

Revealing the unknown aspects of initial steps of prenylated flavin mononucleotide biosynthesis – the role of Lys129 in *PaUbiX*

Szymon Żaczek, Agnieszka Dybala-Defratyka*

Institute of Applied Radiation Chemistry, Faculty of Chemistry, Lodz University of Technology, Zeromskiego 116, 90-924 Lodz, Poland

Highlights:

- Lys129 is required for prenylated flavin mononucleotide (prFMN) biosynthesis via *PaUbiX*
- Lys129 shares a single proton with Glu140, which was previously reported to directly donate a proton to dimethylallyl monophosphate (DMAP) in the initial step of prFMN biosynthesis
- DMAP breakdown might occur without the proton transfer itself – released phosphate group is well stabilized by Lys129 and Glu140
- This behaviour is likely to be conserved within the whole UbiX enzyme family

Corresponding Author

*agnieszka.dybala-defratyka@p.lodz.pl

Abstract

Background

Prenylated flavin mononucleotide (prFMN) is a recently discovered, heavily modified flavin compound. It is the only known cofactor that enables enzymatic 1,3-dipolar cycloaddition reactions. It is produced by enzymes from UbiX family, from flavin mononucleotide and either dimethylallyl mono- or diphosphate. prFMN biosynthesis is currently reported to be initiated by a protonation of the substrate by Glu140.

Methods

Computational chemistry methods are applied herein - mostly different flavors of molecular dynamics MD, such as Constant pH MD, hybrid Quantum-Mechanical / Molecular Mechanical MD, and classical MD.

Results

Glu140 competes for a single proton with Lys129 but it is the latter that adopted a protonated state throughout most of the simulation time. Lys129 plays a key role in the positioning of the DMAP's phosphate group within the *PaUbiX* active site. DMAP's breakdown into a phosphate and a prenyl group can be decoupled from the protonation of the DMAP's phosphate group.

Conclusions

The role of Lys129 in functioning of *PaUbiX* is reported for the first time. The severity of interactions between Glu140, Lys129, and DMAP's phosphate group enables an unusual decoupling of phosphate's protonation from DMAP's breakdown. Those findings are most likely conserved throughout the UbiX family to the structural resemblance of active sites of those proteins.

Significance

Mechanistic insights into a crucial biochemical process, biosynthesis of prFMN, are provided. This study, although purely computational, extends and perfectly complements the knowledge obtained in classical laboratory experiments.

Keywords

UbiX; prenylated flavin mononucleotide; protonation states; proton transfer's decoupling; phosphate

1. Introduction

Flavins are among the most established chemical species that may act as cofactors, enabling various enzymes to perform biocatalysis [1]. They have been mostly known for its capabilities in redox chemistry [2] due to the versatility of the conjugated π electron system within isoalloxazine heterocycle [3]. And yet, recent discoveries show that flavins are capable of much more [4]. Decarboxylation of α,β -unsaturated acids [5], carboxylation reactions [6,7], biosynthesis of tautomycin [8], and ubiquinone [9], degradation of the precursor of phenazine metabolites [10] - those processes are all enabled by prenylated flavin mononucleotide (prFMN). Its uniqueness stems from the fact that the structure of a traditional tricyclic flavin ring (Figure 1A) is extended by an additional non-aromatic ring (Figure 1B), what makes prFMN capable of facilitating an unparalleled covalent catalysis, such as the first known enzymatic 1,3-dipolar cycloaddition [5].

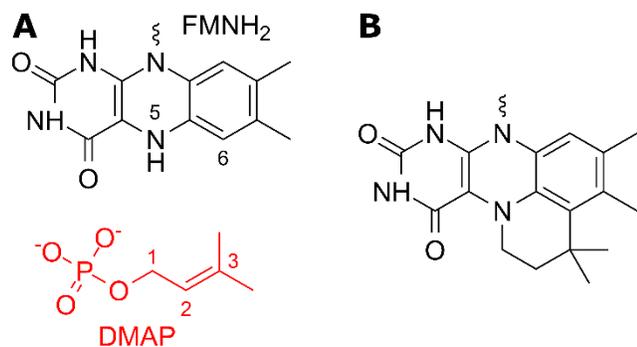


Figure 1. Chemical structures of: A) substrates for prFMN biosynthesis: reduced flavin mononucleotide (FMNH₂) and dimethylallyl monophosphate (DMAP); B) Prenylated flavin mononucleotide.

Enzymes from the UbiX family synthesize prFMN [11,12]. They require flavin mononucleotide and either dimethylallyl pyrophosphate (DMAPP) [13] or dimethylallyl monophosphate (DMAP) [9,12] as substrates. The first major step in the prFMN biosynthesis is the addition of a prenyl moiety from a given phosphate substrate to the flavin by forming the N5 - C1 bond (Figure 1A). For this to occur though, a phosphate group must be eliminated from DMAP or DMAPP, which is reported to be possible via protonation of the phosphate (Figure 2A). [14] Furthermore, in this process, the formation of the fourth non-aromatic ring in flavin by creating the C6 - C1 bond is reported to be initiated by another protonation of the phosphate (Figure 2B). It was proposed that it is Glu140 (*Pseudomonas aeruginosa* UbiX (*PaUbiX*) numbering) which acts as a proton donor in both of those steps [14]. Nonetheless, this mechanistic scenario was proposed based on a structural analysis of available crystal structures - to the best of our knowledge, there are no direct measurements supporting those claims.

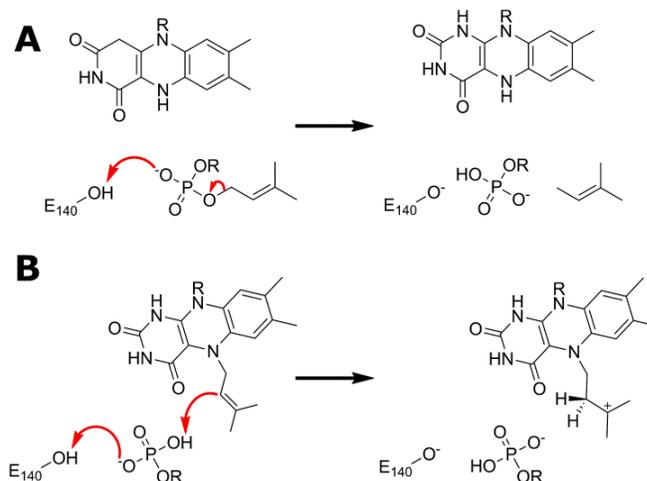


Figure 2. Individual steps of prFMN biosynthesis that are reported to be enabled by Glu140, as reported by Marshal et al. [14] A) The initial step in the prFMN biosynthesis - DMAP or (DMAPP) breakdown coupled with protonation of the phosphate group by Glu140. B) The initial step in terpene cyclase step in the prFMN biosynthesis - protonation of the prenyl moiety coupled with protonation of the phosphate (or pyrophosphate) by Glu140.

Even though most of the residues within active sites are conserved throughout the UbiX family [14] (Figure 3), the selectivity of UbiX enzymes varies. [12] *Aspergillus niger* UbiX (*AnUbiX*) and *Escherichia coli* UbiX (*EcUbiX*) [14] require DMAPP, *PaUbiX* works only with DMAP [9], whereas *Escherichia coli* O157UbiX is capable of utilizing both DMAP and DMAPP [14] in the prFMN biosynthesis. Performing biocatalysis with DMAP is unmet elsewhere - it is not used in any other metabolic pathways. Interestingly, *EcUbiX* might produce prFMN upon incubation with DMAP if an additional phosphate/pyrophosphate is provided to the reaction pot [14]. This means that the phosphate remaining in the active site is crucial for the prFMN biosynthesis, what is achieved with a very strong stabilization provided by a hydrogen-bonding network [15]. In the closest vicinity of DMAPP/DMAP's phosphate group apart from Glu140 there is another titrable amino acid - Lys129. Interestingly, the protonation state analysis based on the crystal structure showcased that Glu140 most likely adopts a deprotonated state [15] - in such a case, this residue would not be able to act as a proton donor, therefore it would not facilitate the prenyl-FMN adduct formation (Figure 2). On the other hand, Lys129 is likely to adopt a protonated state [15] and its location makes it a feasible candidate for donating a proton either directly to the ligand or via prior Glu140 protonation.

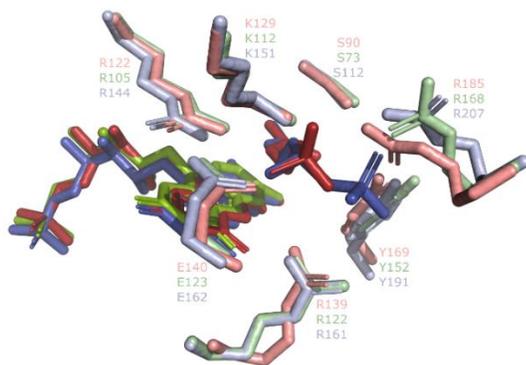


Figure 3. Alignment of active sites of different enzymes within the UbiX family: *PaUbiX* (red), *O157UbiX* (green), and *AnUbiX* (blue) active sites. Amino acid residues are shown in pale tones, whereas ligands are depicted in more saturated tones. All the structures have FMN in their active sites. *AnUbiX* has also DMAPP, *PaUbiX* has also DMAP, whereas *O157UbiX* crystal structure was resolved without any phosphate bound.

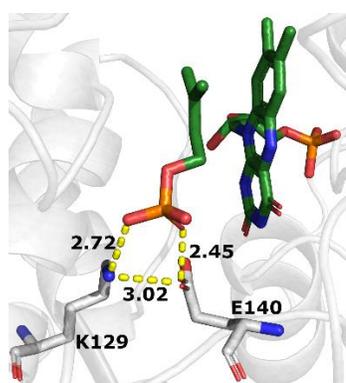


Figure 4. The active site of the *PaUbiX* complex with DMAP and FMN (PDB ID: 4ZAF). Carbon atoms of ligands are colored in green, amino acids carbon atoms of Lys129 and Glu140 are colored in gray. Distances are shown in Å.

Since protonation state analysis based on the crystal structures indicated that Glu140 might not be capable of fulfilling its assumed role in the biosynthesis of prFMN due to its adoption of a deprotonated state [15], we have resorted to verifying it with a much more robust methodology - Constant pH MD simulations. Subsequently, we have considered potential pathways of prenyl-FMN formation, using adaptive string method [16] coupled with the Quantum Mechanics/Molecular Mechanics Molecular Dynamics (QM/MM MD) approach within the AMBER package environment.

Hence the objective of the present study was two-fold. First, we aimed at exploring whether the static picture of events taking place in the active site of PaUbiX can be confirmed by other independent methods allowing for dynamic evolution of the studied system in time. Second, by modeling the initial steps of prFMN formation we investigated and described the role of Lys129 and Glu140 and their mutual relation in PaUbiX functioning.

2. Results and Discussion

2.1. Constant pH Molecular Dynamics

Both explicit and implicit solvent CpH MD simulations have been performed. In the former simulations, pK_a of Lys129 was equal to 7.7, whereas pK_a of Glu140 was equal to 6.4 (Figure 1A, Table S1). Those values indicate that both of those residues heavily interact with their surroundings – their average pK_a values in proteins are equal to 10.7 and 4.1 [17]. Implicit solvation, on the other hand, seems to underestimate the influence of the surroundings on the considered herein titrable amino acids – in fact, no events of changing protonation states of Lys129 had place, thus it was not possible to assess its pK_a values. Additionally, pK_a of Glu140 obtained using implicit solvation was 4.8, being close to the native 4.1 value. Consequently, Glu140 was protonated for only 0.6% of the simulation time during implicit solvation CpH MD (Figure 1A, Table S1). This is in accordance with findings of Swails et al. [18] stating that implicit solvent CpH MD might provide unrealistic results. However, since there are no experimental values of pK_a for *PaUbiX* available, we are unable to draw any further conclusions in that regard. Nonetheless, the fact that in explicit solvent CpH MD Lys129 was protonated for 82.7% of the simulation time and Glu140 was protonated for 20.2% (Figure 1B, Table S1) indicates that their protonation states are strictly correlated and if one of them occurs in a protonated state, the other one occurs in a deprotonated state. Only for 2.5 % of the time they both remained protonated. This means that they compete for a single proton which is exchanged between them, with a strong preference to stay on Lys129 though. However, if a residue is considered a proton donor, the distance which proton would need to travel to the acceptor atom should not be too large. In fact, it turns out that Lys129 formed a hydrogen bond with DMAP whenever it occurred in a protonated state, whereas Glu140 formed hydrogen bond with DMAP for only 12.0% of the simulation time, even though it was protonated for 20.2% of the time (Table S1). Thus, Glu140 would be rarely capable of donating the proton to DMAP during the prFMN biosynthesis if it were necessary for the process to proceed. Those results contradict a current hypothesis in which Glu140 is believed to donate its proton to DMAP [9,14] – rarely it adapts a protonated form, thus the probability of the subsequent proton transfer is modest. Interestingly, this hypothesis was based on the obtained crystal structures of *PaUbiX* enzyme and an assumption that the released phosphate group should not occur in a fully deprotonated state in the active site [9,14]. The latter assumption is obviously justified (Figure S1B), however, it has been shown that the proton transfer onto the released phosphate might be decoupled from the actual release reaction (though to the best of our knowledge, it was only observed in the mutated enzymes so far [19]). On the other hand, water molecules might be directly involved in the phosphate release [20–23] but no water molecules were included in the QM region in this study thus this alternative was not further tested. Considering all of this, by using QM/MM MD simulations, we have further investigated the roles of Lys129 and Glu140 in the initial steps of prFMN biosynthesis – DMAP's breakdown and prenyl-FMN adduct formation.

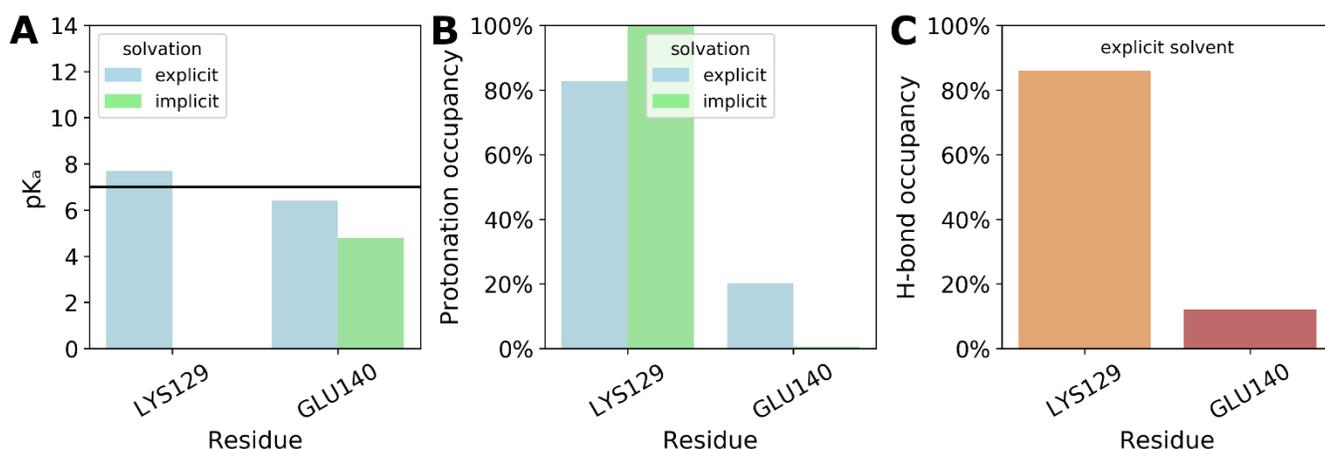


Figure 5. Results of Constant pH MD simulations. A) pK_a values of the considered residues. There are no values available for Lys129 and Tyr169 for implicit solvation because there were no protonation changes observed, thus CpH MD failed to predict pK_a. B) Occupancy of protonated states throughout the simulations. C) Occupancy of hydrogen bond between considered residue and any oxygen atom from DMAP. For Glu140, both oxygen atoms of carboxylic group were considered as probable donors. An interaction was considered a hydrogen bond if the distance between donor and acceptor atoms was below 3.0 Å

2.2. DMAP's breakdown and prenyl-FMN adduct formation

Near neutral pH DMAP tends to be in either deprotonated or a single-protonated state (Figure S1A), whereas phosphates are either single or double-protonated (Figure S1B). Therefore, as suggested by Marshall et al [14], prior to (or concertedly to) the detachment of prenyl moiety from DMAP, the phosphate group should most likely be protonated first from one of the neighboring residues from the active site. Since the results of Constant pH MD (Figure 5, Table S1) clearly indicated that there is a competition between Lys129 and Glu140 for a single proton (with a strong preference towards Lys129), if DMAP were to be protonated via Glu140, an earlier proton transfer from Lys129 to Glu140 would be most likely required. Nonetheless, in the available crystal structure of *PaUbiX* complexed with DMAP and FMN, DMAP is in a very close vicinity of both Lys129 and Glu140 (Figure 4). Those three species form a very tight, square-like structure of atoms capable of exchanging a proton between them. Based on a crystal structure alone, it might also seem as if both Lys129 and Glu140 should occur in a protonated state, since the relevant distances between DMAP's oxygen atoms and either nitrogen from Lys129 or oxygen atom from Glu140 are equal to 2.72 and 2.45 Å, respectively. Nonetheless, as it was recently proved by Huang et al. [24], predicting protonation states of catalytic residues just based on a single static structure might not be accurate. Instead, dynamic behavior of proteins should be considered in such studies, as it is reported herein. Interestingly, the mutual arrangement of Lys129, Glu140 and the phosphate is retained in the crystal structure after DMAP's breakdown and formation of a prenyl-FMN adduct (Figure 6). This stability might suggest that interactions between the phosphate and its neighboring residues remain unaltered throughout the DMAP's breakdown and that the proton transfer could be decoupled from this process (as it was demonstrated by Lopata et al.

[19]). Considering those findings, we have explored the concerted mechanism comprising the DMAP's breakdown (the C – O bond cleavage) and a prenyl-FMN adduct formation (the C – N bond formation (Figure 7A) without any additional proton transfer reactions. Throughout this pathway Lys129 remained protonated, whereas Glu140 was deprotonated (as was established with CpH MD simulations). Also, DMAP was in a deprotonated state.

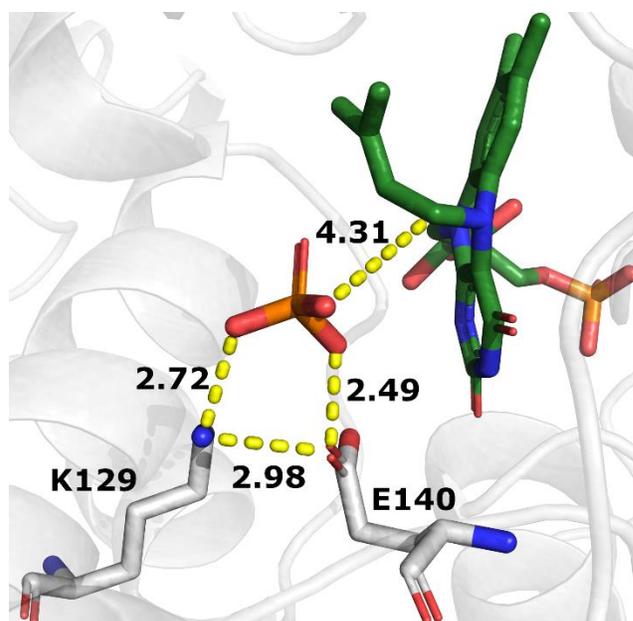


Figure 6. The active site of the *PaUbiX* complex with FMN-prenyl covalent adduct and a phosphate (PDB ID: 4ZAV) [5]. Carbon atoms of ligands are colored in green, amino acids carbon atoms of Lys129 and Glu140 are colored in gray. Distances are shown in Å.

The barrier height for the concerted DMAP's breakdown and the prenyl-FMN adduct formation calculated with DFTB₃ with 3OB:OP parameters is 25.1±0.3 kcal/mol,

though with only a little apparent stabilization provided at the products state (PS) region (about 3 kcal/mol as compared to the transition state, TS, Figure 7). The energy barrier seems to be a bit overestimated, but one should bear in mind that the QM method used in this study – DFTB3 – is only a semiempirical method thus it does not allow for a fully quantitative analysis. However, there are several studies from recent years demonstrating its successful application for the discrimination of reaction mechanisms in the phosphoryl-transfer reactions [25–29]. Nonetheless, PS indeed seems to be stable – additional relaxation simulations started from the PS structure confirm that no further simultaneous process occurs on its own. (Figure S2). Additionally, released phosphate group is very well stabilized via hydrogen bonding by the remainder of active site residues both at TS and PS and perhaps such stabilization is sufficient for the decoupling phosphate protonation from DMAP breakdown (Figure S3). As it was already mentioned, the pathway for a direct DMAP's breakdown and prenyl-FMN adduct formation as the initial step of prFMN biosynthesis, without any prior proton transfer to the phosphate group from DMAP is in contrast with the results from previous studies. The hypothesis that Glu140 acts as a proton donor in the initial step of the prFMN biosynthesis was based solely on the obtained crystal structure of PaUbiX enzyme and not on the probing actual proton transfer reactions. Apart from the results presented herein, we have also attempted to model different pathways of prenyl-FMN adduct formation, DMAP's disassembly and (possibly) accompanying proton exchanges between Lys129, Glu140, and DMAP's phosphate but none of them has provided reliable results within the used methodology (Figure 8). We have managed to obtain viable results for the proton transfer from Lys129 to Glu140 (Figure S4), which was considered as a first step in a few of the studied pathways though. More details regarding those attempts at modelling prFMN-adduct formation can be found in Supporting Information (Table S2). Inability to obtain viable results for the remainder of the attempted pathways may stem from the limitations of the used herein QM methodology (DFTB3/3ob:op) – in fact it was shown, that DFTB3 may yield quite significant errors with regard to hydrogen bonding [30] which are at the very core of the system studied herein. Also, one could always argue whether the size of QM region should not be larger, which we address in computational details section. Therefore, for quantitative analysis of the initial steps of prFMN biosynthesis, one should rather resort to a more precise description of the QM region such as DFT. Despite these shortcomings, Lys129 is unarguably vital for the prFMN biosynthesis. Furthermore, since other UbiX variants have both Lys and Glu residues at the corresponding locations (Lys151 and Glu162 for *An*UbiX, Lys112 and Glu123 for *O157*UbiX), we assume that the interactions within the active site are most likely conserved throughout the whole UbiX family.

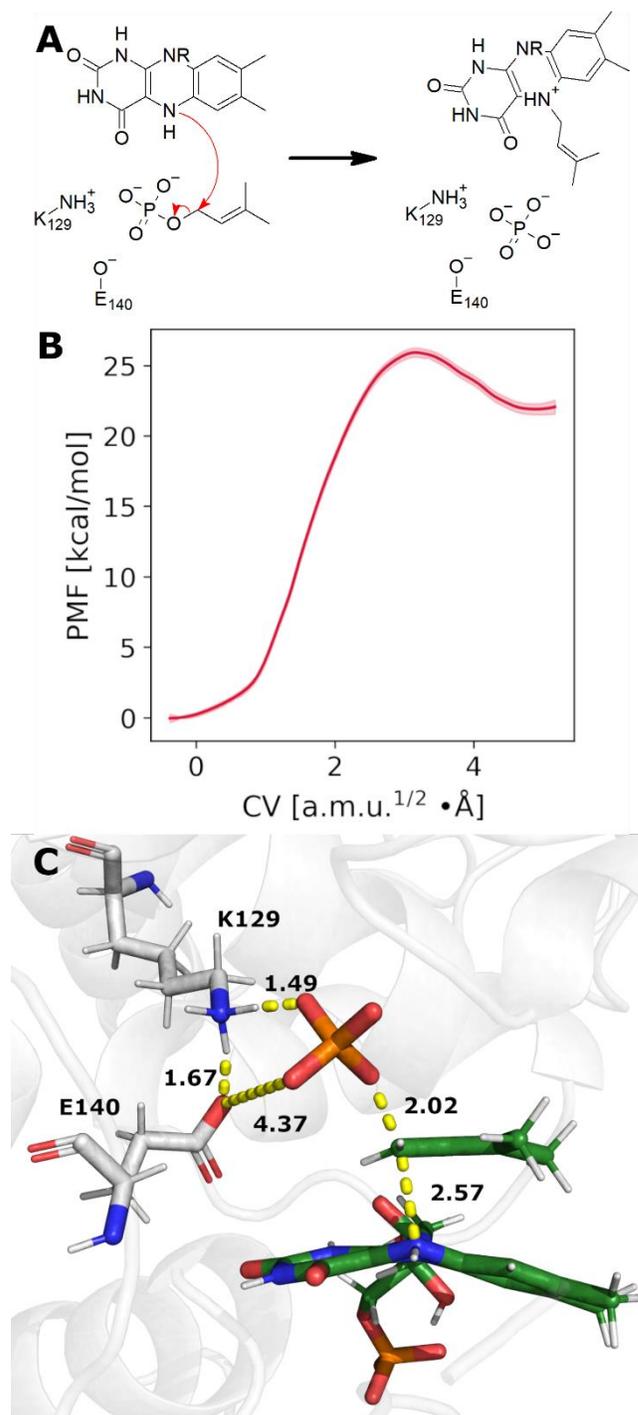


Figure 7. A) The investigated DMAP's breakdown and prenyl-FMN adduct formation reaction. Collective Variables used for this reaction are highlighted by arrows (C – O bond rupture and C – N bond formation). B) Free energy profile for the DMAP's breakdown and the prenyl-FMN adduct formation. C) Representative structure of transition state for the investigated reaction.

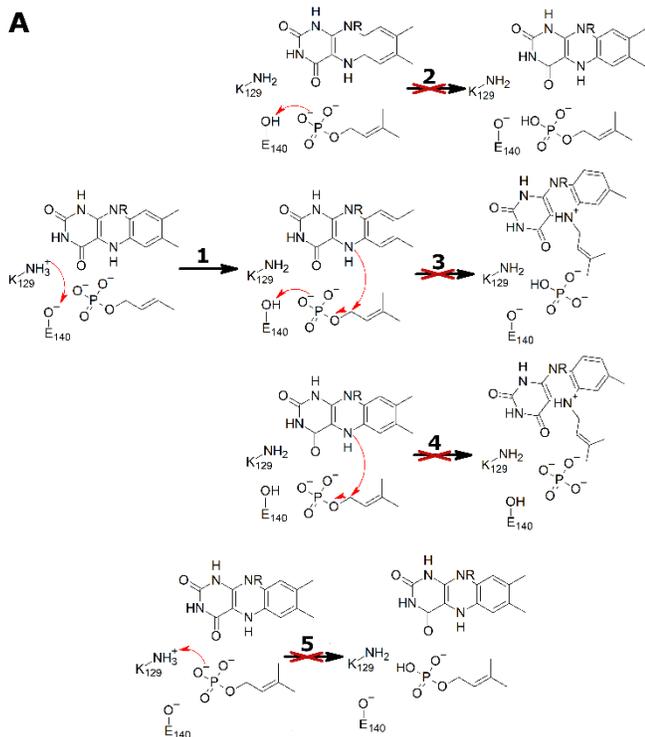


Figure 8. Reactions that were investigated as alternative pathways for the prenyl-FMN adduct formation. A crossed arrow indicates that based on the calculations presented herein the reaction does not occur this way. 1) Proton transfer from Lys129 to Glu140. 2) Proton transfer from Glu140 to DMAP's phosphate. 3) Proton transfer from Glu140 to DMAP's phosphate coupled with DMAP's C-O bond cleavage and prenyl-FMN adduct formation. 4) DMAP's C-O bond cleavage and prenyl-FMN adduct formation with protonated Glu140. 5) Proton transfer from Lys129 to the DMAP's phosphate group.

3. Conclusions

Herein, based on the evidence coming from the protonation state analysis and subsequent free energy surface simulations, we propose that Lys129 is crucial for the functioning of *PaUbiX*. Since it is capable of donating its proton to Glu140 (being itself predominantly protonated though), those two residues together stabilize the position of the phosphate group throughout the biosynthesis of prFMN, both when it is bound to the prenyl group, and when it is released from it. Previously, it was suggested that Glu140 donated its proton to the phosphate group in DMAP, triggering that way the whole prFMN biosynthesis [9,14]. Although we do not explicitly reject this hypothesis, we see the main role of Glu140 differently as our results presented in this work indicate that this residue instead of being the main player on the stage contributes to the hydrogen bonding network within the active site. Moreover, since Glu140 competes for a proton with Lys129 (accordingly to our calculations, the probability of its protonation is 4 times less than the probability of lysine protonation), prFMN biosynthesis would also require a proton transfer from Lys129 to Glu140 as the very first step. What's more, we showed herein, that DMAP's breakdown and FMN-prenyl adduct formation may be decoupled from the

proton transfer onto DMAP's phosphate group. The exact preferences for the whereabouts of proton in each step of the prFMN biosynthesis though should perhaps be established using different methodology than was used herein (higher-level QM methods would be more robust for this purpose) but Lys129 should be considered as a crucial residue for prFMN biosynthesis – it is the source for the proton that either facilitates DMAP's cleavage (if a proton indeed gets transferred) or is required for stabilizing the phosphate in its double-deprotonated form in the active site, which according to our study, is more probable. Therefore, Lys129 is a crucial residue for the functioning of *PaUbiX* and its role is probably conserved throughout the whole UbiX enzyme family, due to the similarity of the active sites of *PaUbiX*, *AnUbiX*, and *O157UbiX*.

4. Computational Methods

4.1. Initial system setup

All of the molecular dynamics calculations were performed for the *PaUbiX* enzyme (PDB ID: 4ZAF) [9] in complex with the reduced flavin mononucleotide (original structure contains oxidized FMN but for the biosynthesis to occur, FMN must be reduced [9]) and dimethylallyl monophosphate (DMAP). Three individual chains were used for simulations, since this is a minimal system with a fully formed active site [9]. Such a structure was obtained according to the information about unit cells provided with the original crystal structure. Symmetry operations were carried out in the PyMOL program [31]. The crystal structure had two residues missing (terminal Asp and Glu) – their initial coordinates were built by using the MODELLER software [32]. No disulfide bridges in the protein were detected by the pdb4amber program, which is a part of the Amber software suite [33]. Initial protonation states of titrable amino acids were determined with both ligands (FMN and DMAP) in the active site using the Propka software with [34,35], implemented at the PDB2PQR webserver [36] at pH 7.0. GAFF2 was used as a force field for ligands [37]. Charges for ligands were derived with the AM1-BCC model [38,39]. Missing force field parameters for ligands were obtained using the Antechamber program [40]. Crystal water molecules were deleted from the system because they caused instabilities in MD simulations. The protein atoms were treated with the ff14SB force field [41]. All flavors of MD simulations performed herein were carried out with the Amber software suite [33].

1.1. Classical Molecular Dynamics

Classical molecular dynamics simulations were performed prior to QM/MM MD. 24 Na⁺ ions were needed to neutralize the system. The protein-ligand complex was solvated with TIP3P water molecules (the initial size of the periodic box was equal to 89.2 Å x 109.8 Å x 98.3 Å), amounting to 80 265 atoms in the model. Classical MD simulations were performed using the GPU (CUDA) version of PMEMD (Particle Mesh Ewald Molecular Dynamics) code [42,43] implemented in the Amber software suite. Prior to molecular dynamics, minimization was performed for 4000 steps with 2000 steps of the steepest descent minimization, followed by another 2000 steps of the conjugate gradient

minimization. The resulting structure was subjected to heating to 298.15 K for 100 ps using an NVT ensemble. Then, preliminary equilibration was run for 5 ns in an NPT ensemble. Both of those steps had a time-step of 1 fs. After that, the second run of equilibration was performed with a time-step of 2 fs for 120 ns. Langevin dynamics [44] with collision frequency equal to 2 ps was used for temperature scaling and Berendsen barostat [45] was used to maintain the pressure with a pressure relaxation time of 2 ps. Periodic boundary conditions and a 12 Å cut-off distance for non-bonded interactions were applied throughout the simulations. A shake algorithm [46] was used to constrain bonds involving hydrogen atoms.

4.3. QM/MM Molecular Dynamics

The resulting structure from classical MD was used as a starting point for QM/MM MD. Sidechains of Lys129 and Glu140, isoalloxazine ring from FMN and the entire DMAP molecule were treated with a QM Hamiltonian. This amounted to 76 real atoms in QM region with QM charge equal to -2. There were 3 covalent bonds on the boundary between QM and MM regions (2 bonds between the C α and C β carbon atoms in Lys129 and Glu140, respectively and the C11 - C12 bond in the tail of FMN) - they were capped with the hydrogen-link atoms. Only the most relevant atoms for the studied processes were included in QM region due to computational resources limitation but since it included all of the reacting species, the outcome of the study is unlikely to change if the QM region was extended [47]. All of those calculations were performed with the Amber QM/MM engine [48,49], using SCC-DFTB3 [50] to represent the QM region of the overall system, along with the DFTB3/3OB:OP parameters [51]. Results of the gas-phase calculations of proton affinities for the considered herein proton donors (Table S2) and DMAP as acceptor (Table S3) are comparable to the values reported for the species which were used for the DFTB3 parametrization [50]. Therefore, DFTB3 seems to be a suitable method for the studied herein reactions. To facilitate convergence of the QM subsystem, electronic temperature was set to 100 K. Next, the system was subjected to minimization (2000 steps in total; 1000 of steepest descent and 1000 steps of conjugate gradient), heating (up to 298.15 K) and a short equilibration for a total of 50 ps, with a cut-off of 8 Å and time-step of 1 fs (other parameters were defined as in the previous case). Shake was turned off for the part of the system treated quantum-mechanically. Switching function was used for long-range electrostatics, with default parameters. The structure obtained after the equilibration was chosen as a Reactant State (RS) for all of the subsequent mechanistic simulations. Potentials of Mean Force (PMF) was calculated with a locally changed installation of Amber18, which included the Adaptive finite temperature string method [16]. To get PMF using this method, it is required to have at least structures of reactants (RS) and products states (PS). For each of the considered reactions (Figure 7, Figure 8, Table S2), PS structures were obtained by harmonically restraining a specified Collective Variable to a value that would make a reasonable structure - for instance, for the reaction considered in Figure 7, the N - C bond was

restricted to 1.9 Å and the O - P bond was restricted to 3.0 Å. Additionally, in order to speed up convergence of the string calculations, 4 additional guess structures (so a total of 6 structures was used for each string: RS, PS and 4 intermediates) were determined for each process in the considered CV space in-between RS and PS. Each of those structures was scaled regarding considered CVs and they acted as initial guesses for how a geometry at this particular point in reaction might look like. So, for the reaction considered in Figure 7, the 2nd initial structure was restrained to have the N - C bond of 3.15 Å, whereas O - P bond was restricted to 1.84 Å. After obtaining all those structures for all the considered reactions, the PMF calculations were performed. When we considered proton transfers onto DMAP, the acceptor atom was always an oxygen atom from the phosphate group of DMAP that was the closest to the considered donor atom. The energy required for the rotation of the DMAP's phosphate group is minimal (Figure S5), thus it should be relatively easy for it to rotate even when it is encompassed by an enzyme. Nonetheless, we believe that considering the closest oxygen atom as an acceptor is the best option, since it does not introduce large distances for the proton to travel. Time-step for performing string simulations was 0.5 fs - other parameters were kept as previously. All the string calculations were performed with 30 nodes (30 distinct points defined along the evolution of CVs). The first actual step in used herein string method is the determination of Minimum Free Energy Path (MFEP). This is done by optimizing values of CVs in each node, until subsequent simulations step does not introduce significant changes of those CVs. It was done with using additional, orthogonal potential of 400 (kcal/mol)/(a.m.u. \cdot Å). In our case, all of the successful calculations reached convergence of MFEP with values below 0.1 a.m.u. $^{1/2}$ ·Å (Figure S6). After determining MFEP, Umbrella Sampling simulations were performed to obtain the actual PMF. More rigorous description of how adaptive string method works may be found elsewhere [16]. We have included additional information on pathways that did not lead to expected products (Figure 8) along with the reasoning why these pathways were considered as "failed" in SI.

4.4. Implicit Solvent Constant pH Molecular Dynamics

Implicit solvent Constant pH MD simulations were performed with the GPU implementation of the PMEMD engine [52] from the Amber software suite [33], along with the Generalized Born solvation model [53,54]. Salt concentration was set to 0.1 M, whereas solvent pH set to 7.0, without any cut-off for the non-bonded interactions. Prior to MD, minimization was performed for 4000 steps with 2000 steps of the steepest descent minimization, followed by another 2000 steps of the conjugate gradient minimization. The resulting structure was subjected to gradual heating to 298.15 K for 1 ns. An equilibration was performed for another 3 ns. Next, the titrable residues located within 15 Å radius from the dimethylallyl monophosphate molecule were allowed to change their protonation states. Such an exchange was attempted every 10 fs, though only a single residue was examined at a single step. Production runs were performed for 500 ns. For all the MD steps, time-step

was equal to 2 fs. Langevin dynamics [44] with collision frequency equal to 5 ps was used for temperature scaling. A shake algorithm [46] was used to constrain bonds involving hydrogen atoms.

4.5. Explicit solvent Constant pH Molecular Dynamics

Explicit solvent Constant pH MD simulations [18] were carried out within the Amber suite, using the GPU implementation of the PMEMD engine. System preparation, minimization, heating, and equilibration were done accordingly as the same steps performed using classical MD. After equilibration, titrable residues located within 15 Å radius from the dimethylallyl monophosphate molecule were allowed to change their protonation states. Such an exchange was attempted every 200 fs with a random single residue examined at each step, followed by 200 fs of solvent equilibration. Salt concentration during an attempt of changing protonation state was equal to 0.1 M, whereas solvent pH was equal to 7. Explicit solvent CpH MD was run with 10 disconnected replicas of the system, amounting to the total time of 1 μs of simulations. Other parameters that were not specified in this paragraph were the same as in the classical MD simulations.

4.6. Semi-empirical calculations

Semi-empirical calculations were carried out using the DFTB₃ [50] Hamiltonian with the DFTB₃/3OB parameters [51] implemented in the ADF program package [55]. Proton affinities (Tables S2 and S3) were calculated using the method described by Gaus et al. [50] according which proton affinity is the negative enthalpy change for the gas-phase reaction $A^- + H^+ \rightarrow AH$, where A^- is considered an anion, H^+ is a proton and AH is a protonated form of A^- . Rotation of the phosphate group (Figure S6) within DMAP was defined as the C₁-O-P-O (C₁ atom is numbered in Figure 1, subsequent oxygen is the oxygen between the C₁ and P atoms, the last oxygen atom is one of the terminal oxygen atoms of DMAP) dihedral angle with calculating energy every 6° degree for a total of 42 steps. This investigation was started from the optimized structure of DMAP. All the semi-empirical calculations were carried out in gas-phase and all the reported energies are potential energies.

ASSOCIATED CONTENT

Supporting Information. Detailed results of CpH MD simulations, additional information about failed pathways for prenyl-FMN adduct formation, proton affinities for amino acids and DMAP, results of predicting protonation states of DMAP and the phosphate, structure after additional PS stabilization from DMAP's breakdown and prenyl-FMN formation, hydrogen-bonding networks in the active site, free energy profile for proton transfer from Lys129 to Glu140, DMAP's rotation, convergence curves for PMF calculations.

ACKNOWLEDGMENT

This work was financially supported from the budgetary sources for years 2017 – 2021, as a research project within the 'Diamontowy Grant' Program (grant agreement no. 0145/DIA/2017/46). This research was supported in part by PLGrid Infrastructure.

AUTHOR CONTRIBUTIONS

Szymon Żaczek: Conceptualization, Methodology, Investigation, Visualization, Writing – Original Draft, Writing – Review & Editing, Funding acquisition; **Agnieszka Dybala-Defratyka:** Conceptualization, Methodology, Supervision, Writing – Review & Editing

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

- [1] R.L. Fagan, B.A. Palfey, Flavin-dependent enzymes, in: *Compr. Nat. Prod. II Chem. Biol.*, Elsevier Ltd, 2010: pp. 37–113. <https://doi.org/10.1016/b978-008045382-8.00135-0>.
- [2] F. Müller, The flavin redox-system and its biological function., *Top. Curr. Chem.* 108 (1983) 71–107. https://doi.org/10.1007/3-540-11846-2_3.
- [3] C.T.W. Moonen, J. Vervoort, F. Mueller, Reinvestigation of the structure of oxidized and reduced flavin: carbon-13 and nitrogen-15 nuclear magnetic resonance study, *Biochemistry.* 23 (1984) 4859–4867. <https://doi.org/10.1021/bio0316a007>.
- [4] V. Piano, B.A. Palfey, A. Mattevi, Flavins as Covalent Catalysts: New Mechanisms Emerge, *Trends Biochem. Sci.* 42 (2017) 457–469. <https://doi.org/10.1016/j.tibs.2017.02.005>.
- [5] K.A.P. Payne, M.D. White, K. Fisher, B. Khara, S.S. Bailey, D. Parker, N.J.W. Rattray, D.K. Trivedi, R. Goodacre, R. Beveridge, P. Barran, S.E.J. Rigby, N.S. Scrutton, S. Hay, D. Leys, New cofactor supports α,β-unsaturated acid decarboxylation via 1,3-dipolar cycloaddition, *Nature.* 522 (2015) 497–501. <https://doi.org/10.1038/nature14560>.
- [6] S.E. Payer, S.A. Marshall, N. Bärland, X. Sheng, T. Reiter, A. Dordic, G. Steinkellner, C. Wuensch, S. Kaltwasser, K. Fisher, S.E.J. Rigby, P. Macheroux, J. Vonck, K. Gruber, K. Faber, F. Himo, D. Leys, T. Pavkov-Keller, S.M. Glueck, Regioselective para-Carboxylation of Catechols with a Prenylated Flavin Dependent Decarboxylase, *Angew. Chemie - Int. Ed.* 56 (2017) 13893–13897. <https://doi.org/10.1002/anie.201708091>.
- [7] K.A.P. Payne, S.A. Marshall, K. Fisher, M.J. Cliff, D.M. Cannas, C. Yan, D.J. Heyes, D.A. Parker, I. Larrosa, D. Leys, Enzymatic Carboxylation of 2-Furoic Acid Yields 2,5-Furandicarboxylic Acid (FDCA), *ACS Catal.* 9 (2019) 2854–2865. <https://doi.org/10.1021/acscatal.8bo4862>.
- [8] T. Annaval, L. Han, J.D. Rudolf, G. Xie, D. Yang, C.-Y. Chang, M. Ma, I. Crnovcic, M.D. Miller, J. Soman, W. Xu, G.N. Phillips, B. Shen, Biochemical and Structural Characterization of TtnD, a Prenylated FMN-Dependent Decarboxylase from the Tautomycin Biosynthetic Pathway, *ACS Chem. Biol.* 13 (2018) 2728–2738. <https://doi.org/10.1021/acscchembio.8bo0673>.
- [9] M.D. White, K.A.P. Payne, K. Fisher, S.A. Marshall, D. Parker, N.J.W. Rattray, D.K. Trivedi, R. Goodacre, S.E.J. Rigby, N.S. Scrutton, S. Hay, D. Leys, UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis, *Nature.* 522 (2015) 502–506. <https://doi.org/10.1038/nature14559>.
- [10] K.C. Costa, L.S. Moskatel, L.A. Meirelles, D.K. Newman, PhdA catalyzes the first step of phenazine-1-carboxylic acid degradation in *Mycobacterium fortuitum*, *J. Bacteriol.* 200 (2018). <https://doi.org/10.1128/JB.00763-17>.
- [11] S.A. Marshall, K.A.P. Payne, D. Leys, The UbiX-UbiD system: The biosynthesis and use of prenylated flavin (prFMN), *Arch. Biochem. Biophys.* 632 (2017) 209–221. <https://doi.org/10.1016/j.abb.2017.07.014>.
- [12] P.H. Wang, A.N. Khusnutdinova, F. Luo, J. Xiao, K. Nemr, R. Flick, G. Brown, R. Mahadevan, E.A. Edwards, A.F. Yakunin, Biosynthesis and Activity of Prenylated FMN Cofactors, *Cell Chem. Biol.* 25 (2018) 560–570.e6.

- <https://doi.org/10.1016/j.chembiol.2018.02.007>.
- [13] N. Arunrattanamook, E.N.G. Marsh, Kinetic Characterization of Prenyl-Flavin Synthase from *Saccharomyces cerevisiae*, *Biochemistry*. 57 (2018) 696–700. <https://doi.org/10.1021/acs.biochem.7b01131>.
- [14] D.L. Stephen A. Marshall, Karl A. P. Payne, Karl Fisher, Mark D. White, Aisling Ní Cheallaigh, Arune Balaikaite, Stephen E. J. Rigby, The UbiX flavin prenyltransferase reaction mechanism resembles Class I terpene cyclase chemistry, *Nat. Commun.* 10 (2019) 6–8. <https://doi.org/10.1038/s41467-019-10220-1>.
- [15] S. Żaczek, J. Kowalska, A. Dybala-Defratyka, Ligand-Driven Conformational Dynamics Influences Selectivity of UbiX, *ChemBioChem*. 19 (2018) 2403–2409. <https://doi.org/10.1002/cbic.201800389>.
- [16] K. Zinovjev, I. Tuñón, Adaptive Finite Temperature String Method in Collective Variables, *J. Phys. Chem. A*. 121 (2017) 9764–9772. <https://doi.org/10.1021/acs.jpca.7b10842>.
- [17] S. Pahari, L. Sun, E. Alexov, PKAD: a database of experimentally measured pKa values of ionizable groups in proteins, *Database*. 2019 (2019). <https://doi.org/10.1093/database/baz024>.
- [18] J.M. Swails, D.M. York, A.E. Roitberg, Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation, *J. Chem. Theory Comput.* 10 (2014) 1341–1352. <https://doi.org/10.1021/ct401042b>.
- [19] A. Lopata, P.G. Jambrina, P.K. Sharma, B.R. Brooks, J. Toth, B.G. Vertessy, E. Rosta, Mutations Decouple Proton Transfer from Phosphate Cleavage in the dUTPase Catalytic Reaction, *ACS Catal.* 5 (2015) 3225–3237. <https://doi.org/10.1021/CS502087F>.
- [20] S. Lu, W. Huang, X. Li, Z. Huang, X. Liu, Y. Chen, T. Shi, J. Zhang, Insights into the Role of Magnesium Triad in myo-Inositol Monophosphatase: Metal Mechanism, Substrate Binding, and Lithium Therapy, *J. Chem. Inf. Model.* 52 (2012) 2398–2409. <https://doi.org/10.1021/CL300172R>.
- [21] R. Gill, F. Mohammed, R. Badyal, L. Coates, P. Erskine, D. Thompson, J. Cooper, M. Gore, S. Wood, IUCr, High-resolution structure of myo-inositol monophosphatase, the putative target of lithium therapy, *Urn:Issn:0907-4449*. 61 (2005) 545–555. <https://doi.org/10.1107/S0907444905004038>.
- [22] G. Schenk, N.Š. Mitić, G.R. Hanson, P. Comba, Purple acid phosphatase: A journey into the function and mechanism of a colorful enzyme, *Coord. Chem. Rev.* 257 (2013) 473–482. <https://doi.org/10.1016/J.CCR.2012.03.020>.
- [23] E.E. Kim, H.W. Wyckoff, Reaction mechanism of alkaline phosphatase based on crystal structures, *J. Mol. Biol.* 218 (1991) 449–464. [https://doi.org/10.1016/0022-2836\(91\)90724-K](https://doi.org/10.1016/0022-2836(91)90724-K).
- [24] Y. Huang, Z. Yue, C.-C. Tsai, J.A. Henderson, J. Shen, Predicting Catalytic Proton Donors and Nucleophiles in Enzymes: How Adding Dynamics Helps Elucidate the Structure–Function Relationships, *J. Phys. Chem. Lett.* 9 (2018) 1179–1184. <https://doi.org/10.1021/acs.jpcllett.8b00238>.
- [25] D. Roston, Q. Cui, Substrate and Transition State Binding in Alkaline Phosphatase Analyzed by Computation of Oxygen Isotope Effects, *J. Am. Chem. Soc.* 138 (2016) 11946–11957. <https://doi.org/10.1021/JACS.6B07347>.
- [26] D. Roston, D. Demapan, Q. Cui, Leaving Group Ability Observably Affects Transition State Structure in a Single Enzyme Active Site, *J. Am. Chem. Soc.* 138 (2016) 7386–7394. <https://doi.org/10.1021/JACS.6B03156>.
- [27] X. Lu, V. Ovchinnikov, D. Demapan, D. Roston, Q. Cui, Regulation and Plasticity of Catalysis in Enzymes: Insights from Analysis of Mechanochemical Coupling in Myosin, *Biochemistry*. 56 (2017) 1482–1497. <https://doi.org/10.1021/ACS.BIOCHEM.7B00016>.
- [28] D. Roston, D. Demapan, Q. Cui, Extensive free-energy simulations identify water as the base in nucleotide addition by DNA polymerase, *Proc. Natl. Acad. Sci.* 116 (2019) 25048–25056. <https://doi.org/10.1073/PNAS.1914613116>.
- [29] R. Recabarren, K. Zinovjev, I. Tuñón, J. Alzate-Morales, How a Second Mg²⁺ Ion Affects the Phosphoryl-Transfer Mechanism in a Protein Kinase: A Computational Study, *ACS Catal.* 11 (2020) 169–183. <https://doi.org/10.1021/ACSCATAL.oC03304>.
- [30] J. Rezáč, Non-Covalent Interactions Atlas Benchmark Data Sets: Hydrogen Bonding, *J. Chem. Theory Comput.* (2020). <https://doi.org/10.1021/ACS.JCTC.9B01265>.
- [31] L. Schrödinger, PyMOL, The PyMOL Molecular Graphics System, (2017).
- [32] B. Webb, A. Sali, Comparative protein structure modeling using MODELLER, *Curr. Protoc. Bioinforma.* 2016 (2016) 5.6.1–5.6.37. <https://doi.org/10.1002/cpbi.3>.
- [33] D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E.I. Cheatham, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R. Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, Y. D.M., P.A. Kollman, Amber 18, (2018).
- [34] M.H.M. Olsson, C.R. Søndergaard, M. Rostkowski, J.H. Jensen, PROPKA₃: Consistent treatment of internal and surface residues in empirical pKa predictions, *J. Chem. Theory Comput.* 7 (2011) 525–537. <https://doi.org/10.1021/ct100578z>.
- [35] C.R. Søndergaard, M.H.M. Olsson, M. Rostkowski, J.H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values, *J. Chem. Theory Comput.* 7 (2011) 2284–2295. <https://doi.org/10.1021/ct200133y>.
- [36] T.J. Dolinsky, J.E. Nielsen, J.A. McCammon, N.A. Baker, PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations, *Nucleic Acids Res.* 32 (2004) W665–W667. <https://doi.org/10.1093/nar/gkh381>.
- [37] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general Amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174. <https://doi.org/10.1002/jcc.20035>.
- [38] A. Jakalian, B.L. Bush, D.B. Jack, C.I. Bayly, Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method, *J. Comput. Chem.* 21 (2000) 132–146. [https://doi.org/10.1002/\(SICI\)1096-987X\(20000130\)21:2<132::AID-JCC5>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2<132::AID-JCC5>3.0.CO;2-P).
- [39] A. Jakalian, D.B. Jack, C.I. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *J. Comput. Chem.* 23 (2002) 1623–1641. <https://doi.org/10.1002/jcc.10128>.
- [40] J. Wang, W. Wang, P.A. Kollman, D.A. Case, Automatic atom type and bond type perception in molecular mechanical calculations, *J. Mol. Graph. Model.* 25 (2006) 247–260. <https://doi.org/10.1016/j.jm gm.2005.12.005>.
- [41] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.* 11 (2015) 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
- [42] R. Salomon-Ferrer, A.W. Götz, D. Poole, S. Le Grand, R.C. Walker, Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald, *J. Chem. Theory Comput.* 9 (2013) 3878–3888. <https://doi.org/10.1021/ct400314y>.
- [43] S. Le Grand, A.W. Götz, R.C. Walker, SPFP: Speed without compromise - A mixed precision model for GPU accelerated molecular dynamics simulations, *Comput. Phys. Commun.* 184 (2013) 374–380. <https://doi.org/10.1016/j.cpc.2012.09.022>.
- [44] J.A. Lzaguirre, D.P. Catarella, J.M. Wozniak, R.D. Skeel, Langevin stabilization of molecular dynamics, *J. Chem. Phys.* 114 (2001) 2090–2098. <https://doi.org/10.1063/1.1332996>.
- [45] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81 (1984) 3684–3690.

- <https://doi.org/10.1063/1.448118>.
- [46] J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* 23 (1977) 327–341. [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5).
- [47] G. Jindal, A. Warshel, Exploring the Dependence of QM/MM Calculations of Enzyme Catalysis on the Size of the QM Region, *J. Phys. Chem. B.* 120 (2016) 9913–9921. <https://doi.org/10.1021/ACS.JPCB.6B07203>.
- [48] R.C. Walker, M.F. Crowley, D.A. Case, The implementation of a fast and accurate QM/MM potential method in Amber, *J. Comput. Chem.* 29 (2008) 1019–1031. <https://doi.org/10.1002/jcc.20857>.
- [49] G.D.M. Seabra, R.C. Walker, M. Elstner, D.A. Case, A.E. Roitberg, Implementation of the SCC-DFTB method for hybrid QM/MM simulations within the Amber molecular dynamics package, in: *J. Phys. Chem. A*, American Chemical Society, 2007: pp. 5655–5664. <https://doi.org/10.1021/jp070071l>.
- [50] M. Gaus, Q. Cui, M. Elstner, DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB), *J. Chem. Theory Comput.* 7 (2011) 931–948. <https://doi.org/10.1021/ct100684s>.
- [51] M. Gaus, X. Lu, M. Elstner, Q. Cui, Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications, *J. Chem. Theory Comput.* 10 (2014) 1518–1537. <https://doi.org/10.1021/ct401002w>.
- [52] D.J. Mermelstein, J.A. McCammon, R.C. Walker, pH-dependent conformational dynamics of beta-secretase 1: A molecular dynamics study, *J. Mol. Recognit.* 32 (2019) e2765. <https://doi.org/10.1002/jmr.2765>.
- [53] A. Onufriev, D. Bashford, D.A. Case, Modification of the Generalized Born Model Suitable for Macromolecules, *J. Phys. Chem. B.* 104 (2002) 3712–3720. <https://doi.org/10.1021/jp994072s>.
- [54] A. Onufriev, D. Bashford, D.A. Case, Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model, *Proteins Struct. Funct. Genet.* 55 (2004) 383–394. <https://doi.org/10.1002/prot.20033>.
- [55] R. Rüger, A. Yakovlev, P. Philippsen, S. Borini, P. Melix, A.F. Oliveira, M. Franchini, T. Soini, M. de Reus, M. Ghorbani Asl, T.Q. Teodoro, D. McCormack, S. Patchkovskii, T. Heine, *AMS DFTB 2019*, (2019).
- [56] Marvin 19.24, (2019).