



Emergence of pyridoxal phosphorylation through a promiscuous ancestor during the evolution of hydroxymethyl pyrimidine kinases



Victor Castro-Fernandez¹, Felipe Bravo-Moraga¹, Cesar A. Ramirez-Sarmiento, Victoria Guixe*

Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Santiago, Chile

ARTICLE INFO

Article history:

Received 7 April 2014

Revised 5 June 2014

Accepted 6 June 2014

Available online 19 June 2014

Edited by Miguel De la Rosa

Keywords:

HMP kinase

Pyridoxal kinase

Bifunctional enzyme

Enzyme evolution

Resurrection ancestral enzyme

ABSTRACT

In the family of ATP-dependent vitamin kinases, several bifunctional enzymes that phosphorylate hydroxymethyl pyrimidine (HMP) and pyridoxal (PL) have been described besides enzymes specific towards HMP. To determine how bifunctionality emerged, we reconstructed the sequence of three ancestors of HMP kinases, experimentally resurrected, and assayed the enzymatic activity of their last common ancestor. The latter has ~8-fold higher specificity for HMP due to a glutamine residue (Gln44) that is a key determinant of the specificity towards HMP, although it is capable of phosphorylating both substrates. These results show how a specific enzyme with catalytic promiscuity gave rise to current bifunctional enzymes.

© 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

Although textbooks and research articles traditionally highlighted the remarkable specificity of enzyme action, the promiscuity of enzymes has received considerable attention in the past few years due to its role in the evolution of new functions. Several observations support the hypothesis that evolutionary progenitors and intermediates display broad specificity or high promiscuity. Therefore, bifunctional or promiscuous proteins could reflect ancestral forms from which two different specificities split and originated through a gene duplication event during evolution. Catalytic promiscuity can be defined as a secondary catalytic activity, since either affinity or intracellular substrate concentrations are not physiologically relevant [1], while bifunctionality refers to catalytic activities that have physiological relevance.

The family of ATP-dependent vitamin kinases from the ribokinase superfamily features a single domain with a conserved Rossmann-like fold [2], thus structurally diverging from other

Abbreviations: HMP, 4-amino-5-hydroxymethyl-2-methylpyrimidine; HMPK, 4-amino-5-hydroxymethyl-2-methylpyrimidine kinase; PL, pyridoxal; PLK, pyridoxal kinase; PLP, pyridoxal-5'-phosphate; THZK, 4-methyl-5-β-hydroxy-ethylthiazole kinase

* Corresponding author. Address: Departamento de Biología, Facultad de Ciencias, Universidad de Chile, Las Palmeras 3425, Ñuñoa, Santiago 7800003, Chile. Fax: +56 2 227 12983.

E-mail address: vguixe@uchile.cl (V. Guixe).

¹ These authors contributed equally to this work.

<http://dx.doi.org/10.1016/j.febslet.2014.06.033>

0014-5793/© 2014 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

enzymes of this superfamily that have, in addition, a small domain composed of a β-sheet that may also include some α-helical insertions [3]. This family includes hydroxymethylpyrimidine kinases (HMPK, EC 2.7.1.49), pyridoxal kinases (PLK, EC 2.7.1.35) and hydroxyethylthiazole kinases (THZK, EC 2.7.1.50), enzymes that participate in the biosynthesis *de novo* of active (phosphorylated) forms of vitamin B1 (thiamine pyrophosphate) and B6 (PLP) in bacteria, essential cofactors for several enzymes involved in the metabolism of amino acids [4] and carbohydrates [5]. Although at first sight it seems that these enzymes can be easily split in different groups based on their substrate specificities, the fact is that several members of this protein family are able to phosphorylate more than one substrate. Some enzymes have been classified based either in enzyme kinetics or in structural comparisons as being specific for HMP, such as *Escherichia coli* and *Salmonella typhimurium* HMPKs derived from the *thiD* gene [6,7], or for pyridoxal (PL), like the protein codified by the *pdxY* gene of *E. coli*, where the PL moiety is so tightly bound to the enzyme that it is not released unless the protein is subjected to denaturation [8]. However, several bifunctional HMPK/PLK enzymes that are able to phosphorylate both PL and HMP have been described, such as the protein encoded by the *pdxK* gene of *E. coli* [9,10], *Trypanosoma brucei* and *Plasmodium falciparum* [11], and the protein encoded by the *thiD* gene of *Bacillus subtilis* [12]. Also, the crystal structures of bifunctional kinases from *B. subtilis* and *Staphylococcus aureus* have been recently solved [13,14]. Both enzymes have been described as possessing significant activity towards both PL and the thiamine

precursor HMP, thus conferring to these enzymes a dual function in both the pyridoxal and thiamine biosynthesis pathways. Moreover, the *thiD* gene of *B. subtilis* has been found not to be essential for bacterial survival [15] and located outside the operon where thiamine biosynthetic genes are usually clustered [13,16]. Although this evidence sheds light about the emergence of PLK activity inside the group of HMPK enzymes as a late event of convergence in the evolution of these kinases, there is still no proof that the trait of their last common ancestor corresponded to the specificity towards HMP.

In order to test the evolutionary history of the ATP-dependent vitamin kinases and the role of bifunctionality in the appearance of the HMPK and PLK activities, we performed phylogenetic analysis, ancestral enzyme reconstruction, molecular modeling and docking, and we also tested experimentally the kinetic features of the resurrected last common ancestor of the HMPK enzymes. Our results show that this ancestor is able to phosphorylate both PL and HMP, but has a 8-fold higher preference for HMP. Also, the expected catalytic rate at near-physiological concentrations is higher for HMP than PL. These results provide strong evidence of how PLK activity emerged during the evolution of this protein family.

2. Materials and methods

2.1. Structural alignment and phylogenetic tree of PLKs and HMPKs

Crystal structures of PLKs and HMPKs (PDB ID: 2I5B, 4C5L, 1JXJ, 1UB0, 1TD2, 2YXT, 1LHP and 2DDM) were structurally aligned using STAMP [17], and the resulting alignment was used to reconstruct a phylogenetic tree with MrBayes 3.1.2 [18]. For the analysis we performed 2 runs with 4 chains per run and heating temperature set to 0.2 for 1×10^6 generations using the mixed analysis of fixed amino acid substitution models with fixed site-specific rates. Samples were collected every 100 steps and the initial 40% of samples were removed before summarizing trees and parameters, ensuring that the average standard deviation of split frequencies was less than 0.01 upon convergence. The consensus tree was processed in Dendroscope 3.2.10 [19].

2.2. Multiple sequence alignment

Sequences were collected from the non-redundant protein database (nr) using a PSI-BLAST algorithm with 3 iterations through Protein BLAST Server (blast.ncbi.nlm.nih.gov), using HMPKs and PLKs with known structure as templates (PDB ID: 2I5B, 1JXH, 1TD2, 2F7K and 2DDM). An initial multiple sequence alignment (MSA) was constructed using MAFFT [20], then redundant sequences were removed using QR Sequence tool of Multiseq in VMD [17] with a PID cutoff of 88%. A second MSA was constructed based on three-dimensional and secondary structure constraints using Promals3D [21]. Misaligned C-terminal, N-terminal and loop positions were manually corrected based on the structural alignment obtained previously. The final MSA used in this work is available in the [Supplementary information](#).

2.3. HMPKs phylogenetic tree

Sequences of HMPK were selected from MSA and their phylogeny was inferred with MrBayes version 3.1.2 [18], where human PLK was used as outgroup. For the analysis we used WAG as the fixed model, based on its posterior probability of 1.0 in the mixed model analysis, and gamma-shaped rate variation across sites with a proportion of invariable sites. Runs were performed as previously described, but increasing the number of generations to 1×10^7 and

only removing the initial 30% of samples before summarizing trees and parameters. The average standard deviation of the split frequencies was less than 0.01 upon convergence.

2.4. Ancestral sequence reconstruction

We used a modified method described by Hall [22]. Briefly, ancestral sequences were inferred through a hierarchical Bayes approach implemented in MrBayes 3.1.2 with model and parameters used for phylogeny inference [23]. Each target node was constrained and probabilities for each amino acid were calculated in each position of alignment, where the amino acid with highest probability in each position were selected. Sequence gaps were inferred as described by Hall [22].

2.5. Gene synthesis, protein expression and purification

The gene of the last common ancestor was codon-optimized for expression in *E. coli* and synthesized by GENSCRIPT (Piscataway, NJ, USA), then cloned into a modified pET-28b vector and verified by DNA sequencing. *E. coli* BL21(DE3) were transformed and grown in LB broth containing 35 $\mu\text{g}/\text{mL}$ kanamycin at 37 °C until OD600 reached ~ 0.8 . Expression of the recombinant protein was induced with 1 mM of isopropyl- β -D-thiogalactopyranoside overnight. Cells were harvested by centrifugation, resuspended in binding buffer (50 mM Tris-HCl pH 7.6, 500 mM NaCl, 20 mM imidazole and 5 mM MgCl_2) and disrupted by sonication. After centrifugation (18514 g for 45 min), the soluble fraction was loaded onto a Ni^{2+} -NTA affinity column (HisTrap HP, GE Healthcare, UK). Protein was eluted with a linear gradient between 20 and 500 mM imidazole and fractions with enzyme activity were pooled and stored at 4 °C with 1 mM ATP and 5% glycerol. Enzyme purity was analyzed by SDS-PAGE stained with Coomassie blue.

2.6. Enzyme activity assays

HMPK activity was measured following the appearance of ADP with a coupled assay containing pyruvate kinase/lactate dehydrogenase. Briefly, enzyme preparation was mixed with reaction buffer containing 25 mM Tris-HCl pH 7.8, 0.8 U/mL of pyruvate kinase, 2.4 U/mL lactate dehydrogenase, 0.3 mM phosphoenolpyruvate (PEP), 125 mM KCl, 0.2 mM NADH, 10 mM ATP, 15 mM HMP and 15 mM MgCl_2 .

PLK activity was measured spectrophotometrically by following PLP formation at 388 nm. The reaction mixture consisted of 10 mM ATP, 15 mM PL, 15 mM MgCl_2 and 25 mM PIPES pH 6.5. An extinction coefficient of $6.22 \text{ mM}^{-1} \text{ cm}^{-1}$ for NADH and $2.886 \text{ mM}^{-1} \text{ cm}^{-1}$ for PLP were used, and the enzymatic unit (U) was defined as $\mu\text{mol min}^{-1}$. Both activities were measured at 37 °C.

2.7. Homology modeling and docking

Fifty models were constructed for each ancestral protein with MODELLER 8 [24]. The best 10 models were chosen based on DOPE potential, and its quality evaluated with PROSA2003 [25], Procheck [26] and VERIFY3D [27].

Docking assays were performed with AutodockVina 1.0 [28], the protonation state of the ionizable residues was calculated using the web server H++ [29] and partial charges were derived with Gasteiger method using AutoDockTools [30]. Hydrogens of HMP and PL substrates were added and optimized with Gaussian [31]. Docking results with the lowest interaction energy and the phosphoryl acceptor hydroxyl oriented towards the GXGD(C) motif were selected, since the aspartic acid (or cysteine) is considered the catalytic base for the phosphate group transfer in all ribokinase superfamily members [32].

3. Results and discussion

3.1. PLK and HMPK activity, evidence for bifunctional enzymes

The relationship between PLKs or HMPKs and bifunctional enzymes was analyzed by constructing a phylogenetic tree based on structural alignment (Fig. 1). The resulting dendrogram shows that enzymes HMPK or PLK appear as separate groups, while bifunctional enzymes cluster inside the HMPK group, as described previously [13,14]. Therefore, the HMPK group can be divided into specific HMPKs (*Thermus thermophilus* and *S. typhimurium*) and bifunctional HMPK/PLKs (*B. subtilis* and *S. aureus*) whereas the PLK group includes enzymes from prokaryotes (*pdxY* and *pdxK*) and eukaryotes (human and sheep). It is interesting to analyze the evolutionary history of the HMPK group, since the ability to phosphorylate PL within the group can be an event of convergent evolution, in which PLK activity reappears in the HMPK group as proposed by Newman et al. [13] or, conversely, the bifunctionality is the ancestral trait [33] and HMP specificity would be a recent event in the evolutionary history of these enzymes. These two hypotheses can be scrutinized by reconstructing the last common ancestor of the HMPK group (Fig. 1). The experimental resurrection of this ancestor and its kinetic characterization, regarding its ability to phosphorylate PL and HMP, will reveal if this ancestor was either specific for HMP (convergent evolution) or bifunctional (ancestral conservation of PLK activity) (see Fig. 2).

3.2. HMPK phylogeny and ancestral sequence reconstruction

In order to reconstruct the last common ancestor of HMPK enzymes, we established the phylogeny of bacterial HMPKs by bayesian method (Fig. S1), using the human PLK as outgroup to root the tree. As expected, these enzymes cluster in two groups; one that contains the sequences of HMPK/PLK from *B. subtilis* and *S. aureus* (putative bifunctional group) with a posterior probability (pp) of 1.0, and other that includes specific HMPK enzymes from *S. typhimurium* and *T. thermophilus* (HMPK specific group) with a support of 0.93. The posterior probability for the node of the last common ancestor for all HMPKs was 0.92. We inferred the ancestral sequences of the last common ancestor of HMPKs enzymes (common ancestor, ancC), of the specific HMPK enzymes (ancS) and of the putative bifunctional group of enzymes (ancB) using the hierarchical bayesian approach [23]. Sequences of ancestors ancS and ancB presented a very high support of posterior probability of reconstruction; ancS (Fig. S2A) shows 94% of residues over pp = 0.90 (254 of 268 aa) (Fig. S2A), while ancB (Fig. S2B) displays 92% of residues over pp = 0.90 (245 of 266 aa) (Fig. S2B). AncC was more ambiguously reconstructed (mean pp = 0.84 overall)

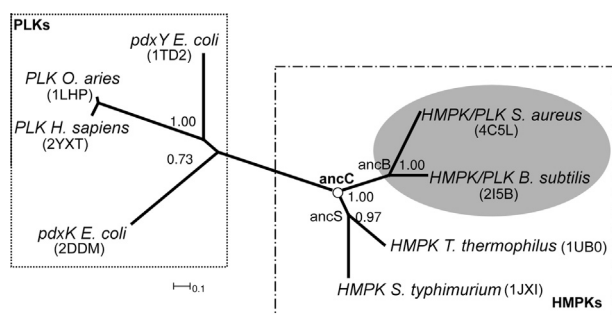


Fig. 1. Dendrogram of PLK and HMPK activities. Enzymes with determined structure were structurally aligned and phylogeny was reconstructed. Enzymes with dual activity are shadowed in gray, while squares indicate PLK and HMPK groups. The last common ancestor of HMPK group (ancC), ancestor of HMPK specific (ancS) and ancestor of bifunctional enzymes (ancB) are labeled.

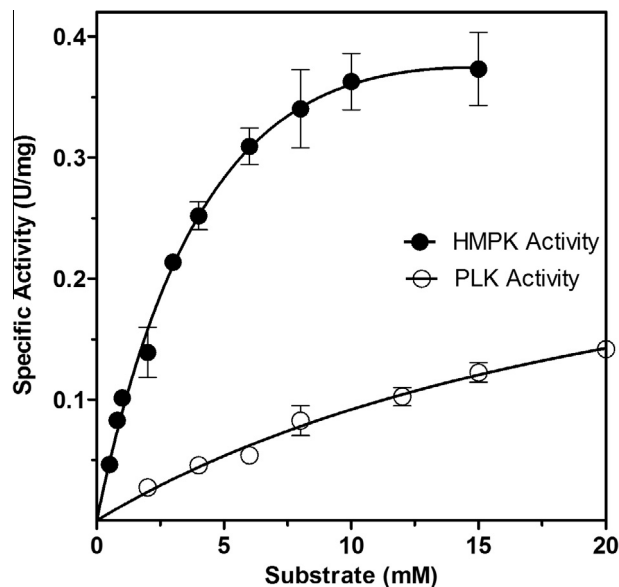


Fig. 2. Dependence of the enzymatic activity of the last common ancestor (ancC) as a function of the concentration of HMP or PL.

(Fig. S2C), with 55% of its residues having pp \geq 0.90 (Fig. S2C). Nevertheless, the reconstruction of ligand-contacting sites was considerably more robust (mean pp = 0.94) (Table S1). This mean posterior probability is a good support for the validity of the ancestral sequence reconstruction, as reported in other publications [34,35].

3.3. Experimental resurrection of the last common ancestor of HMPK enzymes

We expressed and purified the last common ancestor of HMPK enzymes (ancC). The ancestral enzyme showed activity with both substrates (PL and HMP) with an optimum pH of 8.0 and 6.5 for the HMPK and PLK activity, respectively (Fig. S3). The K_m for ATP using HMP as co-substrate was 1.00 mM; this substrate also shows a slight inhibition with a K_i of 44 mM. A K_m of 7 mM was obtained when HMP was used as substrate, a value very similar to those reported for current enzymes, like the one present in *B. subtilis* (K_m of 2 mM) [12]. At HMP concentrations above 10 mM, substrate inhibition was observed with a K_i of 31 mM. When employing PL as substrate a K_m of 28 mM, a k_{cat} of 0.2 s⁻¹ and a K_i of 85 mM for substrate inhibition (Table 1) were obtained. This K_m value for PL is much higher than the ones reported for current enzymes (Table S2); both the *pdxK*-derived enzyme from *E. coli* as well as the *thiD*-derived enzyme from *B. subtilis* presents K_m values for PL in the micromolar range. Comparison of the k_{cat}/K_m values for HMP and PL show that ancC displays a HMP preference of 8 times over PL, which results mainly due to its high K_m for PL. This suggest that the catalytic activity of ancC with PL corresponds to catalytic promiscuity since, due to the high chemical reactivity of this metabolite with proteins [36], its intracellular concentration in

Table 1
Kinetic constants for ancC.

Parameters	HMP	PL	Mg-ATP
K_m (mM)	7 ± 2	28 ± 9	1 ± 0.2
k_{cat} (s ⁻¹)	0.40 ± 0.06	0.20 ± 0.04	0.080 ± 0.005
K_i (mM)	31 ± 17	85 ± 54	44 ± 19
k_{cat}/K_m (M ⁻¹ s ⁻¹)	57	7	80

E. coli is 27 μM total and only 10 μM free, not-protein bound [37], thus suggesting that it would not be physiologically relevant. In fact, in this concentration range, the catalytic rate is significantly higher when HMP is used as substrate compared to PL. This idea is reinforced by the fact that if k_{cat}/K_m values are calculated from initial velocity slopes the difference between HMP versus PL as substrates is 11-fold (Fig. S4). In the case of the bifunctional enzyme from *B. subtilis*, there is a ~ 4 -fold preference of PL over HMP that is accompanied by a K_m for PL in the micromolar range, thus exhibiting a higher catalytic rate for PL for the expected intracellular concentrations (Fig. S5) [12]. Moreover, in vivo covalent modification by analogs of the antibacterial compound rugulactone of the enzyme encoded by the *thiD* gene of *S. aureus*, which has a 65-fold preference for PL over HMP and possess a millimolar K_m for HMP [14], leads to inhibition of thiamine biosynthesis [38], thus providing indirect evidence that the millimolar K_m values observed for these enzymes are of biological relevance. Altogether, these results support that the last common ancestor of HMPK preferred HMP over PL under near-physiological conditions due to significantly different catalytic efficiencies, although it was capable of phosphorylating both substrates. Therefore, the PLK activity must have been tuned during the evolutionary history of the family in order to reappear later in the HMPK group as a convergent evolutionary event, a process that may have been favored in the absence of selective pressure as evidenced by the location of genes encoding current bifunctional enzymes outside the thiamine biosynthesis operon [12,14], and in agreement with the proposal by Newman et al. [13]. The fact that the sequence identity between modern

bifunctional enzymes and HMPKs (on average, 38%) is higher than when compared with PLKs (on average, 17%), along with the noticeable differences in the residue composition of the binding sites of both types of kinases (Fig. 4), reinforces the idea that the ability of phosphorylating pyridoxal emerged in HMPKs as a convergent, independent process.

3.4. Molecular modeling of ancestral HMPKs

In order to find the structural determinants that account for the specificity of ancC by HMP, we built homology models for this ancestor and for the ancestral ancS and ancB sequences (Table S3), and then performed docking assays with PL and HMP. Docking results showed the orientation for both ligands were similar to those present in crystallized enzymes. For ancC, the closest residues to PL and HMP moieties are: Met81, Ser13, Val43, Gln45, His50, His211, Val108 and Cys215 (Fig. 3). For ancB these residues are Met80, Thr12, Val42, Gln44, Val49, His209, Val107 and Cys213, and for ancS with HMP the residues are the following: Met81, Thr12, Val42, Gln44, Val49, His210, Val108 and Cys214. A comparison of the active sites of current bifunctional HMPK/PLK and specific HMPKs enzymes, shows that the main differences between them correspond to residues in positions 44 and 52 where the bifunctional enzyme from *B. subtilis* has Met and His respectively, while in the specific HMPK enzyme from *S. typhimurium* Gln44 and Val49 are found in equivalent positions [13]. In the ancC ancestor we found glutamine in position 45 and histidine in position 50, so that Gln44 can explain the HMP specificity and His50 could

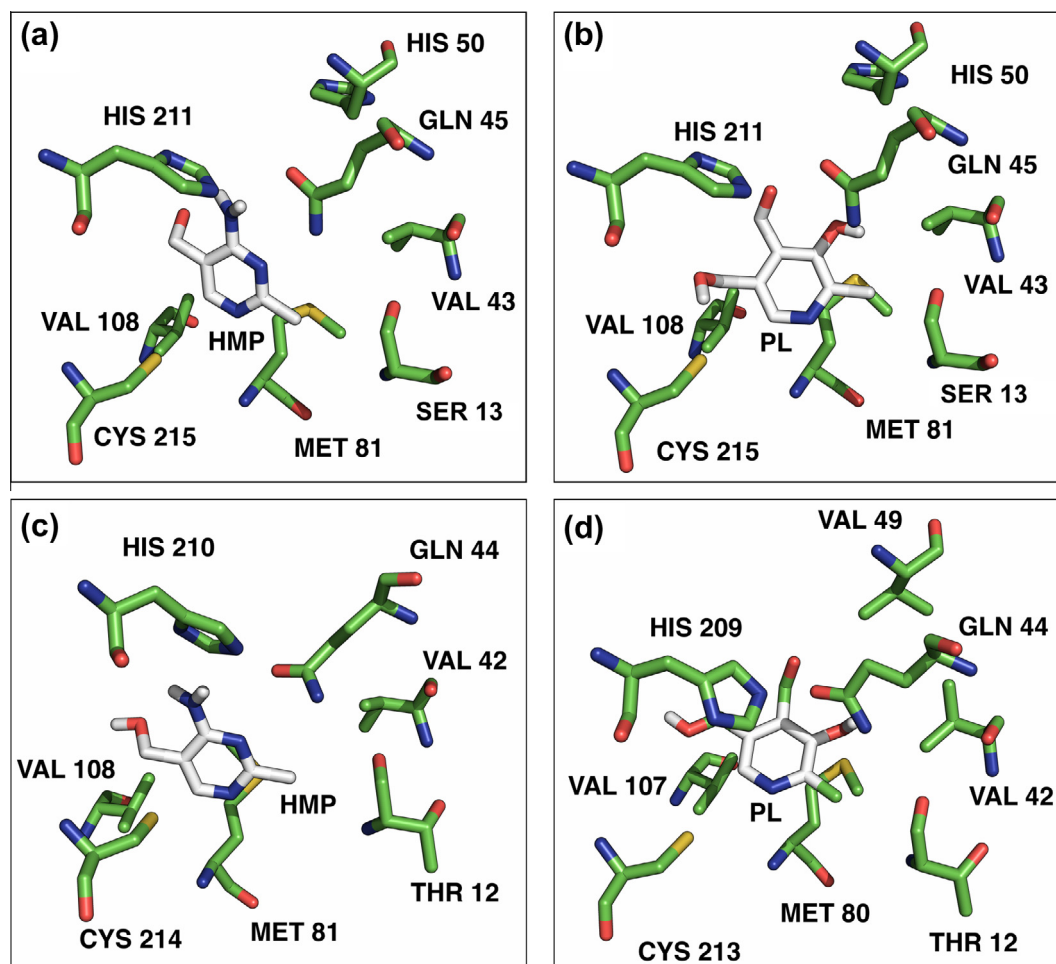


Fig. 3. Docking assays with HMP and PL on the binding site of the ancestral enzymes. (a) AncC with HMP. (b) ancC with PL. (c) ancS with HMP. (d) ancB with PL.

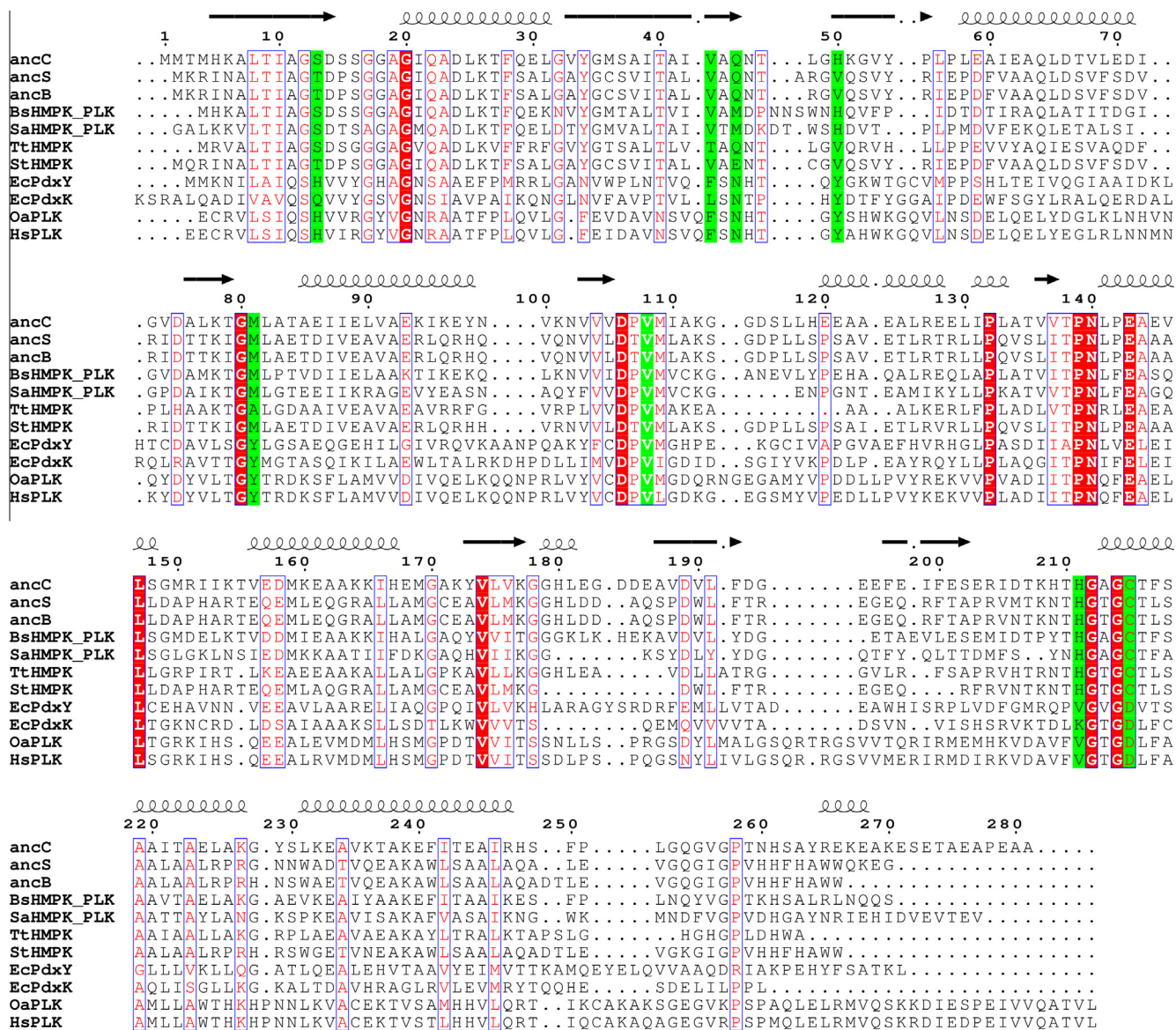


Fig. 4. Structure-based alignment of models of ancestral enzymes and structures of modern protein. BshHMPK_PLK (*B. subtilis*, PDB: 215B), SaHMPK_PLK (*S. aureus*, PDB: 4C5L), TtHMPK (*T. thermophilus*, PDB: 1UB0), StHMPK (*S. thymophilum*, PDB: 1JX1), EcPdxY (*E. coli*, PDB: 1TD2), EcPdxK (*E. coli*, PDB: 2DDM), OaPLK (*O. aries*, PDB: 1LHP), HsPLK (*H. sapiens*, PDB: 2YXT). Alignment was performed with Stamp. The similarity coloring scheme is based on blosum62 matrix. Residues of the active site of ancC are highlighted in green and the secondary structure elements are annotated based on the homology model of ancC.

account for the promiscuous activity with PL, highlighting the role of Gln in HMP specificity. Moreover, ancS presents glutamine (Gln44) and valine (Val49) in the corresponding positions, which also stressed the importance of glutamine for this specificity. Interestingly, ancB presents glutamine (Gln44) and valine (Val49) in the corresponding positions, which does not match with the active sites of current bifunctional enzymes since *S. aureus* and *B. subtilis* enzymes presents Met and His in those positions (Fig. 4). A search for residue conservation at the active site inside the group of bifunctional enzymes shows that some of them (enzymes from *Megamonas funiformis*, *Caldalkalibacillus thermarum*, *Exiguobacterium sp.*, *Acetonema longum*) retain a glutamine residue like the ancestral enzyme. These enzymes are found in different branches of the bifunctional group in the phylogenetic tree (Fig. S1), therefore suggesting some of these members may also use HMP as the preferred substrate.

4. Conclusions

The experimental resurrection of the last common ancestor of the HMPK group showed that probably this protein was not able to use PL under physiological conditions. Therefore, the PLK activity present in the current bifunctional enzymes must have appeared in a convergent event independently of the PLK activity of *pdxY* and *pdxK* genes, as was proposed by Newman et al. [13]. The ability of ancC to phosphorylate pyridoxal, which is 8-times less preferred than the phosphorylation of HMP, was considered by us as a promiscuous activity, since its high K_m value for PL would not be physiologically relevant. The promiscuous activity of enzymes has been proposed as the starting point for new activities [1,39]. In our case, this trait would have allowed the appearance of an activity already present in this family, in a convergent and independent manner.

Acknowledgment

This work was supported by Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt, Chile) Grant 1110137.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.febslet.2014.06.033>.

References

- [1] Allewell, N.M. (2012) Thematic minireview series on enzyme evolution in the post-genomic era. *J. Biol. Chem.* 287, 1–2.
- [2] Sigrell, J.A., Cameron, A.D., Jones, T.A. and Mowbray, S.L. (1998) Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: insights into a new family of kinase structures. *Structure* 6, 183–193.
- [3] Rivas-Pardo, J.A., Herrera-Morande, A., Castro-Fernandez, V., Fernandez, F.J., Vega, M.C. and Guixé, V. (2013) Crystal structure, SAXS and kinetic mechanism of hyperthermophilic ADP-dependent glucokinase from *Thermococcus litoralis* reveal a conserved mechanism for catalysis. *PLoS One* 8, e66687.
- [4] Eliot, A.C. and Kirsch, J.F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* 73, 383–415.
- [5] Pohl, M., Sprenger, G.A. and Müller, M. (2004) A new perspective on thiamine catalysis. *Curr. Opin. Biotechnol.* 15, 335–342.
- [6] Cheng, G., Bennett, E.M., Begley, T.P. and Ealick, S.E. (2002) Crystal structure of 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase from *Salmonella typhimurium* at 2.3 Å resolution. *Structure* 10, 225–235.
- [7] Reddick, J.J., Kinslanda, C., Nicewongera, R., Christianb, T., Downs, D.M., Winkler, M.E. and Begley, T.P. (1998) Overexpression, purification and characterization of two pyrimidine kinases involved in the biosynthesis of thiamin: 4-amino-5-hydroxymethyl-2-methylpyrimidine kinase and 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase. *Tetrahedron* 54, 15983–15991.
- [8] Safo, M.K., Musayev, F.N., Hunt, S., di Salvo, M.L., Scarsdale, N. and Schirch, V. (2004) Crystal structure of the PdxY Protein from *Escherichia coli*. *J. Bacteriol.* 186, 8074–8082.
- [9] Mizote, T. and Nakayama, H. (1989) Purification and properties of hydroxymethylpyrimidine kinase from *Escherichia coli*. *Biochim. Biophys. Acta* 991, 109–113.
- [10] Yang, Y., Tsui, H.C., Man, T.K. and Winkler, M.E. (1998) Identification and function of the pdxY gene, which encodes a novel pyridoxal kinase involved in the salvage pathway of pyridoxal 5'-phosphate biosynthesis in *Escherichia coli* K-12. *J. Bacteriol.* 180, 1814–1821.
- [11] Wrenger, C., Eschbach, M.-L., Müller, I.B., Laun, N.P., Begley, T.P. and Walter, R.D. (2006) Vitamin B1 de novo synthesis in the human malaria parasite *Plasmodium falciparum* depends on external provision of 4-amino-5-hydroxymethyl-2-methylpyrimidine. *Biol. Chem.* 387, 41–51.
- [12] Park, J.-H., Burns, K., Kinslanda, C. and Begley, T.P. (2004) Characterization of two kinases involved in thiamine pyrophosphate and pyridoxal phosphate biosynthesis in *Bacillus subtilis*: 4-amino-5-hydroxymethyl-2-methylpyrimidine kinase and pyridoxal kinase. *J. Bacteriol.* 186, 1571–1573.
- [13] Newman, J.A., Das, S.K., Sedelnikova, S.E. and Rice, D.W. (2006) The crystal structure of an ADP complex of *Bacillus subtilis* pyridoxal kinase provides evidence for the parallel emergence of enzyme activity during evolution. *J. Mol. Biol.* 363, 520–530.
- [14] Nodwell, M.B., Koch, M.F., Alte, F., Schneider, S. and Sieber, S.A. (2014) A sub-family of bacterial ribokinases utilizes a hemithioacetal for pyridoxal phosphate salvage. *J. Am. Chem. Soc.* 136, 4992–4999.
- [15] Kobayashi, K., Ehrlich, S.D., Albertini, A., et al. (2003) Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. USA* 100, 4678–4683.
- [16] Petersen, L.A. and Downs, D.M. (1997) Identification and characterization of an operon in *Salmonella typhimurium* involved in thiamine biosynthesis. *J. Bacteriol.* 179, 4894–4900.
- [17] Roberts, E., Eargle, J., Wright, D. and Luthey-Schulten, Z. (2006) MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* 7, 382.
- [18] Ronquist, F. and Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- [19] Huson, D.H. and Scornavacca, C. (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 61, 1061–1067.
- [20] Katoh, K., Misawa, K., Kuma, K. and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- [21] Pei, J., Kim, B.-H. and Grishin, N.V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* 36, 2295–2300.
- [22] Hall, B.G. (2006) Simple and accurate estimation of ancestral protein sequences. *Proc. Natl. Acad. Sci. USA* 103, 5431–5436.
- [23] Huelsenbeck, J.P. and Bollback, J.P. (2001) Empirical and hierarchical Bayesian estimation of ancestral states. *Syst. Biol.* 50, 351–366.
- [24] Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815.
- [25] Sippl, M.J. (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins Struct. Funct. Genet.* 17, 355–362.
- [26] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.
- [27] Bowie, J.U., Lüthy, R. and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164–170.
- [28] Trott, O. and Olson, A.J. (2009) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461.
- [29] Gordon, J.C., Myers, J.B., Folta, T., Shoja, V., Heath, L.S. and Onufriev, A. (2005) H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* 33, W368–W371.
- [30] Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. and Olson, A.J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791.
- [31] M.J. Frisch et al., *Gaussian09 Revision A.1., Gaussian 09, Revis. A.02, 2009*.
- [32] Gandhi, A.K., Ghatge, M.S., Musayev, F.N., Sease, A., Aboagye, S.O., di Salvo, M.L., Schirch, V. and Safo, M.K. (2009) Kinetic and structural studies of the role of the active site residue Asp235 of human pyridoxal kinase. *Biochem. Biophys. Res. Commun.* 381, 12–15.
- [33] Khersonsky, O. and Tawfik, D.S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* 79, 471–505.
- [34] Eick, G.N., Colucci, J.K., Harms, M.J., Ortlund, E.a. and Thornton, J.W. (2012) Evolution of minimal specificity and promiscuity in steroid hormone receptors. *PLoS Genet.* 8, e1003072.
- [35] Hobbs, J.K., Shepherd, C., Saul, D.J., Demetras, N.J., Haaning, S., Monk, C.R., Daniel, R.M. and Arcus, V.L. (2012) On the origin and evolution of thermophily: reconstruction of functional precambrian enzymes from ancestors of *Bacillus*. *Mol. Biol. Evol.* 29, 825–835.
- [36] Christensen, H.N. (1958) Three schiff base types formed by amino acids, peptides and proteins with pyridoxal and pyridoxal-5-phosphate 1. *J. Am. Chem. Soc.* 80, 99–105.
- [37] Fu, T.F., di Salvo, M. and Schirch, V. (2001) Distribution of B6 vitamers in *Escherichia coli* as determined by enzymatic assay. *Anal. Biochem.* 298, 314–321.
- [38] Nodwell, M.B., Menz, H., Kirsch, S.F. and Sieber, S.A. (2012) Rugulactone and its analogues exert antibacterial effects through multiple mechanisms including inhibition of thiamine biosynthesis. *ChemBioChem* 13, 1439–1446.
- [39] Khersonsky, O., Roodveldt, C. and Tawfik, D.S. (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* 10, 498–508.