UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN

"*Biotecnologie e Bioscienze*"

CICLO XXXIV

Identification and characterization of new enzymes involved in bacterial phosphonate catabolism

Coordinatore:

Chiar.mo Prof. Marco Ventura

Tutore:

Chiar.mo Prof. Alessio Peracchi

Dottorando: Erika Zangelmi

Anni Accademici 2018/2019 – 2020/2021

Abstract

Phosphonic acids, or phosphonates, are a class of organophosphorus compounds characterized by a chemically very stable carbon-phosphorus (C-P) bond. In organisms, the stability of these compounds is exploited to cover a wide range of biological functions. Moreover, many bacteria have acquired the ability to degrade and use phosphonates as a source of C, N and P. Based on current knowledge, the catabolism of phosphonates can be subdivided into three groups of enzymatic pathways: C-P hydrolase, C-P oxidase and C-P lyase. With the exception of C-P lyase, which is as flexible as it is complex, most of them appear to be specialized pathways for the degradation of a specific compound. In this work, we explored through bioinformatics analysis the genomic contexts of genes encoding for C-P hydrolase pathways involved in 2-aminoethylphosphonic acid (2-AEP) degradation, in order to identify additional novel enzymes functionally related with these pathways and characterized their activity and role. Specifically, we detected four genes frequently associated with the hydrolytic *phnW-phnX* and *phnW-phnA-phnY* clusters: one automatically annotated as encoding for a PLP-dependent 4-aminobutyrate transaminase (GABA-T) (PF00202) and three, others that, although different from each other, are annotated as hypothetical FAD-dependent oxidoreductases (PF01266); we labeled them as PbfA, PbfB, PbfC and PbfD, respectively. We report here that the PLP-dependent enzyme PbfA from the marine bacterium *Vibrio splendidus* catalyzes an elimination reaction on the natural compound *R*-1-hydroxy-2-aminoethylphosphonate (*R*-OH-AEP), yielding ammonia and phosphonoacetaldehyde, and that the FAD-dependent enzyme PbfC from the rhizobacterium *Azospirillum lipoferum* catalyzes an oxidation reaction on N-monomethyl-2-aminoethylphosphonate (MMAEP), yielding methylamine and phosphonoacetaldehyde. Although these two aminophosphonates are structurally close to 2-AEP, they cannot be processed properly by the PhnW-PhnX (or PhnW-PhnY-PhnA) pathway. However, the scope of these specialized pathways is extended by the accessory enzymes PbfA and PbfC, which channel phosphonoacetaldehyde into the main pipeline. In addition to filling in some missing pieces of the phosphonate metabolism puzzle, our results contribute to the advancement in the field of phosphonate catabolism by emphasizing that such accessory enzymes increase the utility of the common (but very specific) hydrolytic pathways for 2-AEP degradation.

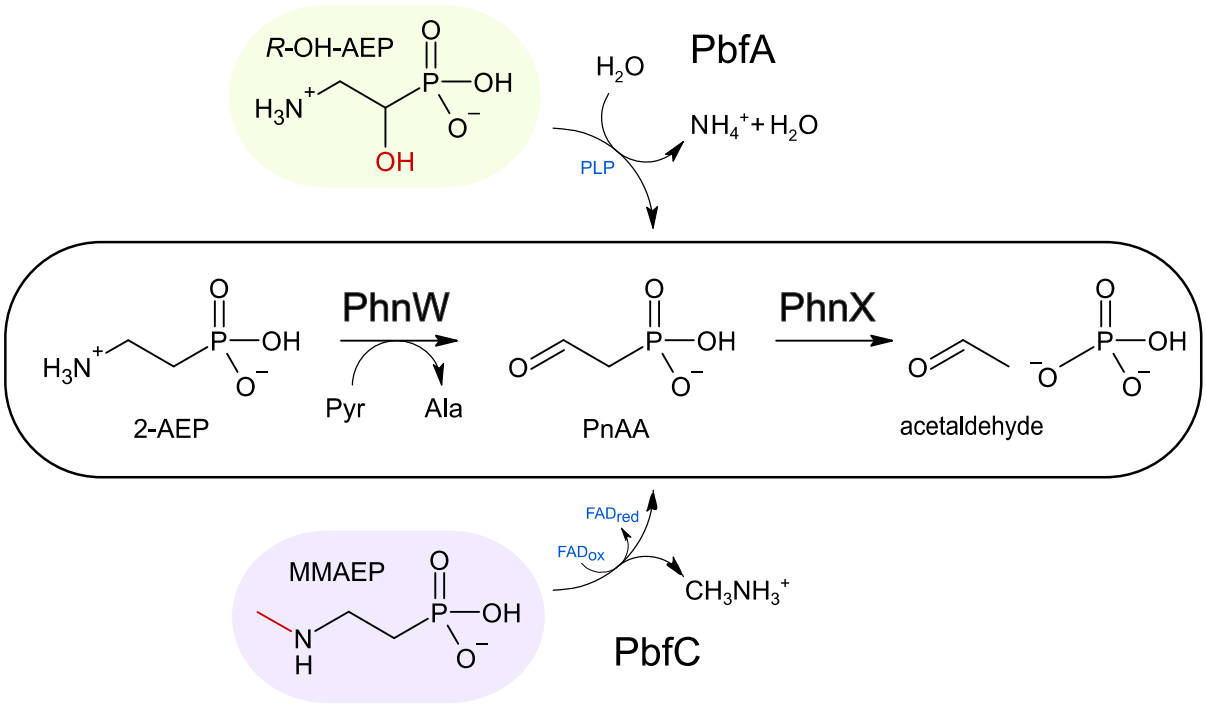

R-OH-AEP
H3N+
P
OH
O−
OH
H2O
PbfA
NH4+ + H2O
PLP
2-AEP
PhnW
Pyr
Ala
PnAA
PhnX
acetaldehyde
MMAEP
N
H
FADred
FADox
CH3NH3+
PbfC

Table of contents

ABBREVIATIONS

Enzymes: **PhnW**, 2-aminoethylphosphonate:pyruvate aminotransferase; **PhnX**, phosphonoacetaldehyde hydrolase; **PhnY**, phosphonoacetaldehyde dehydrogenase; **PhnA**, phosphonoacetate hydrolase; **PalB**, phosphonoalanine aminotransferase; **PalA**, phosphonopyruvate hydrolase; **PhnZ**, *R*-1-hydroxy-2-aminoethylphosphonate dioxygenase; **PhnY***, 2-aminoethylphosphonate dioxygenase; **PepM**, phosphoenolpyruvate mutase; **ADH**, alcohol dehydrogenase; **GDH**, glutamate dehydrogenase; **AlaDH**, alanine dehydrogenase; **BSA**, bovine serum albumin.

Phosphonate compounds: **2-AEP**, ciliatine, 2-aminoethylphosphonate; **PnPyr**, phosphonopyruvate; **PnAA**, phosphonoacetaldehyde; **PnAc**; phosphonoacetate; **PnAla**, phosphonoalanine, 2-amino-3-phosphonopropionate; **OH-AEP**, hydroxyciliatine, 1-hydroxy-2-aminoethylphosphonate; **MMAEP**, N-monomethyl-2-aminoethylphosphonate; **DMAEP**, N-dimethyl-2-aminoethylphosphonate; **TMAEP**, N-trimethyl-2-aminoethylphosphonate; **OH-APP**, 1-hydroxy-2-aminopropylphosphonate; **F-AEP**, fluoro-2-aminoethylphosphonic acid.

General compounds: **PLP**, pyridoxal 5'-phosphate; **Pyr**, pyruvate; **α-KG**, α-ketoglutarate; **Glx**, glyoxylate; **TLC**, thin-layer chromatography; **FAD**, flavin adenine dinucleotide; **DCPIP**, 2,6-dichlorophenolindophenol; **PMS**, phenazine methosulfate.

Chapter 1

1.1 General introduction and overview about phosphonates

Phosphonic acids, or phosphonates, are a class of unusual organophosphorus compounds characterized by the presence of a covalent bond between carbon and phosphorus (P) atoms (C-P bond). The C-P linkage is less reactive than the N-P, S-P or O-P linkages of analogous compounds, conferring to most phosphonates high stability against chemical and enzymatic hydrolysis, thermolysis and photolysis [1]. It was only in 1959 that the first natural phosphonate, 2-aminoethylphosphonate (2-AEP), was identified in an extract of ciliates from sheep's rumen, and hence also called "ciliatine" [2]. Given its structural analogy with phosphoethanolamine, it is not surprising that subsequent investigations confirmed the presence of 2-AEP, usually bound as a polar headgroup of lipids, not only in protozoa but also in a wide variety of bacteria and marine organisms like sea anemones, mollusks, starfish, and to a lesser extent in fungi, insects and mammals; among these, some are able to synthesize 2-AEP, others to assimilate it from diet or gut microbiota [3], [4]. Subsequently, came the discovery of other phosphonic acids that, like 2-AEP, were found incorporated into lipids, glycan and proteins: these included two analogs of the major component of phospholipids (serine and choline), i.e., 2-amino-3-phosphonopropionate (phosphonoalanine; PnAla) and N-methyl derivatives of 2-AEP, respectively, as well as 1-hydroxy-2-aminoethylphosphonate (hydroxyciliatine; OH-AEP) [5], [6].

The physiological role of phosphonate-containing macromolecules is not fully established but based on their occurrence and distribution, they may play an essential role in the organisms that produce them. It was postulated that phosphonates could increase membrane resistance, becoming advantageous during host infection or under harsh conditions (such as in environments rich in hydrolytic enzymes, e.g., phosphatases and phospholipases), contribute to cell signaling, or be phosphorus reservoirs during a phosphate (Pi) starvation response [7], [8]. In addition, many natural phosphonates are free bioactive secondary metabolites. The biological activity of these compounds could be attributed both to the strength of the C-P bond and the chemistry of the phosphonyl group, which can mimic the phosphate and carboxyl groups of normal cellular metabolites, acting as an enzymatic inhibitor [9]. Many of these compounds are antibiotics. These include fosfomycin, produced by both streptomycetes [10] and pseudomonads [11], [12], which inhibits the enzyme involved in the biosynthesis of peptidoglycan by miming its native substrate. Similarly, FR900098 [13] and fosmidomycin interfere with the non-mevalonate pathway of isoprenoid biosynthesis. Some other phosphonates, like the rhizocticins produced by *Bacillus subtilis*, possess an antifungal activity [14], while still others, like the newly-discovered pantaphos produced by *Pantoea ananatis*, have an herbicidal activity [15]. Interestingly, most of them have found therapeutic and agricultural applications and paved the way for the development of new compounds, i.e., drugs [16]–[19], cosmetic and detergent components [20], popular pesticides and herbicides, and efficient catalysts in

chemical research. The most famous example is the synthetic herbicide glyphosate (Roundup), whose phytotoxic activity is due to the inhibition of a shikimate-pathway enzyme, blocking the biosynthesis of aromatic amino acids in plants [21].

Given the current concerns about nature maintenance and global climate, in recent years, phosphonates have gained attention also from an environmental point of view. The growing use of man-made phosphonates and their persistence in the environment increased the interest in their fate and biodegradation [22]. Furthermore, studies on the microbial ability to consume both natural and synthetic phosphonates as a nutrient source revealed their previously unrecognized role in the biogeochemical P cycling (24) and marine "methane paradox" [23], [24], i.e., the phenomenon whereby methane (commonly produced in anaerobic environments) is also formed in an oxygen-rich environment, such as the water surface, as a result of the degradation of certain phosphonates.

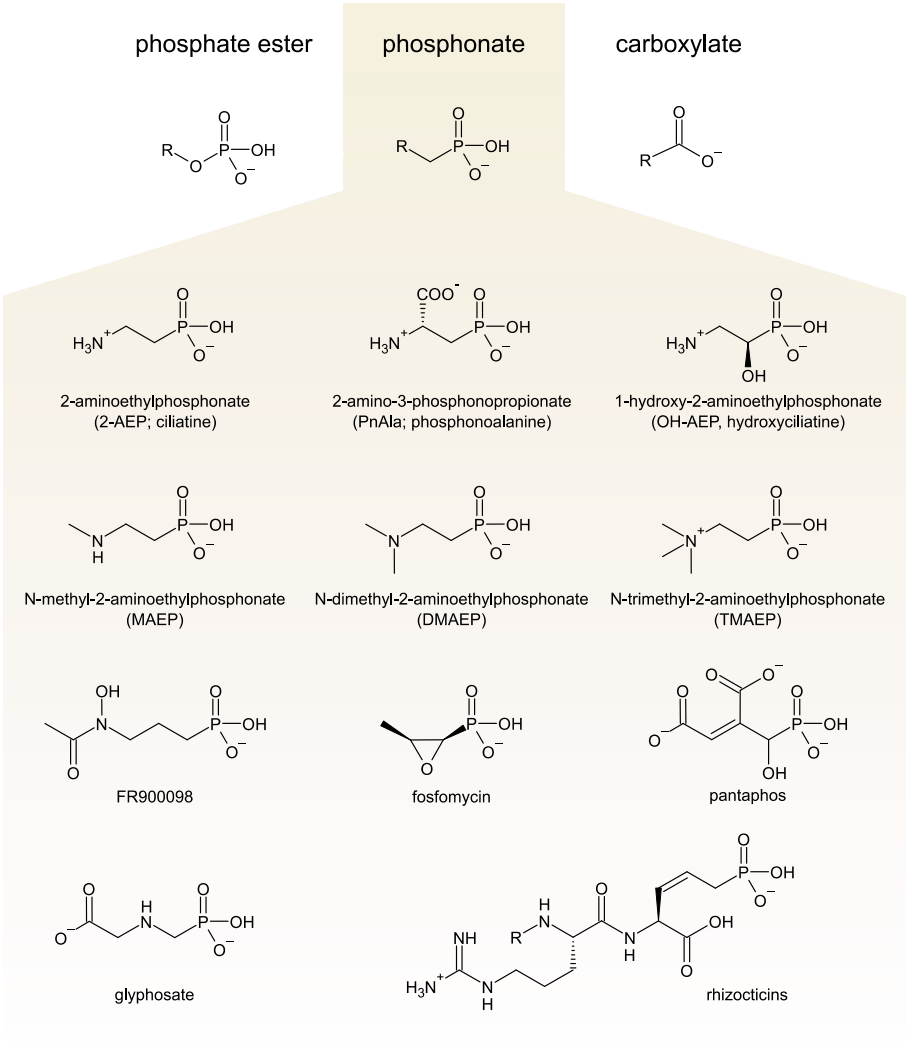


Figure 1. *Phosphonate group mimicry of a phosphate ester and carboxylate and structures of the natural phosphonates mentioned in the text (excluding glyphosate, the only man-made one).*

1.2 Importance of biogenic phosphonates as a phosphorus source

Although phosphorus is one of the most abundant chemical elements on earth, its accessibility by organisms is usually limited in many environments. The microbial mineralization process is one of the main contributors to its bioavailability [25]; indeed, microorganisms evolved different strategies to acquire phosphorus from numerous organic compounds, even phosphonic acids. Although phosphonates occur in a wide variety of soils [26], [27] and some land organisms possess genes for their metabolism [28]–[30], what amount they contribute as a P source in the terrestrial biosphere is not yet established. On the other hand, it was estimated that phosphonates could account for up to 25% of dissolved organic phosphorus (DOP) in the oceans [31]. This abundance is mirrored in the widespread distribution of marine organisms that produce and/or consume phosphonates [32], [33]. Another evidence that phosphonates sustain survival and growth in the marine biosphere is provided by genomic studies, which revealed that in bacterial genomes, genes for phosphonate degradation are more frequent than those for phosphonate biosynthesis, approximately 40% and 10%, respectively [34]. Furthermore, it was observed that the occurrence and variety of phosphonate catabolic genes could change according to both the environmental Pi concentration [35] and the ocean depth [36], maybe as a response to a different composition of phosphonates in the water column. Hence, phosphonate exploitation seems an advantageous strategy in low phosphorus regions where phosphorus availability acts as a dominant selective force [37]. Originally it was believed that the genes involved in the catabolism of phosphonates were only expressed at low Pi level under the Pho regulon control; now we know that many phosphonates, especially aminophosphonates, can also be employed as a source of carbon and nitrogen in a phosphate independent manner, highlighting the importance of their use as a complete nutrient source [38].

1.3 Microbial phosphonate catabolism

Since 1963, when *E. coli* was observed to grow in the presence of phosphonates as the only P-source [39], to the present day, only bacteria and a few fungal species have been described as possessing a "treasure trove of unusual enzymology" [40] to break down the strong C-P bond. Based on current knowledge, phosphonate catabolism can be subdivided into three groups of enzymatic pathways that differ based on the mechanisms by which the C-P bond is ultimately cleaved to release phosphate: these are the radical, hydrolytic and oxidative mechanisms.

1.3.1 The radical mechanism for the C-P bond breakdown

The first mechanism is represented by the C-P lyase pathway, a widespread route by which microorganisms catabolize a great variety of phosphonic acids into the corresponding hydrocarbons and phosphate (Fig. 2) [41]. Its broad specificity could result from the multiplicity of enzymes that constitute the C-P lyase complex. Specifically, the multienzyme complex is produced from an operon composed of 14 genes (*phn*CDEFGHIJKLMNOP) that encode for proteins responsible for both phosphonate transport and utilization:

- PhnC, PhnD and PhnE are ABC phosphonate transporters. Of these, the PhnD gene, due to its distinctive sequence, is often used as a marker to detect the ability of bacteria to catabolize phosphonates [34];
- PhnF is a transcriptional regulator that may provide a more specific control of expression in addition to the Pho regulon, which regulates the transcription of the C-P lyase operon by activating it exclusively during phosphate deficiency;
- PhnG, PhnH, PhnI, PhnJ, PhnK, PhnL and PhnM proteins are essential for the C-P lyase cleavage activity and are thus well conserved among bacteria. In particular, it is the SAM radical enzyme PhnJ that catalyzes the C-P bond cleavage, forming radicals as reactive intermediates upon SAM reduction;
- PhnN, PhnO and PhnP are accessory proteins that are probably involved in bringing other compounds into the pathway and eliminating reaction intermediates.

While in *E. coli* the genes encoding for these proteins are organized in tandem and are transcribed from the same promoter, in other bacterial species they may be arranged differently or have enzymatic variants [5], [42]. Although some steps of the pathway are still unclear since its complexity makes it challenging to study it *in vitro*, we can speculate, to the best of our knowledge, that it plays a significant role in the global P cycle. This is mainly for two reasons: first, because it is widespread in a wide range of organisms, including some very abundant ones such as in the common

cyanobacterium *Trichodesmium erythraeum* [43], and second because it allows the acquisition of phosphate, albeit only during Pi deficiency, from a multitude of compounds.

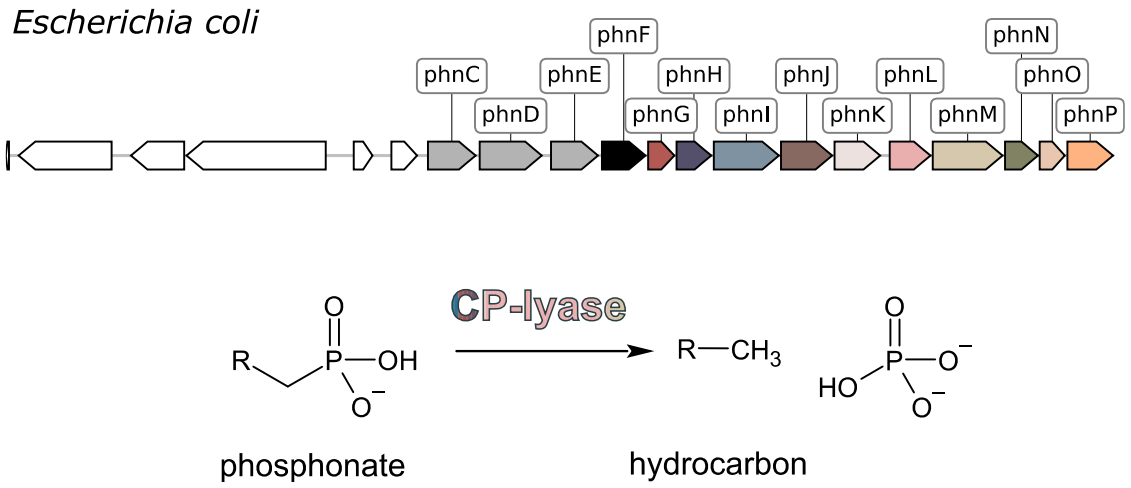


Figure 2. *Visual representation of the gene cluster in E. coli and general reaction scheme of the C-P lyase pathway.*

1.3.2 The hydrolytic mechanism for the C-P bond breakdown

In contrast to the C-P lyase pathway, the enzymes acting through the hydrolytic mechanism are highly specific for substrates with a carbonyl functional group, in beta position relative to the phosphonate, which helps stabilize the carbanion of the leaving group formed after the nucleophilic attack of water on the phosphorus. Accordingly, three enzymes were found to convert phosphonopyruvate (PnPyr), phosphonoacetaldehyde (PnAA), and phosphonoacetate (PnAc) to phosphate; they are the hydrolases PalA, PhnX, and PhnA, respectively. These enzymes are members of three different catabolic pathways, and the expression of their genes can be activated at either low Pi concentration (under the control of the Pho regulon) or in the presence of the coinducer, usually the aminophosphonate substrate from which the carbonyl compound derives.

1.3.2.1 PalA and its role in the phosphonoalanine degradation

The phosphonopyruvate hydrolase PalA (PPH; EC 3.11.1.3) is the second enzyme of the two-step catabolic pathway for phosphonoalanine degradation. In particular, in the first step, the phosphonoalanine aminotransferase PalB transaminates the L-enantiomer of phosphonoalanine (PnAla) forming phosphonopyruvate, which PalA subsequently converts into pyruvate and Pi (Fig. 3 A). Even though PalA activity was isolated and characterized in *Burkholderia cepacia Pal6* [44] and subsequently in *Variovorax sp. Pal2* [45], the complete pathway for phosphonoalanine degradation was clearly described only in the latter [46]. The corresponding gene cluster comprises a gene for the LysR-type transcriptional regulator (*palR*) that induces the transcription, only in the presence of PnAla

or PnPyr [46], of three putative ABC transporter genes (*palE*, *palC* and *palD*) in addition to *palB* and *palA* (Fig. 3 A). This Pi-independent transcriptional regulation allows the bacterium to grow on phosphonoalanine, employing it as a source of carbon and nitrogen as well as phosphorus. In contrast to PalA, the activity of PalB was not confirmed *in vitro* but was assumed based on *Variovorax* gene cluster studies and its sequence similarity to characterized enzymes. PalB is a pyridoxal 5'-phosphate (PLP) -dependent enzyme related to the aspartate aminotransferase superfamily (fold-type I), of which most members catalyze the transfer of ammonia from an amino acid donor to an α-keto acid acceptor. PalA belongs to the isocitrate lyase/phosphoenolpyruvate mutase (ICL/PEPM) superfamily and shares a high percentage of sequence identity (≈40%) with phosphoenolpyruvate mutase (PepM), i.e., the first enzyme in phosphonate biosynthesis that produces phosphonopyruvate from phosphoenolpyruvate. A single substitution of an active site residue (threonine instead of asparagine) can be used to distinguish PalA from PepM (T118 in PalA vs. N122 PepM) [47]. The catalytic parameters of PalA from *Variovorax* (k_{cat} = 105 s^{-1}, K_M = 19 µM and k_{cat}/K_M = 5.5 × 10^6 s^{-1} M^{-1}) and its structure were also determined [47].

1.3.2.2 PhnX and PhnA and their role in the 2-aminoethylphosphonate degradation

The phosphonoacetaldehyde hydrolase PhnX (EC 3.11.1.1) and the phosphonoacetate hydrolase PhnA (EC 3.11.1.2) are both involved in the 2-aminoethylphosphonate (2-AEP; ciliatine) degradation by two separate pathways that, however, share the first enzymatic step, which, as in the degradation of phosphonoalanine, consists of a transamination reaction. The 2-aminoethylphosphonate:pyruvate aminotransferase PhnW (EC 2.6.1.37), a PLP-dependent enzyme belonging to the aspartate aminotransferase superfamily (not closely related to PalB), catalyzes the transamination of 2-AEP by transferring the amine group to pyruvate, yielding alanine and phosphonoacetaldehyde (PnAA) [48]. The PnAA produced is the substrate of either PhnX, which through Mg^{2+}-dependent hydrolytic cleavage of the C-P bond releases acetaldehyde and Pi (PhnW-PhnX pathway), or of the phosphonoacetaldehyde dehydrogenase PhnY (EC 1.2.1.3), which generates phosphonoacetate that in turn is converted to acetate and Pi by PhnA (PhnW-PhnY-PhnA pathway) (Fig. 3B). Since the transamination reaction of PhnW is reversible, this enzyme takes part both in the phosphonate biosynthesis and catabolism. When involved in degradation, its gene can be clustered in bacterial genomes with PhnX (*phnWX* cluster) [49], [50] or PhnY and PhnA (*phnWAY* cluster) [51] (Fig. 3B). The Pho regulon can regulate their expression or, more frequently, they are activated in the presence of the substrate by members of the LysR family of transcriptional regulators [49], again highlighting the advantage of using these pathways not solely for phosphorus acquisition.

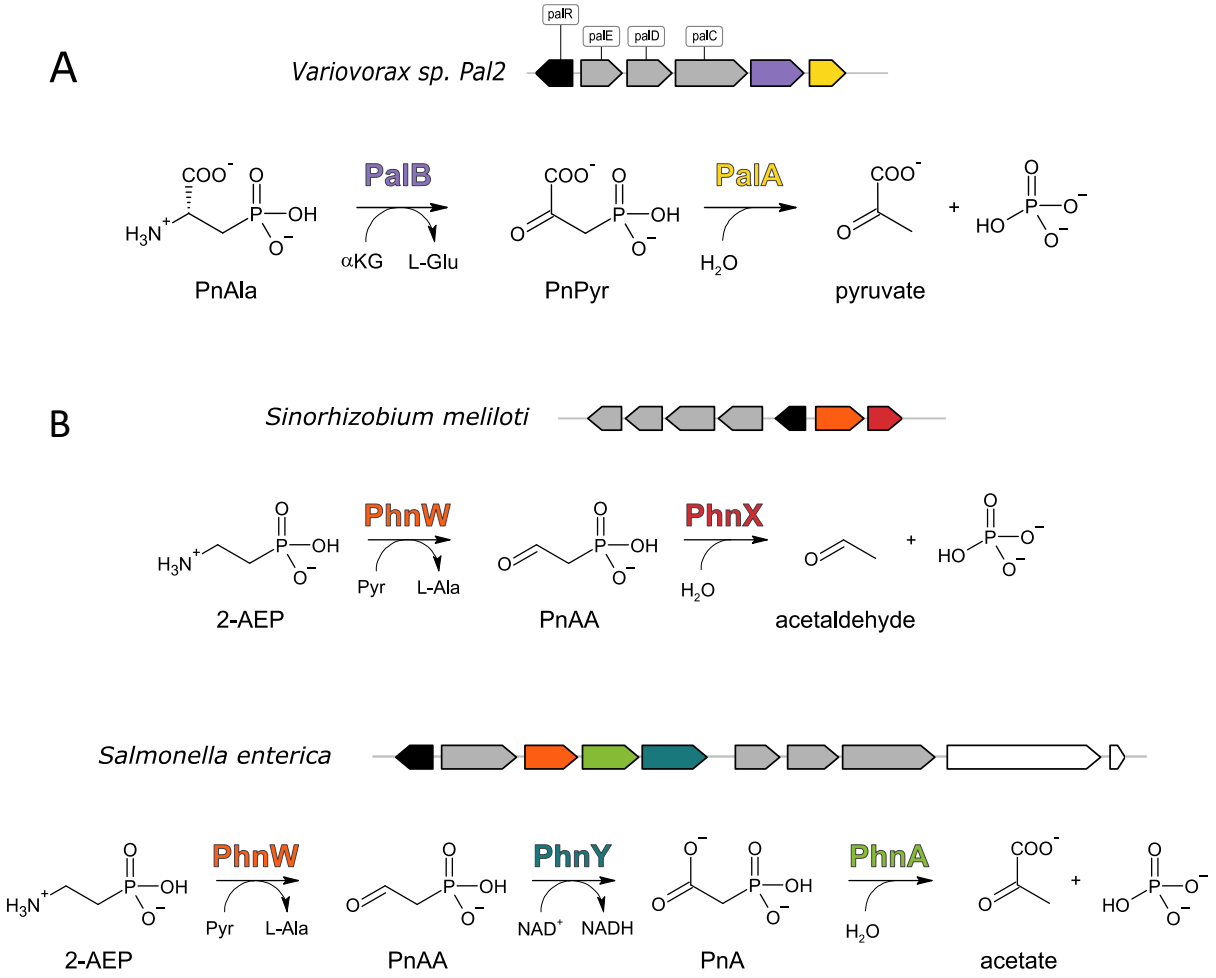


Figure 3. *Genes and enzymes of the hydrolytic phosphonate degradation pathways*. (**A**) Visual representation of phosphonoalanine degradative gene cluster of *Variovorax sp.* Pal2 and reaction scheme of PalA-PalB pathway. (**B**) Visual representation of 2-AEP degradative gene clusters *phnWX* and *phnWAY* and reaction schemes of the PhnW-PhnX and PhnW-PhnY-PhnA pathways.

1.3.3 The oxidative mechanism for the C-P bond breakdown

The third described mechanism of C-P bond breaking is the oxidative mechanism, which occurs through the enzyme 1-hydroxy-2-aminoethylphosphonate dioxygenase PhnZ (EC 1.13.11.78). Although PhnZ belongs to the HD superfamily of phosphohydrolases, it is a monooxygenase that converts *R*-1-hydroxy-2-aminoethylphosphonate (*R*-OH-AEP) to glycine and Pi using Fe(II) and dioxygen [52]. The *phnZ* gene was first isolated clustered with the *phnY** gene through functional screening of a marine metagenomic DNA library that aimed at searching for enzymes that allow bacterial growth on 2-AEP as a P source [33]. The 2-aminoethylphosphonate dioxygenase PhnY* (EC 1.14.11.46) is a non-heme Fe(II)/α-ketoglutarate-dependent dioxygenase that hydroxylates the α-carbon of 2-AEP to form *R*-OH-AEP, which PhnZ subsequently breaks down (PhnY*-PhnZ pathway) (Fig.

4) [53]. Phylogenetic analyses and sequence similarity networks of PhnY* and PhnZ homologs revealed their clusterization into several subclades [54], [55]. Although generally *phnZ* sequences are encoded in genomes in tandem with *phnY**, *phnZ* also occurs in association with other genes for phosphonate catabolism, such as the *phnWX* cluster or that encoding for the C-P lyase complex. Otherwise, *phnY** may also be found in the cluster for phosphonate biosynthesis [54]. Recently, homologs of PhnY* and PhnZ have been characterized as having different substrate specificities:

- PhnY* and PhnZ from *Gimesia maris* DSM8797 do not degrade 2-AEP, but *Gm*PhnY* (EC 1.14.11.71) is specific to hydroxylate the methylphosphonate to hydroxymethylphosphonate, which in turn is oxidized to formic acid and Pi by *Gm*PhnZ (EC 1.13.11.89) [54];
- TmpA (EC 1.14.11.72), a Fe(II)/α-ketoglutarate-dependent oxygenase moderately similar to PhnY*, and TmpB (EC 1.13.11.90), a homolog of PhnZ (sequence identity ≈34%), from *Leisingera caerulea* are specific for the catabolism of N-trimethyl-2-aminoethylphosphonate (TMAEP), a methylated analog of 2-AEP [56]. Compared to the PhnY-PhnZ pathway, TmpA converts stereospecifically TMAEP to *R*-1-hydroxy-N-trimethyl-2-aminoethylphosphonate (*R*-OH-TMAEP), which subsequently TmpB cleaves into glycine betaine and Pi.

These studies also suggest that the substrate specificity of the pathway is primarily determined by PhnY* (or TmpA) rather than by its oxygenase partner (PhnZ or TmpB), which is a more promiscuous enzyme [56].

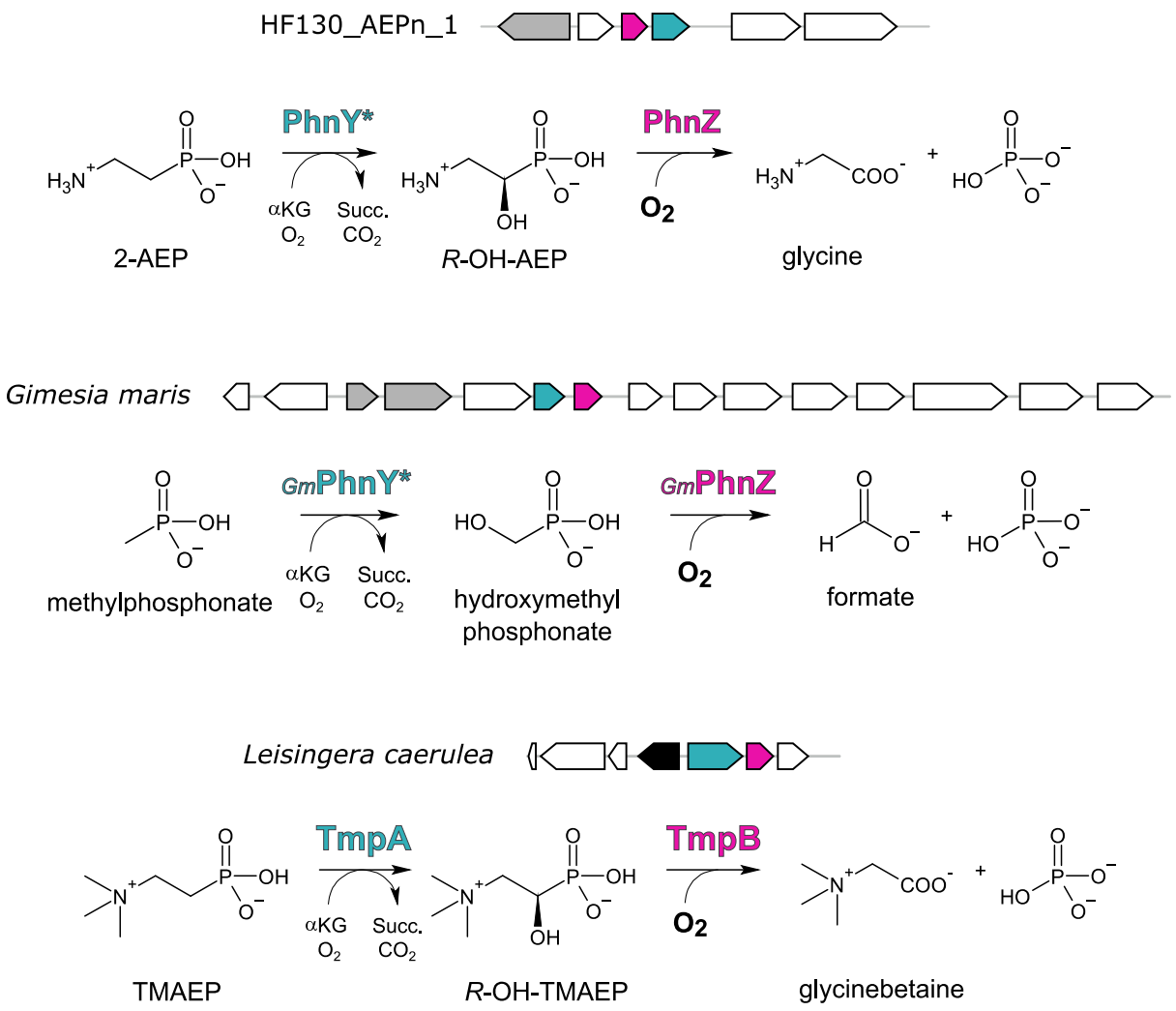


Figure 4. *Visual representation of oxidative gene clusters phnY*Z, GmphnY*Z and tmpAB, and the reaction schemes of the corresponding pathways.*

1.3 The widespread mineralization of 2-aminoethylphosphonate

Based on the currently characterized pathways for phosphonate catabolism, bioinformatic analyses in bacterial genomes have identified a large presence of enzymes involved in the uptake and breakdown of 2-AEP. Although the available microbial genome sequencing data are mostly biased toward pathogens and some microbial groups of human interest, in 2011 Villarreal-Chiu and co-workers showed that among the phosphonate degradation pathways detected in complete genome sequences deposited in the NCBI GenBank database, about 59% are specific pathways for 2-AEP degradation (42% PhnW-PhnX, 16% PhnW-PhnY-PhnA, 1% PhnY*-PhnZ), 0.5% for PnAla degradation (PalA-PalB), and the other 40% are the promiscuous one of the C-P lyase complex [34]. In addition, recently, Murphy and co-workers characterized three Pi-insensitive transporters (AepXVW, AepP and AepSTU)

specific for 2-AEP that are widespread and highly expressed throughout the global ocean, reinforcing the observation that 2-AEP mineralization is dominant in this habitat [57]. However, there are many cases of organisms catabolizing phosphonates in a still unknown way, suggesting that novel pathways await to be characterized [58]. Furthermore, as a result of evolution and lateral gene transfers [59], [60], additional uncharacterized enzymes associated with known pathways or genetic variants of them are often found. These can be expected to provide variations in substrate specificity or pathway functionality, enabling the use of other naturally occurring phosphonates whose catabolism may be more relevant in the environment than appreciated today [34].

1.5 Purpose of the research

Since the initial discovery of phosphonates the interest in their biosynthesis and degradation has increased primarily for two reasons: because of their versatile stability and mimicry, and because of the unusual repertoire of enzymes that utilize them. It is now well established that phosphonates are widespread molecules, especially in bacteria. However, estimating with certainty the extent of their physiological and ecological importance is still tricky as there is not yet a complete understanding of their metabolism. Recent developments in analytical techniques have made it easier to isolate and identify phosphonates, but it remains often challenging to study or characterize the enzymes involved in their metabolism due to the lack of potential phosphonate substrates, especially in the pure enantiomeric form, which implies the advance of dedicated synthetic methodologies [61]. The gap in the knowledge of phosphonate catabolism has aroused our interest in investigating this field in search of unexplored enzymatic reactions for two main reasons: (i) to contribute to figuring out how and which phosphonates are degraded by organisms; (ii) to expand the scope of known enzymatic chemistry with the prospect of paving the way for future applications in biocatalysis and biotechnology. Because phosphonate catabolism genes in bacteria are often organized in clusters, we analyzed these genomic contexts looking for promising hypothetical enzymes. In particular, the work in this dissertation focused on inspecting the clusters encoding pathways that employ the hydrolytic mechanisms, as they are the most abundant in nature and relatively simple to study *in vitro*. Our findings confirmed the observations of other authors who reported a frequent presence of accessory enzymes of unknown function associated with these operons. By applying a combination of bioinformatics and biochemical approaches, we attempted to hypothesize the reaction mechanism and the potential substrate of some of these enzymes and then tested our predictions *in vitro*.

1.6 References

[1] N. G.Ternan, J. W. Mc Grath, G. Mc Mullan, and J. P. Quinn, "Review: Organophosphonates: occurrence, synthesis and biodegradation by microorganisms.," *World Journal of Microbiology and Biotechnology*, vol. 14, no. 5, pp. 635–647, 1998, doi: 10.1023/A:1008848401799.

[2] M. HORIGUCHI and M. KANDATSTU, "Isolation of 2-Aminoethane Phosphonic Acid from Rumen Protozoa," *Nature*, vol. 184, no. 4690, pp. 901–902, Sep. 1959, doi: 10.1038/184901b0.

[3] E. Y. Kostetsky, P. v. Velansky, and N. M. Sanina, "Phospholipids of the Organs and Tissues of Echinoderms and Tunicates from Peter the Great Bay (Sea of Japan)," *Russian Journal of Marine Biology*, vol. 38, no. 1, pp. 64–71, Jan. 2012, doi: 10.1134/S1063074012010099.

[4] J. S. Kittredge and E. Roberts, "A Carbon-Phosphorus Bond in Nature," *Science*, vol. 164, no. 3875, pp. 37–42, Apr. 1969, doi: 10.1126/science.164.3875.37.

[5] G. P. Horsman and D. L. Zechel, "Phosphonate Biochemistry," *Chemical Reviews*, vol. 117, no. 8, pp. 5704–5783, Apr. 2017, doi: 10.1021/acs.chemrev.6b00536.

[6] Kh. S. Mukhamedova and A. I. Glushenkova, "Natural Phosphonolipids," *Chemistry of Natural Compounds*, vol. 36, no. 4, pp. 329–341, 2000, doi: 10.1023/A:1002804409503.

[7] W. W. Metcalf and W. A. van der Donk, "Biosynthesis of phosphonic and phosphinic acid natural products," *Annual Review of Biochemistry*, vol. 78. pp. 65–94, 2009. doi: 10.1146/annurev.biochem.78.091707.100215.

[8] A. Kim, J. Kim, B. M. Martin, and D. Dunaway-Mariano, "Isolation and Characterization of the Carbon–Phosphorus Bond-forming Enzyme Phosphoenolpyruvate Mutase from the MolluskMytilus edulis," *Journal of Biological Chemistry*, vol. 273, no. 8, pp. 4443–4448, Feb. 1998, doi: 10.1074/jbc.273.8.4443.

[9] L. Azema, R. Baron, and S. Ladame, "Targeting Enzymes with Phosphonate-Based Inhibitors: Mimics of Tetrahedral Transition States and Stable Isosteric Analogues of Phosphates," *Current Enzyme Inhibition*, vol. 2, no. 1, pp. 61–72, Feb. 2006, doi: 10.2174/157340806775473526.

[10] D. Hendlin *et al.*, "Phosphonomycin, a New Antibiotic Produced by Strains of Streptomyces," *Science*, vol. 166, no. 3901, pp. 122–123, Oct. 1969, doi: 10.1126/science.166.3901.122.

[11] M. A. Simon, C. Ongpipattanakul, S. K. Nair, and W. A. van der Donk, "Biosynthesis of fosfomycin in pseudomonads reveals an unexpected enzymatic activity in the metallohydrolase superfamily," *Proceedings of the National Academy of Sciences*, vol. 118, no. 23, p. e2019863118, Jun. 2021, doi: 10.1073/pnas.2019863118.

[12] J. SHOJI *et al.*, "Production of fosfomycin (phosphonomycin) by Pseudomonas syringae.," *The Journal of Antibiotics*, vol. 39, no. 7, pp. 1011–1012, 1986, doi: 10.7164/antibiotics.39.1011.

[13] M. OKUHARA *et al.*, "Studies on new phosphonic acid antibiotics. I. FR-900098, isolation and characterization.," *The Journal of Antibiotics*, vol. 33, no. 1, pp. 13–17, 1980, doi: 10.7164/antibiotics.33.13.

[14] S. A. Borisova, B. T. Circello, J. K. Zhang, W. A. van der Donk, and W. W. Metcalf, “Biosynthesis of Rhizocticins, Antifungal Phosphonate Oligopeptides Produced by Bacillus subtilis ATCC6633,” *Chemistry and Biology*, vol. 17, no. 1, pp. 28–37, Jan. 2010, doi: 10.1016/j.chembiol.2009.11.017.

[15] A. L. A. Polidore, L. Furiassi, P. J. Hergenrother, and W. W. Metcalf, “A phosphonate natural product made by pantoea ananatis is necessary and sufficient for the hallmark lesions of onion center rot,” *mBio*, vol. 12, no. 1, pp. 1–18, Jan. 2021, doi: 10.1128/mBio.03402-20.

[16] S. J. Hecker and M. D. Erion, “Prodrugs of phosphates and phosphonates,” *Journal of Medicinal Chemistry*, vol. 51, no. 8. pp. 2328–2345, Apr. 24, 2008. doi: 10.1021/jm701260b.

[17] R. G. G. Russell, “Bisphosphonates: The first 40 years,” *Bone*, vol. 49, no. 1. pp. 2–19, Jul. 2011. doi: 10.1016/j.bone.2011.04.022.

[18] T. Umeda, N. Tanaka, Y. Kusakabe, M. Nakanishi, Y. Kitade, and K. T. Nakamura, “Molecular basis of fosmidomycin’s action on the human malaria parasite Plasmodium falciparum,” *Scientific Reports*, vol. 1, 2011, doi: 10.1038/srep00009.

[19] S. S. Patel, J. A. Balfour, and H. M. Bryson, “Fosfomycin Tromethamine,” *Drugs*, vol. 53, no. 4, pp. 637–656, Apr. 1997, doi: 10.2165/00003495-199753040-00007.

[20] W. E. Gledhill and T. C. J. Feijtel, “Environmental Properties and Safety Assessment of Organic Phosphonates Used for Detergent and Water Treatment Applications,” in *Detergents*, Berlin, Heidelberg: Springer Berlin Heidelberg, 1992, pp. 261–285. doi: 10.1007/978-3-540-47108-0_8.

[21] H. C. Steinrücken and N. Amrhein, “The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimic acid-3-phosphate synthase,” *Biochemical and Biophysical Research Communications*, vol. 94, no. 4, pp. 1207–1212, Jun. 1980, doi: 10.1016/0006-291X(80)90547-1.

[22] B. Nowack, “Environmental chemistry of phosphonates,” *Water Research*, vol. 37, no. 11. Elsevier Ltd, pp. 2533–2546, 2003. doi: 10.1016/S0043-1354(03)00079-4.

[23] D. J. Repeta *et al.*, “Marine methane paradox explained by bacterial degradation of dissolved organic matter,” *Nature Geoscience*, vol. 9, no. 12, pp. 884–887, Dec. 2016, doi: 10.1038/ngeo2837.

[24] A. Paytan and K. McLaughlin, “The Oceanic Phosphorus Cycle,” *Chemical Reviews*, vol. 107, no. 2, pp. 563–576, Feb. 2007, doi: 10.1021/cr0503613.

[25] K. R. M. Mackey and A. Paytan, “Phosphorus Cycle,” in *Encyclopedia of Microbiology*, Elsevier, 2009, pp. 322–334. doi: 10.1016/B978-012373944-5.00056-0.

[26] B. L. Turner, R. Baxter, N. Mahieu, S. Sjögersten, and B. A. Whitton, “Phosphorus compounds in subarctic Fennoscandian soils at the mountain birch (Betula pubescens)—tundra ecotone,” *Soil Biology and Biochemistry*, vol. 36, no. 5, pp. 815–823, May 2004, doi: 10.1016/j.soilbio.2004.01.011.

[27] B. L. Turner, J. A. Chudek, B. A. Whitton, and R. Baxter, “Phosphorus composition of upland soils polluted by long-term atmospheric nitrogen deposition,” *Biogeochemistry*, vol. 65, no. 2, pp. 259–274, 2003, doi: 10.1023/A:1026065719423.

[28] I. D. E. A. Lidbury *et al.*, "Comparative genomic, proteomic and exoproteomic analyses of three Pseudomonas strains reveals novel insights into the phosphorus scavenging capabilities of soil bacteria," *Environmental Microbiology*, vol. 18, no. 10, pp. 3535–3549, Oct. 2016, doi: 10.1111/1462-2920.13390.

[29] Y. Tapia-Torres *et al.*, "How To Live with Phosphorus Scarcity in Soil and Sediment: Lessons from Bacteria," *Applied and Environmental Microbiology*, vol. 82, no. 15, pp. 4652–4662, Aug. 2016, doi: 10.1128/AEM.00160-16.

[30] Y. Park, M. Solhtalab, W. Thongsomboon, and L. Aristilde, "Strategies of organic phosphorus recycling by soil bacteria: acquisition, metabolism, and regulation," *Environmental Microbiology Reports*, Jan. 2022, doi: 10.1111/1758-2229.13040.

[31] L. L. Clark, E. D. Ingall, and R. Benner, "Marine phosphorus is selectively remineralized," *Nature*, vol. 393, no. 6684, pp. 426–426, Jun. 1998, doi: 10.1038/30881.

[32] S. T. Dyhrman, C. R. Benitez-Nelson, E. D. Orchard, S. T. Haley, and P. J. Pellechia, "A microbial source of phosphonates in oligotrophic marine systems," *Nature Geoscience*, vol. 2, no. 10, pp. 696–699, Oct. 2009, doi: 10.1038/ngeo639.

[33] A. Martinez, G. W. Tyson, and E. F. DeLong, "Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses," *Environmental Microbiology*, vol. 12, no. 1, pp. 222–238, Jan. 2010, doi: 10.1111/j.1462-2920.2009.02062.x.

[34] J. F. Villarreal-Chiu, "The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment," *Frontiers in Microbiology*, vol. 3, 2012, doi: 10.3389/fmicb.2012.00019.

[35] O. A. Sosa, D. J. Repeta, E. F. DeLong, M. D. Ashkezari, and D. M. Karl, "Phosphate-limited ocean regions select for bacterial populations enriched in the carbon–phosphorus lyase pathway for phosphonate degradation," *Environmental Microbiology*, vol. 21, no. 7, pp. 2402–2414, Jul. 2019, doi: 10.1111/1462-2920.14628.

[36] H. Luo, H. Zhang, R. Long, and R. Benner, "Depth distributions of alkaline phosphatase and phosphonate utilization genes in the North Pacific Subtropical Gyre," *Aquatic Microbial Ecology*, vol. 62, no. 1, pp. 61–69, Jan. 2011, doi: 10.3354/ame01458.

[37] M. L. Coleman and S. W. Chisholm, "Ecosystem-specific selection pressures revealed through comparative population genomics," *Proceedings of the National Academy of Sciences*, vol. 107, no. 43, pp. 18634–18639, Oct. 2010, doi: 10.1073/pnas.1009480107.

[38] J. P. Chin, J. P. Quinn, and J. W. McGrath, "Phosphate insensitive aminophosphonate mineralisation within oceanic nutrient cycles," *The ISME Journal*, vol. 12, no. 4, pp. 973–980, Apr. 2018, doi: 10.1038/s41396-017-0031-7.

[39] L. D. Zeleznick, T. C. Myers, and E. B. Titchener, "Growth of Escherichia coli on methyl- and ethylphosphonic acids," *Biochimica et Biophysica Acta*, vol. 78, no. 3, pp. 546–547, Nov. 1963, doi: 10.1016/0006-3002(63)90921-1.

[40] S. C. Peck and W. A. van der Donk, "Phosphonate biosynthesis and catabolism: a treasure trove of unusual enzymology," *Current Opinion in Chemical Biology*, vol. 17, no. 4, pp. 580–588, Aug. 2013, doi: 10.1016/j.cbpa.2013.06.018.

[41] N. Stosiek, M. Talma, and M. Klimek-Ochab, “Carbon-Phosphorus Lyase—the State of the Art,” *Applied Biochemistry and Biotechnology*, vol. 190, no. 4, pp. 1525–1552, Apr. 2020, doi: 10.1007/s12010-019-03161-4.

[42] S. Y. Yunusov, K. S. Mukhamedova, and A. I. Glushenkova, “Natural phosphonolipids,” *Chemistry of Natural Compounds*, vol. 36, no. 4, Feb. 2000, doi: 10.1023/a:1002804409503

[43] S. T. Dyhrman *et al.*, “Phosphonate utilization by the globally important marine diazotroph Trichodesmium,” *Nature*, vol. 439, no. 7072, pp. 68–71, Jan. 2006, doi: 10.1038/nature04203.

[44] N. G. Ternan, J. T. G. Hamilton, and J. P. Quinn, “Initial in vitro characterisation of phosphonopyruvate hydrolase, a novel phosphate starvation-independent, carbon-phosphorus bond cleavage enzyme in Burkholderia cepacia Pal6,” *Archives of Microbiology*, vol. 173, no. 1, pp. 35–41, Jan. 2000, doi: 10.1007/s002030050005.

[45] A. N. Kulakova, G. B. Wisdom, L. A. Kulakov, and J. P. Quinn, “The Purification and Characterization of Phosphonopyruvate Hydrolase, a Novel Carbon-Phosphorus Bond Cleavage Enzyme from Variovorax sp. Pal2,” *Journal of Biological Chemistry*, vol. 278, no. 26, pp. 23426–23431, Jun. 2003, doi: 10.1074/jbc.M301871200.

[46] A. N. Kulakova, L. A. Kulakov, J. F. Villarreal-Chiu, J. A. Gilbert, J. W. McGrath, and J. P. Quinn, “Expression of the phosphonoalanine-degradative gene cluster from Variovorax sp. Pal2 is induced by growth on phosphonoalanine and phosphonopyruvate,” *FEMS Microbiology Letters*, vol. 292, no. 1, pp. 100–106, Mar. 2009, doi: 10.1111/j.1574-6968.2008.01477.x.

[47] C. C. H. Chen *et al.*, “Structure and Kinetics of Phosphonopyruvate Hydrolase from Voriovorax sp. Pal2: New Insight into the Divergence of Catalysis within the PEP Mutase/Isocitrate Lyase Superfamily,” *Biochemistry*, vol. 45, no. 38, pp. 11491–11504, Sep. 2006, doi: 10.1021/bi061208l.

[48] A. D. Kim, A. S. Baker, D. Dunaway-Mariano, W. W. Metcalf, B. L. Wanner, and B. M. Martin, “The 2-Aminoethylphosphonate-Specific Transaminase of the 2-Aminoethylphosphonate Degradation Pathway,” *Journal of Bacteriology*, vol. 184, no. 15, pp. 4134–4140, Aug. 2002, doi: 10.1128/JB.184.15.4134-4140.2002.

[49] J. P. Quinn, A. N. Kulakova, N. A. Cooley, and J. W. McGrath, “New ways to break an old bond: the bacterial carbon?phosphorus hydrolases and their role in biogeochemical phosphorus cycling,” *Environmental Microbiology*, vol. 9, no. 10, pp. 2392–2400, Oct. 2007, doi: 10.1111/j.1462-2920.2007.01397.x.

[50] J. C. Errey and J. S. Blanchard, “Functional Annotation and Kinetic Characterization of PhnO from Salmonella enterica,” *Biochemistry*, vol. 45, no. 9, pp. 3033–3039, Mar. 2006, doi: 10.1021/bi052297p.

[51] S. A. Borisova *et al.*, “Genetic and Biochemical Characterization of a Pathway for the Degradation of 2-Aminoethylphosphonate in Sinorhizobium meliloti 1021,” *Journal of Biological Chemistry*, vol. 286, no. 25, pp. 22283–22290, Jun. 2011, doi: 10.1074/jbc.M111.237735.

[52] L. M. van Staalduinen *et al.*, “Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 14, pp. 5171–5176, Apr. 2014, doi: 10.1073/pnas.1320039111.

[53] F. R. McSorley, P. B. Wyatt, A. Martinez, E. F. DeLong, B. Hove-Jensen, and D. L. Zechel, "PhnY and PhnZ Comprise a New Oxidative Pathway for Enzymatic Cleavage of a Carbon–Phosphorus Bond," *Journal of the American Chemical Society*, vol. 134, no. 20, pp. 8364–8367, May 2012, doi: 10.1021/ja302072f.

[54] S. R. Gama *et al.*, "An Oxidative Pathway for Microbial Utilization of Methylphosphonic Acid as a Phosphate Source," *ACS Chemical Biology*, vol. 14, no. 4, pp. 735–741, Apr. 2019, doi: 10.1021/acschembio.9b00024.

[55] B. Worsdorfer *et al.*, "Organophosphonate-degrading PhnZ reveals an emerging family of HD domain mixed-valent diiron oxygenases," *Proceedings of the National Academy of Sciences*, vol. 110, no. 47, pp. 18874–18879, Nov. 2013, doi: 10.1073/pnas.1315927110.

[56] L. J. Rajakovich *et al.*, "A New Microbial Pathway for Organophosphonate Degradation Catalyzed by Two Previously Misannotated Non-Heme-Iron Oxygenases," *Biochemistry*, vol. 58, no. 12, pp. 1627–1647, Mar. 2019, doi: 10.1021/acs.biochem.9b00044.

[57] A. R. J. Murphy *et al.*, "Transporter characterisation reveals aminoethylphosphonate mineralisation as a key step in the marine phosphorus redox cycle," *Nature Communications*, vol. 12, no. 1, p. 4554, Dec. 2021, doi: 10.1038/s41467-021-24646-z.

[58] J. W. McGrath, J. P. Chin, and J. P. Quinn, "Organophosphonates revealed: new insights into the microbial metabolism of ancient molecules," *Nature Reviews Microbiology*, vol. 11, no. 6, pp. 412–419, Jun. 2013, doi: 10.1038/nrmicro3011.

[59] M. N. Price, A. P. Arkin, and E. J. Alm, "The Life-Cycle of Operons," *PLoS Genetics*, vol. 2, no. 7, p. e126, 2006, doi: 10.1371/journal.pgen.0020126.

[60] J. Huang, Z. Su, and Y. Xu, "The Evolution of Microbial Phosphonate Degradative Pathways," *Journal of Molecular Evolution*, vol. 61, no. 5, pp. 682–690, Nov. 2005, doi: 10.1007/s00239-004-0349-4.

[61] K. Pallitsch, T. Kalina, and T. Stanković, "Synthetic Phosphonic Acids as Potent Tools to Study Phosphonate Enzymology," *Synlett*, vol. 30, no. 07, pp. 770–776, Apr. 2019, doi: 10.1055/s-0037-1611460.

Chapter 2

2.1 Challenges in the discovery of new enzymatic functions

The recent increase in genomic data, obtained from next-generation sequencing (NGS) technologies and metagenomic libraries, provides the scientific community with a huge amount of records potentially exploitable for biochemical discovery. Unfortunately, determining reliable functions for unknown proteins from these genome projects is a major challenge in today's biology. The low similarity of these uncharacterized proteins to the known ones makes it difficult to deduce their activity only through homology-based function prediction, so their automatic annotation in databases is often erroneous or misleading [1]. Several strategies have been devised in response to these issues, e.g., the Enzyme Function Initiative (EFI) addressed the establishment of a multidisciplinary workflow for an efficient enzyme function assignment [2]. In prokaryotes, a combination of genomic context and homology searches can be a valuable tool to generate hypotheses about the enzyme's reaction mechanism or its substrate/product structure, facilitating and speeding up experimental validation [3]. In bacteria, the genes encoding the enzymes of a given metabolic pathway are frequently organized into an operon or cluster, allowing for a relationship of co-regulation or connected biochemical functions; for example, in many cases, the product of one enzyme is the substrate of another one. Association is more likely to reflect a functional constraint if the gene cluster composition is maintained across multiple genomes. However, less frequent but conserved variations are also of interest because they may indicate an adaptive response to the requirements of particular environments [4]. In the current research, genomic context exploitation is most commonly used to study biosynthetic gene clusters (BGCs), where the compound generated can be tentatively predicted based on the analysis and logic of enzymes within the cluster [5]; nevertheless, it can also be applied to catabolic operons [3].

2.1.1 The genomic context for exploring phosphonate metabolism

The identification of the first phosphonate in 1959 has been followed by the discovery of a large and diverse repertoire of phosphonic acids and of enzymes involved in their metabolism. The traditional method to study phosphonate metabolism entails, first, the isolation of microorganisms capable of producing or degrading these compounds, and only afterward the characterization of the enzymes that confer such an ability. Efforts to complete the global metabolism of phosphonates have benefited from the advent of the genomic era. The accessibility to genomic information and the organization in clusters of previously described enzymes have made it possible to understand how widespread the use of phosphonates is in nature, while identifying new enzymatic activities and metabolites. This approach has led to significant results, especially in investigating gene clusters for the synthesis of phosphonates, with a greater focus on those for the synthesis of secondary bioactive metabolites. The

methodology required for the gene-based discovery of these clusters is well established and consists of using the phosphoenolpyruvate mutase (PepM) coding gene as a probe to establish phosphonate producers organisms. Indeed, the biosynthetic pathways of all but two known phosphonates share the initial step, i.e., the PepM-catalyzed conversion of phosphoenolpyruvate to phosphonopyruvate (Fig. 5). Searching for *pepM* in bacterial genomes reveals that 5% of them possess this gene; furthermore, the genome contexts where it occurs are highly diverse, reflecting the diversification of phosphonate products. These gene neighborhoods can be grouped by similarity and include candidates for phosphonolipids and phosphonoglycans synthesis, and others for bioactive phosphonates. This kind of research made it possible to identify numerous novel phosphonate compounds, the enzymes involved in their production, and to estimate their abundance [6]–[11].

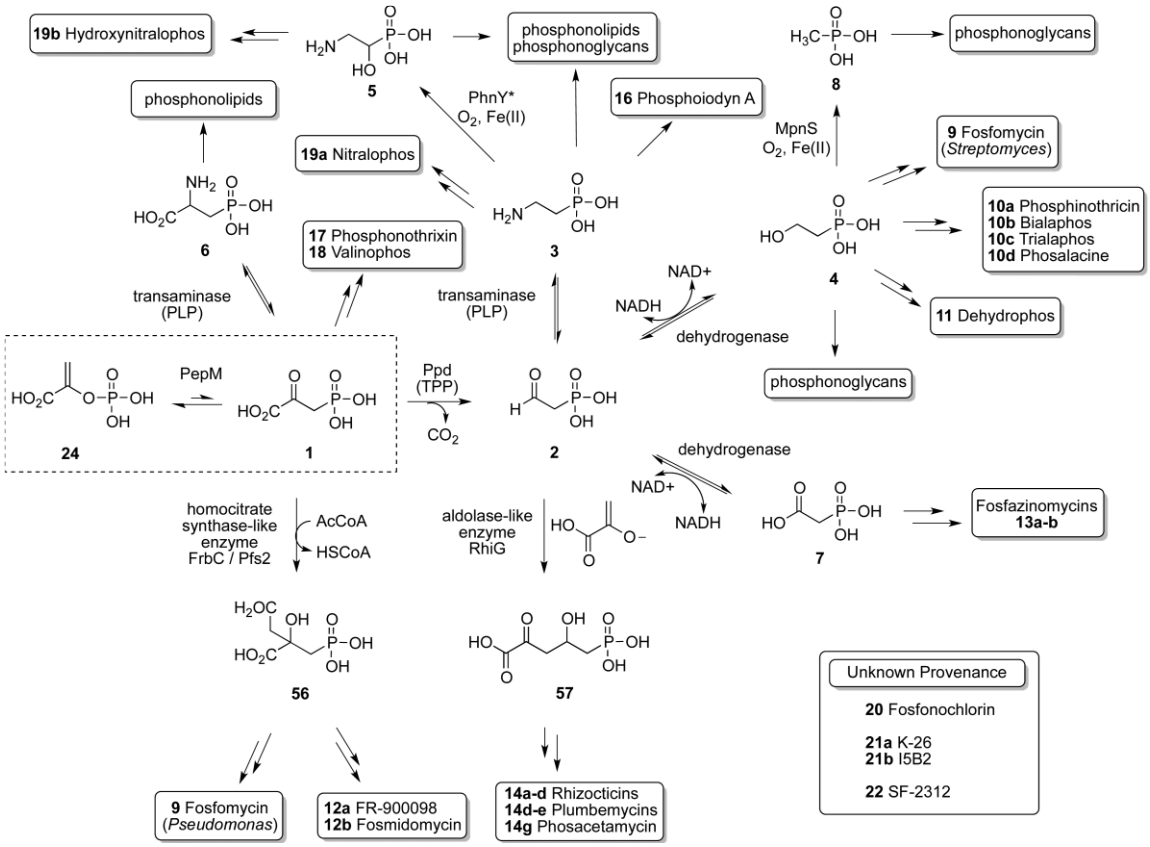


Figure 5. *Overview of phosphonate biosynthesis.* The first reaction common to most biosynthetic pathways, i.e., the conversion of phosphoenolpyruvate to phosphonopyruvate by PepM, is highlighted in the dashed box. Scheme from G. P. Horsman and D. L Zechel, Chem. Rev. 2017 [12].

Unlike the biosynthetic routes, which all share the PepM reaction and then diversify depending on the generated compound, those involved in degradation, based on the best of our knowledge, can be grouped according to the mechanism for the C-P bond cleavage (see Section 1.3 of Chapter 1). Consequently, the phosphonate transporter genes (*phnD*, *aepX* and *aepP)* and those for the enzymes

that catalyze the C-P bond breaking (*phnJ* of C-P lyase, *phnX, phnA* and *palA* of the hydrolytic mechanism and *phnZ* of the oxidative one) can be used as molecular markers for their recognition. The results of several studies reported that genes for phosphonate degradation are more frequent than those for synthesis and are distributed in different organisms depending on the resources offered by the environment in which they live [13]–[16]. There is also high variability in the composition of well-established clusters, i.e., there are often additional unknown genes in their gene neighborhood, or some of the core genes are missing, replaced by new ones, or in unusual combinations (e.g., *phnZ* can be found together with *phnWX* without its *phnY** partner). These observations emphasize that we are far from having a complete view of phosphonate catabolism, motivating further investigation in this area. Although identifying new catabolic enzymes is relatively straightforward, hypothesizing their mechanism of action and substrate specificity is a bit tricky, especially given their poor similarity to known proteins and our incomplete understanding of what phosphonates can be subject to catabolism. In contrast to what has been done so far, we looked for new hypothetical enzymes in the hydrolytic clusters (*phnWX*, *phnWAY* and *palBA*), and based on the gene contexts in which they recur and their homology to know enzymes, we tried to predict their potential substrate and reaction mechanism assuming that the product is functional for the pathway to which it seems associated (Fig. 6). Evidently, this approach is not practical when the new enzyme is linked to the main ones just for a co-regulation benefit.

Figure 6. *The rationale of the computational procedure adopted.* (**1**) As a first step, we searched for an unknown gene associated with one of the hydrolytic clusters. (**2**) Second, we analyzed the frequency of this association and its distribution in bacterial genomes. (**3**) Third, we compared the sequence of the hypothetical enzyme with those of characterized ones to predict its reaction type. (**4**) Fourth, we envisage its potential substrate by assuming that it is naturally occurring and converted to a useful intermediate for the catabolic pathway to which it belongs. (**5**) Finally, we moved forward with the experimental test of our hypotheses.

2.2 Methods

2.2.1 Gene context analysis of *phnWX*, *phnWAY* and *palBA*

To identify the hydrolytic clusters, we searched for homologs of the enzymes determined experimentally (PhnW, PhnX, PhnY, PhnA, PalB and PalA) in different bacterial genomes; the queries used for this analysis are shown in Table 1. Then, we considered only those clusters in which the genes position suggests a single transcriptional unit or a coregulated region (i.e., those where genes are arranged unidirectionally or diverging from each other), and in which all genes of the core pathway are present (thus excluding variants in which even one of them is missing). In particular, the clusters were visualized and inspected manually through the "Conserved Neighborhood" tool in the Integrated Microbial Genomes & Microbiomes website (IMG/M: https://img.jgi.doe.gov/m/) and the "Tree-based Genome Browser" tool of Microbesonline (http://www.microbesonline.org/). In addition, these websites and STRING (https://string-db.org/) were also used to assess whether the newly detected gene association within this surrounding neighborhood is preserved in multiple genomes.

Table 1. *Accession number list of enzymes used as queries for bioinformatic analysis.*

Enzyme	*Accession number*	*Organism*
PhnW	NP_459426.1	*Salmonella enterica*
PhnX	NP_459427.1	*Salmonella enterica*
PhnY	RMI18174.1	*Sinorhizobium meliloti*
PhnA	AAC15507.1	*Pseudomonas fluorescens*
PalB	ABR13824.1	*Variovorax sp. Pal2*
PalA	AAO24736.1	*Variovorax sp. Pal2*

2.2.2 Multiple sequence alignment (MSA)

For each of the identified hypothetical enzymes, we performed a homology search (setting an E-value threshold of 1e-5 and ≈30% sequence identity over the entire sequence length) with BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) in the UniProtKB/Swiss-Prot and PDB databases, and with Hidden Markovian Models (HMMs) in the B6 database (http://bioinformatics.unipr.it/cgi-bin/bioinformatics/B6db/home.pl); the B6 database is dedicated to the analysis of PLP-dependent enzymes. The sequences of validated enzymes most similar to the hypothetical one were aligned with ClustalX2 [17], and the multiple sequence alignment was visualized with ESPript 3.0 (https://espript.ibcp.fr/ESPript/ESPript/). The ClustalX2 program also generated a Neighbor-Joining

phylogenetic tree, visualized and edited with Figtree (http://tree.bio.ed.ac.uk/software/figtree/). To find orthologs of the hypothetical enzymes identified, we used as queries the protein sequences of PbfA and PbfB from *V. splendidus* 12B01 (NCBI Reference Sequence: WP_004730150.1 and WP_004730145.1, respectively), of PbfC from *Azospirillum lipoferum* B510 (NCBI Reference Sequence: WP_012976454.1), of PbfD1 of *Acinetobacter baumannii* (NCBI Reference Sequence: WP_079548425.1) and of PbfD2 from *Mariniblastus fucicola* (NCBI Reference Sequence: WP_075082418.1). Orthology was confirmed by checking for the presence of these genes in the context of phosphonate degradation clusters.

2.2.3 Frequency of the hypothetical gene recurrence in catabolic clusters

In order to quantitate the association recurrence of the hypothetical gene with the clusters for phosphonate catabolism, we automated the analysis process with python3 scripts. In detail, after the download of the protein FASTA assemblies (.faa) of complete bacterial genomes in the NCBI Assembly database, we proceeded through these steps:

- We created a BLAST database from the downloaded proteomes to perform a local homology search with DIAMOND BLASTP [18] (setting an E-value threshold of 1e-5 and using as queries the sequences of the validated enzymes PhnY* (ACU83549.1), PhnZ (ACU83550.1), the phosphonopyruvate decarboxylase Ppd (WP_011202932.1) and the hypothetical ones under investigation, in addition to those in Table 1);

- We labeled each homolog resulting from the BLASTP analysis with the respective protein name (e.g., we named "PhnW" all the PhnW homologs) to facilitate extraction and organization of the results;

- We assigned a number in ascending order to each sequence in the .faa file to extrapolate with DBSCAN [19] all the proteins (at least two hits, min_samples = 2) resulting from the homology analysis that are "located" at a maximum distance of 8 sequences from each other (eps = 8). Since the order of the sequences in the .faa file reflects those of genes in the genome, we considered only the homologs that are close to each other and, therefore, most likely belong to the same cluster.

- We filtered the obtained putative clusters considering only those that respect the arrangement of the known PhnW-PhnX, PhnW-PhnY-PhnA, PalA-PalB, PhnY*-PhnZ, and their corresponding variants that include the hypothetical enzyme. Since PalA is a close homolog of PepM enzyme, which forms with Ppd the most common pathway for phosphonate biosynthesis, we excluded all the PalA-Ppd cases from the results.

2.3 Results and discussion

We visually inspected the bacterial genomic contexts of the C-P hydrolase pathways coding genes, starting with the first enzyme of 2-aminoethylphosphonate (2-AEP; ciliatine) degradation route, i.e., the PLP-dependent aminotransferase PhnW, which converts 2-AEP to phosphonoacetaldehyde (PnAA). Because it is known that PhnW can play a role in both biosynthesis and degradation of 2-AEP (see Section 1.3.2.2 of Chapter 1), we excluded from the analysis all cases in which homologs of PhnW (with a minimum of 35% sequence identity to the characterized enzyme of *Salmonella enterica* [20]) clustered with *pepM* and *ppd* genes involved in phosphonate biosynthesis. In contrast, we focused on PhnW homologs whose genes clustered with either *phnX* or the *phnY-phnA* pair (*phnWX* and *phnWAY* clusters). While by doing so we found many clusters to analyze, the search of PalA and PalB homologs was less productive, primarily because: (i) PepM shares a high percentage of sequence identity with PalA, so many of the results obtained were related to biosynthetic pathways; (ii) it is difficult to discern the true PalB homologs from the wrong ones without any structural information (PalB is a PLP-dependent enzyme, and these enzymes, despite being often structurally similar, perform a wide range of different biochemical reactions, so it is complicated to assign a biochemical function based on homology alone).

In particular, we detected two frequently found genes associated with the *phnWX* and *phnWAY* clusters; one automatically annotated as coding for a PLP-dependent 4-aminobutyrate transaminase (GABA-T) (PF00202) and one coding for a hypothetical FAD-dependent oxidoreductase (PF01266). Because of the recurrent association of the genes with these clusters for phosphonate degradation, we provisionally labeled the encoded protein as "PbfX" (X = a letter in alphabetical order), which is the abbreviation for "phosphonate breakdown factor X".

2.3.1 Bioinformatic analysis of PbfA: a putative GABA-T (PF00202)

2.3.1.1 Occurrence and composition of PbfA's gene clusters

Given the long tradition in the study of PLP-dependent enzymes in our laboratory, the first candidate we decided to analyze was the PLP-dependent enzyme annotated as 4-aminobutyrate transaminase (GABA-T) or, more generically, as a member of the "class-III aminotransferase" family (PF00202). We termed this enzyme PbfA. PbfA's orthologs were found mainly in γ-proteobacteria belonging to *Aeromonadales*, *Oceanospirillales* and *Vibrionales* orders, clustered within the *phnWX* cluster. However, it was also found in some α-proteobacteria of the *Rhizobiales* and *Rhodospirillales* orders, grouped with *phnWAY* (Fig. 7). This co-occurrence strongly suggests that PbfA could be involved in 2-AEP biodegradations, maybe forming 2-AEP or phosphonoacetaldehyde since they are common intermediates of both the PhnW-PhnX and PhnW-PhnY-PhnA pathways.

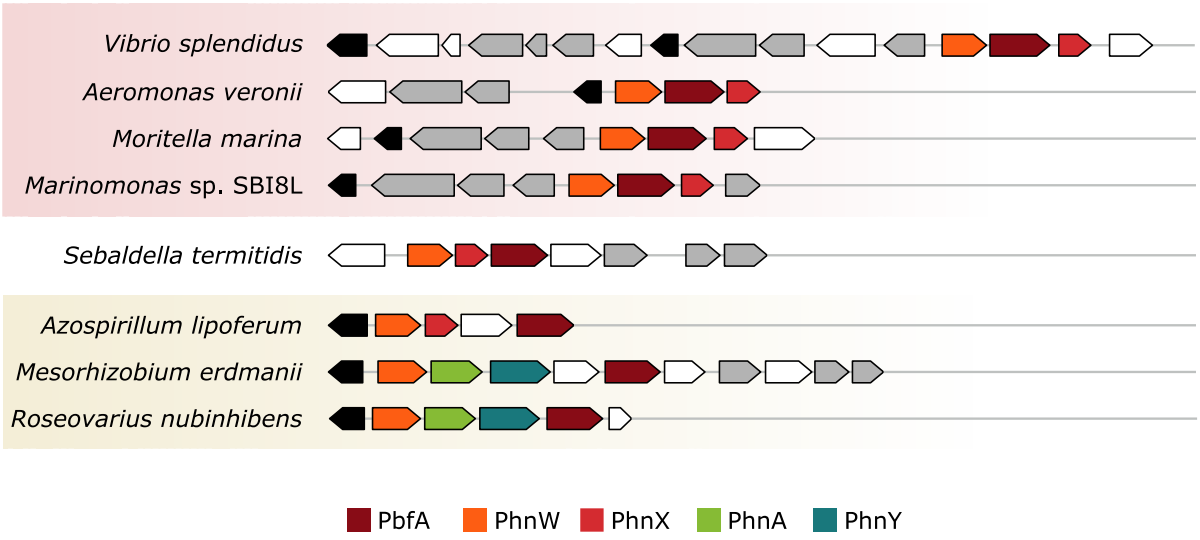
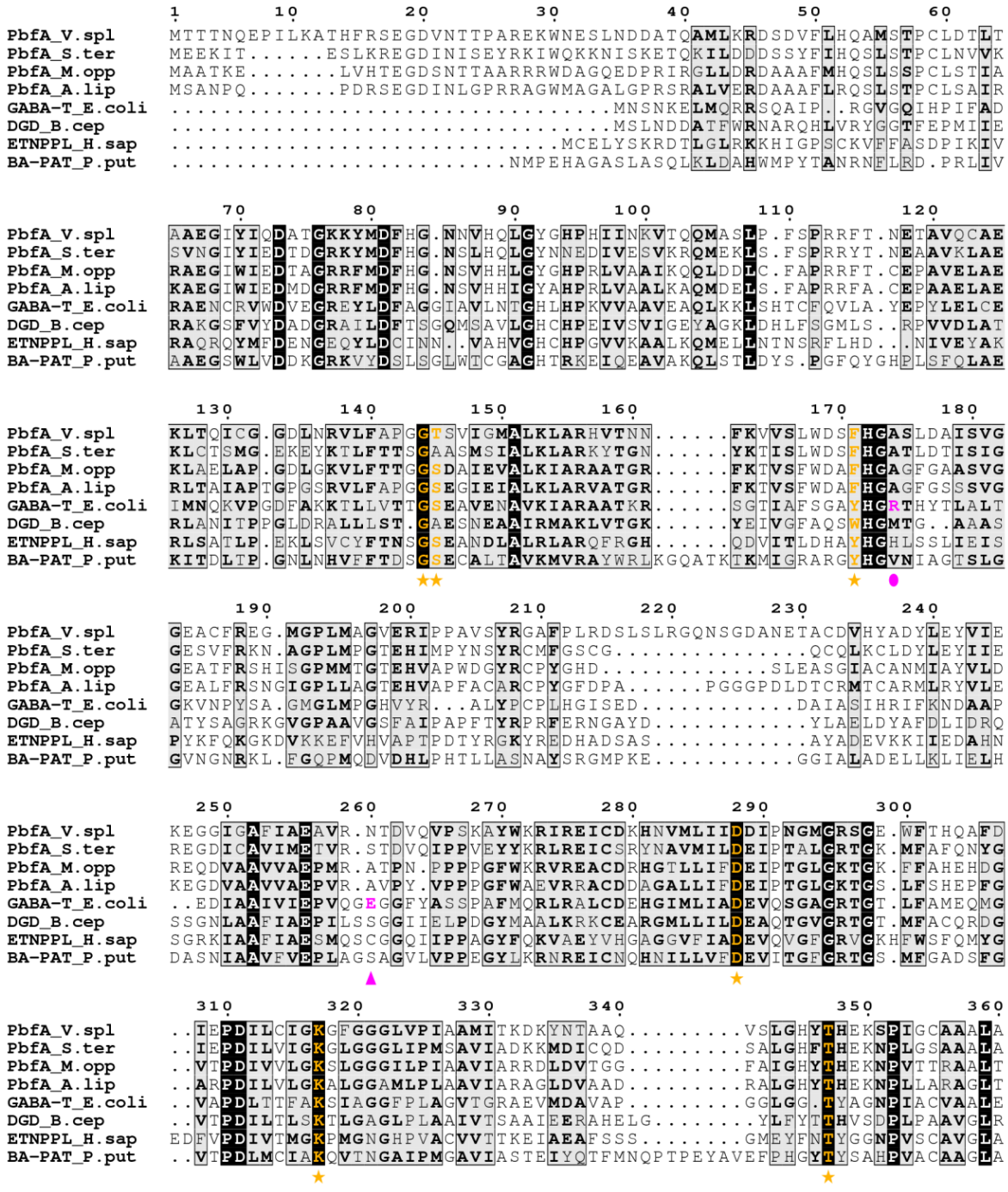


Figure 7. *Recurrent presence of pbfA within different bacterial genomes*. The *pbfA* gene is found associated with both clusters for 2-AEP degradation, *phnWX* and *phnWAY*. Genes encoding for phosphonate transporters are shown in gray and the transcriptional regulators in black. These clusters are particularly abundant in γ-proteobacteria (highlighted by a burgundy background) and α-proteobacteria (highlighted by an ochre background), but are also found in very distant bacterial lineages, such as *Sebaldella termitidis*.

2.3.1.2 PbfA similarity to functionally validated PLP-dependent enzymes

In addition to homology searches in the Uniprot and PDB databases, the similarity of PbfA with functionally validated PLP-dependent enzymes was also verified by consulting the B6 database. This analysis shows that PbfA belongs to the fold-type I structural family of PLP-dependent enzymes [21], more specifically to the subgroup of "class-III aminotransferase", which generally act on ω-amines whose amino group is not adjacent to a carboxylate [22]; examples are the 4-aminobutyrate transaminase from *Escherichia coli* (30% identity to PbfA), β-alanine:pyruvate aminotransferase from *Pseudomonas aeruginosa* (29% identity to PbfA) and 2,2-dialkylglycine decarboxylase from *Burkholderia cepacia* (30% identity to PbfA). Most of the experimentally characterized enzymes in the "class-III aminotransferase" subgroup are aminotransferases, with the exception of 2,2-dialkylglycine decarboxylase [23] and of two lyases (such as *O*-phosphoethanolamine phospho-lyase [24]). A multiple sequence alignment of PbfA sequences with these enzymes revealed conservation of PLP-binding residues and of the catalytic lysine, whereas there was no conservation of residues known to be essential for substrate and reaction specificity (Fig. 8). Specifically, if we take the PbfA sequence of *Vibrio splendidus* (PbfA_V.spl) as a reference, the PLP-binding lysine is K317 and the coordinating residues T347, G144, and T145 bind the phosphate group, D288 interacts with the pyridine nitrogen and F171 stacks on the PLP ring. In contrast, PbfA lacks residues typical of aminotransferases, i.e., the

glutamate and two arginines characteristic of GABA-T and transaminases that use α-KG as the amino group acceptor, and the arginine characteristic of β-alanine:pyruvate transaminase and transaminases that use pyruvate as the amino acceptor (Fig. 8). Nevertheless, PbfA possesses a lysine instead of the arginine of β-alanine:pyruvate transaminase. This substitution could be compatible with the possibility that PbfA catalyzes transamination of 2-AEP, which is a structural analog of β-alanine. This hypothesis is contradicted by the fact that PhnW, which performs this reaction, is already present in these clusters; therefore, 2-AEP transamination by PbfA would be a redundant activity.



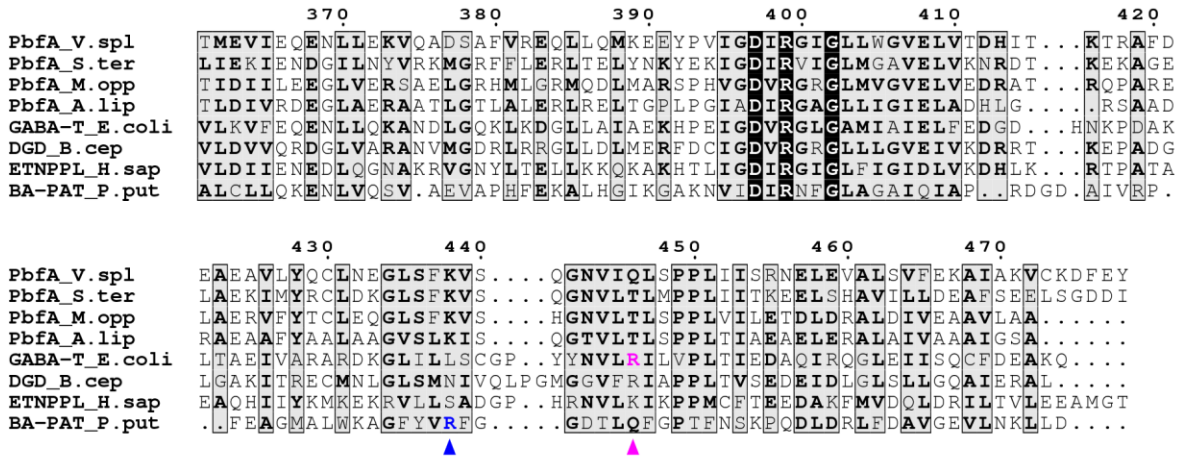


Figure 8. *Multiple sequence alignment of PbfA homologs with "class-III aminotransferases"*. The orange stars indicate residues that interact with the PLP cofactor and the catalytic Lys residues. The color pink denotes the residues typical of aminotransferases that use α-ketoglutarate (α -KG) as the amino acceptor: the pink triangles indicate the Arg and the Glu of the 'gateway system' (R398 and E211 in GABA-T_E.coli) that prevent the reaction of PLP with α-amino groups instead of ω-amino groups and ensure the correct positioning of the α-KG; the Arg residue indicated with a pink circle coordinates the α-KG (R141 in GABA-T_E.coli). If the substrate has not a distal carboxylic group, this Arg is substituted by other residues, e.g., a His in the lyase ETNPPL_H.sap and a Met in DGD_B.cep. In PbfA, the R141 of GABA-T_E.coli is replaced by an Ala, and in BA-PAT_P.put by a Val. PbfA sequences also contain a conserved lysine (K438 in PbfA_V.slp) at the same position where BA-PAT_P.put and other transaminases have an Arg to use the pyruvate as amino acceptor (position marked by a blue triangle). Accession codes: PbfA_V.sple, WP_004730150.1; PbfA_S.ter, WP_012862841.1; PbfA_M.opp, WP_013896952.1; PbfA_A.lip, WP_014188790.1; 4-aminobutyrate transaminase GABA-T_E.coli, P22256; 2,2-dialkylglycine decarboxylase DGD_B.cep, P16932; *O*-phosphoethanolamine phospho-lyase ETNPPL_H.sap, Q8TBG4; β-alanine:pyruvate transaminase BA-PAT_P.put, P28269. To compare PbfA with "class-III aminotransferase", we consulted these papers [22], [25].

2.3.1.3 Hypotheses about the possible PbfA's activity

If we relied only on the information gained from homology analysis, we would have insufficient data to make reliable speculations regarding the activity and substrate specificity of PbfA. This analysis concluded that PbfA belongs to the subgroup of "class-III aminotransferase", but it does not share any conserved residues with the validated enzymes in this subfamily, except those residues that bind the cofactor. In particular, PbfA lacks the residues involved in the binding of α-keto acid substrates (pyruvate and α-KG), typical of a transamination reaction. We, therefore, searched in the literature for a potential aminophosphonate substrate considering that (i) it should contain a primary amino

group as is the norm for PLP-dependent enzymes, (ii) it should be quite abundant in nature to justify the frequent presence of PbfA in bacterial genomes, and (iii) it should be converted into 2-AEP or phosphonoacetaldehyde to feed into the PhnW-PhnX or PhnW-PhnY-PhnA pathway. We also considered the different reaction mechanisms of "class-III aminotransferase" enzymes, namely transamination, decarboxylation, 1,2-elimination. On the basis of all these aspects, two hypotheses seemed to be the most probable:

1) PbfA could act on phosphonoalanine (PnAla) through a decarboxylation reaction producing 2-AEP, or through a decarboxylation reaction followed by transamination to yield PnAA (similar to the reaction mechanism of 2,2-dialkylglycine decarboxylase).

2) PbfA could act on 1-hydroxy-2-aminoethylphosphonate (OH-AEP) through a 1,2-elimination reaction producing PnAA (similar to the reaction mechanism of O-phosphoethanolamine phospho-lyase) (Fig. 9).

These hypotheses are also reinforced by the fact that in the complete genomes in which PbfA homologs are present, the enzymes known to degrade PnAla (PalB-PalA) and that to degrade OH-AEP (PhnZ) are absent. The only exception is *Mesorhizobium erdmanii* that possess both *phnWAY-pbfA* and *palAB* clusters.

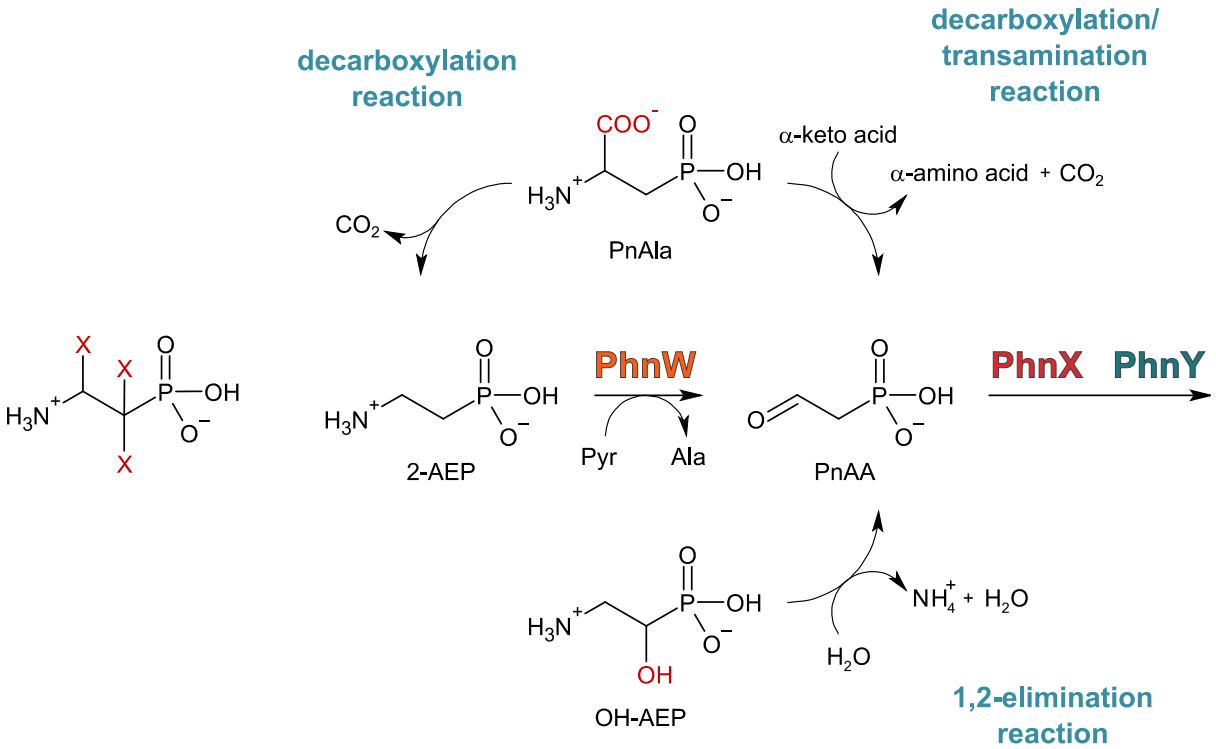


Figure 9. *Reaction scheme of the hypothesized mechanisms of PbfA's reactions and substrates*.

To test PbfA activity *in vitro*, we decided to recombinantly produce in *E. coli* the enzyme from the marine *Vibrio splendidus* 12B01, along with the predicted PhnW and PhnX from the same organism.

2.3.2 Bioinformatic analysis of three putative FAD-dependent oxidoreductases (PF01266)

The frequent presence of genes apparently encoding for a FAD-dependent protein in *phnXY* clusters also emerged quite clearly from our research. In particular, we were struck by the fact that, although they all belong to the FAD-dependent oxidoreductase family (PF01266), from a sequence perspective, these gene products are not closely related to each other but fall into at least three subgroups, which we named PbfB, PbfC, and PbfD (Fig. 10). Specifically, an analysis in the NCBI Conserved Domain Database (CDD) shows that PbfB belongs to the prokaryotic protein family annotated as TIGR03329 (Phn_aa_oxid) and PbfD to that annotated as TIGR03364 (HpnW_proposed), both of which are frequently observed in a genomic context that includes genes for phosphonate import and catabolism. Although their possible role is proposed in these annotations, no experimental follow-up has been done.

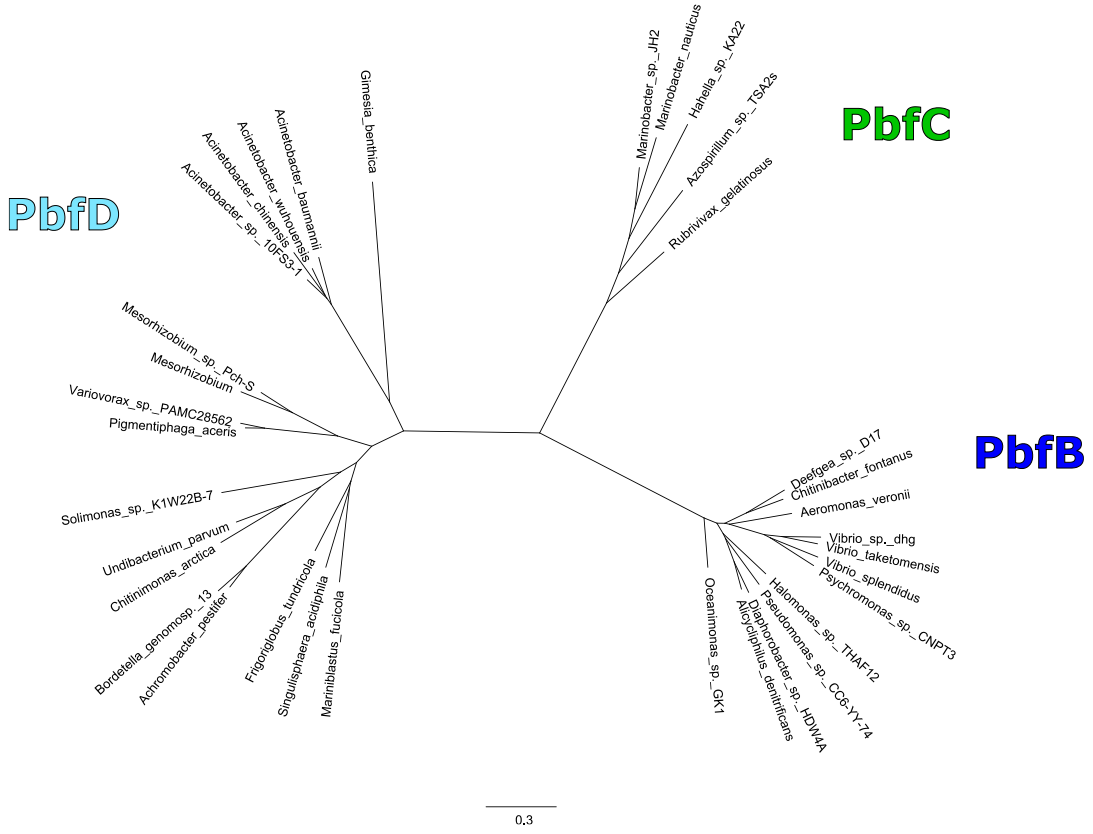


Figure 10. *Phylogenetic tree of representative FAD-dependent enzymes encoded in the gene clusters of the hydrolytic pathways (phnWX and phnWAY)*. The alignment excludes the sequences with an identity >80% to avoid overrepresentation of neighboring sequences. The sequences are identified by the organism's name from which they are derived.

2.3.2.1 Occurrence and composition of PbfB-C-D's gene clusters

PbfB was initially identified as being positioned in a divergent direction to the *phnWX-pbfA* cluster of *Vibrio splendidus*. PbfB orthologs are found distributed mainly in γ-proteobacteria as *Vibrionales*, *Aeromonadales*, and *Pseudomonadales*, and in β-proteobacteria as *Burkholderiales*. PbfC was found instead in α/β/γ-proteobacteria such as *Rhodospirillales*, *Burkholderiales* and *Oceanospirillales*, respectively. PbfD is also found like PbfC in several α/β/γ-proteobacteria, for example, it occurs in *Hyphomicrobiales* (such as *Mesorhizobium erdmanii*, which we mentioned above in the text for the concomitant presence of *phnWAY-pbfA* and *palAB* clusters), *Burkholderiales* and *Pseudomonadales*, as well as in some organisms of the *Planctomycetia* class. Although *pbfD* recurs in many *Acinetobacter* strains clustered with *phnWX*, it is more often found in taxonomically diverse organisms associated with *phnZ* and *phnX*, in the absence of *phnW* (Fig. 11). Like *pbfA*, it is shown by the figure that *pbfB*, *pbfC*, and *pbfD* also co-occur in both well-known clusters for 2-AEP degradation, suggesting that their reaction product may be either 2-AEP or PnAA, common intermediates to both hydrolytic pathways. In addition, the fact that *pbfD* is often combined with only *phnX* suggests that much more likely its product may be PnAA, and that in this case, it may replace the *phnW* reaction by acting on 2-AEP through an oxidative deamination.

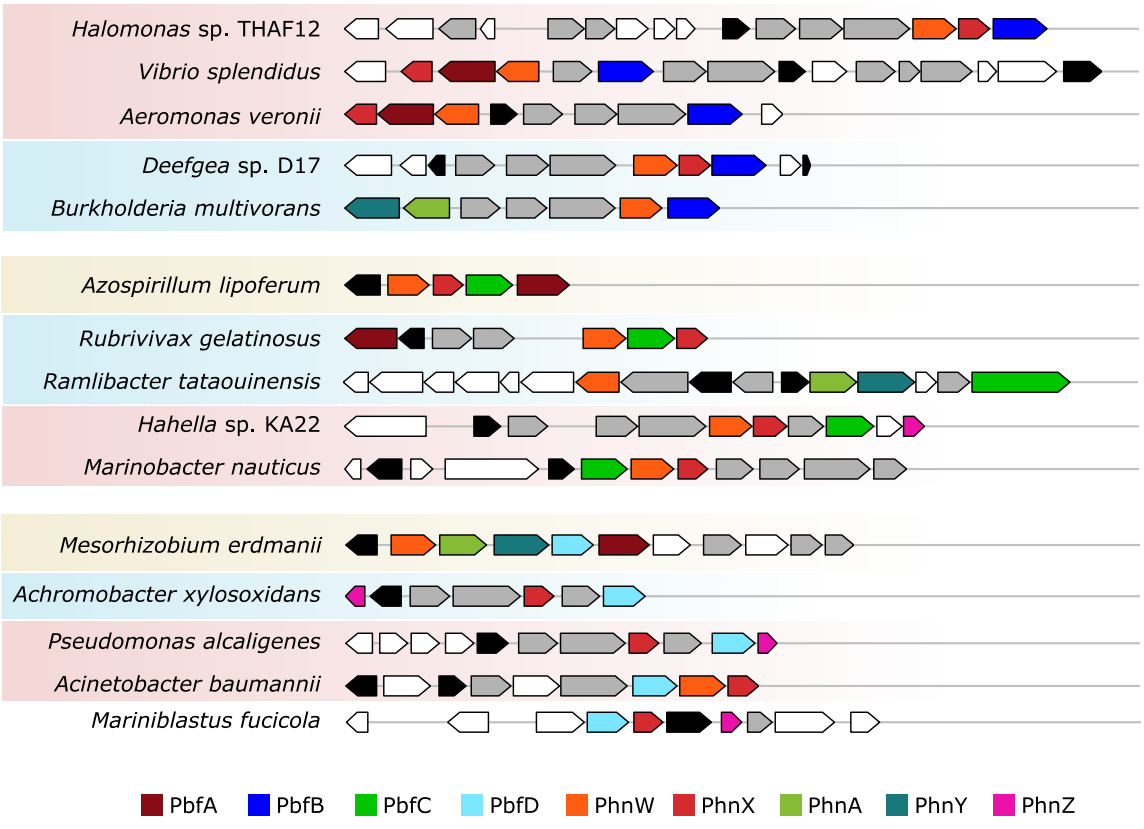


Figure 11. *Recurrent presence of pbfB, pbfC and pbfD within different bacterial genomes*. The *pbfB, pbfC and pbfD* genes are associated with both clusters for 2-AEP degradation, *phnWX* and *phnWAY*.

The *pbfD* gene is often associated with *phnX* and *phnZ*, without *phnW*. Genes encoding for phosphonate transporters are shown in gray and the transcriptional regulators in black. These clusters are found in α-proteobacteria (highlighted by an ochre background), β-proteobacteria (highlighted by a blue background) and γ-proteobacteria (highlighted by a burgundy background).

2.3.2.2 PbfB-C-D similarity to functionally validated FAD-dependent enzymes

Despite their limited reciprocal similarities, all these enzymes belong to the D-amino acid oxidase/sarcosine oxidase structural family (PF01266) [26], composed of both amino acid oxidases (AO-oxidases) and dehydrogenases, such as D-amino acid oxidase (DAAO; EC 1.4.3.3), sarcosine oxidase (SOX; EC 1.5.3.1), glycine oxidase (GOX; EC 1.4.3.19) and proline dehydrogenase (LPDH; EC 1.5.5.2). These enzymes catalyze the oxidative cleavage of C-N bond of a primary, secondary or tertiary amine through a proton abstraction from the amine group followed by a hydride transfer to the flavin cofactor. This type of reaction leads to the formation of intermediate imine species, which then hydrolyzes spontaneously in aldehyde and ammonia (or amine for secondary and tertiary amine substrates). Meanwhile, the FAD reoxidation occurs by different mechanisms; in the case of the oxidases it reacts directly with dioxygen to yield H_2O_2, whereas in dehydrogenases the reduced FAD can transfer its electrons to other cofactors, such as an iron-sulfur center or ubiquinone, or to another protein, such as the electron transfer flavoprotein. A sequence similarity search in the UniProtKB/Swiss-Prot database shows few results with a significant E-value threshold for PbfB and PbfD. PbfB shares 29% sequence identity with γ-glutamylputrescine oxidoreductase of *E. coli,* and PbfD shares 24% sequence identity with the monomeric sarcosine oxidase of *Bacillus sp.* B-0618. In contrast, sequences related to PbfC are the glycine oxidase of *Geobacillus kaustophilus* HTA426 (32% identity), D-arginine dehydrogenase of *Pseudomonas aeruginosa* (29% identity), 4-methylaminobutanoate oxidase of *Paenarthrobacter nicotinovorans* (28% identity) and dimethylglycine oxidase of *Arthrobacter globiformis* (27% identity). They are all members of the DAAO family that act on α-amino acids and N-methylated amino acids, but not all of them break the C-N bond between the α-carbon and the amino group. For example, DAAO oxidizes the Cα-N bond of sarcosine producing glyoxylate and methylamine, while SOX oxidizes the bond between the N and the C methyl group of sarcosine into glycine and formaldehyde (Fig. 12).

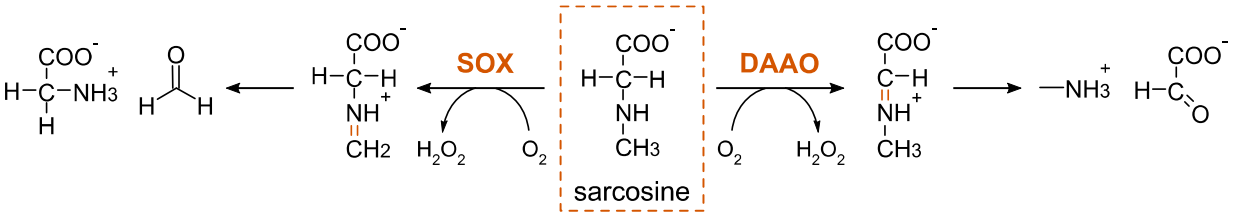


Figure 12. *Scheme reaction of sarcosine oxidation catalyzed by SOX and DAAO*.

However, DAAOs and SOXs belong to the same structural family, they can be distinguished by specific structural and biochemical features [27]. A multiple sequence alignment of PbfB, PbfC and PbfD enzymes with validated D-amino acid oxidase/sarcosine oxidase family enzymes shows the conservation of the N-terminal FAD-binding region ($GXGXXG(X)_{17\text{-}19}D/E$ motif) [28] (Fig. 13). Binding motifs for the FAD cofactor w also assessed by the use of the Rossmann-toolbox (https://lbs.cent.uw.edu.pl/rossmann-toolbox) and Cofactory 1.0 (https://services.healthtech.dtu.dk/service.php?Cofactory-1.0) web tools. By contrast, no residue known to covalently bind FAD was apparently preserved. Furthermore, the unique binding motifs that discriminate DAAOs from SOXs were not identified (such as the R, Y and K residues involved in the proton related system in SOX or the VAAGL motif of DAAO) [27], suggesting that these new hypothetical FAD-dependent enzymes do not belong to either family but only share the overall structure with them.

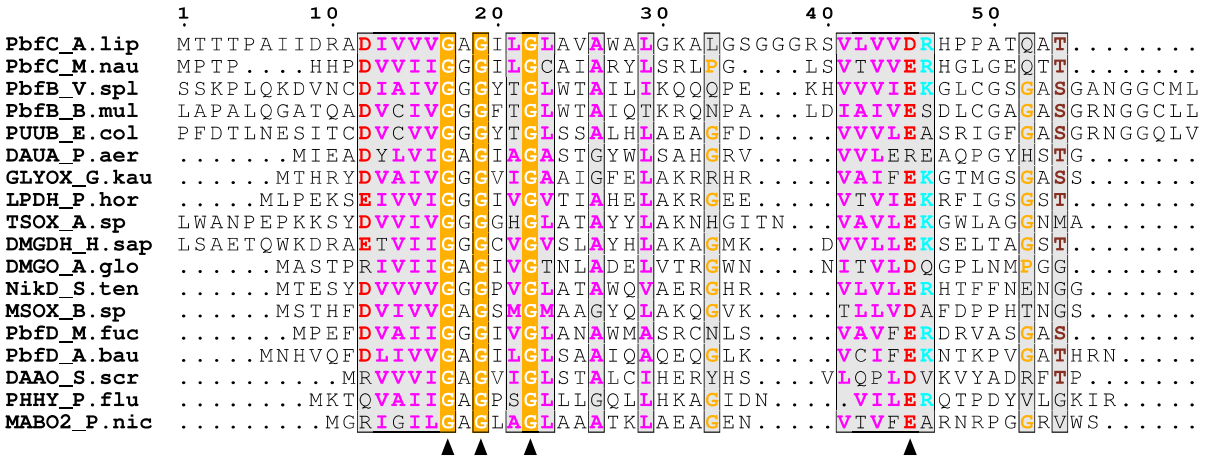


Figure 13. *Multiple sequence alignment of PbfB-C-D homologs with validated D-amino acid oxidase/sarcosine oxidase family enzymes*. The black triangles indicate the residues involved in the FAD-binding ($GXGXXG(X)_{17\text{-}19}D/E$ motif). Accession codes: γ-glutamylputrescine oxidoreductase (PUUB_E.col), P37906; Dimethylglycine oxidase (DMGO_A.glo), Q9AGP8; Glycine oxidase (GLYOX_G.kau), Q5L2C2; Monomeric sarcosine oxidase (MSOX_B.sp), P40859; D-arginine dehydrogenase (DAUA_P.aer) Q9HXE3; 4-methylaminobutanoate oxidase (MABO2_P.nic), Q8GAJ0; D-amino acid oxidase (DAAO_S.scr), 1VE9; p-hydroxybenzoate hydroxylase (PHHY_P.flu), 1BF3_1; L-proline dehydrogenase (LPDH_P.hor), WP_010885454.1; Dimethylglycine dehydrogenase (DMGDH_H.sap), XP_006714660.1; Sarcosine oxidase (TSOX_A.sp), Q9AGP3; NikD (NikD_S.ten), 2OLN; PbfC_A.lip, WP_012976454.1; PbfC_M.nau, WP_014420916.1; PbfB_V.spl, WP_004730145.1; PbfB_B.mul, WP_012214375.1; PbfD_A.bau, WP_079548425.1; PbfD_M.fuc, WP_075082418.1.

2.3.2.3 Hypotheses about the possible PbfB-C-D's activity

By analogy to the reaction mechanisms and amino-substrates on which the validated enzymes act, we looked for naturally occurring phosphonates containing a primary or secondary amino group, always assuming that the reaction product should presumably be 2-AEP or PnAA in order to feed the PhnW-PhnX or PhnW-PhnY-PhnA pathways. Based on these considerations, our initial working hypothesis was that all three types of FAD-dependent enzymes could catalyze the oxidative deamination of some 2-AEP derivative(s), alkylated on the amino group (Fig. 14). In particular, they could perform the oxidation of the N-monomethyl derivative of 2-AEP (MMAEP), which has been repeatedly isolated in nature, especially in the marine environment [29]–[31]. The reaction would produce methylamine, H_2O_2 and phosphonoacetaldehyde (PnAA, substrate for PhnX). If the substrate were 2-AEP itself, the enzyme would generate PnAA and release ammonia. This reaction, although biologically acceptable, seems redundant with the PhnW reaction, which also generates PnAA from 2-AEP; however, it might be useful in those cases where *phnW* is absent or in the case of pyruvate deficiency.

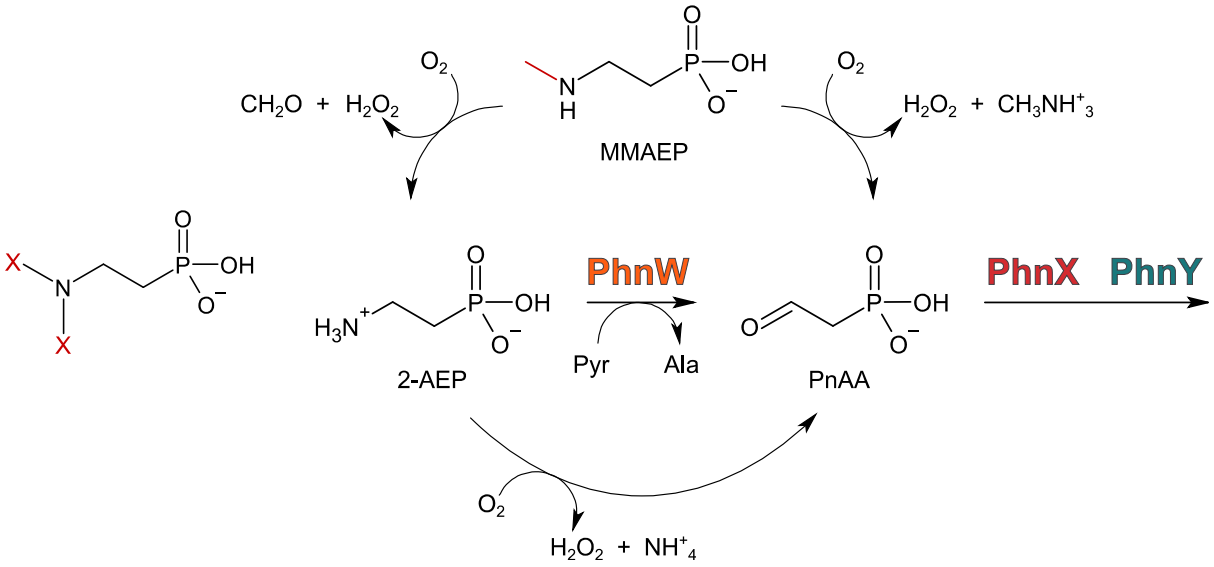


Figure 14. *Reaction scheme of the hypothesized mechanisms of PbfB-C-D's reactions and substrate*.

No other enzymes are known to degrade these N-methyl derivatives of 2-AEP (MMAEP and DMAEP), only the recently discovered iron-dependent oxygenase TmpB is known to degrades TMAEP [32], which is, however, a quaternary amine.

To test our hypotheses, we decided to recombinantly produce in *E. coli* PbfB from *Vibrio splendidus* (along with the *phnWX-pbfA* cluster of which it belongs), PbfC from the soil bacteria *Azospirillum lipoferum* B510 (clusterized with *phnWX*), and two PbfDs one from the pathogen *Acinetobacter baumannii* (PbfD1) associated with *phnWX* and the other from the planctomycete *Mariniblastus fucicola* (PbfD2) associated with *phnZ* and *phnX*.

2.3.3 Frequency estimation of *pbfA*, *pbfB*, *pbfC* and *pbfD* genes within *phnWX* or *phnWAY* clusters among bacteria

To achieve a survey of aminophosphonate degradative gene clusters associated to *pbfA*, *pbfB*, *pbfC* and *pbfD*, we downloaded 19,425 complete bacterial genomes deposited in the NCBI Assembly database at the end of November 2020 (see Section 2.2.3) (Table 2). In addition, to predict PbfA function, we checked whether *palAB* and *phnY*Z* clusters were present in *pbfA*-containing genomes, and we did not find them. The results show a similar frequency (about 14%) of finding *pbfA* or *pbfB* associated with *phnWX*, moreover, 62% of *phnWX-pbfB* clusters also possess *pbfA* (data not shown). The recurrence of *pbfC* or *pbfD* with *phnWX* is much lower, especially we observed few *pbfC* cases, about half of which also possess *pbfA*. PbfD gene was found in appreciable amounts and in taxonomically diverse bacteria in cluster with *phnZ* and *phnX* (without *phnW*), whereas it associated with *phnWX* mainly in the *Acinetobacter* genus. Although analyzing only the bacteria complete genomes in NCBI may be insufficient to estimate the frequency of these enzymes in nature (since most complete genomes are those of agricultural or pharmaceutical interest), these frequencies support the hypothesis that they are related to phosphonate biodegradation.

Table 2. – *Census and composition of aminophosphonate degradative gene clusters in complete bacterial genomes.*

Cluster	***total***	***+ pbfA***	***+ pbfB***	***+ pbfC***	***+ pbfD***
phnWX	1186	160 (13.5%)	169 (14.3%)	9 (0.8%)	48 (4.1%)
phnWAY	75	17 (22.7%)	10 (13.3%)	1 (1.6%)	2 (2.6%)
*phnY*Z*	7	0	-	-	-
palAB	50	0	-	-	-

2.4 References

[1] E. J. Richardson and M. Watson, "The automatic annotation of bacterial genomes," *Briefings in Bioinformatics*, vol. 14, no. 1, pp. 1–12, Jan. 2013, doi: 10.1093/bib/bbs007.

[2] J. A. Gerlt *et al.*, "The Enzyme Function Initiative," *Biochemistry*, vol. 50, no. 46, pp. 9950–9962, Nov. 2011, doi: 10.1021/bi201312u.

[3] D. J. Kountz and E. P. Balskus, "Leveraging Microbial Genomes and Genomic Context for Chemical Discovery," *Accounts of Chemical Research*, vol. 54, no. 13, pp. 2788–2797, Jul. 2021, doi: 10.1021/acs.accounts.1c00100.

[4] M. A. Huynen and B. Snel, "Gene and context: Integrative approaches to genome analysis," *Advances in protein chemistry*, vol. 54, pp. 345–379, 2000, doi: 10.1016/S0065-3233(00)54010-8.

[5] J. C. Navarro-Muñoz *et al.*, "A computational framework to explore large-scale biosynthetic diversity," *Nature Chemical Biology*, vol. 16, no. 1, pp. 60–68, Jan. 2020, doi: 10.1038/s41589-019-0400-9.

[6] T. Shiraishi and T. Kuzuyama, "Biosynthetic pathways and enzymes involved in the production of phosphonic acid natural products," *Bioscience, Biotechnology, and Biochemistry*, vol. 85, no. 1, pp. 42–52, Jan. 2021, doi: 10.1093/bbb/zbaa052.

[7] S. Li and G. P. Horsman, "An inventory of early branch points in microbial phosphonate biosynthesis," *Microbial Genomics*, vol. 8, no. 2, pp: 000781, Feb. 2022, doi: 10.1099/mgen.0.000781.

[8] K. Rice *et al.*, "The predominance of nucleotidyl activation in bacterial phosphonate biosynthesis," *Nature Communications*, vol. 10, no. 1, p. 3698, Dec. 2019, doi: 10.1038/s41467-019-11627-6.

[9] K.-S. Ju *et al.*, "Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes," *Proceedings of the National Academy of Sciences*, vol. 112, no. 39, pp. 12175–12180, Sep. 2015, doi: 10.1073/pnas.1500873112.

[10] X. Yu *et al.*, "Diversity and abundance of phosphonate biosynthetic genes in nature," *Proceedings of the National Academy of Sciences*, vol. 110, no. 51, pp. 20759–20764, Dec. 2013, doi: 10.1073/pnas.1315107110.

[11] K.-S. Ju, J. R. Doroghazi, and W. W. Metcalf, "Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways," *Journal of Industrial Microbiology and Biotechnology*, vol. 41, no. 2, pp. 345–356, Feb. 2014, doi: 10.1007/s10295-013-1375-2.

[12] G. P. Horsman and D. L. Zechel, "Phosphonate Biochemistry," *Chemical Reviews*, vol. 117, no. 8, pp. 5704–5783, Apr. 2017, doi: 10.1021/acs.chemrev.6b00536.

[13] J. Huang, Z. Su, and Y. Xu, "The Evolution of Microbial Phosphonate Degradative Pathways," *Journal of Molecular Evolution*, vol. 61, no. 5, pp. 682–690, Nov. 2005, doi: 10.1007/s00239-004-0349-4.

[14] A. Martinez, G. W. Tyson, and E. F. DeLong, "Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses," *Environmental Microbiology*, vol. 12, no. 1, pp. 222–238, Jan. 2010, doi: 10.1111/j.1462-2920.2009.02062.x.

[15] J. F. Villarreal-Chiu, “The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment,” *Frontiers in Microbiology*, vol. 3, 2012, doi: 10.3389/fmicb.2012.00019.

[16] A. R. J. Murphy *et al.*, “Transporter characterisation reveals aminoethylphosphonate mineralisation as a key step in the marine phosphorus redox cycle,” *Nature Communications*, vol. 12, no. 1, p. 4554, Dec. 2021, doi: 10.1038/s41467-021-24646-z.

[17] M. A. Larkin *et al.*, “Clustal W and Clustal X version 2.0,” *Bioinformatics*, vol. 23, no. 21, pp. 2947–2948, Nov. 2007, doi: 10.1093/bioinformatics/btm404.

[18] B. Buchfink, C. Xie, and D. H. Huson, “Fast and sensitive protein alignment using DIAMOND,” *Nature Methods*, vol. 12, no. 1, pp. 59–60, Jan. 2015, doi: 10.1038/nmeth.3176.

[19] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” 1996. [Online]. Available: www.aaai.org

[20] A. D. Kim, A. S. Baker, D. Dunaway-Mariano, W. W. Metcalf, B. L. Wanner, and B. M. Martin, “The 2-Aminoethylphosphonate-Specific Transaminase of the 2-Aminoethylphosphonate Degradation Pathway,” *Journal of Bacteriology*, vol. 184, no. 15, pp. 4134–4140, Aug. 2002, doi: 10.1128/JB.184.15.4134-4140.2002.

[21] N. v. Grishin, M. A. Phillips, and E. J. Goldsmith, “Modeling of the spatial structure of eukaryotic ornithine decarboxylases,” *Protein Science*, vol. 4, no. 7, pp. 1291–1304, Jul. 1995, doi: 10.1002/pro.5560040705.

[22] D. Schiroli and A. Peracchi, “A subfamily of PLP-dependent enzymes specialized in handling terminal amines,” *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1854, no. 9, pp. 1200–1211, Sep. 2015, doi: 10.1016/j.bbapap.2015.02.023.

[23] W. Liu, C. J. Rogers, A. J. Fisher, and M. D. Toney, “Aminophosphonate Inhibitors of Dialkylglycine Decarboxylase: Structural Basis for Slow Binding Inhibition,” *Biochemistry*, vol. 41, no. 41, pp. 12320–12328, Oct. 2002, doi: 10.1021/bi026318g.

[24] M. Veiga-da-Cunha, F. Hadi, T. Balligand, V. Stroobant, and E. van Schaftingen, “Molecular Identification of Hydroxylysine Kinase and of Ammoniophospholyases Acting on 5-Phosphohydroxy-l-lysine and Phosphoethanolamine,” *Journal of Biological Chemistry*, vol. 287, no. 10, pp. 7246–7255, Mar. 2012, doi: 10.1074/jbc.M111.323485.

[25] F. Steffen-Munsberg *et al.*, “Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications,” *Biotechnology Advances*, vol. 33, no. 5, pp. 566–604, Sep. 2015, doi: 10.1016/j.biotechadv.2014.12.012.

[26] P. F. Fitzpatrick, “Oxidation of amines by flavoproteins,” *Archives of Biochemistry and Biophysics*, vol. 493, no. 1, pp. 13–25, Jan. 2010, doi: 10.1016/j.abb.2009.07.019.

[27] M. Lahham, S. Jha, D. Goj, P. Macheroux, and S. Wallner, “The family of sarcosine oxidases: Same reaction, different products,” *Archives of Biochemistry and Biophysics*, vol. 704, p. 108868, Jun. 2021, doi: 10.1016/j.abb.2021.108868.

[28] O. Dym and D. Eisenberg, “Sequence-structure analysis of FAD-containing proteins,” *Protein Science*, vol. 10, no. 9, pp. 1712–1728, Sep. 2001, doi: 10.1110/ps.12801.

[29] L. D. Quin and G. S. Quin, “Screening for carbon-bound phosphorus in marine animals by high-resolution 31P-NMR spectroscopy: coastal and hydrothermal vent invertebrates,” *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, vol. 128, no. 1, pp. 173–185, Jan. 2001, doi: 10.1016/S1096-4959(00)00310-9.

[30] Kh. S. Mukhamedova and A. I. Glushenkova, “Natural Phosphonolipids,” *Chemistry of Natural Compounds*, vol. 36, no. 4, pp. 329–341, 2000, doi: 10.1023/A:1002804409503.

[31] J. S. Kittredge, A. F. Isbell, and R. R. Hughes, “Isolation and Characterization of the N-Methyl Derivatives of 2-Aminoethylphosphonic Acid from the Sea Anemone, Anthopleura xanthogrammica*,” *Biochemistry*, vol. 6, no. 1, pp. 289–295, Jan. 1967, doi: 10.1021/bi00853a045.

[32] L. J. Rajakovich *et al.*, “A New Microbial Pathway for Organophosphonate Degradation Catalyzed by Two Previously Misannotated Non-Heme-Iron Oxygenases,” *Biochemistry*, vol. 58, no. 12, pp. 1627–1647, Mar. 2019, doi: 10.1021/acs.biochem.9b00044.

Chapter 3

3.1 PLP-dependent enzymes

The pyridoxal 5'-phosphate (PLP) coenzyme, i.e., the active form of vitamin B6, is valued for its impressive polyvalence in enzymatic catalysis. Indeed, PLP-dependent activities belong to all but two classes defined by the Enzyme Commission, covering about 4% of all classified activities [1]. PLP is covalently bound to the enzyme active site through the reversible formation of a Schiff base (C=N) between the aldehyde group of the cofactor and the ε-amino group of the catalytic lysine, yielding the so-called "internal aldimine". When a substrate enters the active site, its amino group (necessarily primary) replaces that of the lysine through a transaldimination reaction leading to a "*gem*-diamine" species, which subsequently converts to the "external aldimine" (Fig. 15). Despite the multitude of possible reactions, this starting point of catalysis is common to all PLP-dependent enzymes. The only exception is the glycogen phosphorylase, in which PLP remains always bound as internal aldimine and the phosphate group of the cofactor is involved in catalysis [2]. After this first step, the specificity of the reaction depends on the stereochemistry of the external aldimine formed and on the enzymatic scaffold that forces the correct orientation of the substrate. The bond to be broken is assumed to lie in a plane orthogonal to the cofactor pyridine ring, which acts as an electron sink, as explored in detail by Dunathan [3]; after the bond cleavage, the PLP cofactor stabilizes several carbanionic reaction intermediates by resonance [4]. This permits a wide range of reactions to be catalyzed by PLP-dependent enzymes, namely, transamination, racemization, α- and β-decarboxylation, β- and γ-deletion, β- and γ-substitution, retro-aldol cleavage, and oxidation. Despite this functional variety, all PLP-dependent enzymes are now classified based on structural information into seven distinct groups (known as "fold-types") [5], [6]. Each fold-type was associated with the name of the first enzyme identified within that fold, thus: fold-type I is the aspartate aminotransferase superfamily, fold-type II is the tryptophan β-synthase superfamily, fold-type III is the alanine racemase superfamily, fold-type IV is the D-amino acid aminotransferase superfamily, fold-type V is the glycogen phosphorylase superfamily, fold-type VI is the lysine-5,6-aminomutase superfamily, and fold-type VII is the lysine-2,3-aminomutase superfamily. Most of the known structures and best characterized PLP-dependent enzymes belong to fold-type I, which includes mainly aminotransferases and some decarboxylases and lyases. They are obligatorily homodimers (or higher-order oligomers) since both subunits contribute to the formation of the active site. As mentioned in the previous chapters of this dissertation, the 2-aminoethylphosphonate:pyruvate aminotransferase (PhnW), phosphonoalanine aminotransferase (PalB) and the hypothetical PLP-dependent enzyme PbfA belong to the fold-type I. However, because enzymes of the same fold-type possess conserved structural features and interactions with PLP but can show very different reactions and substrate specificity, it is particularly difficult to infer the function of a new PLP-dependent enzyme based only on similarity [5], [7]. In our specific case, to

identify the function of PbfA, we relied on the logic of its gene context and were guided by our familiarity in the study of PLP-dependent enzymes and their reaction mechanism.

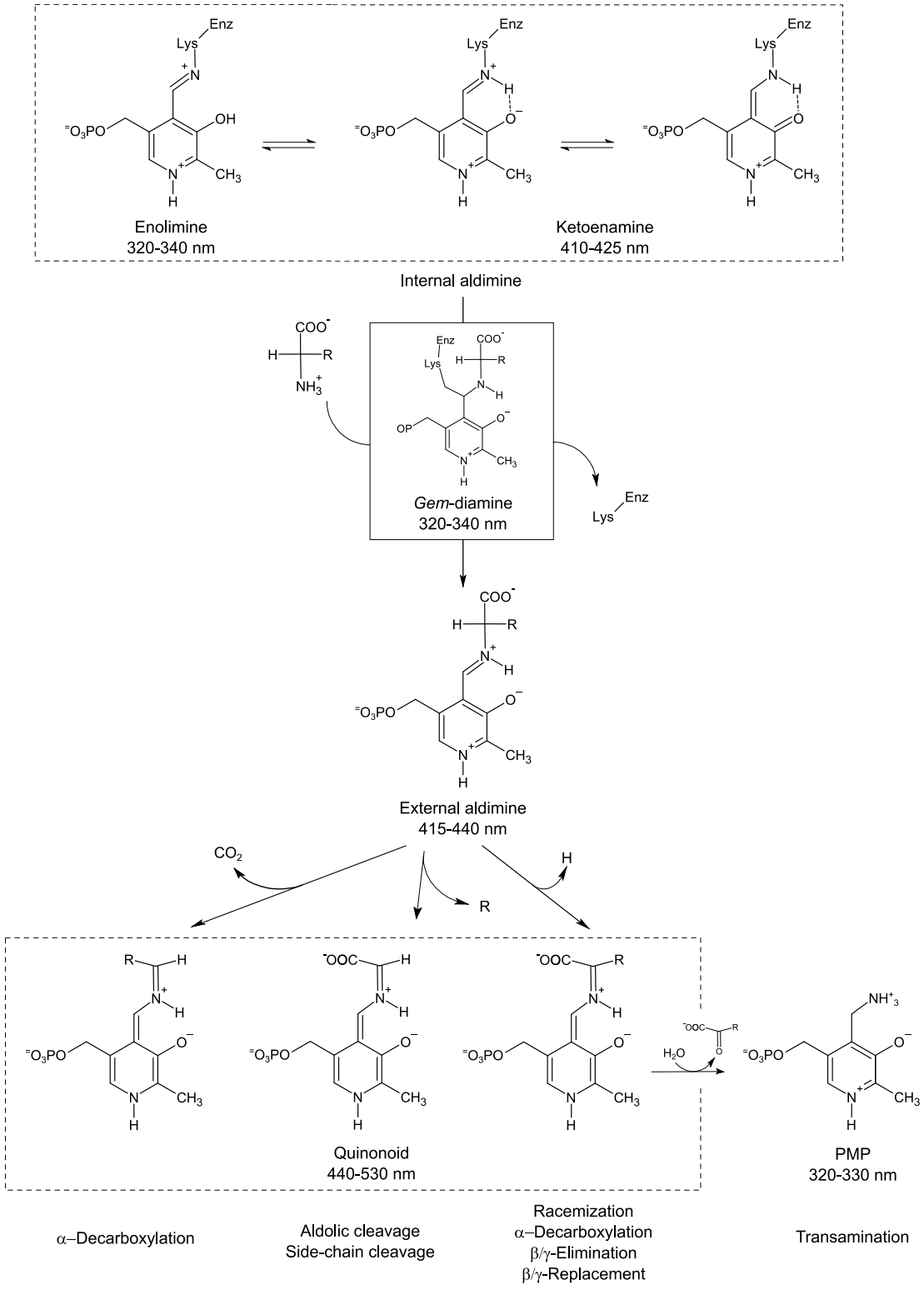


Figure 15. *Scheme of the early common steps of the reaction mechanism of PLP-dependent enzymes*. In the resting state, PLP is covalently bound to the enzyme as an internal aldimine (dashed box, upper), which can exist in two tautomeric forms depending on the protonation state: the ketoenamine and

enolimine tautomers. When the substrate enters the active site forms the external aldimine. From here, depending on which group is removed, the corresponding quinonoid intermediate is generated and the different reaction types follow (dashed box, lower); for example, the step of the first half of a transamination reaction (PMP formation) is shown at the bottom right. PLP confers a characteristic absorption spectrum to the protein that possesses it. The different covalent intermediates formed during catalysis can exhibit characteristic UV-visible spectra, for some of which the wavelength range at which they absorb is given.

3.2 Methods

3.2.1 Materials

Triethanolamine (TEA), pyridoxal-5'-phosphate (PLP), bovine serum albumin (BSA) were from Sigma-Aldrich. NADH and ninhydrin were from Alfa Aesar. TLC plates (silica gel 60 F254) were from Merck Millipore. Deuterium oxide was from VWR International and deuterated trimethylsilyl propionic acid (TSP) from Stohler Isotope Chemicals. All other reagents were from Fluka or Sigma–Aldrich.

Enzymes: alcohol dehydrogenase (ADH) from yeast, L-glutamate dehydrogenase (GDH) from bovine liver and L-alanine dehydrogenase (AlaDH) from *Bacillus subtilis* were from Sigma, the CpoI (RsrII) restriction enzyme was from Takara.

Phosphonic acids: D,L-2-Amino-3-phosphonopropionic acid (D,L-phosphonoalanine; D,L-PnAla) was from Sigma. 2-aminoethylphosphonic acid (ciliatine; 2-AEP) was from Wako chemicals. The chemical synthesis of 1-hydroxy-2-aminoethylphosphonic acid (hydroxyciliatine; OH-AEP) was essential to perform these studies. Specifically, racemic OH-AEP and the pure enantiomers *R*- and *S*- were synthesized by the laboratory of Dr. Katharina Pallitsch at the Institute of Organic Chemistry, University of Vienna. See the article [8] for details on their synthesis. The same group also synthesized some analogs of OH-AEP: 1-hydroxy-2-aminopropyl-phosphonic acid (OH-APP) and *R,S*-1-fluoro-2-aminoethylphosphonic acid (*R,S*-F-AEP).

3.2.2 Plasmid constructs and protein purification

We purchased from BaseGene BV (Leiden, the Netherlands) the *Vibrio splendidus* 12B01 genes coding for PhnW (WP_004730149.1) and PhnX (WP_004730152.1) cloned in the NdeI/NotI sites of a pET24a-C-HIS and pET28a-N-HIS, respectively. In contrast, the PbfA gene was purchased in the pBluescript II KS+ vector with the CpoI cleavage sites (CGGTCCG) at the ends. After reaction with the CpoI restriction enzyme, the gene for PbfA (WP_004730150.1) was subcloned into the pET28-CpoI expression vector [9] to obtain an N-terminal His6-tagged protein. The subcloned *pbfA* gene was sequence-verified, and the plasmid was used to transform by electroporation *E. coli* Tuner™ (DE3) cells (EMD Biosciences) for

protein expression. PhnX was also overexpressed in Tuner™ (DE3) cells, while PhnW was expressed in *E. coli* BL21 Star™ (DE3) cells. For protein expression, the bacterial cultures were grown at 37 °C in 1 L of LB medium supplemented with kanamycin (50 µg/mL) until OD_{600} reached 0.7-0.8. At this point, the temperature was reduced to 20°C and the induction was started by adding 0.3 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). 20 hours after induction, the cells were harvested by centrifugation (7,200 × *g* for 10 minutes at 4°C). The cell pellets were washed once with PBS, resuspended, centrifuged, and finally stored at -20°C. For protein purification, pellets were resuspended in the appropriate lysis buffer (see Lysis buffer in Table 3), sonicated and centrifuged (26,200 × *g* for 40 minutes at 4°C). Cleared cell lysates were loaded on a His-Select® cobalt affinity resin (Sigma-Aldrich) and the recombinant proteins were purified following the manufacturer's instructions. Protein purity was assessed by SDS-PAGE and fractions with purity higher than 90% were pooled and extensively dialyzed against their respective storage buffer (see Dialysis buffer in Table 3) to remove imidazole. After dialysis, 10% (v/v) glycerol was added to the enzyme stocks and the enzyme concentration was measured spectrophotometrically. Protein concentrations were determined using the extinction coefficient at 280 nm estimated from the amino acid sequence by the ProtParam tool (http://web.expasy.org/protparam/) (see Table 3). Each stock was subdivided into aliquots (about 0.5 ml each), frozen in liquid nitrogen and stored at -80 °C. The protein yield was 22 mg, 50 mg and 47 mg per liter of culture for PhnW, PhnX and PbfA, respectively.

Table 3. *The buffer solutions composition used during purification and the physical and chemical parameters of the His6-tagged proteins calculated by the ProtParam tool*.

Enzymes	Lysis buffer	Dialysis buffer	ε280nm value; molecular weight; theoretical pI
PhnW	50 mM Na-phosphate buffer pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol, 1 mM PMSF, 1 mM benzamidine, 1 mg/ml lysozyme, 10 µM PLP	50 mM Na-phosphate pH 7.5, 300 mM NaCl, 1 mM DTT, 5 µM PLP	48,360 M^{-1} cm^{-1}; 43.1 kDa; 5.69
PhnX	50 mM HEPES pH 7.5, 200 mM NaCl, 5 mM β-mercaptoethanol, 1 mM PMSF, 1 mM benzamidine, 1 mg/ml lysozyme, 10 mM $MgCl_2$	50 mM HEPES pH 7.5, 1 mM DTT, 10 mM $MgCl_2$	32,430 M^{-1} cm^{-1}; 31.6 kDa, 5.13
PbfA	50 mM Na-phosphate buffer pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol, 1 mM PMSF, 1 mM benzamidine, 1 mg/ml lysozyme, 10 µM PLP	50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT, 5 µM PLP	46,870 M^{-1} cm^{-1}; 55.0 kDa, 5.71

3.2.3 Qualitative detection of enzyme activity by thin-layer chromatography (TLC)

For qualitative inspection of the enzymatic activity of PhnW and PbfA, we coupled the TLC chromatographic technique, which separates mixtures of primary amines as α-amino acids, to subsequent staining of the plate with ninhydrin, which reacts rapidly with the amine groups giving colored spots (Ruhemann's purple). PhnW and PbfA were typically incubated with the substrates at 37 °C in 50 mM TEA-HCl buffer pH 8.0. After incubation was completed, two microliters of each reaction mixture were spotted onto a TLC plate and the samples were run with a solution of acetic acid:1-propanol:water in a ratio of 1:3:1. After chromatographic separation, the plate was sprayed with 0.3% ninhydrin solubilized in methanol, dried, and heated for about 5 min. Control experiments indicated that this assay could detect the formation of amino acid products at concentrations as low as 0.2 mM.

3.2.4 ADH-coupled assay for 2-AEP or PnAA detection

To test the hypotheses that PbfA produced 2-AEP or PnAA, we set up a continuous spectrophotometric assay in the presence of the PhnW-PhnX pair and of alcohol dehydrogenase (ADH) or with only PhnX and ADH, respectively. We initially verified the system's functionality by following the PhnW's 2-AEP transamination by monitoring the disappearance of NADH at 340 nm on a Varian Cary 400 UV-VIS thermostatted spectrophotometer. ADH-coupled assay for monitoring PhnW or PbfA activity, conducted in 150 µl final volume, typically contained 50 mM buffer TEA-HCl pH 8.0, 0.5 mg/ml BSA, 0.25 mM NADH, 5 mM $MgCl_2$, 5 µM PLP, an excess of PhnX and ADH, and the substrates. Initial velocities were calculated using the extinction coefficient of NADH at 340 nm of 6,220 $M^{-1}cm^{-1}$. Each reaction was repeated in triplicate and the kinetic parameters were determined by nonlinear least-squares fitting to the Michaelis-Menten equation using Sigma Plot (Systat Software Inc.).

3.2.5 GDH-coupled assay for ammonia detection

The hypothesis that PbfA can produce PnAA through an elimination reaction on OH-AEP was also tested with a GDH-coupled assay. The PbfA reaction was expected to release ammonia, which in turn, allows glutamate dehydrogenase (GDH) to react with α-ketoglutarate to produce L-glutamate by oxidizing NADH. The reaction was followed by monitoring the disappearance of NADH at 340 nm at the spectrophotometer. Usually, the reaction mixture contained 50 mM TEA-HCl buffer pH 8.0, 0.25 mM NADH, 5 mM $MgCl_2$, 5 µM PLP, 1 mM α-KG and an excess of GDH.

3.2.6 Colorimetric assay for phosphate release

The BIOMOL® Green kit (Enzo Life Sciences) was used following the manufacturer's instructions to detect phosphate release from enzymatic or spontaneous hydrolysis of phosphonates. Since PhnW

was stored in phosphate buffer, before testing its activity, the enzyme was dialyzed against a phosphate-free buffer (30 mM Hepes pH 7.5, 300 mM NaCl, 1 mM DTT, 5 µM PLP). A typical reaction was carried out in tubes containing 50 mM TEA-HCl buffer pH 8.0, the enzymes PhnW or PbfA and PhnX, the cofactors 5 µM PLP and 5 mM $MgCl_2$ and 1 mM of substrates. Reactions were conducted at room temperature, and in order to stop the enzyme activity, 18 µL of mixture reaction was transferred into 200 µL of BIOMOL® Green reagent. Before analyzing the collected samples, we waited 30 min for the color change (the greening of the solution) of the phosphate-containing reaction mixture. Then, the absorbance at 620 nm (due to the of the molecular complex between inorganic phosphate and the BIOMOL green reagent) was measured with a UV-Vis Cary 50 spectrophotometer (Varian).

3.2.7 Structure prediction and docking procedure

Structure prediction of the PbfA enzyme was obtained by exploiting the novel machine learning approach of AlphaFold2 [10] through the Colab Notebook, which supports the modeling of protein complexes (both homo and hetero-oligomers) starting from the FASTA sequences (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/beta/AlphaFold2_advanced.ipynb). Specifically, in this case, a model of the PbfA structure (.pdb file) was obtained using its FASTA sequence (WP_004730150.1), setting as two the homo-oligomers to model and leaving the other default parameters unchanged. This is because PLP-dependent enzymes belonging to fold-type I, such as PbfA, are obligate dimers because their active site spans both the two subunits. The docking analysis was performed with AutoDock Flexible Receptor (ADFRsuite 1.0) [11], which allows selective flexing of the receptor. Then the .pdb file obtained was converted in .pdbqt format with the prepare_receptor4.py utility command line of ADFRsuite. The ligands, *R*-OH-AEP and *S*-OH-AEP, were manually constructed with Avogadro [12] as the external aldimine, i.e., already covalently bound to the aldehyde group of PLP; the energies of the structures were minimized using Ghemical as the force field and finally saved as .mol2 file. The .pdbqt files of the ligand were prepared with the prepare_ligand4.py utility command line of ADFRsuite, preserving the previously assigned charges. The affinity maps were obtained with the agfr command using as search space a cube of 20 Å centered to the ε-amino group of the catalytic lysine (K317) in the chain A of the receptor .pdbqt. Docking results, i.e., .pdbqt files with the predicted binding modes and the .dlg docking log file containing clustering information (table S1), were obtained with the adfr command, keeping the default parameters unchanged and the catalytic lysine (K317) flexible, as well as the ligand. The .pdbqt files were visualized through PyMol software (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

3.2.8 UV-absorbance measurements

Absorption spectra were collected in a quartz microcuvette (l=1 cm) using a V-750 UV-Visible spectrophotometer (Jasco Inc.) and corrected for buffer contribution.

3.2.9 AlaDH-coupled assay and phosphate release upon PhnW reaction with OH-HAEP

The reaction of PhnW with *R*-OHAEP and pyruvate was monitored by following the production of alanine and phosphate. A reaction mixture of 1,350 µl was prepared with 20 mM Hepes pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 5 mM pyruvate, 5 mM *R*-OH-AEP and 0.8 µM PhnW to test whether the two compounds were produced in equimolar amounts. Every 10 min after addition of the enzyme, an aliquot of 150 µL was taken to quantify alanine by AlaDH-coupled assay and an aliquot of 18 µL to quantify phosphate by BIOMOL green kit. Each 18 µL aliquot was added to 200 µL of BIOMOL green reagent; the absorption at 620 nm of this sample was measured after 30 min to estimate the amount of phosphate. Aliquots for alanine quantification were immediately heated at 100 °C for 5 min, then frozen and subsequently analyzed by the AlaDH-coupled assay to determine the amount of alanine formed as described by Williamson [13]. Briefly, 100 µl of the sample to be analyzed were supplemented with 51 µl of freshly prepared hydrazine-tris buffer (40 mM Tris-HCl pH 10, 1.4 mM EDTA and 1 M hydrazine). After 30 min at room temperature, 0.8 mM NAD+ and finally 40 mU/mL alanine dehydrogenase were added. The sample's absorption at 340 nm was measured after 50, 100, and 240 minutes. The amounts of phosphate and alanine formed were estimated by referring to a standard curve obtained using known concentrations of phosphate and alanine, respectively.

3.2.10 NMR experiments

^{1}H NMR spectra were recorded at 25°C on a JEOL ECZ600R spectrometer equipped with a 5 mm Royal probe, using the DANTE presat sequence to suppress the water signal. In a 5 mm NMR tube, the reaction mixture contained 10% D_2O, 1 mM TSP (used as an internal chemical shift reference; δ=0.00 ppm), 50 mM potassium phosphate pH 8.0, 100 mM KCl, 5 mM $MgCl_2$, the enzymes (0.8 µM PhnW or 0.8 µM PbfA, and 1 µM PhnX) and 5 mM of substrates, in a final volume of 500 µL. Enzyme stocks containing glycerol and organic buffers were dialyzed prior to their use. NMR spectra were processed with MestReNova v12.0.4 (Mestrelab Research) software.

3.3 Results and discussion

This section will present the results obtained from the functional characterization of the PLP-dependent enzyme PbfA. To test our hypotheses, we made use of several techniques, including:

- Thin-layer chromatography (TLC) to observe substrate consumption or product formation;
- Absorption spectroscopy analyses of the enzyme alone or in the presence of potential substrates (this analysis is typical in the study of PLP-dependent enzymes);
- Development of continuous spectrophotometer coupled assays to follow product production;
- Detection of phosphate release through a colorimetric assay;
- Nuclear magnetic resonance (NMR) spectroscopy to characterize formed products.

The application of some of these methods are summarized in the figure below (Fig. 16), and part of the work about the PbfA characterization was published in this article [14].

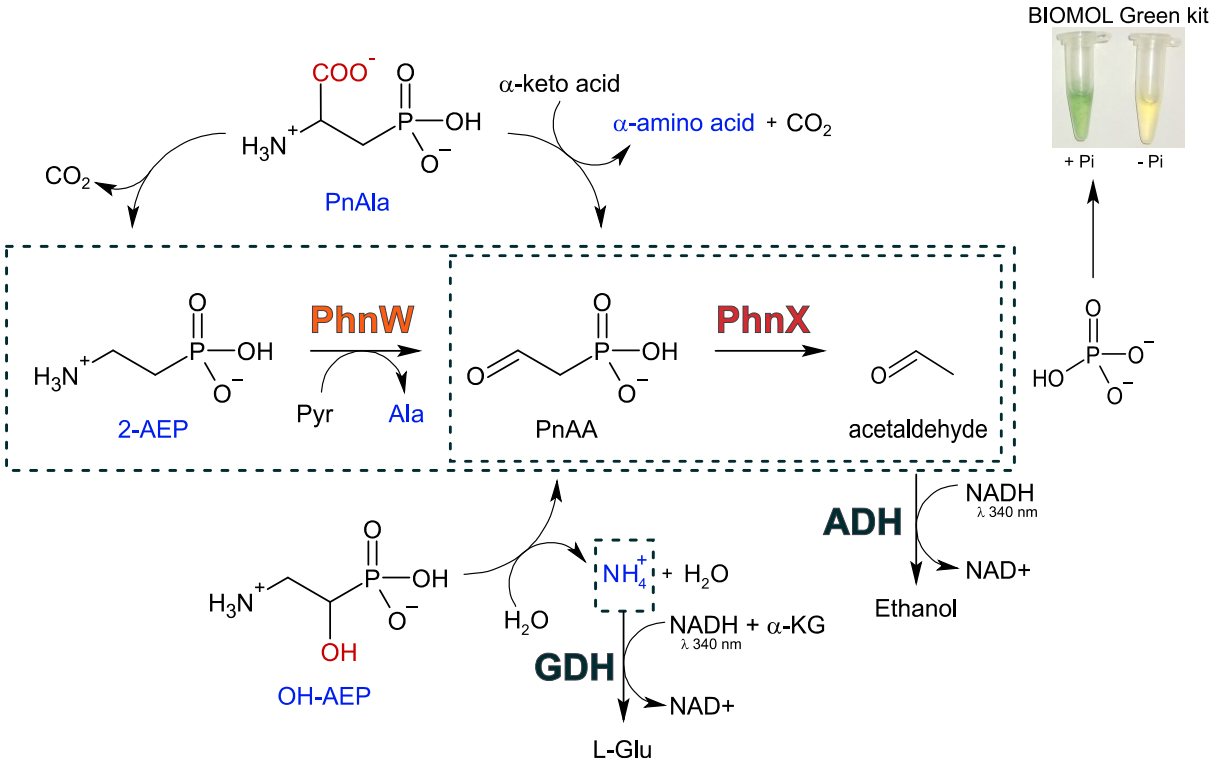


Figure 16. *Schematic representation of the techniques used to study PbfA activity*. The compounds that TLC can detect are in blue. The reactions that were coupled to alcohol dehydrogenase (ADH) and glutamate dehydrogenase (GDH) are highlighted in the dashed box. Phosphate released from the PhnX reaction was established using the BIOMOL green reagent.

3.3.1 Initial screening of phosphonates compounds by TLC

Once we obtained the three enzymes of the *Vibrio splendidus phnWX-pbfA* cluster, we proceeded with a rapid screening by TLC of different compounds to visually inspect the PhnW and PbfA activity. We initially tested the promising phosphonates, namely 2-AEP, but especially phosphonoalanine (PnAla) and hydroxyciliatine (OH-AEP) (the latter two in the racemic form), alone or in the presence of potential amino group acceptors (pyruvate, glyoxylate and α-ketoglutarate) (Fig. 17). This qualitative analysis confirmed the expected activity of PhnW towards 2-AEP and pyruvate (alanine formation) (Fig. 17 D). However, it showed that the enzyme can also use glyoxylate as an amino acceptor (glycine formation) (Fig. 17 D), and that it was also capable of a transamination reaction on *R,S*-OH-AEP, a close analog of 2-AEP, with both pyruvate and glyoxylate (Fig. 17 F). On the other hand, incubation of PbfA with these phosphonates failed to show the production of the amino acids expected from a transamination reaction (Fig. 17 A,B,C), ruling out the possibility that this enzyme could be a transaminase acting on 2-AEP (like PhnW), or that it could be a decarboxylase (similar to dialkylglycine decarboxylase) acting on PnAla. However, after incubation with racemic OH-AEP, a faint orange spot (similar to that obtained with ammonia) could be seen. Although *R,S*-OH-AEP was not all consumed, this spot appeared in all PbfA-containing samples, suggesting the possible elimination reaction previously hypothesized. These experiments were performed multiple times to rule out artifacts.

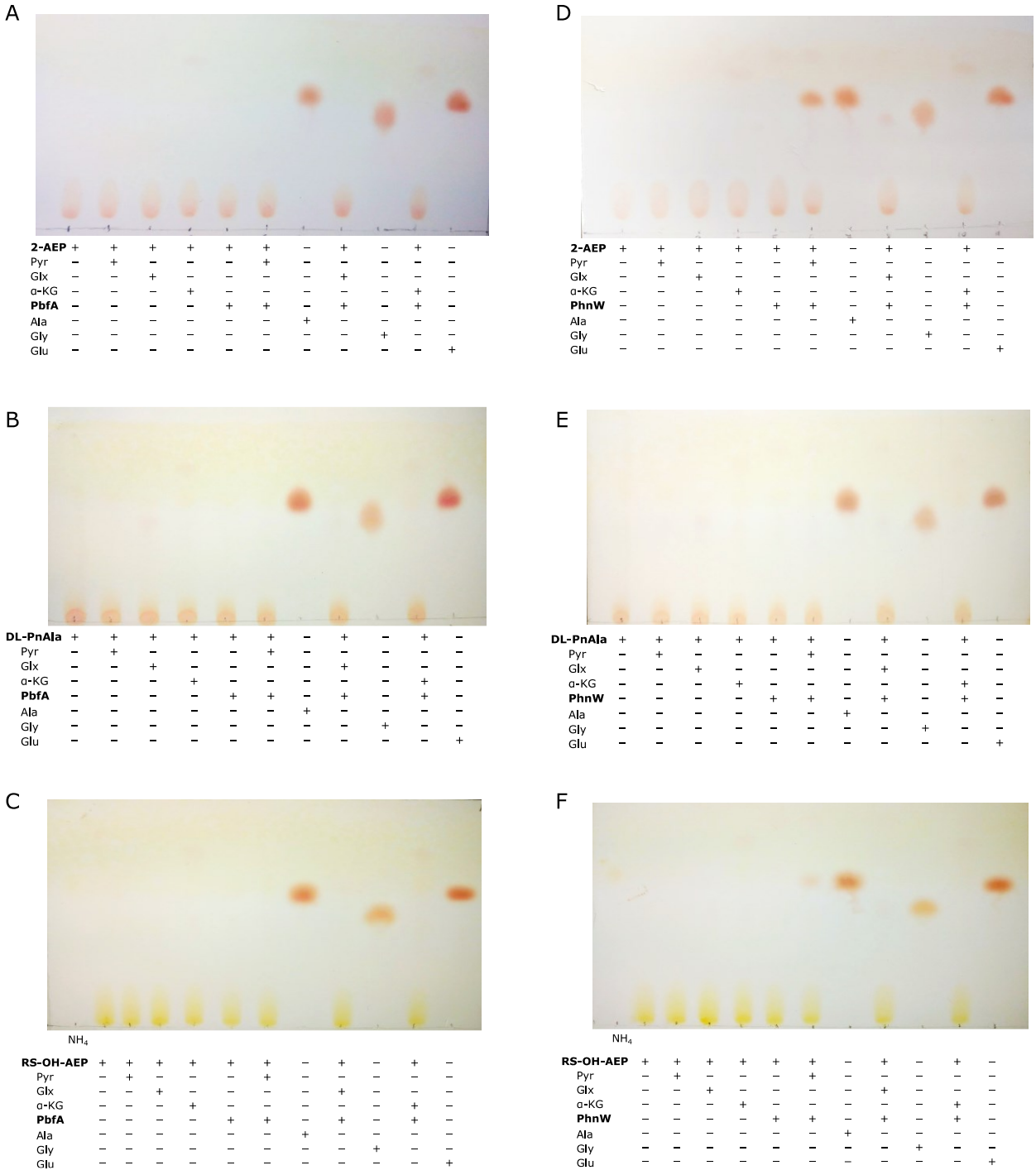


Figure 17. *The activity of PbfA and PhnW towards 2-AEP, PnAla and OH-AEP*. All reactions were triggered by enzyme addition and carried out at 37 °C for one hour in 150 µl of TEA-HCl buffer pH 8.0 and 5 mM $MgCl_2$. 5 mM alanine, 5 mM glycine, 5 mM glutamate and 10 mM ammonia were also loaded as control samples. (**A**), (**B**) and (**C**) Incubation of 1 µM PbfA and 10 mM 2-AEP, 20 mM D,L-PnAla, 20 mM *R,S*-OH-AEP, respectively, with and without the amino acceptors pyruvate, glyoxylate, or α-ketoglutarate. (**D**), (**E**) and (**F**) Incubation of 1 µM PhnW and 10 mM 2-AEP, 20 mM D,L-PnAla, 20 mM *R,S*-OH-AEP, respectively, with and without the amino acceptors pyruvate, glyoxylate, or α-

ketoglutarate. At the end of the incubation, the reactions were spotted on a TLC plate as described in Methods. Abbreviations: Pyr, pyruvate; Glx, glyoxylate; α-KG, α-ketoglutarate; Ala, alanine; Gly, glycine; Glu, glutamate.

3.3.2 PbfA acts neither as a transaminase nor a decarboxylase, but is a lyase acting on OH-AEP

The results obtained by TLC were confirmed through a spectrophotometric assay in the presence of PhnX and alcohol dehydrogenase (PhnX-ADH-coupled assay). The rationale of the assay is as follows: upon production, phosphonoacetaldehyde (PnAA) is converted by PhnX to acetaldehyde, a substrate of ADH that in turn reduces it to ethanol while oxidizing NADH to NAD+. Since NADH has a strong absorption at 340 nm (while NAD+ does not), the rate of PbfA reaction can be directly monitored by following the decrease in absorbance at that wavelength. In case the reaction product should be 2-AEP instead of PnAA, PhnW was added to the PhnX-ADH-coupled assay. However, by incubating PbfA in a reaction mixture containing 2-AEP, an amino group acceptor (pyruvate, glyoxylate or α-ketoglutarate), NADH, and the enzymes PhnX and ADH, no acetaldehyde formation was detected. Similarly, under the above conditions, there was no reaction by incubating PbfA with D,L-PnAla. The PbfA inactivity on PnALA was also confirmed with the addition of PhnW (and pyruvate) in the reaction mixture, which would detect the possible generation of 2-AEP. Unlike the results above, we observed PnAA production upon incubating PbfA with *R,S*-OH-AEP (without α-keto acids) in the PhnX-ADH-coupled assay. In addition to PnAA, ammonia was also released by the PbfA reaction as expected in a 1,2-elimination reaction. Ammonia production was seen from the initial TLC but it was also inferred by coupling the PbfA reaction with that of glutamate dehydrogenase (GDH-coupled assay), which catalyzes the production of L-glutamate from ammonia and α-ketoglutarate oxidizing NADH. Furthermore, we noticed that PbfA acted on only one of the two enantiomers in the racemic mixture, since when the reaction contained a concentration of *R,S*-OH-AEP lower than that of NADH, only 50% of the substrate was consumed (Fig. 18). We synthesized both the pure enantiomeric forms of OH-AEP to find out which of the two enantiomers was the PbfA substrate.

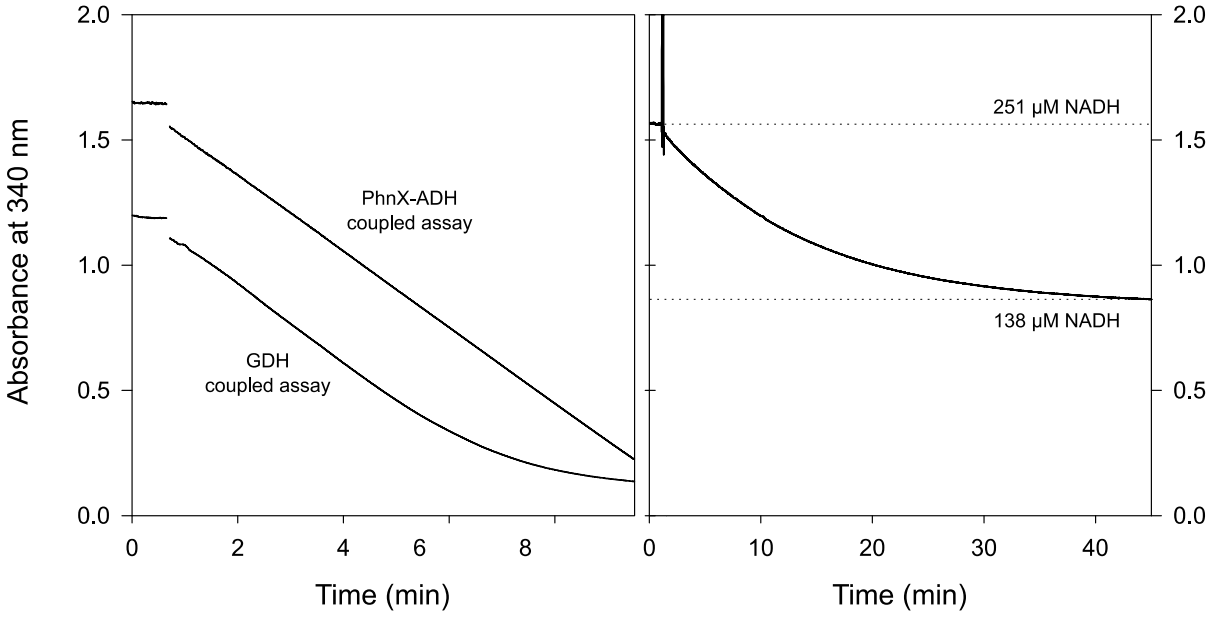


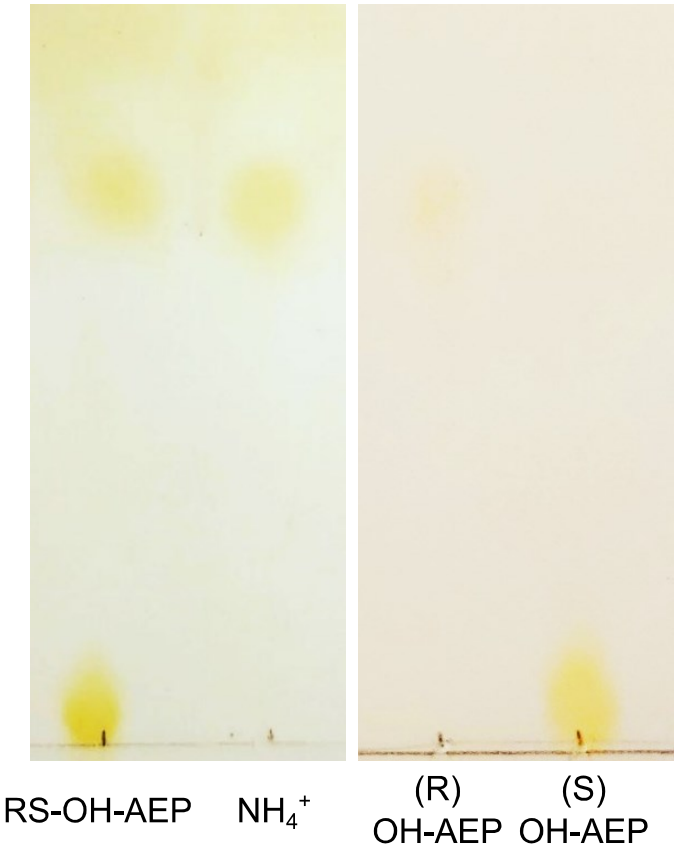
Figure 18. *PbfA reaction with R,S-OH-AEP*. The left panel shows the kinetics obtained with the PhnX-ADH-coupled assay and GDH-coupled assay used to monitor PnAA and ammonia production, respectively. Both the reactions were in 50 mM TEA-HCl pH 8.0 with 5 mM $MgCl_2$, 5 µM PLP, 0.5 mg/ml BSA, 10 mM *R,S*-OH-AEP, NADH and 2.9 µM PbfA. In addition, the PhnX-ADH-coupled assay contained 4.5 µM PhnX and 9 U ADH, instead, the GDH-coupled assay contained 1 mM α-KG and 6.7 U GDH. The sloped obtained through the two coupled assays were comparable, underlining that the process was rate-limiting by the reaction of PbfA. The right panel shows the PbfA reaction measured with the PhnX-ADH-coupled assay in the presence of 250 µM NADH and 200 µM racemic OH-AEP. From the graph, it is possible to estimate that PbfA consumes about 100 µM of NADH, suggesting that only one enantiomer is its substrate.

3.3.3 PbfA substrate specificity

Once the pure enantiomers (*R*-OH-AEP and *S*-OH-AEP) were synthesized, they were tested against PbfA. The TLC experiments and the spectrophotometric coupled assays showed that the enzyme efficiently consumes *R*-OH-AEP, while *S*-OH-AEP is completely unreactive (Fig. 19 A). We also tested some commercially available OH-AEP analogs without detecting reaction. For example, we tested D,L-isoserine (3-amino-2-hydroxypropane), which possesses a carboxylate instead of phosphonate group, and ethanolamine, bromoethanolamine, and cysteamine in which the phosphonate group is missing and the OH group is replaced by other good leaving groups, such as bromine or thiol. However, when the leaving group replaced the hydroxyl group and the phosphonate group was still present (such as in 1-fluoro-2-aminoethylphosphonate, F-AEP), production of PhnAA (PhnX-ADH-coupled assay) and ammonia (GDH-coupled assay) were detected. Furthermore, PbfA acted weakly on 1-hydroxy-2-

aminopropylphosphonic acid (OH-APP), releasing ammonia (GDH-coupled assay) and the corresponding phosphonate aldehyde, which in turn was found to be a substrate of PhnX (PhnX activity was confirmed with the BIOMOL green assay since its product is not an ADH substrate and no signal could be expected with the PhnX-ADH-coupled assay) (Fig. 19 B). As for OH-AEP, PbfA acted on the racemic mixtures of *R,S*-F-AEP and (*R,S*),(*S,R*)-OH-APP consuming only half of the substrate (GDH-coupled assay), again suggesting that it recognizes only one of the two enantiomers present in the reaction. In particular, based on the observations made with OH-AEP, it seems reasonable to assume that it acts on the *R*-F-AEP enantiomer, and on the (*R,S*)-OH-APP (this because no reaction was seen on the pure (*R,R*)-OH-APP, and on the racemic mixture (*R,R*),(*S,S*)-OH-APP). These data prove that proper phosphonate and hydroxyl group placement is critical for PbfA catalysis, indicating a very high substrate specificity towards *R*-OH-AEP. The stereospecific activity of PbfA is consistent with the fact that the oxygenase PhnZ also degrades only the *R*-OH-AEP isomer (although it is converted to glycine and phosphate) [8], and that this isomer is naturally occurring, since the OH-AEP isolated from the polysaccharides of *Acanthamoeba castellanii* has been confirmed as the *R*- enantiomer [15].

A



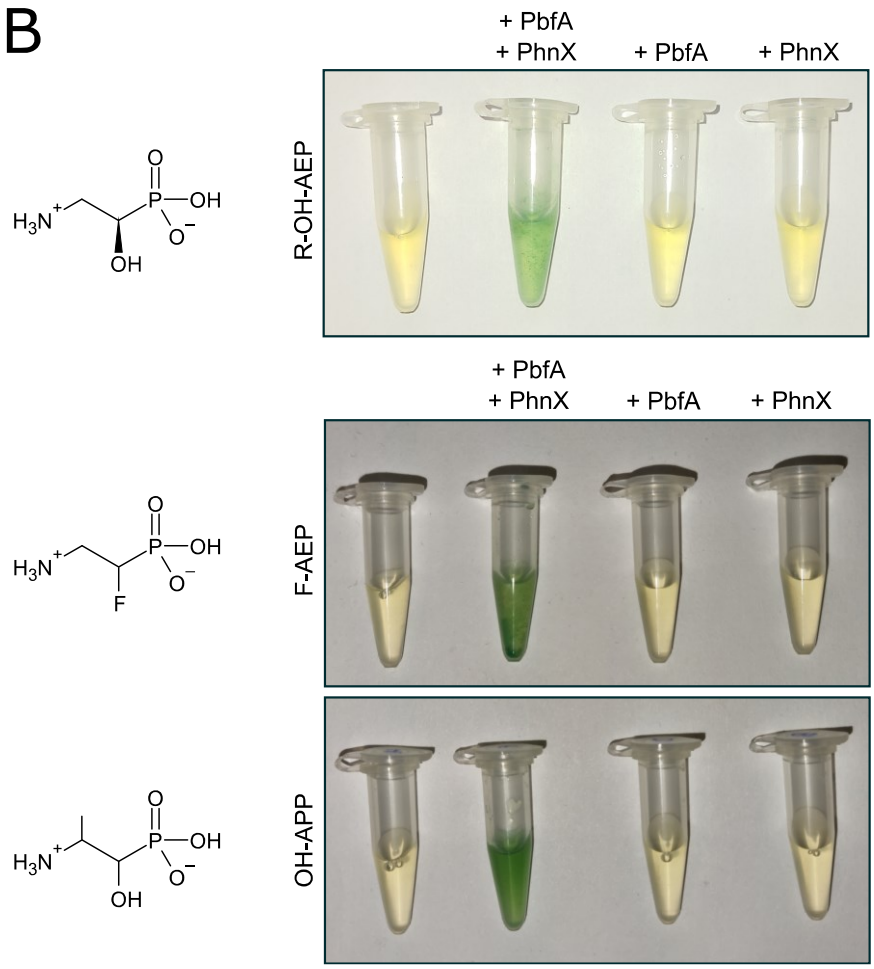
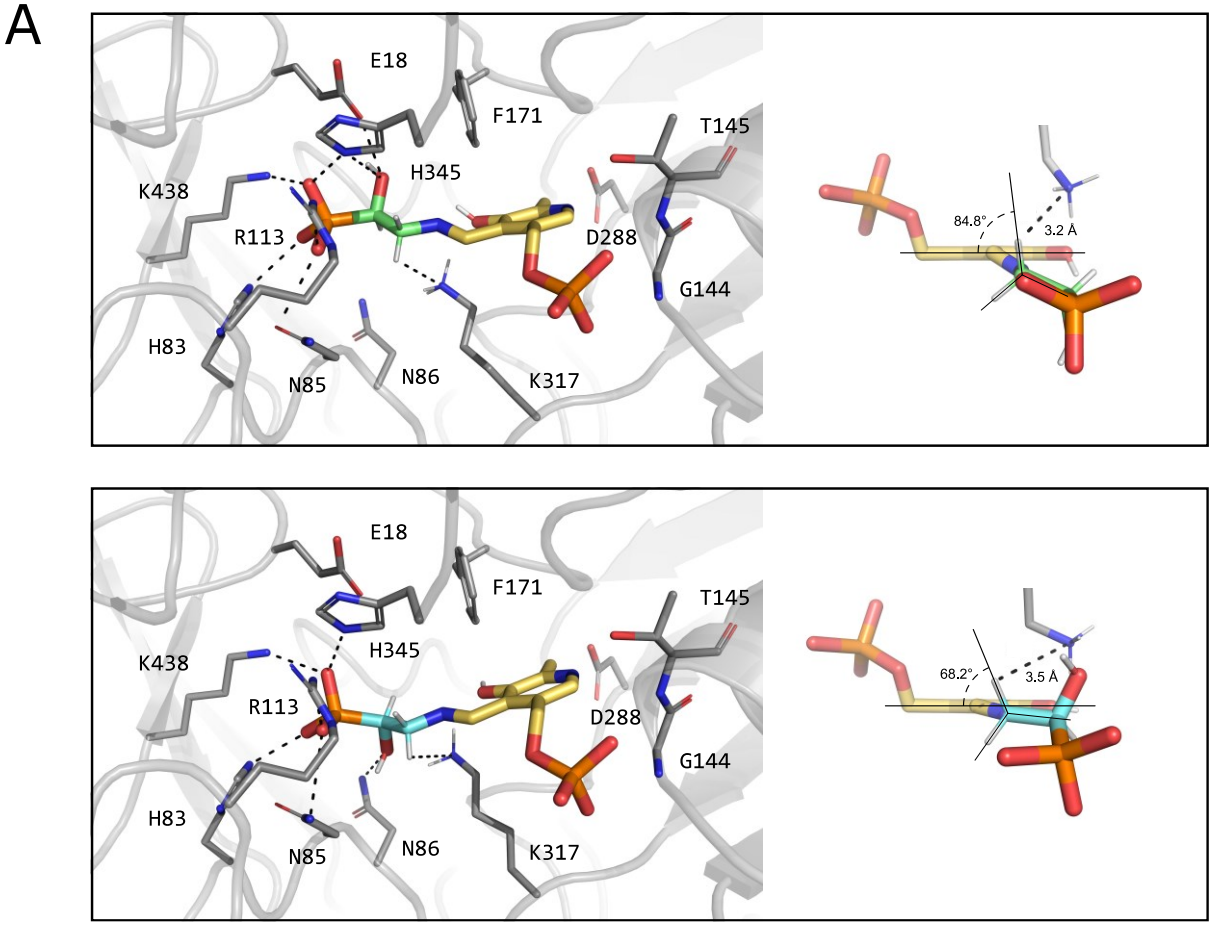


Figure 19. (**A**) *Enantioselective PbfA-catalyzed reaction on R,S-OH-AEP*. Reactions were carried out at 25 °C for one hour in 150 µl of TEA-HCl buffer pH 8.0, 5 mM $MgCl_2$, 5 µM PLP, 20 mM *R,S*-OH-AEP or 10 mM *R*-OH-AEP or 10 mM *S*-OH-AEP and 2 µM PbfA; 10 mM ammonia were also loaded as control samples. When the racemic mixture of OH-AEP was used, the substrate is not fully consumed and ammonia is released. Vice versa, when the two enantiomers were used separately, the *R*-enantiomer is degraded completely and ammonia is released, while the *S*-enantiomer is not consumed. (**B**) *PbfA and PhnX activity on two phosphonate analogs of OH-AEP*. Four reactions for each substrate analyzed were carried out at 25 °C for one hour in a final volume of 150 µl. They all contained TEA-HCl buffer pH 8.0, 5 mM $MgCl_2$, 5 µM PLP, 0.5 mg/ml BSA and 3 mM *R*-OH-AEP or 20 mM *R,S*-F-AEP or 20 mM (*R,S*),(*S,R*)-OH-APP. The first aliquot contained only the substrate as a control, in the second 2 µM of PbfA was added, in the third 2 µM of PbfA and 2 µM of PhnX, and only 2 µM of PhnX were present in the last one.

The kinetic investigation was complemented by computational analysis of the PbfA active site in the presence of both OH-AEP enantiomers (obtained as explained in Methods). As shown in figure 20, the proton on C2 of the *R*-OH-AEP, i.e., the one that should leave in the first reaction step (see intermediate EA1 in Fig. 20 B), is positioned close to the catalytic lysine (Lys317) and it is approximately perpendicular to the plane of the PLP, which would favor bond breakage as explained in the Dunathan's hypothesis [3]. The catalytic lysine could be involved in the abstraction of this proton, as suggested for the *O*-phosphoethanolamine phospho-lyase of *Arthrobacter aurescens* TC1 [16]. Furthermore, Glu18 and His345 are close enough to the hydroxyl group to suggest that they may be important for proper substrate positioning or to participate in the proton transfer promoting OH elimination (see intermediate Q in Fig. 20 B). Contrarily, in *S*-OH-AEP, the residue closest to the OH group is an Asn85, which cannot perform general acid catalysis and is not conserved in the multiple sequence alignment of PbfA homologs. Additionally, the proton on C2 results to be not perpendicular to the PLP plane. In a scenario where instead of the *R*-OH-AEP the (*R,R*)-OH-APP is placed in the active site, the methyl group on C2 would replace the proton that in the *R*-OH-AEP is close to Lys317, and this would explain the inactivity of the enzyme toward this substrate, implying that this may be the proton removed in the first step. Based on these results and in analogy to 1,2-elimination reactions catalyzed by other PLP-dependent enzymes, we also proposed the reaction mechanism of PbfA provisionally (Fig. 20 B).



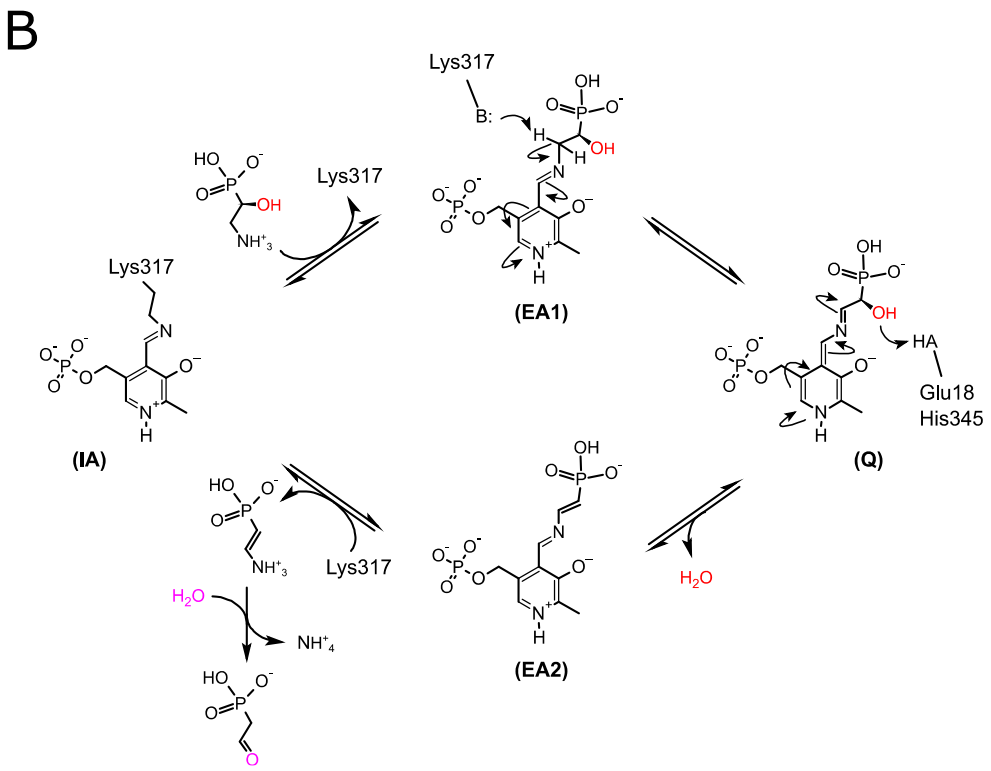


Figure 20. (**A**) *Representation of the active site of PbfA with R-OH-AEP (top panel) and S-OH-AEP (bottom panel)*. The phosphonate group of both enantiomers is recognized by a cluster of positively

charged residues, as in [16]. H345 could interact with the phosphonate group and be involved in proton transfer to the hydroxyl group of *R*-OH-AEP; E18 is also close to OH group. On the left of the first panel, the 3D structure of the external aldimine (PLP + *R*-OH-AEP) is shown. The angle between the proton interacting with lysine and the plane of PLP is 85°. In the lower panel, on the left are shown the expected contacts between the active site residues and the substrate, *S*-OH-AEP. On the right is shown the external aldimine (PLP + *S*-OH-AEP), in which the angle between the proton interacting with lysine and the PLP plane of 68° is highlighted. All the interactions highlighted with a black dashed line are < 3.5 Å. (**B**) *Reaction mechanism of PbfA*. The PbfA reaction mechanism was proposed in [14] and adapted from observations obtained by docking results. Figure adapted from Zangelmi et al., 2021 [14].

3.3.4 Kinetic characterization and proposed mechanism of PbfA reaction

Based on previous observations, we estimated the kinetic parameters of PbfA-catalyzed 1,2-elimination of *R*-OH-AEP using the PhnX-ADH-coupled assay since, albeit slightly more complex to set up, the GDH-coupled assay shows an initial lag phase hard to eliminate. At pH 8.0 and 25°C, we obtained kinetic parameters comparable to those reported for other PLP-dependent lyases [17], [18]: k_{cat} = 5.3 ± 0.3 s^{-1}, K_M = 0.43 ± 0.06 mM, and k_{cat}/K_M =12,300 ± 1,200 $M^{-1}s^{-1}$ (Fig. 21). Nevertheless, these parameters may be underestimated, as the specific activity of purified PbfA stocks tends to decrease significantly within hours of thawing.

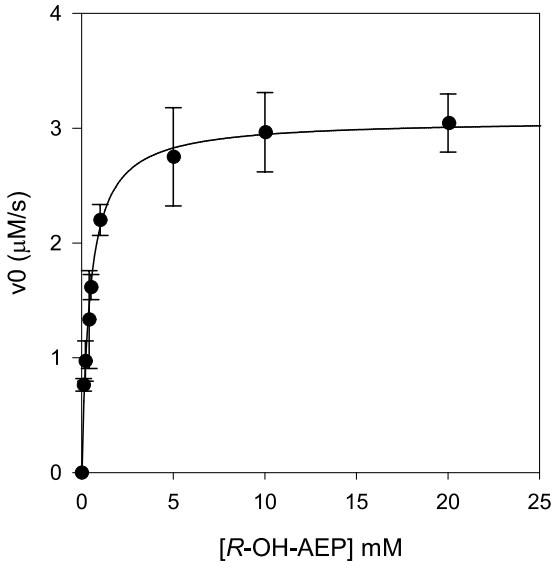
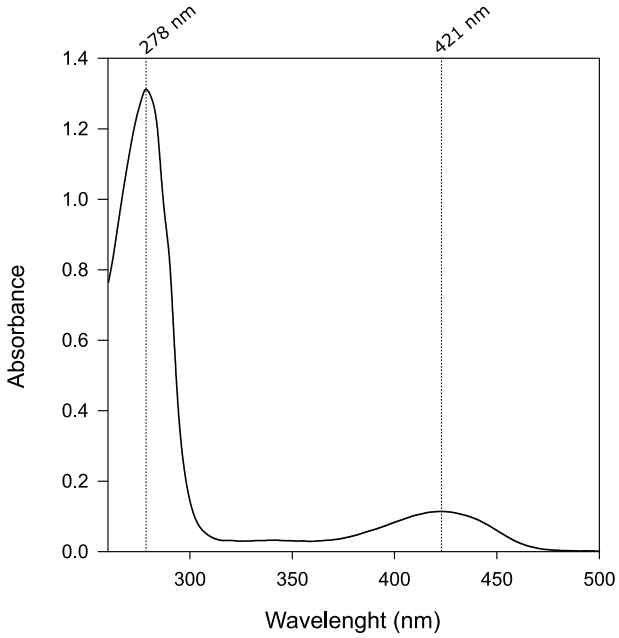


Figure 21. *Dependence of the initial velocity of PbfA reaction on R-OH-AEP (0.09-17 mM)*. The kinetic parameters were obtained using the PhnX-ADH-coupled assay at 25°C, pH 8.0. Each assays contain 50 mM TEA-HCl pH 8.0, 0.25 mM NADH, 5 µM PLP, 5 mM $MgCl_2$, 0.5 mg/ml BSA, 4 µM PhnX, 9 U ADH and were triggered by the adding of 0.58 µM PbfA. The triplicates of initial rates were fitted to the

Michaelis-Menten equation and the calculated k_{cat}, k_{cat}/K_M and K_M are reported as +/- standard error of the regression.

This loss of activity appears to be accompanied by a change in the absorption spectrum of the enzyme. The spectrum of the native enzyme in the visible region usually shows two maximum peaks at 278 nm and 421 nm and a minor one at 335 nm (Fig. 22 A). The last two peaks should correspond to the two tautomeric forms of PLP bound to the enzyme as internal aldimine, i.e., the tautomer with the protonated imino group (ketoenamine) and the tautomer with the deprotonated one (enolimine), respectively. Sometimes a minor and nearly buried peak at 325-340 nm has also been attributed to a bound PLP enzymatically inactive [19]. During enzyme inactivation, there is a decrease in the peak at 421 nm, followed by an increase at 335 nm. The absorption spectrum of PbfA, even though it changes slowly over time, was also examined during the reaction with *R*-OH-AEP to detect possible changes in absorption due to substrate interaction or the formation of reaction intermediates. Immediately after the addition of *R*-OH-AEP, a shoulder appears at 435 nm concomitant with an increase of the signal at 280 nm. When the shoulder disappears, the 280 nm signal stops increasing (Fig. 22 B). Unlike the temporary shoulder at 435 nm that could be associated with the formation of a reaction intermediate (possibly the external aldimine, which is usually accompanied by a red shift of the main PLP band) [20], the increase at 280 nm persists over time, suggesting that it could be due to the formation of the PnAA product. Although this assumption was not verified with a spectrum of authentic PnAA, it seems to be supported by the fact that the addition of PhnX in the solution decreases the signal at 280 nm.

A



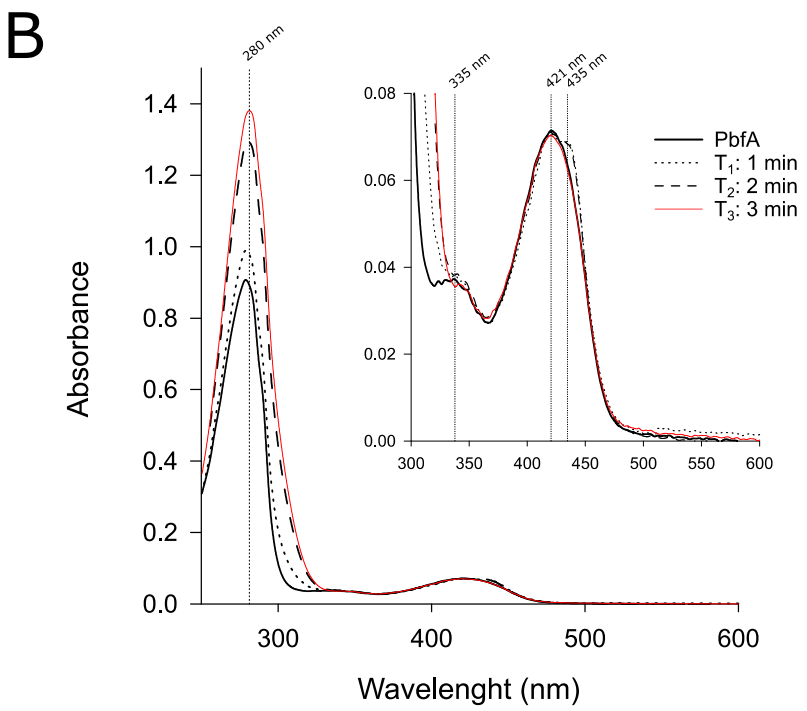
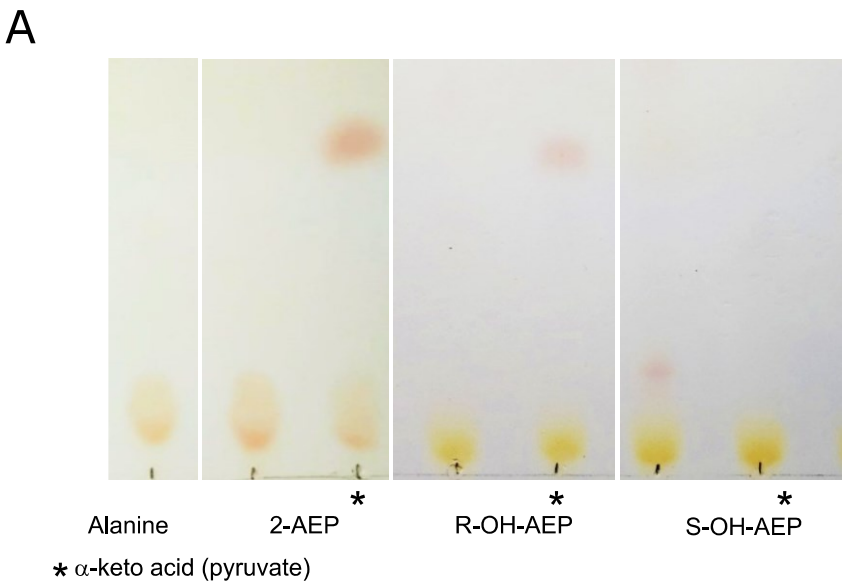


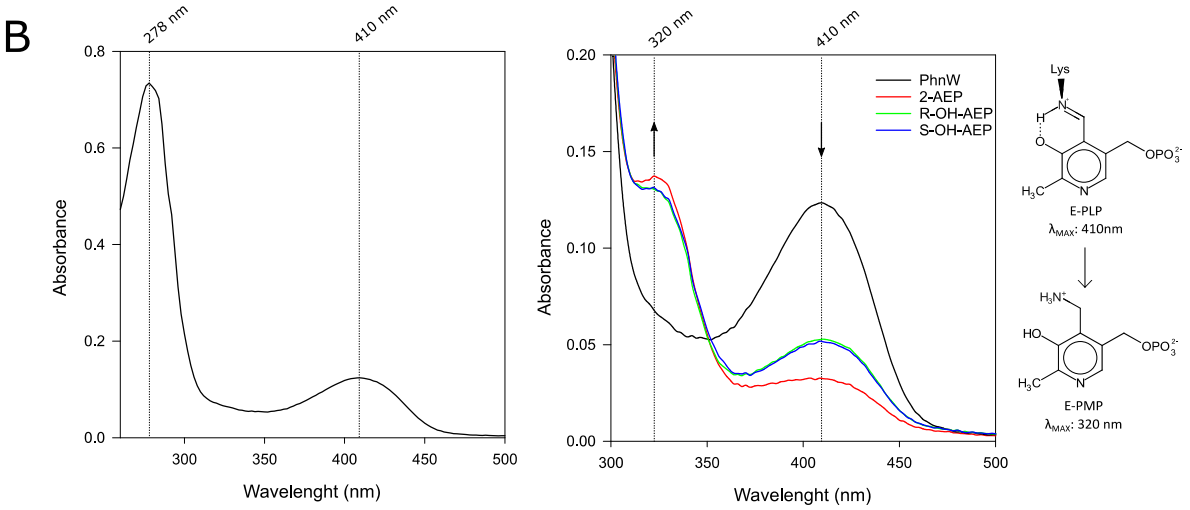
Figure 22. *UV-vis absorption spectrum of PbfA before and after R-OH-AEP addition*. (**A**) Absorption spectrum of 28 µM PbfA in a solution containing 50 mM TEA-HCl buffer pH 8.0, 100 mM KCl, 5 µM PLP, 1 mM DTT. The spectrum was corrected for solution contribution. (**B**) Absorption spectrum of 19 µM PbfA before and after adding 6 mM *R*-OH-AEP in a solution containing 50 mM TEA-HCl buffer pH 8.0, 100 mM KCl, 5 µM PLP, 1 mM DTT. A spectrum was collected every 30 seconds; the figure shows a spectrum every minute until the signal stabilization at 280 nm and the concomitant disappearance of the shoulder at 435 nm. The spectra were corrected for solution contribution.

2.3.4 The PhnW-PhnX pathway does not efficiently consume OH-AEP

As mentioned in Section 3.3.1 of this Chapter, we observed during the TLC screening a side reaction of PhnW toward *R,S*-OH-AEP (Fig. 23 F). After obtaining both OH-AEP enantiomers in the pure form, we tested them separately to reveal if PhnW could discriminate between the two. The results show that, albeit with lower efficiency compared to the reaction with 2-AEP, PhnW can catalyze a transamination reaction with both enantiomers. Not only alanine is detectable by TLC after incubation of PhnW with OH-AEP and pyruvate (Fig. 23 A), but after the addition of OH-AEP a peak at 320 nm corresponding to pyridoxamine 5'-phosphate (PMP), i.e., the typical intermediate of the first half-transamination reaction, appears in the absorption spectrum of the enzyme (Fig. 23 B). Interestingly, the reaction product of PhnW with OH-AEP, presumably "hydroxy-phosphonoacetaldehyde" (OH-PnAA), is not a substrate of PhnX; in fact, neither the PhnX-ADH-coupled assay nor the BIOMOL green

assay detected activity. However, the BIOMOL green assay shows that a small amount of Pi is released from the reaction of PhnW with OH-AEP and pyruvate, even in the absence of PhnX (Fig. 23 C, left panel). When the PhnW reaction was investigated by monitoring in parallel the release of both alanine (AlaDH-coupled assay) and phosphate (BIOMOL green assay), it was evident that the release of phosphate occurs more slowly and independently of that of alanine, suggesting a spontaneous degradation of the putative OH-PnAA product (Fig. 23 C, right panel).





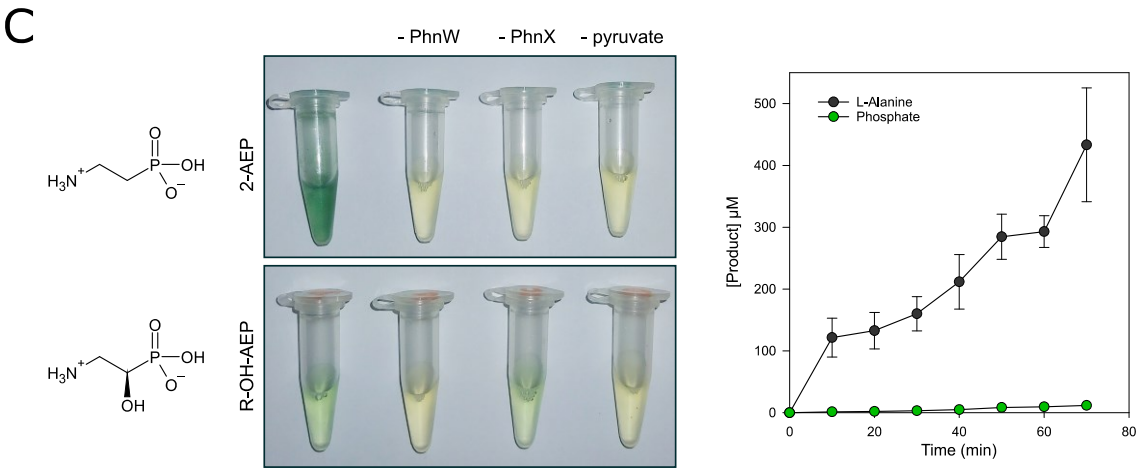
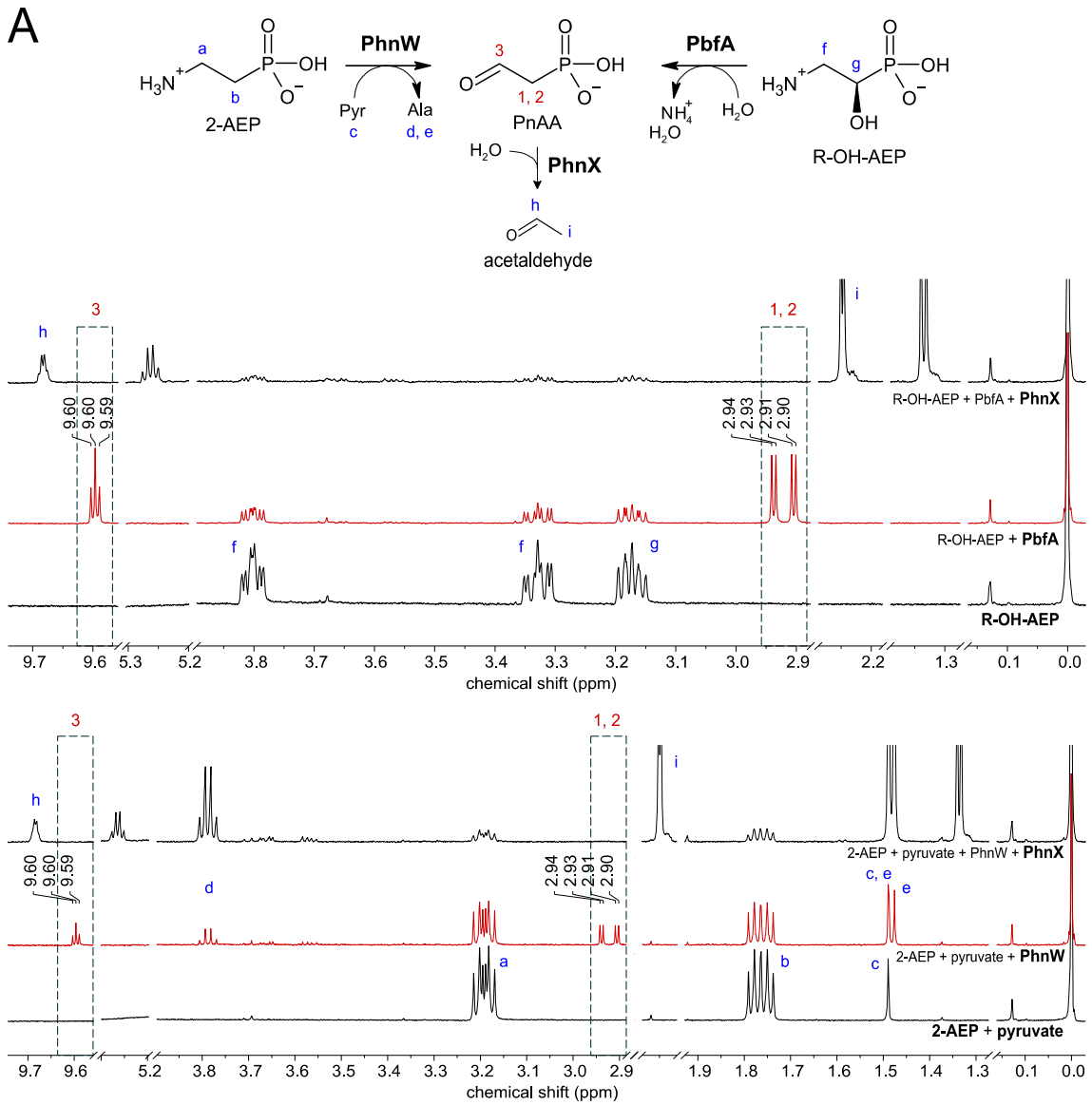


Figure 23. *Overview of PhnW activity against OH-AEP*. (**A**) Reactions for the TLC analysis were carried out at 25 °C for one hour in 150 µl of TEA-HCl buffer pH 8.0, 5 mM $MgCl_2$, 5 µM PLP, 6 mM 2-AEP or 6 mM *R*-OH-AEP or 6 mM *S*-OH-AEP and 2 µM PhnW; 5 mM of L-alanine was loaded as control. (**B**) Absorption spectra of 15 µM PhnW before (black line) and after adding 6 mM of 2-AEP, *R*-OH-AEP or *S*-OH-AEP in a solution containing 50 mM TEA-HCl buffer pH 8.0, 100 mM KCl, 5 µM PLP, 1 mM DTT. The spectra were corrected for solution contribution and recorded after 9 minutes. Immediately after substrate addition, the peak at 410 nm (internal aldimine signal) gradually decreased while the peak at 320 nm (that of PMP) increased. (**C**) Four reactions for each substrate analyzed (2-AEP and *R*-OH-AEP) were carried out at 25 °C for one hour in a final volume of 150 µl. The first aliquot on the left contained TEA-HCl buffer pH 8.0, 5 mM $MgCl_2$, 5 µM PLP, 100 mM KCl, 1 mM DTT, 1 mM pyruvate, 1 mM 2-AEP or *R*-OH-AEP, 2 µM PhnX and 2 µM PhnW. As a control, the other three aliquots do not contain PhnW, PhnX and pyruvate, respectively. A slight greenish color indicates weak Pi release in the first and third aliquots containing *R*-OH-AEP. The right panel shows the lack of correlation between alanine and Pi production. The experiment is described in Methods.

To further confirm the nature of the compounds generated by the non-specific activity of PhnW and the highly-specific one of PbfA on *R*-OH-AEP, we followed the enzymatic reactions at ^{1}H NMR spectroscopy. The spectra recorded confirm our previous findings: (i) the 1,2-elimination catalyzed by PbfA on *R*-OH-AEP generates phosphonoacetaldehyde; indeed, the signals obtained were consistent with those reported by Lacoste et al. [21] and from those observed in the reaction of PhnW with 2-AEP (Fig. 24 A); (ii) in the reaction of PhnW with *R*-OH-AEP and pyruvate, peaks appear corresponding to alanine and one typical of aldehyde protons (chemical shift 9.5-10.5 ppm). Although we do not have the pure OH-PnAA compound to test as control, the doublet at 9.73 ppm (*J*= 1.6 Hz) could be consistent with this assumption (Fig. 24 B).


A
2-AEP
PhnW
Pyr
Ala
PnAA
PbfA
R-OH-AEP
PhnX
acetaldehyde
R-OH-AEP + PbfA + PhnX
R-OH-AEP + PbfA
R-OH-AEP
chemical shift (ppm)
2-AEP + pyruvate + PhnW + PhnX
2-AEP + pyruvate + PhnW
2-AEP + pyruvate
chemical shift (ppm)

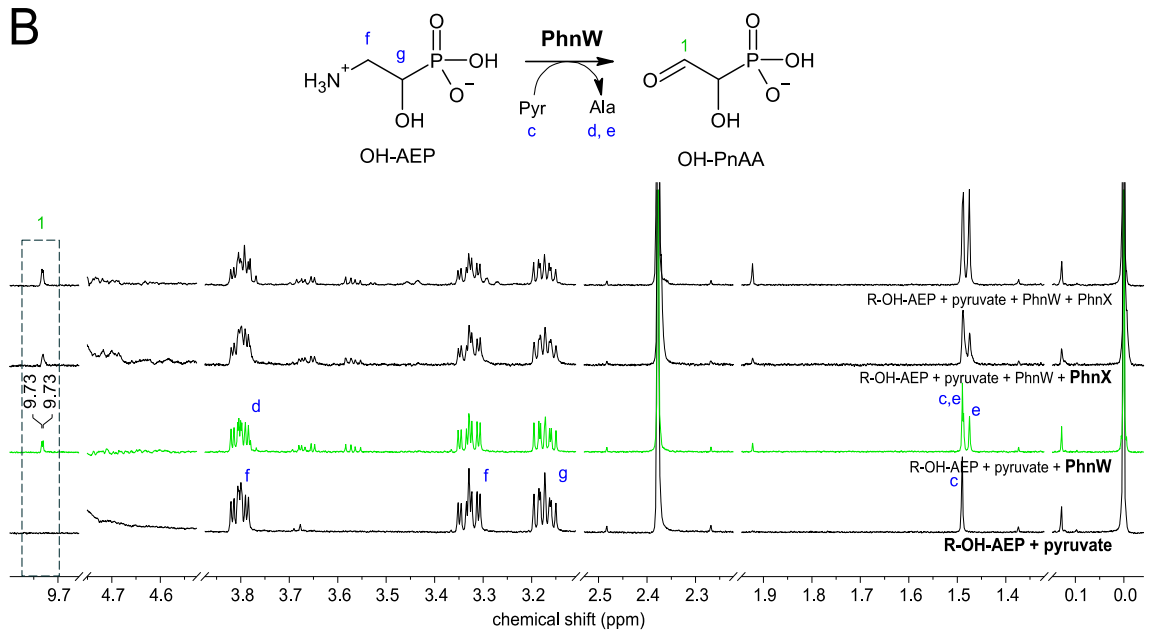

B
OH-AEP
PhnW
OH-PnAA
R-OH-AEP + pyruvate + PhnW + PhnX
R-OH-AEP + pyruvate + PhnW + PhnX
R-OH-AEP + pyruvate + PhnW
R-OH-AEP + pyruvate
chemical shift (ppm)

Figure 24. *^{1}H NMR spectra of PbfA and PhnW reactions towards R-OH-AEP*. (**A**) Comparison between 1,2-elimination of *R*-OH-AEP by PbfA (upper panel) and transamination of 2-AEP by PhnW (bottom panel). A scheme of the reactions analyzed is shown above the two panels. The blue letters (a to h) mark the signals of each group of hydrogen atoms of the structures. The red numbers and dashed boxes indicate the proton signals of PnAA; in detail, the signal of the aldehydic proton (3) is a triplet at 9.60 ppm (*J*= 4.25 Hz), and the protons on the C2 (1, 2) are a doublet of doublets at 2.90-2.94 ppm (*J*= 19.90 and 4.25 Hz). Red spectra were collected after the addition of the PbfA or PhnW enzymes. By adding PhnX, the PnAA peaks disappear, and acetaldehyde peaks, in equilibrium with its hydrated form (CH_3, dd at 1.33–1.34 ppm; CH, m at 5.25–5.28 ppm), appear. (**B**) Transamination reaction of PhnW with *R*-OH-AEP. The green spectrum was collected after adding the PhnW enzyme, followed by the formation of a doublet at 9.73 ppm (*J*= 1.6 Hz), maybe attributable to the hypothetical OH-PnAA product. After the PhnX addition, this signal does not disappear, indicating the inactivity of the enzyme on this compound. After an o/n incubation of the reaction (the last spectrum above), the alanine signals increase non-proportionally to the increase of the presumed OH-PnAA peak.

Given the inability to obtain the catalytic parameters of the PhnW transamination reaction on *R*-OH-AEP, we obtained those with the physiological substrate 2-AEP, using, as for PbfA, the PhnX-ADH-coupled assay. At pH 8.0 and 25°C, we achieved values comparable to those reported for PhnW from *Salmonella enterica* [22]: k_{cat} = 15.5 ± 0.3 s^{-1}, 2-AEP K_M = 3.16 ± 0.15 mM and k_{cat}/K_M =4,900 ± 300 $M^{-1}s^{-1}$, pyruvate K_M = 0.58 ± 0.05 mM and k_{cat}/K_M =26,700 ± 2,800 $M^{-1}s^{-1}$ (Fig. 25).

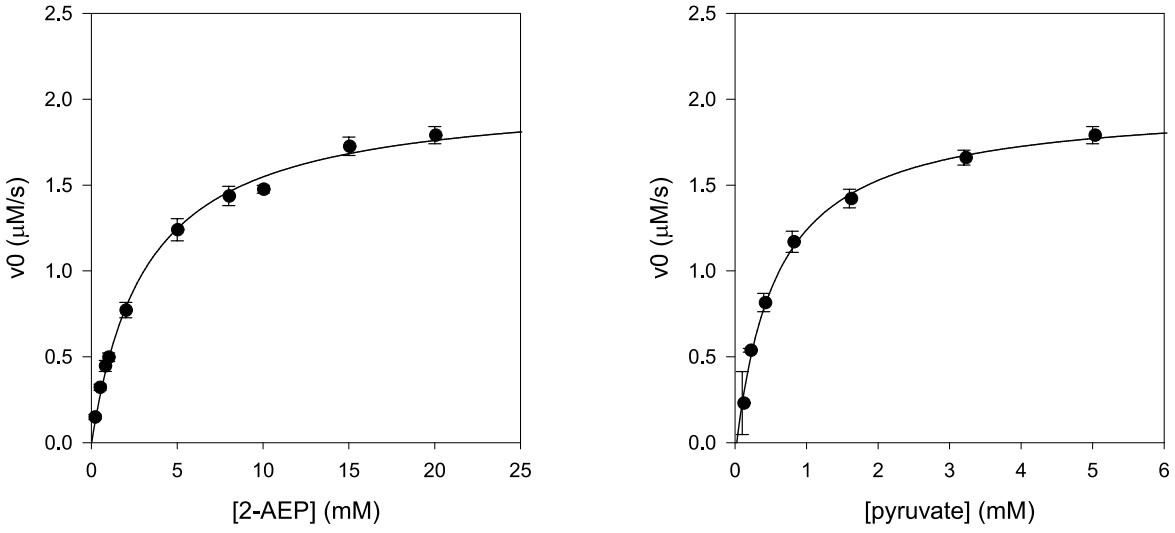


Figure 25. *Dependence of the initial velocity of PhnW reaction on 2-AEP and pyruvate concentration, respectively*. The kinetic parameters were obtained using the PhnX-ADH-coupled assay at 25°C, pH 8.0. Each assays contain 50 mM TEA-HCl pH 8.0, 0.25 mM NADH, 5 µM PLP, 5 mM $MgCl_2$, 0.5 mg/ml BSA, 1 mM DTT, 4 µM PhnX, 9 U ADH and were triggered by the adding of 0.13 µM PhnW. 2-AEP k_{cat}

and K_M were measured at various concentrations of 2-AEP (0.2-20 mM) in the presence of a saturating concentration of pyruvate (5 mM). Pyruvate k_{cat} and K_M were measured at various concentrations of pyruvate (0.1-5 mM) in the presence of a saturating concentration of 2-AEP (20 mM). The triplicates of initial rates were fitted to the Michaelis-Menten equation and the calculated k_{cat}, k_{cat}/K_M and K_M are reported as +/- standard error of the regression.

3.4 Conclusions

3.4.1 PbfA expands the scope and utility of the PhnW-PhnX pathway

In this chapter, we reported the several analyses that allowed us, following the initial postulated hypotheses, to characterize the function of PbfA, i.e., a PLP-dependent enzyme automatically annotated in public databases as 4-aminobutyrate aminotransferase whose gene often recurs in *phnWX* and *phnWAY* clusters involved in the degradation of 2-AEP. PbfA from *Vibrio splendidus* catalyzes a previously undescribed elimination reaction on 1-hydroxy-2-aminoethylphosphonate (OH-AEP), generating ammonia and phosphonoacetaldehyde (PnAA). Furthermore, PbfA is very substrate-specific since it is highly stereospecific for the *R*- enantiomer of OH-AEP and fails to act on natural analogs. The use of synthetic analogs and a docking prediction allowed us to inspect the arrangement of the substrate in the active site and speculate on the reasons for its stereospecificity. In contrast, *R*-OH-AEP, despite being structurally close to 2-AEP, cannot be adequately processed by the PhnW-PhnX pair. In fact, PhnW can indistinctly transaminate both enantiomers of OH-AEP, processing them into a compound that PhnX does not recognize and, therefore, is apparently useless to the organism. Furthermore, the uptake of OH-AEP in some organisms could have a toxic effect, as reported for *E. coli* [23]. Thus, the existence of a dedicated degradative enzyme could benefit the organism that possesses it, not only by scavenging this compound but also employing it as a valuable food source, allowing *R*-OH-AEP to be "broken-down" into a carbon, nitrogen, and phosphate supply, by funneling it into a pre-existing pathway. Although there is limited information on the abundance of OH-AEP in nature and it is primarily recognized as a reaction intermediate of the PhnY*-PhnZ pathway [24] in phosphonate catabolism, the frequency with which PbfA is found in bacterial genomes suggests that its diffusion may be greater than previously recognized. Based on these observations, the physiological role of PbfA seems to be strictly dedicated to the *R*-OH-AEP degradation, so we proposed the name 1-hydroxy-2-aminoethylphosphonate ammonia-lyase [14].

3.5 Supporting information

Table S1. *Information on clustering and interaction energies of the docking outputs*. Highlighted in light green and light blue are the binding modes of *R*-OH-AEP and *S*-OH-AEP, respectively, in the PbfA active site, which were analyzed with PyMol.

Mode	Affinity (kcal/mol)	RMSD	Size	RSMD STDV	Energy STDV	Best Run
1	-9.4	0	20	0.4	0.5	3
2	-7.6	2.9	15	0.5	0.5	41
3	-7.4	3.7	1	NA	NA	21
4	-7.4	3.2	1	NA	NA	1
5	-6.9	6.8	6	0.7	0.4	28
6	-6.7	2.4	3	0.5	0	48
7	-5.8	3.8	1	NA	NA	39
8	-5.7	6.8	1	NA	NA	30
9	-5.6	6.5	1	NA	NA	42
10	-5.4	3.5	1	NA	NA	15

Mode	Affinity (kcal/mol)	RMSD	Size	RSMD STDV	Energy STDV	Best Run
1	-9.5	0	22	0.4	0.7	6
2	-8.2	4	5	0.6	0.3	25
3	-7.6	2.5	6	0.8	0.6	5
4	-7.3	2.7	3	0.8	0.2	4
5	-7	3.9	3	0.5	0	41
6	-7	2.1	3	0.5	0.1	50
7	-5.7	3.4	1	NA	NA	17
8	-5.6	7.4	3	0.4	0.1	44
9	-5.3	4.4	2	0.5	0	42
10	-4.2	3.7	1	NA	NA	34
11	-4.1	4.5	1	NA	NA	12

3.6 References

[1] R. Percudani and A. Peracchi, "A genomic overview of pyridoxal-phosphate-dependent enzymes," *EMBO reports*, vol. 4, no. 9, pp. 850–854, Sep. 2003, doi: 10.1038/sj.embor.embor914.

[2] D. Palm, H. W. Klein, R. Schinzel, M. Buehner, and E. J. M. Helmreich, "The role of pyridoxal 5'-phosphate in glycogen phosphorylase catalysis," *Biochemistry*, vol. 29, no. 5, pp. 1099–1107, Feb. 1990, doi: 10.1021/bi00457a001.

[3] H. C. Dunathan, "Conformation and reaction specificity in pyridoxal phosphate enzymes.," *Proceedings of the National Academy of Sciences*, vol. 55, no. 4, pp. 712–716, Apr. 1966, doi: 10.1073/pnas.55.4.712.

[4] M. D. Toney, "Controlling reaction specificity in pyridoxal phosphate enzymes," *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, vol. 1814, no. 11, pp. 1407–1418, Nov. 2011, doi: 10.1016/j.bbapap.2011.05.019.

[5] R. Percudani and A. Peracchi, "The B6 database: a tool for the description and classification of vitamin B6-dependent enzymatic activities and of the corresponding protein families," *BMC Bioinformatics*, vol. 10, no. 1, p. 273, Dec. 2009, doi: 10.1186/1471-2105-10-273.

[6] N. v. Grishin, M. A. Phillips, and E. J. Goldsmith, "Modeling of the spatial structure of eukaryotic ornithine decarboxylases," *Protein Science*, vol. 4, no. 7, pp. 1291–1304, Jul. 1995, doi: 10.1002/pro.5560040705.

[7] J. Srivastava, P. v Balaji, and J. Srivastava Lab, "Clues to reaction specificity in PLP-dependent fold type I aminotransferases of monosaccharide biosynthesis," *Proteins*, Jan. 2021, doi: 10.1002/prot.26305.

[8] L. M. van Staalduinen *et al.*, "Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates," *Proceedings of the National Academy of Sciences*, vol. 111, no. 14, pp. 5171–5176, Apr. 2014, doi: 10.1073/pnas.1320039111.

[9] S. Donini *et al.*, "Recombinant production of eight human cytosolic aminotransferases and assessment of their potential involvement in glyoxylate metabolism," *Biochemical Journal*, vol. 422, no. 2, pp. 265–272, 2009, doi: 10.1042/BJ20090748.

[10] J. Jumper *et al.*, "Highly accurate protein structure prediction with AlphaFold," *Nature*, vol. 596, no. 7873, pp. 583–589, Aug. 2021, doi: 10.1038/s41586-021-03819-2.

[11] P. A. Ravindranath, S. Forli, D. S. Goodsell, A. J. Olson, and M. F. Sanner, "AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility," *PLOS Computational Biology*, vol. 11, no. 12, p. e1004586, Dec. 2015, doi: 10.1371/journal.pcbi.1004586.

[12] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison, "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform," *Journal of Cheminformatics*, vol. 4, no. 1, p. 17, Dec. 2012, doi: 10.1186/1758-2946-4-17.

[13] D. H. Williamson, "l-Alanine Determination with Alanine Dehydrogenase," in *Methods of Enzymatic Analysis*, Elsevier, 1974, pp. 1679–1685. doi: 10.1016/B978-0-12-091304-6.50012-4.

[14] E. Zangelmi, T. Stanković, M. Malatesta, D. Acquotti, K. Pallitsch, and A. Peracchi, "Discovery of a New, Recurrent Enzyme in Bacterial Phosphonate Degradation: (R)-1-Hydroxy-2-

aminoethylphosphonate Ammonia-lyase," *Biochemistry*, vol. 60, no. 15, pp. 1214–1225, Apr. 2021, doi: 10.1021/acs.biochem.1c00092.

[15] F. Hammerschmidt and H. Völlenkle, "Absolute Konfiguration der (2-Amino-1-hydroxyethyl)phosphonsäure ausAcanthamoeba castellanii (Neff) - Darstellung der Phosphonsäure-Analoga von (+)- und (-)-Serin," *Liebigs Annalen der Chemie*, vol. 1989, no. 6, pp. 577–583, Jun. 1989, doi: 10.1002/jlac.1989198901101.

[16] A. Cuetos *et al.*, "Structural Basis for Phospholyase Activity of a Class III Transaminase Homologue," *ChemBioChem*, vol. 17, no. 24, pp. 2308–2311, Dec. 2016, doi: 10.1002/cbic.201600482.

[17] A. Husain, G. Jeelani, D. Sato, V. Ali, and T. Nozaki, "Characterization of two isotypes of l-threonine dehydratase from Entamoeba histolytica," *Molecular and Biochemical Parasitology*, vol. 170, no. 2, pp. 100–104, Apr. 2010, doi: 10.1016/j.molbiopara.2009.11.004.

[18] M. Veiga-da-Cunha, F. Hadi, T. Balligand, V. Stroobant, and E. van Schaftingen, "Molecular Identification of Hydroxylysine Kinase and of Ammoniophospholyases Acting on 5-Phosphohydroxy-l-lysine and Phosphoethanolamine," *Journal of Biological Chemistry*, vol. 287, no. 10, pp. 7246–7255, Mar. 2012, doi: 10.1074/jbc.M111.323485.

[19] L. Davis and D. E. Metzler, "2 Pyridoxal-Linked Elimination and Replacement Reactions," *Biochimie*, vol. 58, no. 1–2, pp. 5-17, Mar. 1976, doi: 10.1016/S1874-6047(08)60446-1.

[20] D. Schiroli, L. Ronda, and A. Peracchi, "Kinetic characterization of the human O-phosphoethanolamine phospho-lyase reveals unconventional features of this specialized pyridoxal phosphate-dependent lyase," *FEBS Journal*, vol. 282, no. 1, pp. 183–199, Jan. 2015, doi: 10.1111/febs.13122.

[21] A.M. Lacoste, C. Dumora, L. Balas, F. Hammerschmidt, and J. Vercauteren, "Stereochemistry of the reaction catalysed by 2-aminoethylphosphonate aminotransferase. A 1H-NMR study," *European Journal of Biochemistry*, vol. 215, no. 3, pp. 841–844, Aug. 1993, doi: 10.1111/j.1432-1033.1993.tb18100.x.

[22] A. D. Kim, A. S. Baker, D. Dunaway-Mariano, W. W. Metcalf, B. L. Wanner, and B. M. Martin, "The 2-Aminoethylphosphonate-Specific Transaminase of the 2-Aminoethylphosphonate Degradation Pathway," *Journal of Bacteriology*, vol. 184, no. 15, pp. 4134–4140, Aug. 2002, doi: 10.1128/JB.184.15.4134-4140.2002.

[23] J. P. Cioni *et al.*, "Cyanohydrin Phosphonate Natural Product from Streptomyces regensis," *Journal of Natural Products*, vol. 77, no. 2, pp. 243–249, Feb. 2014, doi: 10.1021/np400722m.

[24] F. R. McSorley, P. B. Wyatt, A. Martinez, E. F. DeLong, B. Hove-Jensen, and D. L. Zechel, "PhnY and PhnZ Comprise a New Oxidative Pathway for Enzymatic Cleavage of a Carbon–Phosphorus Bond," *Journal of the American Chemical Society*, vol. 134, no. 20, pp. 8364–8367, May 2012, doi: 10.1021/ja302072f.

Chapter 4

4.1 FAD-dependent enzymes

The FMN and FAD cofactors are derivatives of vitamin B2 and compete for the versatility master title with pyridoxal 5'-phosphate (PLP). Indeed, although most flavoenzymes are oxidoreductases, flavin activities also belong to all but two classes defined by the Enzyme Commission [1]. These enzymes can be grouped according to their structure into 23 different clans, many more than the seven fold-types of PLP-dependent enzymes. The comparison of these clans shows no correlation between the folding architecture and function, although some features are conserved among proteins that employ the same reaction mechanism [2]. In flavin-dependent enzymes, the functional and mechanistic complexity is certainly increased by the addition to the cofactor's redox-active isoalloxazine ring-system of auxiliary redox groups, such as iron-sulfur clusters and heme. The biochemical utility of the FAD cofactor is based primarily on its ability for electron transfer reactions and dioxygen activation. It can exchange one or two electrons with the substrate at a time, thus adopting different redox (oxidized, quinone; one-elector reduced, semiquinone; two-elector reduced, hydroquinone) and protonation states (Fig. 25). Furthermore, the protein scaffold regulates the redox potential and determines the substrate and reaction specificity. Several attempts to classify these enzymes can be found in the literature [3]–[6] however, the most common is the one that groups them by the reaction mechanism and reactivity with oxygen: oxidases, dehydrogenases, monooxygenases, reductases, and redox-neutral flavin-dependent enzymes [7]. The catalytic mechanism of FAD-dependent enzymes generally includes two semi-reactions in which FAD alternates in the oxidized and reduced states. In flavin-dependent enzymes such as oxidases and dehydrogenases, the substrate enters the active site and reacts with FAD, by reducing it, and exits as a product (reductive half-reaction); subsequently, the cofactor is reoxidized by oxygen, as in oxidases, or by an electron acceptor (e.g., NAD(P)+ or benzoquinone derivatives), as in dehydrogenases, (oxidative half-reaction) (Fig. 25). Conversely, in monooxygenases, flavin reduction occurs before substrate entry and its oxidation is concomitant with product release. Sequence analysis of the hypothetical FAD-dependent enzymes PbfB, PbfC and PbfD identified during bioinformatics analysis establishes all three enzymes as belonging to the oxidase/dehydrogenase group, (Pfam: PF01266).

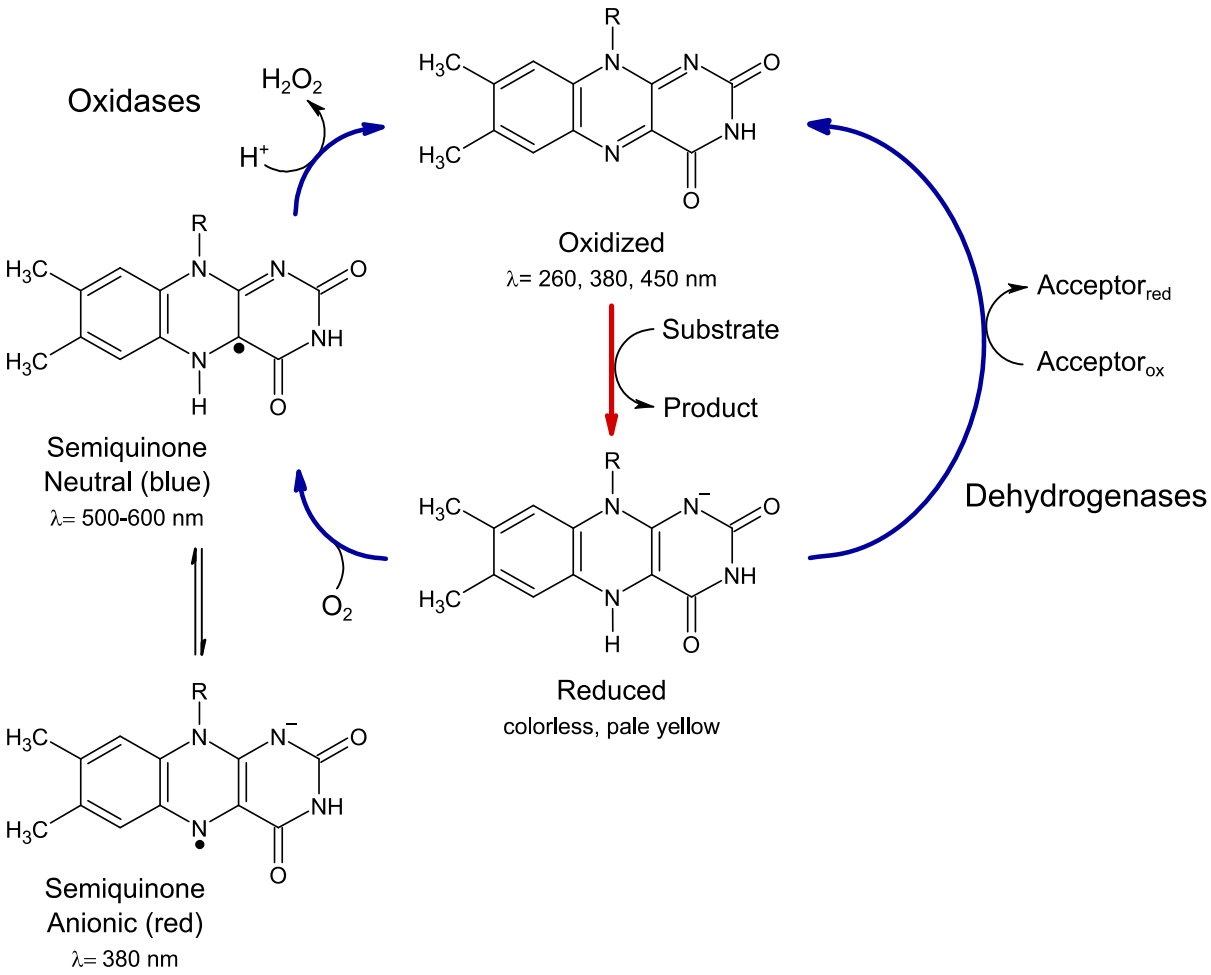


Figure 25. *Scheme of the two half-reactions of the catalytic cycle of oxidases and dehydrogenases.* FAD can either be covalently bound to the enzyme or not, and is typically found in the resting state of the enzyme in oxidized form. When the substrate enters the site, there is an oxidative half-reaction (red arrow), where FAD is reduced and the product released. The reaction cycle is completed by the reoxidation of FAD in the reductive half-reaction (blue arrow) by oxygen, in the case of oxidases, or by another electron acceptor (natural, such as NAD(P)+, or synthetic, such as DCPIP) in dehydrogenases. The FAD cofactor gives a characteristic absorption spectrum to the protein to which it is bound and this feature can be used to identify different reaction intermediates, due to the different wavelength at which they maximally absorb.

4.2 Methods

4.2.1 Materials

Triethanolamine (TEA), bovine serum albumin (BSA), flavin adenine dinucleotide (FAD), *o*-dianisidine dihydrochloride, 2,6-dichlorophenolindophenol (DCPIP), phenazine methosulfate (PMS) were from Sigma-Aldrich. NADH was from Alfa Aesar. All other reagents were from Fluka or Sigma-Aldrich.

Enzymes: alcohol dehydrogenase (ADH) from yeast, L-glutamate dehydrogenase (GDH) from bovine liver and peroxidase from horseradish were from Sigma.

Phosphonic acids: N-(Phosphonomethyl)glycine (glyphosate) and D,L-2-Amino-3-phosphonopropionic acid (D,L-phosphonoalanine; D,L-PnAla) were from Sigma-Aldrich. 2-aminoethylphosphonic acid (ciliatine; 2-AEP) was from Wako chemicals. The 1-hydroxy-2-aminoethylphosphonic acid (hydroxyciliatine; OH-AEP), the N-methylated derivatives of OH-AEP (MMOH-AEP, DMOH-AEP, TMOH-AEP) and AEP (MMAEP, DMAEP, TMAEP) were synthesized by the laboratory of Dr. Katharina Pallitsch at the Institute of Organic Chemistry, University of Vienna.

4.2.2 Plasmid constructs and protein purification

The genes of PbfB from *V. splendidus* 12B01 (WP_004730145.1), PbfC from *Azospirillum lipoferum* B510 (WP_012976454.1), PbfD1 from *Acinetobacter baumannii* (WP_079548425.1) and PbfD2 from *Mariniblastus fucicola* (WP_075082418.1) were codon-optimized for expression in *E. coli*, synthesized and cloned in the NdeI/NotI sites of a pET28a-N-HIS by BaseGene BV (Leiden, the Netherlands). The plasmids were used to transform by electroporation *E. coli* Chaperone Competent Cells pGro7/BL21 (Takara Bio Inc.) and *E. coli* Tuner™ (DE3) cells (EMD Biosciences) for PbfB and PbfC expression, respectively. The two synthetic genes for PbfD enzymes have just arrived, and only preliminary expression tests in *E. coli* Tuner™ (DE3) cells have been done. PbfB and PbfC enzymes have already been overexpressed and their purification has been pursued following two procedures.

A PbfB's overnight preculture (Chaperone Competent Cells pGro7/BL21) of 10 mL (LB and antibiotics) was inoculated into 1 L of LB autoinduction medium (0.5 g/L glucose and 2 g/L lactose) supplemented with kanamycin (50 µg/mL), chloramphenicol (20 µg/mL) and 0.5 mg/mL L-arabinose and grown at 20 °C for 20 hours. After induction, the cells were harvested by centrifugation (7,200 × *g* for 10 minutes at 4°C). The cell pellets were washed once with PBS, resuspended, centrifuged, and finally stored at -20°C. For protein purification, pellets were resuspended in a lysis buffer containing 50 mM sodium phosphate (NaP) pH 8.0, 100 mM NaCl, 1 mM PMSF, 1 mM benzamidine, 10% glycerol and 10 µM FAD. Then the cell suspension was sonicated and centrifuged (26,200 × *g* for 40 minutes at 4°C). Cleared cell lysate was loaded into a 50 mL super loop of AKTA Pure System FPLC and purified by affinity

chromatography (IMAC) using a HisTrap™ Fast Flow (Cytiva) column of 5 mL. Once the supernatant was loaded, the column was washed with 30 mL of washing buffer to remove chaperones (50 mM NaP pH 7. 5, 100 mM KCl, 10% glycerol, 500 mM sucrose, 20 mM $MgCl_2$ and 5 mM ATP). The proteins were eluted with a gradient of elution buffer (30 mM NaP pH 8.0, 60 mM NaCl, 400 mM imidazole). SDS-PAGE assessed protein purity and PbfB fractions were pooled and extensively dialyzed against 50 mM NaP pH 7.6 and 150 mM NaCl to remove imidazole. After dialysis, the enzyme solution was frozen in liquid nitrogen and stored at -80 °C. The solution containing PbfB was thawed and dialyzed against 50 mM Hepes pH 8.0 to lower the salt concentration present in the storage buffer to 10 mM NaCl, and then loaded to the cation exchanger column HiTrap™ SP Fast Flow (Cytiva). Proteins were eluted with a gradient of elution buffer (50 mM Hepes pH 8.0 and 1 M NaCl). Protein purity was assessed by SDS-PAGE. PbfB fractions were pooled, supplemented with 10% glycerol, frozen in liquid nitrogen and stored at -80 °C.

PbfC bacterial cultures (Tuner cells) were grown at 37 °C in LB medium supplemented with kanamycin (50 µg/mL) until OD_{600} reached 0.7-0.8. At this point, the temperature was reduced to 20°C and the induction was started by adding 0.25 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). 20 hours after induction, the cells were harvested by centrifugation (7,200 × *g* for 10 minutes at 4°C). The cell pellets were washed once with PBS, resuspended, centrifuged, and finally stored at -20°C. For protein purification, pellets were resuspended in a lysis buffer containing 50 mM sodium pyrophosphate (NaPPi) pH 8.5, 150 mM NaCl, 1 mM PMSF, 1 mM benzamidine, 10% glycerol and 10 µM FAD. Then the cell suspensions were sonicated and centrifuged (26,200 × *g* for 40 minutes at 4°C). The cleaned cell lysate was continuously looped (flow rate: 1 mL/min) on a 5 mL HisTrap™ Fast Flow column for 2 hours at 4°C through a peristaltic pump and then purified on the AKTA Pure System FPLC. Proteins were eluted with a gradient of elution buffer (50 mM NaPPi pH 8.5, 150 mM NaCl, 208 mM imidazole). PbfC fractions were pooled and extensively dialyzed against a dialysis buffer (50 mM NaPPi pH 8.5, 150 mM NaCl, 10% glycerol, 10 µM FAD). Then the protein solution volume was reduced to 3 mL using an Amicon® Ultra-2 3K device (Merck), loaded onto a 5 mL loop of AKTA Pure System FPLC and purified by size exclusion chromatography (SEC) with a HiPrep™ 16/60 Sephacryl S-300 HR (Cytiva) column. The SEC buffer contains 50 mM NaPPi pH 8.5, 150 mM NaCl, 10% glycerol. PbfC fractions were collected, concentrated by Amicon centrifugation and subdivided into aliquots (about 0.5 mL each). They were frozen in liquid nitrogen and stored at −80 °C. Enzyme concentration was calculated spectrophotometrically using the extinction coefficient at 280 nm estimated by the ProtParam tool (http://web.expasy.org/protparam/) (Table 4). Protein concentration was also verified using Bradford's (Bio-Rad protein assay dye reagent concentrate) method with bovine serum albumin as the

standard protein [8], following the "microtiter plate protocol" of the manufacturer's instructions. The protein yield was 3 mg per liter of culture.

Table 4. *Physical and chemical parameters of the His-tagged proteins were calculated by the ProtParam tool*.

Enzymes	ε280nm value	molecular weight	theoretical pI
PbfB	86,290 $M^{-1}cm^{-1}$	54.7 kDa	8.9
PbfC	47,440 $M^{-1}cm^{-1}$	44.2 kDa	7.8
PbfD1	43,320 $M^{-1}cm^{-1}$	45.7 kDa	6.0
PbfD2	53,970 $M^{-1}cm^{-1}$	43.4 kDa	5.9

4.2.3 ADH-coupled assay for 2-AEP or PnAA detection

To test the hypotheses that PbfB-C-D1-D2 may generate 2-AEP or PnAA, we used a continuous spectrophotometric assay in the presence of PhnW-PhnX pair and alcohol dehydrogenase (ADH) or with PhnX and ADH, respectively, as done in the study of PbfA. Briefly, the reactions were conducted in a final volume of 150 µl (typically containing 50 mM buffer TEA-HCl pH 8.0, 0.25 mM NADH, 3 mM $MgCl_2$, an excess of PhnX and ADH, and the substrates) by monitoring the disappearance of NADH at 340 nm on a UV-Vis Cary 50 spectrophotometer (Varian).

4.2.4 GDH-coupled assay for ammonia detection

To detect an oxidative deamination activity on 2-AEP and OH-AEP, we coupled the reaction with that of GDH, since ammonium would be released together with the corresponding phosphonate aldehydes. The GDH-coupled assay mixture usually contained 50 mM TEA-HCl buffer pH 8.0, 0.25 mM NADH, 1 mM α-KG and an excess of GDH and was followed by monitoring the disappearance of NADH at 340 nm on a UV-Vis Cary 50 spectrophotometer (Varian).

4.2.5 UV-absorbance measurements

Absorption spectra were collected in a microcuvette (l=1 cm) using a thermostatted V-750 UV-Vis spectrophotometer (Jasco Inc.) and corrected for buffer contribution.

4.2.6 Peroxidase-*o*-dianisidine-coupled assay for H_2O_2 detection

A discontinuous colorimetric plate (cell culture plate 96-well, SARSTEDT) assay using peroxidase and chromogenic *o*-dianisidine was developed to detect the release of hydrogen peroxide when PbfC

reacted with different phosphonates and phosphonate analogs. In detail, 2 µM PbfC was incubated in 50 mM NaPPi pH 8.5, 150 mM NaCl, 10% glycerol, 10 µM FAD, with 5 mM or 10 mM (for racemic mixtures) substrate. The reactions started by the addition of the enzyme and after 1 hour of incubation at room temperature, they were supplemented with 3 U of peroxidase and 1 mM *o*-dianisidine. Finally, after 10 minutes, 3.5 M sulfuric acid was added to each well (final volume of 100 µL) and the plate was scanned with a plate reader (λ= 490 nm).

4.2.7 DCPIP-coupled assay for reoxidation of reduced FAD

Similar to the peroxidase assay, a continuous colorimetric plate assay was set up using the blue dye DCPIP as the final electron acceptor to detect the activity of PbfC against various phosphonates and phosphonate analogs. In this case, 0.11 µM PbfC was incubated in 50 mM TEA-HCl pH 8.0, 100 µM DCPIP, and with 5 mM or 10 mM (for racemic mixtures) substrate; final volume 200 µL. Reactions were started by the addition of the enzyme, and the plate was scanned with a plate reader (λ= 595 nm) every 5 min at room temperature.

4.2.8 DCPIP-PMS-coupled assay

A coupled assay with the two artificial electron acceptors DCPIP and PMS was developed, similar to that reported by Augustin and coworkers [9], to obtain the kinetic parameters of the PbfC reaction with MMAEP and 2-AEP. DCPIP and PMS were prepared and stored as described by Jahn and coworkers [10]. Assays to establish kinetic parameters were performed at 25°C on a thermostatted V-750 UV-Vis spectrophotometer (Jasco Inc.). Reaction mixtures contained 50 mM TEA-HCl pH 8.0, 3 mM PMS, 80 µM DCPIP, the substrate (MMAEP or 2-AEP), and 0.21 µM PbfC. Reactions were triggered with the enzyme and DCPIP reduction rates were followed at 600 nm. The extinction coefficient of DCPIP at 600 nm (18,400 M^{-1} cm^{-1}) was determined under the same assay conditions. Each reaction was repeated in triplicate and the kinetic parameters were computed by nonlinear least-squares fitting to the Michaelis-Menten equation using Sigma Plot (Systat Software Inc.).

4.3 Results and discussion

In the last section of this dissertation, preliminary results from the *in vitro* study of PbfB, PbfC, PbfD1 and PbfD2 will be presented. We used different techniques, in addition to those used in the characterization of PbfA, to test our initial hypotheses:

- Development of discontinuous coupled assays to screen the activity of these FAD-dependent enzymes against various phosphonate compounds and their analogs.

- Development of continuous spectrophotometer-coupled assays to follow product production or FAD reduction.

The application of some of these methods is summarized in the figure below (Fig. 26).

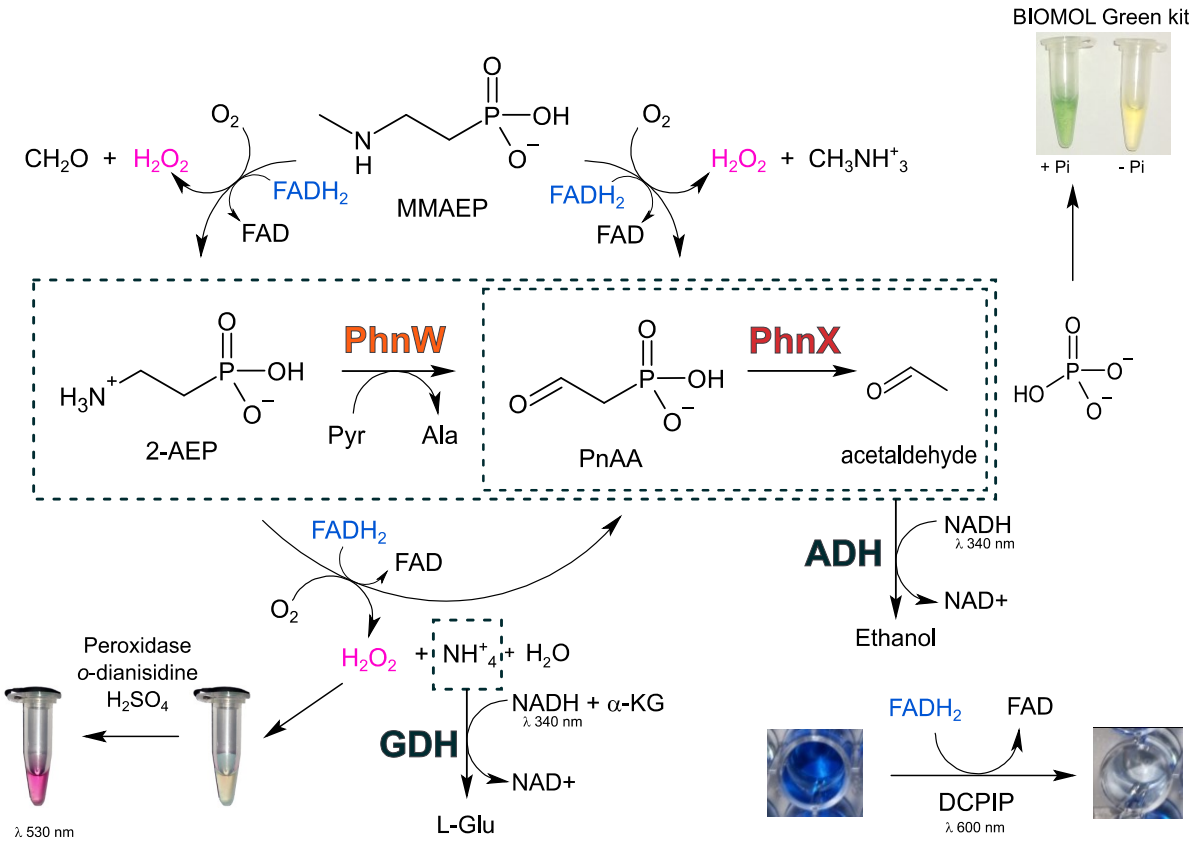


Figure 26. *Schematic representation of the techniques used to study PbfB-C-D1-D2 activity*. The reactions that were coupled to alcohol dehydrogenase (ADH) and glutamate dehydrogenase (GDH)

are highlighted in the dashed box. Phosphate released from the PhnX reaction can be established using BIOMOL green reagent. The peroxidase-*o*-dianisidine-coupled assay can detect hydrogen peroxide release. The DCPIP-coupled assay can follow FAD reoxidation.

4.3.1 Optimization of enzyme expression and purification

Unlike PbfD1 and PbfD2, whose genes have recently arrived and only a preliminary expression tests have been performed, PbfB and PbfC were both overexpressed (Fig. 27 A). However, the two recombinant proteins were poorly soluble and, due to weak binding to the affinity chromatography resin, they were also hard to purify. Therefore, only PbfC was successfully purified. Low expression of PbfB was detected under all tested conditions, i.e., in different *E. coli* expression strains (BL21 star, Tuner, and Chaperone Competent Cells pGro7/BL21), culture media (LB, LB autoinduction, and M9), optical density (OD_{600} 0.6-0.9) and IPTG concentrations (1 mM and 0.25 mM). PbfB was also mainly insoluble after sonication; nevertheless, it was partially present in the soluble fraction when expressed in Chaperone Competent Cells pGro7/BL21). After protein expression in this strain, a two-step purification was attempted: first, on immobilized metal affinity chromatography (IMAC), exploiting the 6xHis-tag at the N-terminus of PbfB, then on cation exchange chromatography (CIEX), exploiting the high isoelectric point (pI) of the protein predicted by ProtParam (Fig. 27 B). These steps were insufficient to obtain a pure solution of PbfB, so we plan to optimize the procedure for better results, e.g., by testing other cell lysis methods, improving the attachment of the enzyme to the affinity resin, addressing other purification steps. Regarding PbfC, we overexpressed the enzyme in Tuner cells, and after lysis we purified it in two steps: first through an IMAC and then through a size exclusion chromatography (SEC) column. PbfC, like PbfB, is mainly in the insoluble fraction after lysis and binds scarcely to the affinity resin. However, it was expressed in larger amounts and its binding was improved by looping the supernatant in the nickel column for a couple of hours with a peristaltic pump (Fig. 28 A). In doing so, we obtained a highly pure PbfC solution with an absorption spectrum typical of FAD-dependent enzymes, showing a maximum at 387 and 465 nm (Fig. 28 B).

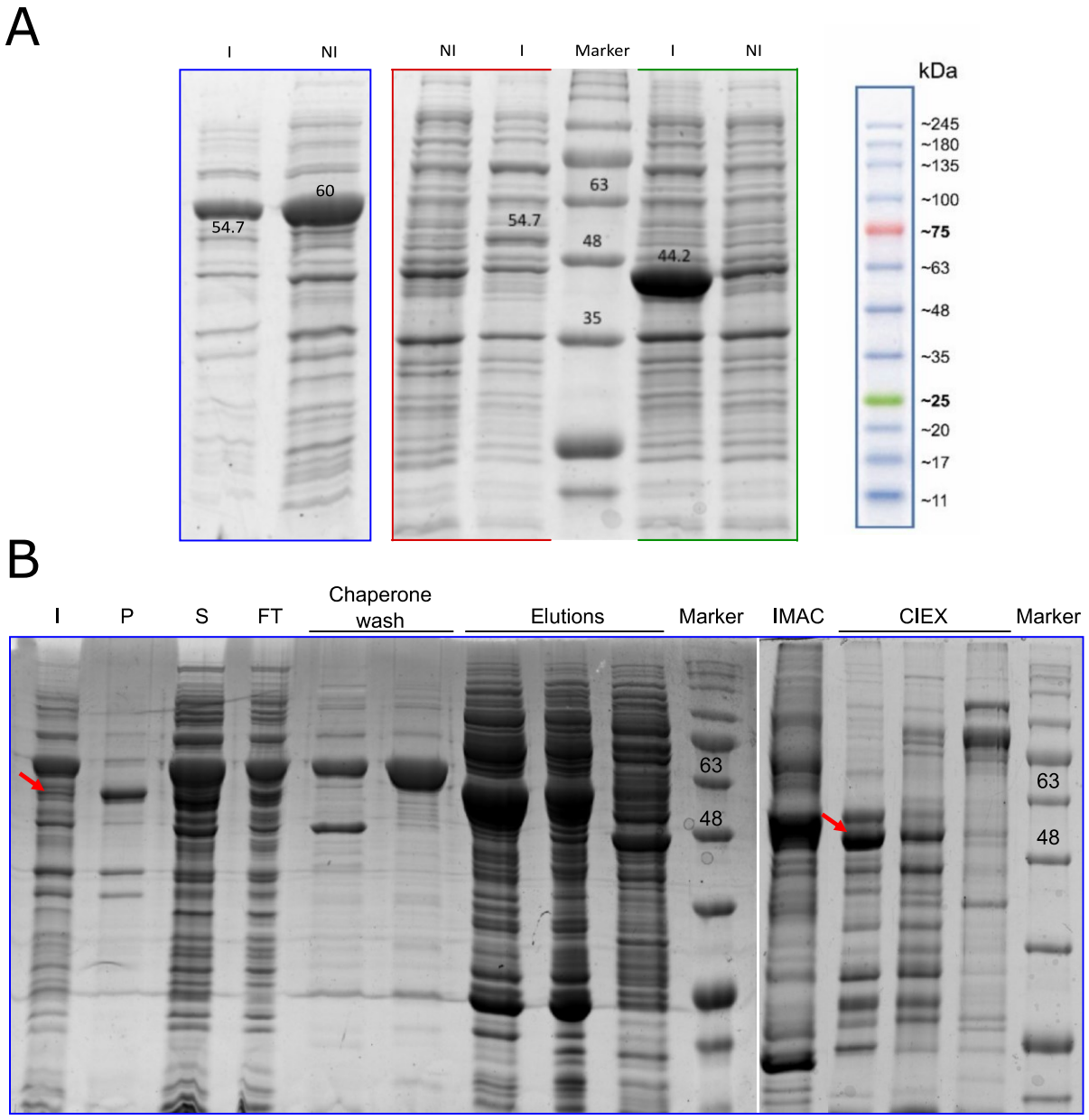


Figure 27. (**A**) *PbfB and PbfC expression trials*. 12% SDS-PAGE gel analysis of PbfB (54.7 kDa) expression assays in the Chaperone Competent Cells pGro7/BL21 (GroEL, 60 kDa; GroES, 10 kDa) (blue box) and Tuner strain (red box); cultures were grown in LB at 37°C and induced at OD_{600} 0.7 with 1 mM IPTG for 2 hours. Low expression levels were observed in both. PbfC (44.2 kDa) was expressed in Tuner strain (green box); culture was grown in LB at 37°C and induced at OD_{600} 0.7 with 1 mM IPTG for 2 hours. Abbreviations: I, induced; NI, not induced. (**B**) *PbfB purification*. 12% SDS-PAGE gel analysis of PbfB purification by IMAC (left gel) followed by CIEX (right gel). The protein band corresponding to PbfB (red arrow) is also found in the pellet (insoluble fraction) and flow-through, left gel. However, even after the CIEX the protein solution is not sufficiently pure, right gel. Abbreviations: I, induced; P, pellet; S, supernatant; FT, flow-through; IMAC, immobilized metal affinity chromatography; CIEX, cation

exchange chromatography; marker (Protein marker BlueStar Prestained, © 2022 NIPPON Genetics Europe).

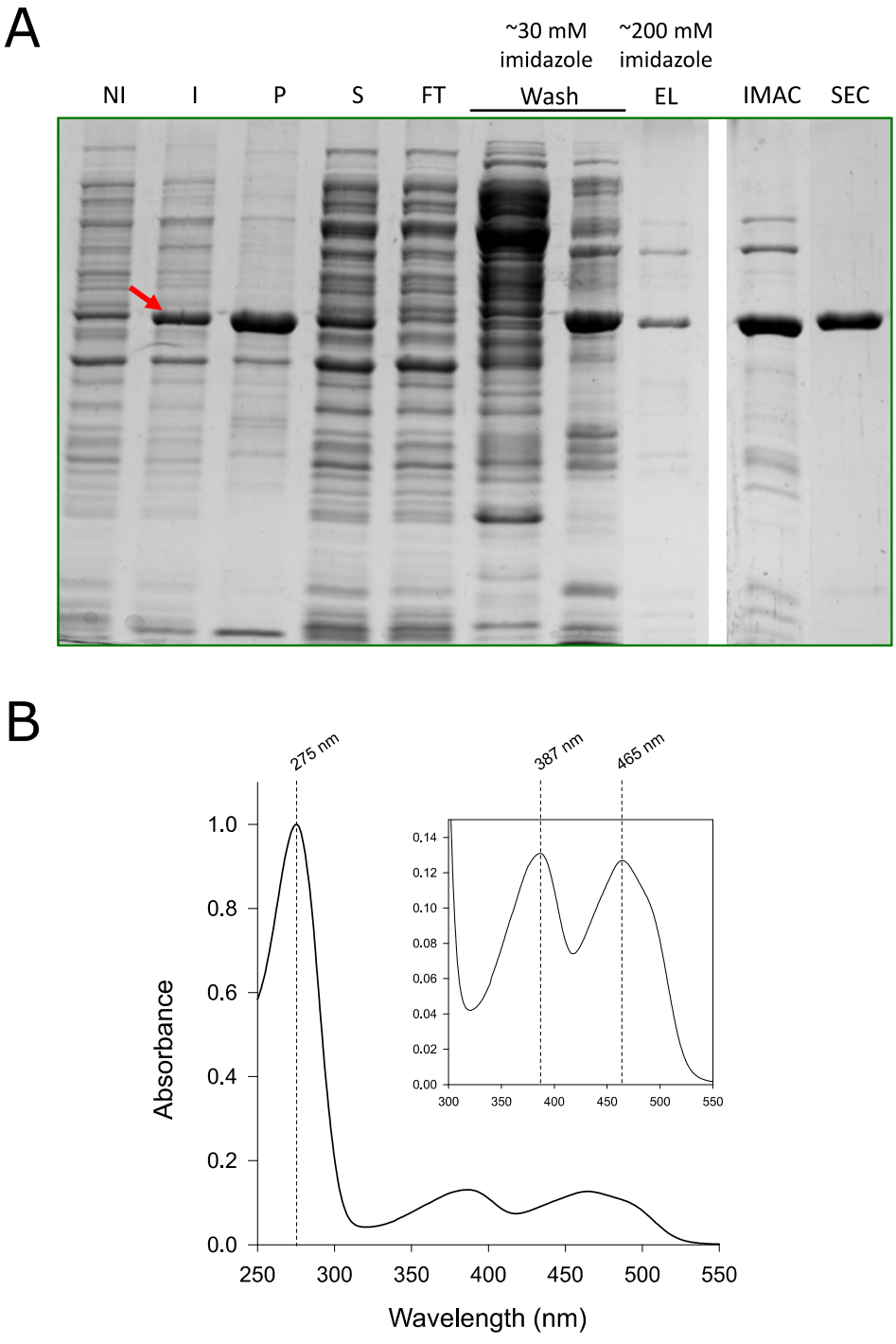


Figure 28. (**A**) *PbfC purification*. 12% SDS-PAGE gel analysis of PbfC purification by IMAC (left gel) followed by SEC (right gel). The protein band corresponding to PbfC (red arrow) is also found in the pellet (insoluble fraction), left gel. After SEC, protein purity was assessed >95%, right gel.

Abbreviations: NI, not induced; I, induced; P, pellet; S, supernatant; FT, flow-through; IMAC, immobilized metal affinity chromatography; SEC, size exclusion chromatography. (**B**) *UV-vis absorption spectrum of PbfC*. Absorption spectrum of 21 µM PbfC in a solution containing 50 mM TEA-HCl pH 8.0, 150 mM NaCl, 10% glycerol. The spectrum was corrected for solution contribution.

4.3.2 Initial screening of phosphonates compounds by peroxidase-*o*-dianisidine-coupled assay

After optimizing the expression and purification procedure of PbfC, we tested its ability to oxidize potential phosphonate substrates and analog compounds. In particular, we first investigated its ability to produce hydrogen peroxide, as done by FAD-dependent oxidases that use oxygen as the final electron acceptor in the half-oxidative reaction (FAD reoxidation). A peroxidase-*o*-dianisidine plate assay was developed, allowing high-throughput screening in a microplate reader. In the assays, hydrogen peroxide release was revealed by exploiting horseradish peroxidase, which oxidizes the chromogenic substrate *o*-dianisidine into a product that absorbs at 450 nm (brown color). Sulfuric acid was subsequently added to increase signal stability and intensity, resulting in a shift of the *o*-dianisidine absorption peak to 530 nm (pink color) (Fig. 29 A). Albeit qualitative, the findings revealed strong reactivity of PbfC with AEP and, most importantly, with its N-monomethylated derivative MMAEP (Fig. 29 B). Furthermore, hydrogen peroxide production was accompanied by that of phosphonoacetaldehyde, which was verified through the PhnX-ADH-coupled assay by starting the kinetics with part of each reaction solution.

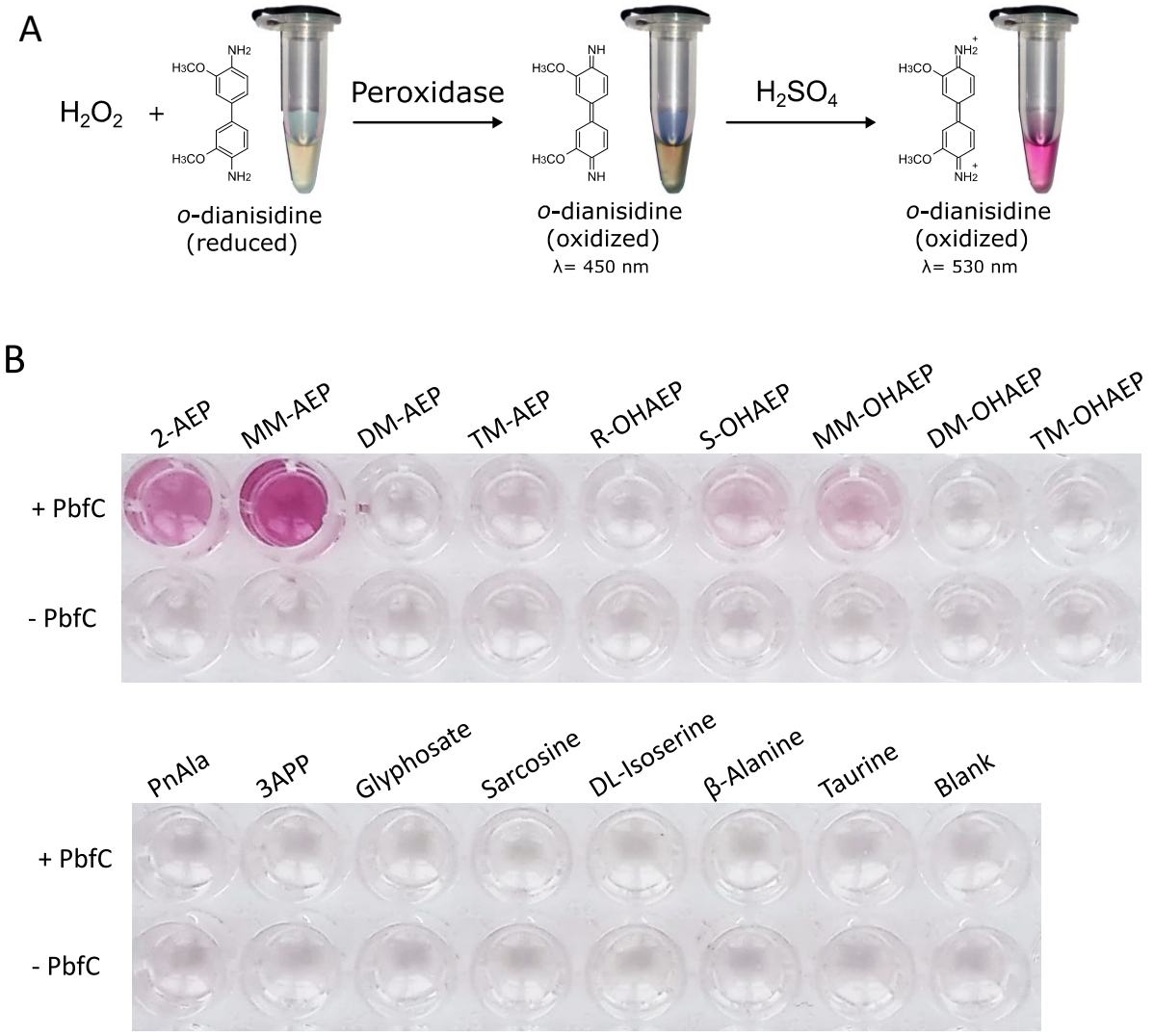


Figure 29. *Investigation of PbfC activity by peroxidase-o-dianisidine-coupled assay*. (**A**) Schematic representation of the peroxidase-*o*-dianisidine-coupled assay. (**B**) Results obtained from peroxidase-o-dianisidine plate assay. 2 µM of PbfC was incubated for 1 hr at room temperature with 5 mM (or 10 mM for racemic solutions) of the substrate. After 1 hour, peroxidase, *o*-dianisidine, and finally sulfuric acid were added. Controls containing all reagents except the enzyme were run for each reaction (see the bottom row indicated by "- PbfC"). The plate was scanned with a plate reader (λ= 490 nm). Abbreviations: 3APP, 3-aminopropylphosphonate.

4.3.3 PbfC is much more active with artificial electron acceptors than with molecular oxygen

When we attempted to observe the reaction of PbfC toward these two compounds with a continuous spectrophotometric assay (using the PhnX-ADH-coupled assay, for 2-AEP and MMAEP, and the GDH-coupled assay, for 2-AEP oxidation), we realized that the reactions were excessively slow and difficult to follow, despite using high concentrations of substrates and enzyme. We obtained comparable

results by employing oxygen-saturated buffers to increase dissolved oxygen concentration in the solution. These findings suggested that oxygen may not be the physiological electron acceptor of PbfC, so we investigated its activity by employing synthetic electron acceptors that are often used in oxidase and dehydrogenase enzymatic assays [11]. Specifically, we adapted the plate screening assay using the redox dye DCPIP (2,6-dichlorophenol-indophenol). This artificial elector acceptor is blue with an absorption maximum at 600 nm when oxidized and becomes colorless when reduced. Contrary to the peroxidase assay, this experiment was performed continuously, and already after 5 minutes, the reaction containing MMAEP had become completely transparent, whereas the others had not (Fig. 30). However, after approximately 1 hour, a slow reoxidation of the reduced DCPIP by oxygen could be observed, implying that upon long incubation times the results of this assay are not reliable anymore [10].

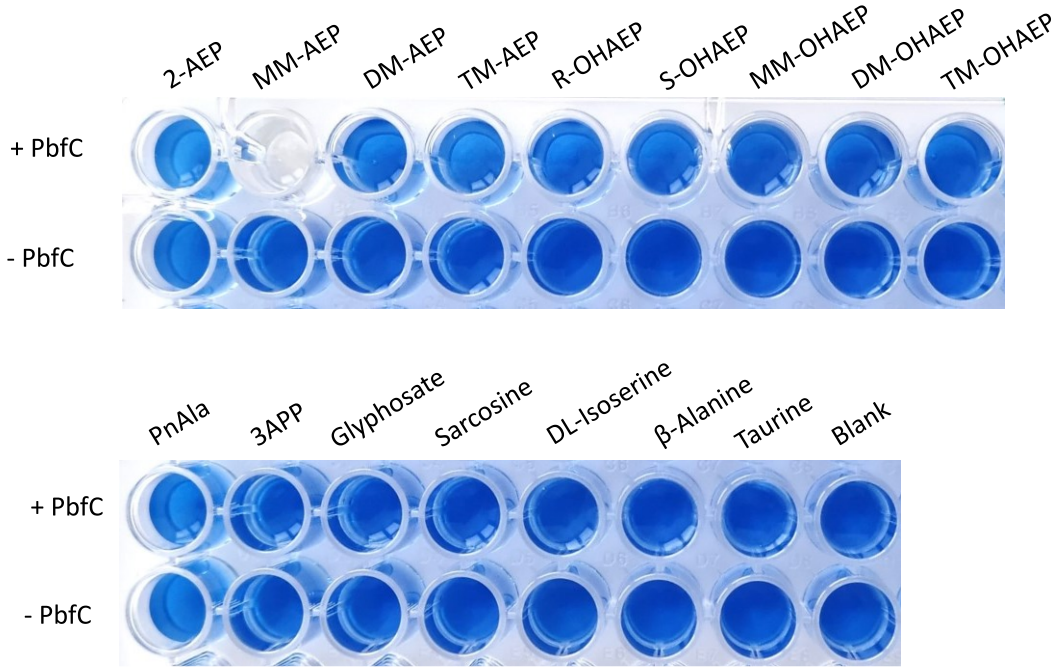


Figure 30. *In plate investigation of PbfC activity by DCPIP-coupled assay*. The reactions were triggered by the addition of 0.11 µM PbfC and contained 50 mM TEA-HCl pH 8.0, 100 µM DCPIP, 5 mM or 10 mM (for racemic mixtures) substrate; final volume 200 µL. This image was collected before a reoxidation of DCPIP by oxygen. Controls containing all reagents except the enzyme were run for each reaction (see the bottom row indicated by "- PbfC"). Abbreviations: 3APP, 3-aminopropylphosphonate.

While the peroxidase assay was more robust in detecting PbfC activity (especially towards 2-AEP) than the DCPIP-based assay, this latter emphasizes the efficiency of the enzyme in catalyzing a specific oxidation reaction with MMAEP. Therefore, although we captured traces of oxidase activity, the dehydrogenase one appear to be the predominant activity of PbfC. In particular, we estimated that

the reaction rate is more than 50 times higher with DCPIP-PMS than in the presence of oxygen as an electron acceptor (Fig. 31). This behavior is characteristic of FAD-dependent dehydrogenases that may not react at all (true dehydrogenase) or react very slowly with oxygen to form H_2O_2 or some superoxide anion [12].

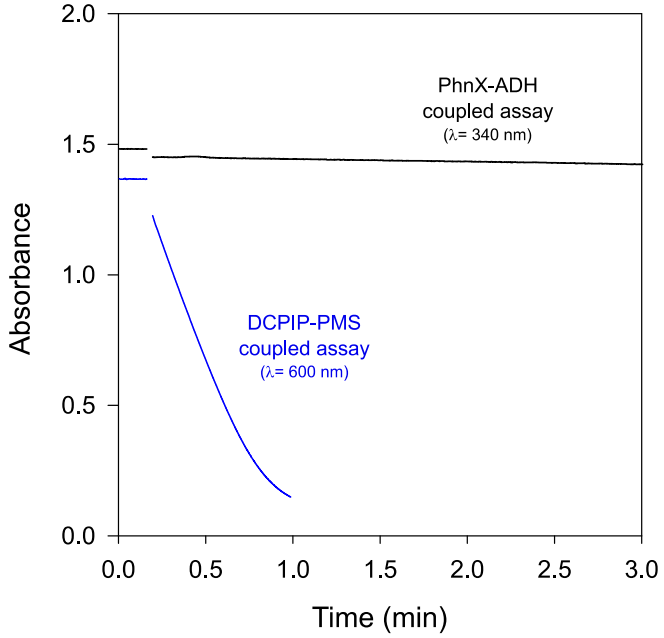


Figure 31. *PbfC-catalyzed oxidation of MMAEP in the presence of different electron acceptors.* Both reactions were carried out at room temperature in 50 mM TEA-HCl, 3 mM MgCl2, 2 mM MMAEP and 0.72 µM PbfC enzyme. The PhnX-ADH-coupled assay reaction was supplemented with 0.25 mM NADH and an excess of PhnX and ADH enzymes (4.5 µM and 9 U, respectively); it was monitored following the oxidation of NADH at 340 nm (black line). The DCPIP-PMS-coupled assay reaction was supplemented with 80 µM DCPIP and 3 mM PMS; it was monitored following the reduction of DCPIP at 600 nm (blue line).

4.3.4 PbfC functional parameters and substrate specificity

The DCPIP-PMS-coupled assay was used to estimate the kinetic parameters of the PbfC-catalyzed oxidation of MMAEP and 2-AEP. Applying this assay at pH 8.0 and 25°C, we obtained k_{cat} = 4.93 ± 0.06 s^{-1}, K_M = 0.19 ± 0.01 mM, and k_{cat}/K_M =25,900 ± 1,500 $M^{-1}s^{-1}$. In contrast, k_{cat}/K_M for 2-AEP was approximately 400-fold slower (k_{cat} = 1.28 ± 0.03 s^{-1}, K_M = 20.3 ± 1.7 mM, and k_{cat}/K_M = 60 ± 6 $M^{-1}s^{-1}$) (Fig. 32). This marked difference supports the idea that the enzyme is specialized to oxidize MMAEP, whereas the oxidation of 2-AEP could be considered a promiscuous reaction, whose rate can approach the rate observed with MMAEP only when the turnover is limited by the slow transfer of electrons from reduced FAD to an inefficient acceptor such as oxygen.

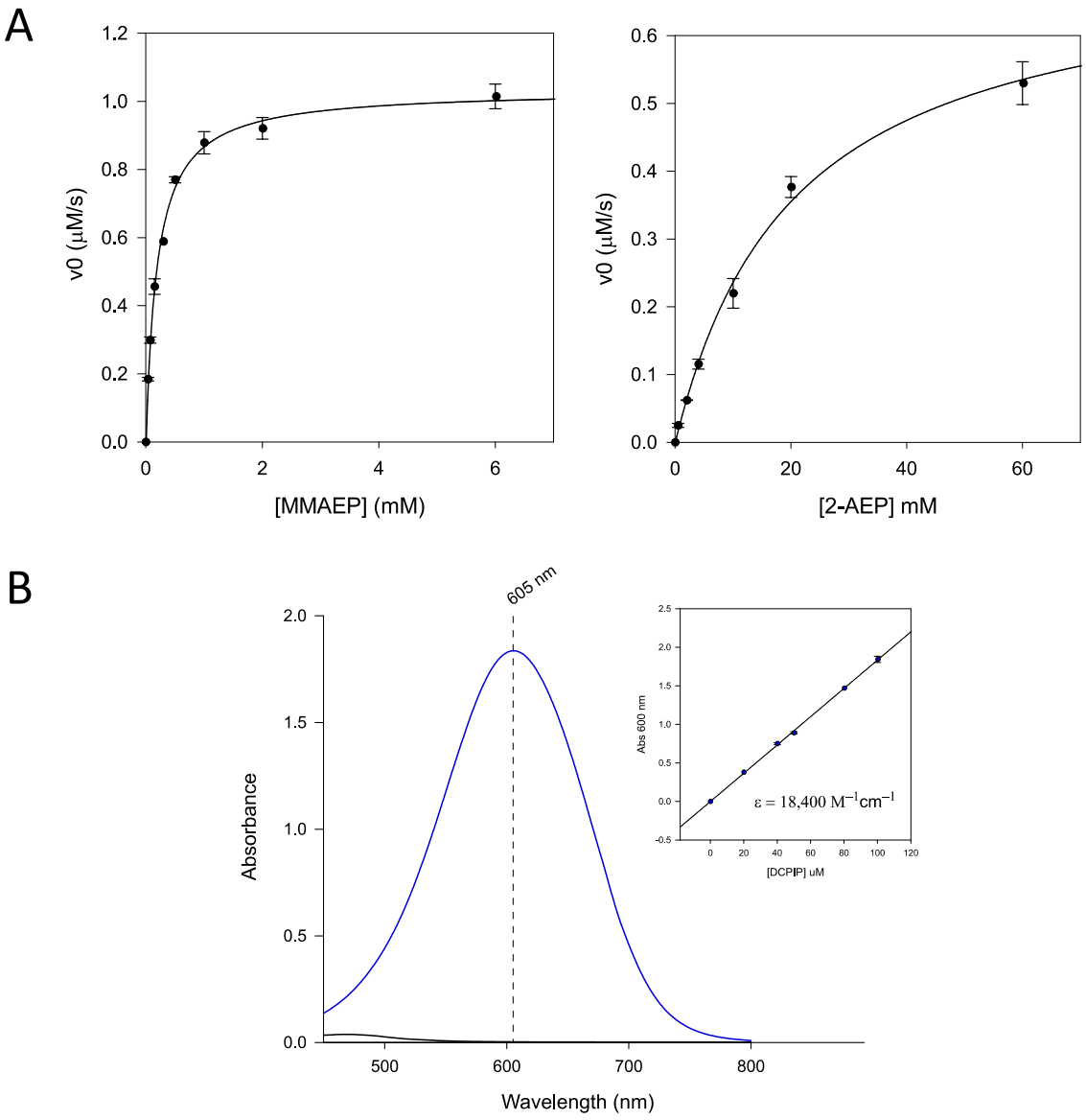


Figure 32. *Dependence of the initial velocity of PbfC reaction on MMAEP (0.04-6 mM) and 2-AEP (0.5-60 mM)*. (**A**) The kinetic parameters were obtained using the DCPIP-PMS-coupled assay at 25°C, pH 8.0. Each assay contained 50 mM TEA-HCl pH 8.0, 3 mM PMS, 80 µM DCPIP, the substrate MMAEP or 2-AEP, and were triggered by adding 0.21 µM and 0.56 µM PbfC, respectively. The triplicates of initial rates were fitted to the Michaelis-Menten equation and the calculated k_{cat}, k_{cat}/K_M and K_M are reported as +/- standard error of the regression. (**B**) UV-visible spectral properties of DCPIP in 50 mM TEA-HCl pH 8.0. The spectrum of oxidized DCPIP (blue line) and that of DCPIP reduced with 1 mM DTT (black line) are shown in the figure. Inset: absorption dependence at 600 nm on DCPIP concentration. Linear fitting of the data indicates an extinction coefficient of 18,400 M^{-1} cm^{-1}. The spectra were corrected for solution contribution.

4.4 Conclusions

4.4.1 PbfC expands the scope and utility of the PhnW-PhnX pathway

In this chapter, our research focused on the biochemical characterization of the three FAD-dependent enzymes identified in the previous bioinformatic analysis (termed PbfB, PbfC, and PbfD). Although we are still far from the conclusion of our study, the evidence that PbfC from *Azospirillum lipoferum* B510 oxidizes the N-monomethyl-2-AEP (that PhnW cannot process because it lacks a primary amino group), producing phosphonoacetaldehyde and methylamine was in agreement with our initial expectations. This is also consistent with the finding that PbfA similarly consumes an aminophosphonate that the PhnW-PhnX pair cannot degrade by generating PnAA and channeling it into the main pipeline. While MMAEP is quite common in the marine environment, for example, it has been isolated in lipids, glycans and proteins of mollusks, shellfish, anemones, and marine snails [13]–[18], PbfC represents the first enzyme dedicated to its specific degradation. In contrast, the biological source of phosphonates in terrestrial environments is still unclear; however, *Azospirillum lipoferum* is a soil bacterium, closely related to aquatic habitat bacteria [19], that plays a key role in the dynamics of N and P by making them available to plants (as a plant growth-promoting bacterium) [20]. Based on the results obtained, we can state that the presence of PbfC and that of a PbfA homolog in this organism allow it to exploit a broader repertoire of aminophosphonates (2-AEP, *R*-OH-AEP and MMAEP) by including them in the specialized PhnW-PhnX pathway, thereby increasing its ability to acquire N and P in a convenient manner. Furthermore, its capacity to convert methylamine to ammonium is also known [21]. Regarding the PbfB and PbfD function, the issue about their exact function is still open, and it will be interesting to understand if they participate in the channeling of other aminophosphonates in the main pathway or they are also specific for MMAEP, and, if so, whit what efficiency.

4.5 References

[1] P. Macheroux, B. Kappes, and S. E. Ealick, "Flavogenomics - a genomic and structural view of flavin-dependent proteins," *FEBS Journal*, vol. 278, no. 15, pp. 2625–2634, Aug. 2011, doi: 10.1111/j.1742-4658.2011.08202.x.

[2] M. W. Fraaije and A. Mattevi, "Flavoenzymes: diverse catalysts with recurrent features," *Trends in Biochemical Sciences*, vol. 25, no. 3, pp. 126–132, Mar. 2000, doi: 10.1016/S0968-0004(99)01533-9.

[3] M. L. Mascotti, M. Juri Ayub, N. Furnham, J. M. Thornton, and R. A. Laskowski, "Chopping and Changing: the Evolution of the Flavin-dependent Monooxygenases," *Journal of Molecular Biology*, vol. 428, no. 15, pp. 3131–3146, Jul. 2016, doi: 10.1016/j.jmb.2016.07.003.

[4] W. P. Dijkman, G. de Gonzalo, A. Mattevi, and M. W. Fraaije, "Flavoprotein oxidases: classification and applications," *Applied Microbiology and Biotechnology*, vol. 97, no. 12, pp. 5177–5188, Jun. 2013, doi: 10.1007/s00253-013-4925-7.

[5] A. H. Merrill, J. D. Lambeth, D. E. Edmondson, and D. B. McCormick, "FORMATION AND MODE OF ACTION OF FLAVOPROTEINS," *Annual Review of Nutrition*, vol. 1, no. 1, pp. 281–317, Jul. 1981, doi: 10.1146/annurev.nu.01.070181.001433.

[6] R. L. Fagan and B. A. Palfey, "Flavin-Dependent Enzymes," in *Comprehensive Natural Products II*, Elsevier, 2010, pp. 37–113. doi: 10.1016/B978-008045382-8.00135-0.

[7] P. Pimviriyakul and P. Chaiyen, "Overview of flavin-dependent enzymes," *The Enzymes*, vol. 47, pp. 1–36, Aug. 2020, doi: 10.1016/bs.enz.2020.06.006.

[8] M. M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," *Analytical Biochemistry*, vol. 72, no. 1–2, pp. 248–254, May 1976, doi: 10.1016/0003-2697(76)90527-3.

[9] P. Augustin, A. Hromic, T. Pavkov-Keller, K. Gruber, and P. Macheroux, "Structure and biochemical properties of recombinant human dimethylglycine dehydrogenase and comparison to the disease-related H109R variant," *The FEBS Journal*, vol. 283, no. 19, pp. 3587–3603, Oct. 2016, doi: 10.1111/febs.13828.

[10] B. Jahn *et al.*, "Understanding the chemistry of the artificial electron acceptors PES, PMS, DCPIP and Wurster's Blue in methanol dehydrogenase assays," *JBIC Journal of Biological Inorganic Chemistry*, vol. 25, no. 2, pp. 199–212, Mar. 2020, doi: 10.1007/s00775-020-01752-9.

[11] E. Rosini, L. Caldinelli, and L. Piubelli, "Assays of D-Amino Acid Oxidase Activity," *Frontiers in Molecular Biosciences*, vol. 4, Jan. 2018, doi: 10.3389/fmolb.2017.00102.

[12] E. Romero, J. R. Gómez Castellanos, G. Gadda, M. W. Fraaije, and A. Mattevi, "Same Substrate, Many Reactions: Oxygen Activation in Flavoenzymes," *Chemical Reviews*, vol. 118, no. 4, pp. 1742–1769, Feb. 2018, doi: 10.1021/acs.chemrev.7b00650.

[13] L. D. Quin and G. S. Quin, "Screening for carbon-bound phosphorus in marine animals by high-resolution 31P-NMR spectroscopy: coastal and hydrothermal vent invertebrates," *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, vol. 128, no. 1, pp. 173–185, Jan. 2001, doi: 10.1016/S1096-4959(00)00310-9.

[14] J. S. Kittredge, A. F. Isbell, and R. R. Hughes, "Isolation and Characterization of the N-Methyl Derivatives of 2-Aminoethylphosphonic Acid from the Sea Anemone, Anthopleura xanthogrammica [*]," *Biochemistry*, vol. 6, no. 1, pp. 289–295, Jan. 1967, doi: 10.1021/bi00853a045.

[15] Kh. S. Mukhamedova and A. I. Glushenkova, "Natural Phosphonolipids," *Chemistry of Natural Compounds*, vol. 36, no. 4, pp. 329–341, 2000, doi: 10.1023/A:1002804409503.

[16] B. Eckmair, C. Jin, D. Abed-Navandi, and K. Paschinger, "Multistep Fractionation and Mass Spectrometry Reveal Zwitterionic and Anionic Modifications of the N- and O-glycans of a Marine Snail," *Molecular & Cellular Proteomics*, vol. 15, no. 2, pp. 573–597, Feb. 2016, doi: 10.1074/mcp.M115.051573.

[17] T. Matsubara, M. Morita, and A. Hayashi, "Determination of the presence of ceramide aminoethylphosphonate and ceramide N-methylaminoethylphosphonate in marine animals by fast atom bombardment mass spectrometry," *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism*, vol. 1042, no. 3, pp. 280–286, Feb. 1990, doi: 10.1016/0005-2760(90)90154-P.

[18] D. S. Kirkpatrick and S. H. Bishop, "Phosphonoprotein. Characterization of aminophosphonic acid rich glycoproteins from sea anemones," *Biochemistry*, vol. 12, no. 15, pp. 2829–2840, Jul. 1973, doi: 10.1021/bi00739a009.

[19] F. Wisniewski-Dyé *et al.*, "Azospirillum Genomes Reveal Transition of Bacteria from Aquatic to Terrestrial Environments," *PLoS Genetics*, vol. 7, no. 12, p. e1002430, Dec. 2011, doi: 10.1371/journal.pgen.1002430.

[20] J. Fukami, P. Cerezini, and M. Hungria, "Azospirillum: benefits that go far beyond biological nitrogen fixation," *AMB Express*, vol. 8, no. 1, p. 73, Dec. 2018, doi: 10.1186/s13568-018-0608-1.

[21] F. Wisniewski-Dyé *et al.*, "Genome Sequence of Azospirillum brasilense CBG497 and Comparative Analyses of Azospirillum Core and Accessory Genomes provide Insight into Niche Adaptation," *Genes*, vol. 3, no. 4, pp. 576–602, Sep. 2012, doi: 10.3390/genes3040576.

General conclusions

While the work in this dissertation does not provide an entirely new pathway for phosphonate degradation, it does report the identification of two novel "ancillary" enzymes, PbfA and PbfC, whose function had never been described or proposed before. Their activities enhance the versatility and utility of the common (but very specific) PhnW-PhnX and PhnW-PhnY-PhnA pathways for 2-aminoethylphosphonate catabolism, and their recurrent presence in these clusters highlights the evolutionary advantage of an organism in consuming multiple distinct compounds through this single catabolic pathway. Although further works are needed to understand whether and how useful PbfA and PbfC are to organisms that possess them, their discovery opens the way for future investigations in the field of aminophosphonate catabolism, such as suggesting the need for further investigation in the study of transport and transcriptional regulators of these clusters that have been analyzed primarily for their specificity for 2-AEP.

Acknowledgements

I want to sincerely thank all the people who collaborated and supported this work: Tamara Dinhof, Toda Stanković, Katharina Pallitsch, Marco Malatesta, Domenico Acquotti and my supervisor Alessio Peracchi.