FEBS Letters xxx (2014) xxx–xxx

journal homepage: www.FEBSLetters.org

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

Victor Castro-Fernandez [1], Felipe Bravo-Moraga [1], 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 xxxx

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.

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 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

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/

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, 1JXI, 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 aminoacid 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 μg/mL kanamicin 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 mM^{-1} cm^{-1} for NADH and 2.886 mM^{-1} cm^{-1} for PLP were used, and the enzymatic unit (U) was defined as μmol 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].

Please cite this article in press as: Castro-Fernandez, V., et al. Emergence of pyridoxal phosphorylation through a promiscuous ancestor during the evolution of hydroxymethyl pyrimidine kinases. FEBS Lett. (2014), http://dx.doi.org/10.1016/j.febslet.2014.06.033