-breaking distances, and the hybridization of the C1 atom of PRPP were chosen to follow the reaction progress. The schematic representation of this mechanism is displayed in Figure 7a. Following the procedure described in

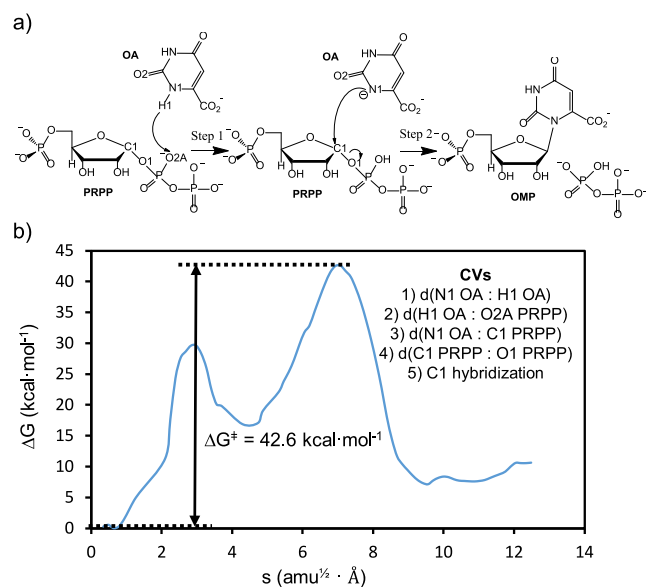


Figure 7. (a) Reaction mechanism of the proton transfer to the α -phosphate O2A atom of the pyrophosphoryl group of PRPP and nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP and formation of OMP and pyrophosphate and (b) corresponding PMF computed at the M06-2X/6-311+G(2df,pd):PM6/MM level and the set of CVs that defines the s coordinate.

the Materials and Methods section, the corresponding PMF was obtained using PM6 to describe the QM region and free energy corrections were computed using the M06-2X density functional with the 6-311+G(2df,pd) basis set with the Gaussian09 program. The PMF along the path collective variable (called s coordinate) and the set of CVs that defines the s coordinate are depicted in Figure 7b. As shown in the figure, the obtained free energy barrier is too high (42.6 kcal·

mol^{-1}) for an enzymatic reaction to occur. Thus, this reaction mechanism was deemed not feasible energetically.

Another possible mechanism that was considered (mechanism 3 in Figure 2) begins with an intramolecular proton transfer from the N1 atom of OA to a carboxylic oxygen of OA. From there, the proton is transferred from the carboxylic group of OA to a water molecule and from the water molecule to the α -phosphate O2A atom of the pyrophosphoryl group of PRPP, and the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP proceeding next. A set of coordinates, the bond-breaking and bond-forming distances (7 distances in total), and the hybridization of the C1 atom were used to obtain the PMF along the path collective variable. The reaction mechanism and the corresponding PMF computed at the M06-2X/6-311+G(2df,pd):PM6/MM level with the set of CVs that defines the s coordinate are displayed in Figure 8. The free energy barrier obtained, $33.8 \text{ kcal}\cdot\text{mol}^{-1}$, is still too high for a plausible enzymatic reaction, therefore concluding that this reaction mechanism is not viable.

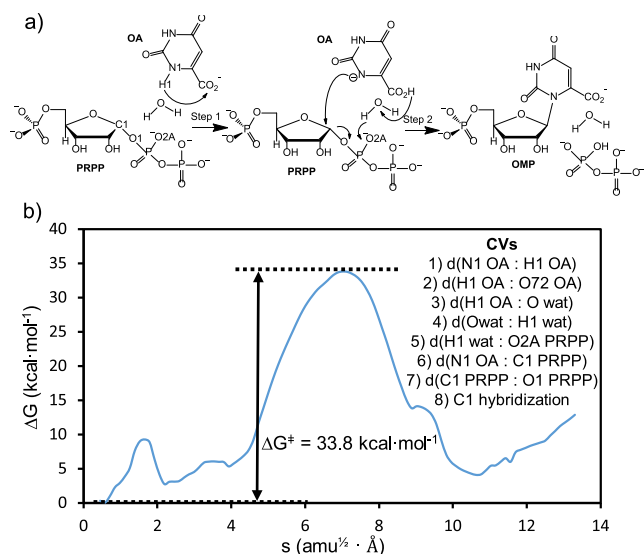


Figure 8. (a) Reaction mechanism of the proton transfer through the carboxylic group of OA and a water molecule and hereafter to the α -phosphate O2A atom of the pyrophosphoryl group of PRPP. Last, the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP and formation of OMP and pyrophosphate takes place. (b) Corresponding PMF computed at the M06-2X/6-311+G(2df,pd):PM6/MM level and the set of CVs that defines the s coordinate.

The last reaction mechanism explored (mechanism 1 in Figure 2) consists of the proton transfer to a water molecule and from the water molecule to the α -phosphate O2A atom of the pyrophosphoryl group of PRPP. Then, the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP yields OMP and pyrophosphate. The distances involved in the bond breaking and bond forming (6 distances in total) and the hybridization of the C1 atom of PRPP were employed to perform the PMF along the s coordinate. The reaction mechanism and the corresponding PMF at the M06-2X/6-311+G(2df,pd):PM6/MM level with the set of CVs are depicted in Figure 9. From the PMF, it can be observed that the reaction mechanism takes place in three steps. The first one consists in the proton transfer from the N1 atom of OA to a water molecule and the formation of a transient hydronium

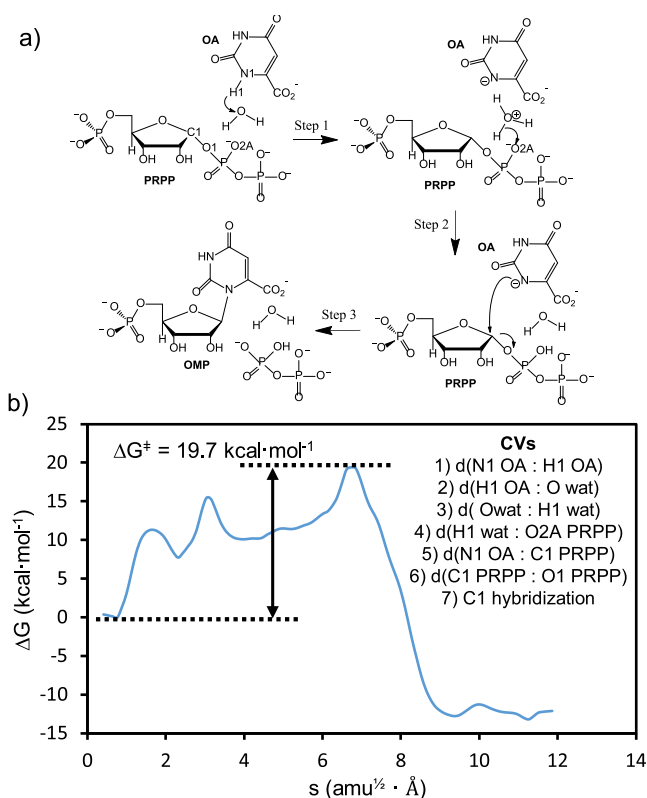


Figure 9. (a) Reaction mechanism of the proton transfer from the N1 atom of OA to a water molecule and from the water to the α -phosphate O2A atom of the pyrophosphoryl group of PRPP and finally the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP and formation of OMP and pyrophosphate. (b) Corresponding PMF computed at the M06-2X/6-311+G(2df,pd):PM6/MM level and the set of CVs that defines the s coordinate.

cation. This intermediate is not very stable, and consecutively, the proton from the hydronium is transferred to the α -phosphate O2A atom of PRPP. The last step consists in the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP, the TS of this process being the rate-limiting one. The free energy barrier obtained ($19.7 \text{ kcal}\cdot\text{mol}^{-1}$) is plausible for an enzymatic reaction. The TS structures localized for the three steps of this reaction mechanism are displayed in Figure S8 in the Supporting Information. From the trajectory file of the string nodes of reactants and the highest energy TS, a set of 10 TS and reactant structures are localized and characterized to compute the ZPE correction to the activation free energy. The corrected free energy barrier is $16.7 \text{ kcal}\cdot\text{mol}^{-1}$, in good agreement with the value derived from the experimental rate constant provided in this work (k_{cat} of 26.4 s^{-1} , corresponding to a free energy barrier of $15.5 \text{ kcal}\cdot\text{mol}^{-1}$ at a temperature of 298 K).

In conclusion, we propose this reaction mechanism as the more adequate to explain the reaction process in EcOPRTase.

From the PMF, we can obtain the value of the s coordinate that corresponds to reactants ($0.5 \text{ amu}^{1/2}\cdot\text{Å}$) and the TS of the rate-limiting step ($6.7 \text{ amu}^{1/2}\cdot\text{Å}$). We analyzed some key distances between the substrates (OA and PRPP) and their surrounding residues at the reactant state and the TS. The values of the distances that present the largest average variation between the reactant state and the TS are shown in Table 2. These values have been averaged over the MD trajectories

Table 2. Average Values of the Distances (with Their Standard Deviations) That Present the Largest Variation between the Reactant State and TS along the MD Simulation of the PMF^a

distances	R	TS
<i>d</i> (N1 OA, C1 PRPP)	3.38 ± 0.18	2.34 ± 0.10
<i>d</i> (C1 PRPP, O1 PRPP)	1.43 ± 0.03	2.04 ± 0.12
<i>d</i> (O1 PRPP, Mg ²⁺)	2.22 ± 0.10	2.02 ± 0.07
<i>d</i> (O2 PRPP, OD2 Asp125)	3.05 ± 0.12	2.73 ± 0.11
<i>d</i> (O3B PRPP, N Lys73)	3.6 ± 0.20	3.43 ± 0.20
<i>d</i> (O2B PRPP, N Tyr72)	3.02 ± 0.11	2.93 ± 0.10
<i>d</i> (O1B PRPP, NH2 Arg99*)	2.98 ± 0.10	2.79 ± 0.10
<i>d</i> (O2B PRPP, NH1 Arg99*)	2.95 ± 0.11	2.81 ± 0.10
<i>d</i> (O1B PRPP NZ Lys103*)	2.85 ± 0.10	2.70 ± 0.09
<i>d</i> (O3A PRPP NZ Lys103*)	3.50 ± 0.22	2.86 ± 0.16

^aThe distances are in angstroms.

from US windows corresponding to the reactant state and the TS. The residues participating in the interaction distances provided in Table 2 are displayed in the TS structures (in Figure S8). Comparing the distances at the reactant state and the TS, it is shown that the distance between the O1 atom of PRPP and the Mg²⁺ ion is shorter at the TS because the charge on the O1 atom at the TS is more negative than in the reactant state. Another interaction that presents a significant difference between the reactant state and the TS is the distance between the O2 atom of the ribose hydroxyl group and Asp125 residue, being shorter at the TS because of the development of a positive charge at the C1 atom of PRPP (the formation of the oxocarbenium). Furthermore, some residues that interact with the oxygen atoms of the α - and β -phosphates of the pyrophosphate leaving group (Lys73, Arg99*, and Lys103*) show shorter distances at the TS, as a result of the increase of negative charge on the leaving group. From these results, we can conclude that the Mg²⁺ ion and key and conserved residues of the active site, such as Lys73, Asp125, Arg99*, and Lys103*, play an essential role in catalysis stabilizing the charge distribution created at the TS. This finding is supported by mutagenesis studies^{25,26} wherein mutations of Asp125, Lys73, and Lys103 residues produced a decrease in the catalytic rate constant, resulting in a substantial loss of enzymatic activity. Our results allow to rationalize the functional defects caused by these mutations as these residues play an essential role in catalysis by electrostatically stabilizing the TS. In addition to that, shorter hydrogen bond interactions between the residues of the adjacent subunit (Arg99* and Lys103*) at the TS sustain a closed conformation of the flexible catalytic loop of the adjacent subunit during the enzymatic process. Overall, these key residues provide a suitable environment to reach the TS and lead the system to the formation of products.

Finally, it is important to point out that the interactions between these key and conserved residues (Lys73, Asp125, Arg99*, and Lys103*) and the substrates (OA and PRPP) must be kept in mind in the design of new TSA as inhibitors to block the OPRTase action.

If we compare the free energy barriers of the rate-limiting step of the three mechanisms, the main difference arises from the reactant and TS structures. It is important to mention, as displayed in Figures 2 and 8a, that the rate-limiting step of mechanism 3 involves not only the nucleophilic attack of the N1 atom to C1 atom but also the proton transfer from the carboxylic group of the OA substrate to one oxygen atom of

the α -phosphate group of PRPP through a water molecule. While the nucleophilic N1 atom approaches the C1 atom for the nucleophilic attack to occur, the proton atom is transferred. Thus, the free energy barrier of this step becomes higher.

The average values of some key distances and the angle that defines the nucleophilic attack at the reactant state and TS for the three mechanisms are shown in Table S1. Based on these data, the distance between the nucleophilic N1 atom and acceptor C1 atom is shorter in the reactant state of mechanism 1 (Figure 2), mechanism that shows lower free energy barrier (Figure 9b). Moreover, the angle that defines the chemical reaction of the nucleophilic attack (N1–C1–O1) is much closer to linearity in the reactant state of mechanism 1, and then this reactant state is more similar to the TS structure than the rest of the mechanisms. With regard to the TS structure, the angle (N1–C1–O1) is closer to linearity in mechanism 1, and then the system at the TS in mechanism 1 exhibits a better orientation for the nucleophilic attack. Furthermore, the interaction distances between the neighboring residues and the PRPP substrate that electrostatically stabilize the TS are in general shorter at the TS of mechanism 1. Accordingly, the TS of the rate-limiting step of mechanism 1 is electrostatically more stabilized than the TSs of the other mechanisms.

Thus, an enzyme active site, which is prepared to stabilize the TS, favors those reactive conformations geometrically closer to the TS, avoiding the energetic penalty associated with the deformation of the enzyme–substrate system along the enzymatic process and providing a lower free energy barrier. Consequently, substrate preorganization and the electrostatic stabilization of the TS by favorable interactions with the protein provide a rate enhancement of the chemical reaction in mechanism 1 (Figures 2 and 9).

4. CONCLUSIONS

In this contribution, we have performed a combined experimental and computational characterization of the structure, dynamics, and reaction mechanism of EcOPRTase. We have reported several high-resolution X-ray structures of the dimeric EcOPRTase with an empty active site and two different substrate complexes (OA and OA/SO₄²⁻) and a kinetic characterization of the catalytic mechanism. Starting from the X-ray structure, a long classical MD simulation and a detailed analysis of the structure, with particular emphasis in the active site region, have been carried out. Comparing the average structure obtained from MD simulations with the X-ray structures and other crystallized structures in the literature, the interactions between OA and the 5'-phosphate group of PRPP with their neighboring residues are mostly maintained because of the presence of a hydrogen bond network that keeps the OA and PRPP substrates attached to the active site. Conversely, in the region of the pyrophosphoryl group of PRPP, conformational changes have been observed compared to the X-ray structures reported in this study and other crystallized structures. From this structural analysis, one might conclude that the catalytic loop that interacts with the oxygen atoms of α - and β -phosphates of the pyrophosphate leaving group is more flexible than the binding regions of OA and the 5'-phosphate moiety of PRPP, and this flexibility is necessary for catalysis.

It was postulated that the OA substrate could present an equilibrium between its tautomeric forms, amide and imidic acid forms, and that the second one could be preferred for catalysis. Free energy perturbation calculations were performed

to discern which tautomeric form is most stable to start the reaction mechanism study. From these calculations, we conclude that the stabilization that the enzymatic environment provides to the imidic form is not sufficient to offset the large free energy difference that favors the amide form in the gas phase. Thus, as the amide form is substantially more stable, the OPRTase reaction is likely to start from the amide form of OA.

An exhaustive mechanistic study has been reported by means of QM/MM MD simulations using a combination of the string method and a path collective variable for the exploration of the free energy landscape. All plausible reaction mechanisms catalyzed by EcOPRTase have been explored and the most likely one consists in three steps. In the first step of the revisited mechanism, a proton is transferred from the N1 atom of OA to a water molecule with the concomitant formation of a transient hydronium cation. Subsequently, the proton from the hydronium cation is transferred to the α -phosphate O2A atom of PRPP, and finally, the nucleophilic attack of the N1 atom of OA to the C1 atom of PRPP proceeds to obtain OMP and pyrophosphate. Reaching the TS for the nucleophilic attack is the rate-limiting step. The activation free energy obtained ($16.7 \text{ kcal}\cdot\text{mol}^{-1}$) is in good agreement with the value derived from the experimental catalytic constant provided in this study ($15.5 \text{ kcal}\cdot\text{mol}^{-1}$ at a temperature of 298 K).

We identified and characterized the TS structure of the rate-limiting step from the QM/MM MD simulations and calculated average distances between the substrates (OA and PRPP) and their surrounding residues at the reactant state and TS. Analysis of the distances with the largest average variation between the reactant state and the TS identified critical catalytic residues (Lys73, Asp125, Arg99*, and Lys103*) that, in addition to the Mg^{2+} ion, electrostatically favor the charge distribution attained at the TS and maintain the flexible catalytic loop of the adjacent subunit (Arg99* and Lys103*) in a closed conformation, thereby providing the adequate environment to perform the chemical reaction. Consequently, the interactions with these residues must be taken into account in the design of new TSA inhibitors to block the OMP biosynthesis for the control of diseases such as malaria, tuberculosis, and cancer.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscatal.9b05294>.


Computational details and distance analysis along the MD simulations; rmsd for the amide and imidic acid forms of OA along 100 ns of MD simulation; close-up of the water-mediated interactions observed between the sulfate ion at the entry gate to the active site and the substrate OA; time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the OA and the surrounding residues along 100 ns of MD simulations; time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the 5'-phosphate group of PRPP and the neighboring residues along 100 ns of MD simulations; time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the oxygen atoms of the pyrophosphoryl group and the ribose hydroxyl groups


of PRPP and the surrounding residues along 100 ns of MD simulations; time evolution of the distances between donor and acceptor atoms of hydrogen bond interactions between the oxygen atoms of the pyrophosphoryl group of PRPP and the surrounding residues of the adjacent subunit and the hydrogen bond interaction between a water molecule and the O2A atom of PRPP along 100 ns of MD simulations; 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 the Lys73 residue along 100 ns of MD simulations; localized TS structures in the active site of the OPRT enzyme of the most favorable mechanism by means of PM6/MM energy optimizations; and average values of some key distances and the angle that defines the chemical reaction of the nucleophilic attack at the reactant state (R), reactant state previous to the rate-limiting step (R'), and TS along the MD simulation of the PMF (PDF)

■ AUTHOR INFORMATION


Corresponding Authors

Maite Roca – *Universitat Jaume I, Castelló, Spain;*

 orcid.org/0000-0003-0937-4722; Phone: (+34) 964 728069; Email: mroca@uji.es; Fax: (+34) 964 728066

M. Cristina Vega – *Center for Biological Research (CIB-CSIC), Madrid, Spain;*  orcid.org/0000-0003-0628-8378; Phone: (+34) 911098070; Email: cvega@cib.csic.es, cristina.vega@strubicib.org; Fax: (+34) 915360432

Íñaki Tuñón – *Universitat de València, Burjassot, Spain;*

 orcid.org/0000-0002-6995-1838; Phone: (+34) 963544880; Email: Ignacio.Tunon@uv.es; Fax: (+34) 963544564

Other Authors

Sergio Navas-Yuste – *Center for Biological Research (CIB-CSIC), Madrid, Spain*

Kirill Zinovjev – *University of Bristol, Bristol, U.K.*

Miguel López-Esteva – *Center for Biological Research (CIB-CSIC), Madrid, Spain*

Sara Gómez – *Center for Biological Research (CIB-CSIC), Madrid, Spain*

Francisco J. Fernández – *Center for Biological Research (CIB-CSIC), Madrid, Spain;*  orcid.org/0000-0002-5015-1849

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acscatal.9b05294>

Author Contributions

M.C.V., I.T., and M.R. conceived and coordinated the project. S.N.-Y., M.L.-E., S.G., and F.J.F. prepared the protein. S.N.-Y., M.L.-E., F.J.F., and M.C.V. crystallized the proteins and collected X-ray data. F.J.F. and M.C.V. processed the X-ray data and solved the crystal structures. S.N.-Y., M.L.-E., F.J.F., and S.G. measured the enzyme kinetics. M.R., K.Z., and I.T. designed and performed the computational chemistry calculations and collected and analyzed the data. M.R., M.C.V., and I.T. wrote the paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Spanish Ministerio de Economía y Competitividad (2015) and Ministerio de Ciencia, Innovación y Universidades (2018) through grants cofinanced with European Union ERDF funds PGC2018-094852-B-C22 to IT and CTQ2015-66206-C2-2-R and RTI2018-1018-102242-B-I00 to M.C.V., by Generalitat Valenciana AICO/2018/238 project and by Universitat Jaume I UJI-B2016-28 project to IT and M.R., respectively, and CSIC intramural grant 2016E064 to M.C.V. M.R. thanks Ministerio de Economía y Competitividad for a “Ramón y Cajal” contract (RYC-2014-16592). S.N.-Y. and M.L.-E. acknowledge the support of the official Ph.D. program in Biochemistry, Molecular Biology and Biomedicine of the Universidad Complutense de Madrid (UCM). The authors gratefully acknowledge the computational facilities of Universitat Jaume I and Universitat de València (Tirant Supercomputer) and the access to the ALBA Synchrotron Light Source, Barcelona, Spain, for provision of synchrotron radiation facilities at the BL13-XALOC beamline.

■ REFERENCES

- (1) Musick, W. D. L.; Nyhan, W. L. Structural Features of the Phosphoribosyl-Transferases and Their Relationship to the Human Deficiency Disorders of Purine and Pyrimidine Metabolism. *CRC Crit. Rev. Biochem.* **1981**, *11*, 1–34.
- (2) Krungkrai, S. R.; Aoki, S.; Palapac, N. M. Q.; Sato, D.; Mitamura, T.; Krungkrai, J.; Horii, T. Human Malaria Parasite Orotate Phosphoribosyltransferase: Functional Expression, Characterization of Kinetic Reaction Mechanism and Inhibition Profile. *Mol. Biochem. Parasitol.* **2004**, *134*, 245–255.
- (3) Breda, A.; Rosado, L. A.; Lorenzini, D. M.; Basso, L. A.; Santos, D. S. Molecular, Kinetic and Thermodynamic Characterization of Mycobacterium Tuberculosis Orotate Phosphoribosyltransferase. *Mol. Biosyst.* **2012**, *8*, 572–586.
- (4) Bhatia, M. B.; Vinitsky, A.; Grubmeyer, C. Kinetic Mechanism of Orotate Phosphoribosyltransferase from Salmonella-Typhimurium. *Biochemistry* **1990**, *29*, 10480–10487.
- (5) Victor, J.; Greenberg, L. B.; Sloan, D. L. Studies of the Kinetic Mechanism of Orotate Phosphoribosyltransferase from Yeast. *J. Biol. Chem.* **1979**, *254*, 2647–2655.
- (6) Wittmann, J. G.; Heinrich, D.; Gasow, K.; Frey, A.; Diederichsen, U.; Rudolph, M. G. Structures of the Human Orotidine-5'-Monophosphate Decarboxylase Support a Covalent Mechanism and Provide a Framework for Drug Design. *Structure* **2008**, *16*, 82–92.
- (7) White, N. J. Antimalarial Drug Resistance. *J. Clin. Invest.* **2004**, *113*, 1084–1092.
- (8) Almeida Da Silva, P. E.; Palomino, J. C. Molecular Basis and Mechanisms of Drug Resistance in Mycobacterium Tuberculosis: Classical and New Drugs. *J. Antimicrob. Chemother.* **2011**, *66*, 1417–1430.
- (9) Donini, S.; Ferraris, D. M.; Miggiano, R.; Massarotti, A.; Rizzi, M. Structural Investigations on Orotate Phosphoribosyltransferase from Mycobacterium Tuberculosis, a Key Enzyme of the De Novo Pyrimidine Biosynthesis. *Sci. Rep.* **2017**, *7*, 1180.
- (10) Stout, J. T.; Caskey, C. T. *The Metabolic Basis of Inherited Disease*; McGraw-Hill: New York, 1989; p 1.
- (11) Chu, E.; Callender, M. A.; Farrell, M. P.; Schmitz, J. C. Thymidylate Synthase Inhibitors as Anticancer Agents: from Bench to Bedside. *Cancer Chemother. Pharmacol.* **2003**, *52*, S80–S89.
- (12) Christopherson, R. I.; Lyons, S. D.; Wilson, P. K. Inhibitors of De Novo Nucleotide Biosynthesis as Drugs. *Acc. Chem. Res.* **2002**, *35*, 961–971.
- (13) Tao, W.; Grubmeyer, C.; Blanchard, J. S. Transition State Structure of Salmonella typhimurium Orotate Phosphoribosyltransferase. *Biochemistry* **1996**, *35*, 14–21.
- (14) Zhang, Y.; Luo, M.; Schramm, V. L. Transition States of Plasmodium Falciparum and Human Orotate Phosphoribosyltransferases. *J. Am. Chem. Soc.* **2009**, *131*, 4685–4694.
- (15) Zhang, Y.; Schramm, V. L. Pyrophosphate Interactions at the Transition States of Plasmodium falciparum and Human Orotate Phosphoribosyltransferases. *J. Am. Chem. Soc.* **2010**, *132*, 8787–8794.
- (16) Abdo, M.; Zhang, Y.; Schramm, V. L.; Knapp, S. Electrophilic Aromatic Selenylation: New OPRT Inhibitors. *Org. Lett.* **2010**, *12*, 2982–2985.
- (17) Zhang, Y.; Evans, G. B.; Clinch, K.; Crump, D. R.; Harris, L. D.; Fröhlich, R. F. G.; Tyler, P. C.; Hazleton, K. Z.; Cassera, M. B.; Schramm, V. L. Transition State Analogues of Plasmodium Falciparum and Human Orotate Phosphoribosyltransferases. *J. Biol. Chem.* **2013**, *288*, 34746–34754.
- (18) Henriksen, A.; Aghajari, N.; Jensen, K. F.; Gajhede, M. A Flexible Loop at the Dimer Interface is a Part of the Active Site of the Adjacent Monomer of Escherichia Coli Orotate Phosphoribosyltransferase. *Biochemistry* **1996**, *35*, 3803–3809.
- (19) González-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.
- (20) 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.
- (21) Wang, G. P.; Cahill, S. M.; Liu, X.; Girvin, M. E.; Grubmeyer, C. Motional Dynamics of the Catalytic Loop in OMP Synthase. *Biochemistry* **1999**, *38*, 284–295.
- (22) Shi, W.; Li, C. M.; Tyler, P. C.; Furneaux, R. H.; Grubmeyer, C.; Schramm, V. L.; Almo, S. C. The 2.0 Å Structure of Human Hypoxanthineguanine Phosphoribosyltransferase in Complex with a Transition-State Analog Inhibitor. *Nat. Struct. Biol.* **1999**, *6*, 588–593.
- (23) Grubmeyer, C.; Segura, E.; Dorfman, R. Active-Site Lysines in Orotate Phosphoribosyltransferase. *J. Biol. Chem.* **1993**, *268*, 20299–20304.
- (24) Goitein, R. K.; Chelsky, D.; Parsons, S. M. Primary C-14 and Alpha-Secondary H-3 Substrate Kinetic Isotope Effects for Some Phosphoribosyltransferases. *J. Biol. Chem.* **1978**, *253*, 2963–2971.
- (25) Ozturk, D. H.; Dorfman, R. H.; Scapin, G.; Sacchettini, J. C.; Grubmeyer, C. Locations and Functional Roles of Conserved Lysine Residues in Salmonella-Typhimurium Orotate Phosphoribosyltransferase. *Biochemistry* **1995**, *34*, 10755–10763.
- (26) Bhatia, M. Orotate Phosphoribosyltransferase: Kinetics and Structure Function Studies. Ph.D. Dissertation, Department of Chemistry, New York University, 1991.
- (27) Vanden-Eijnden, E.; Venturoli, M. Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies. *J. Chem. Phys.* **2009**, *130*, 194103.
- (28) Zinovjev, K.; Tuñón, I. Reaction coordinates and transition states in enzymatic catalysis. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1329.
- (29) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. String Method in Collective Variables: Minimum Free Energy Paths and Isocommittor Surfaces. *J. Chem. Phys.* **2006**, *125*, 024106.
- (30) Maragliano, L.; Vanden-Eijnden, E. On-the-fly String Method for Minimum Free Energy Paths Calculation. *Chem. Phys. Lett.* **2007**, *446*, 182–190.
- (31) Zinovjev, K.; Tuñón, I. Adaptive Finite Temperature String Method in Collective Variables. *J. Phys. Chem. A* **2017**, *121*, 9764–9772.
- (32) Poulsen, P.; Jensen, K. F.; Valentin-Hansen, P.; Carlsson, P.; Lundberg, L. G. Nucleotide-Sequence of the Escherichia-Coli Pyre Gene and of the DNA in front of the Protein-Coding Region. *Eur. J. Biochem.* **1983**, *135*, 223–229.

- (33) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132.
- (34) Kabsch, W. Integration, Scaling, Space-group Assignment and Post-refinement. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 133–144.
- (35) Evans, P. R.; Murshudov, G. N. How Good Are my Data and What Is the Resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204–1214.
- (36) Hough, M. A.; Wilson, K. S. From Crystal to Structure with CCP4. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2018**, *74*, 67.
- (37) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- (38) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501.
- (39) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.refine. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352–367.
- (40) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B., III; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci.* **2018**, *27*, 293–315.
- (41) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; 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.
- (42) Walker, R. C.; Crowley, M. F.; Case, D. A. The Implementation of a Fast and Accurate QM/MM Potential Method in Amber. *J. Comput. Chem.* **2008**, *29*, 1019–1031.
- (43) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (44) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (45) 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.
- (46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Wallingford, CT, 2009.
- (47) Klimovich, P. V.; Shirts, M. R.; Mobley, D. L. Guidelines for the Analysis of Free Energy Calculations. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 397–411.
- (48) Bennett, C. H. Efficient Estimation of Free-Energy Differences from Monte-Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (49) Torrie, G. M.; Valleau, J. P. Non-Physical Sampling Distributions in Monte-Carlo Free-Energy Estimation - Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (50) Martí, S.; Moliner, V.; Tuñón, I. QMcube (QM3): An All-Purpose Suite for Multi Scale QM/MM Calculations. To be published.
- (51) Minasov, G.; Shuvalova, L.; Dubrovskaya, I.; Winsor, J.; Papazisi, L.; Anderson, W. F. 1.9 Å Crystal Structure of Orotate Phosphoribosyltransferase (pyrE) *Francisella tularensis*. To be published.
- (52) Lukacs, C. M.; Abendroth, J.; Edwards, T. E.; Lorimer, D. Crystal Structure of Orotate Phosphoribosyltransferase (OPRTase) from *Burkholderia cenocepacia*. To be published.
- (53) Halavaty, A. S.; Shuvalova, L.; Minasov, G.; Winsor, J.; Dubrovskaya, I.; Kwon, K.; Anderson, W. F. 1.75 Å Resolution Crystal Structure of an Orotate Phosphoribosyltransferase from *Bacillus anthracis* str. “Ames Ancestor”. To be published.
- (54) Halavaty, A. S.; Minasov, G.; Shuvalova, L.; Dubrovskaya, I.; Winsor, J.; Kwon, K.; Anderson, W. F. 2.1 Å Resolution Crystal Structure of an Orotate Phosphoribosyltransferase (pyrE) from *Vibrio cholerae* O1 Biovar eltor str. N16961. To be published.
- (55) Kanagawa, M.; Baba, S.; Kuramitsu, S.; Yokoyama, S.; Kawai, G.; Sampei, G. Crystal Structure of Orotate Phosphoribosyltransferase from *Aeropyrum pernix*. To be published.
- (56) Liu, C.-P.; Xu, R.; Gao, Z.-Q.; Xu, J.-H.; Hou, H.-F.; Li, L.-Q.; She, Z.; Li, L.-F.; Su, X.-D.; Liu, P.; Dong, Y.-H. Structure of orotate phosphoribosyltransferase from the caries pathogen *Streptococcus mutans*. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2010**, *66*, 498–502.
- (57) Chang, C.; Li, H.; Collart, F.; Joachimiak, A. Crystal Structure of Orotate Phosphoribosyltransferase from *Streptococcus pyogenes*. To be published.
- (58) Kumar, S.; Krishnamoorthy, K.; Mudeppa, D. G.; Rathod, P. K. Structure of Plasmodium falciparum Orotate Phosphoribosyltransferase with Autologous Inhibitory Protein-Protein Interactions. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2015**, *71*, 600–608.
- (59) 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.
- (60) McClard, R. W.; Holets, E. A.; MacKinnon, A. L.; Witte, J. F. Half-of-Sites Binding of Orotidine 5'-Phosphate and Alpha-D-5-Phosphorylribose 1-Diphosphate to Orotate Phosphoribosyltransferase from *Saccharomyces cerevisiae* Supports a Novel Variant of the Theorell-Chance Mechanism with Alternating Site Catalysis. *Biochemistry* **2006**, *45*, 5330–5342.
- (61) Shimosaka, M.; Fukuda, Y.; Murata, K.; Kimura, A. Purification and Properties of Orotate Phosphoribosyltransferases from *Escherichia coli*-K-12, and its Derivative Purine-Sensitive Mutant. *J. Biochem.* **1985**, *98*, 1689–1697.