International Journal of Biology and Life Sciences
ISSN: 2957-9511 | Vol. 13, No. 3, 2026

Bioinformatics Analysis of Pyruvate Kinase of Rhodococcus Opacus

Suhang Yu *

School of Biology Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei, China
*** Corresponding author Email:** 1421151192@qq.com

Abstract: In order to deeply explore the key regulation mechanism of carbon metabolism of Rhodococcus, the pyruvate kinase (PyK) of Rhodococcus opacus was systematically analyzed by bioinformatics. The results show that the enzyme has a typical PK-like superfamily domain, which can form a homotetramer spatial conformation, and its ADP binding site with metal ions is highly conserved in evolution. Multi-sequence alignment and phylogenetic analysis further revealed that it was conserved in many species, and suggested that it belonged to Cluster II PyK, and its activity did not depend entirely on potassium ions. In addition, the prediction of phosphorylation sites showed that the enzyme had 44 potential modification sites, indicating that its activity might be regulated by complex post-translational modifications. This study provides an important theoretical basis and a new transformation strategy for optimizing the lipid synthesis and environmental remediation ability of Rhodococcus through metabolic engineering.

Keywords: Rhodococcus Opacus; Pyruvate Kinase; Protein Structure; Bioinformatics Analysis.

1. Introduction

Rhodococcus opacus is a kind of Gram-positive actinomycete widely distributed in various environments, which has attracted much attention for its excellent metabolic diversity and strong environmental adaptability.[1]. As an important environmental remediation agent, the strain can degrade a variety of organic pollutants and is a kind of "cell factory" with great biosynthetic potential [2, 3]. Previous studies have shown that Rhodococcus opacus has the ability to efficiently synthesize and accumulate triacylglycerol (TAG), and its lipid content can reach more than 80% of the dry weight of bacteria. It shows broad application prospects in the fields of microbial production of biodiesel, high-value oil chemicals, and special structure lipids [2]. Therefore, it is of great significance to deeply analyze and optimize the central carbon metabolism network of Rhodococcus for improving its bio-manufacturing efficiency and expanding its industrial application.

Pyruvate Kinase (PyK), as a key rate-limiting step in the glycolytic pathway, catalyzes the irreversible transformation from phosphoenolpyruvate (PEP) to pyruvate, accompanied by the formation of ATP [4]. This reaction not only provides energy and precursors for cells, but also directly regulates the distribution of carbon metabolic flow in the center through its catalytic efficiency, which deeply affects the balance between energy metabolism and biosynthesis [5]. Therefore, PyK is often identified as the core transformation target in microbial metabolic engineering. PyK can directionally optimize the flow direction of metabolites by precisely regulating its gene expression or manipulating its allosteric regulatory network, thus significantly improving the synthesis efficiency of target products (such as lipids, organic acids, etc.) [6]. To sum up, the in-depth analysis of the enzyme structure and regulation mechanism of this strain will lay an important theoretical foundation for building an efficient microbial cell factory.

This study focuses on PyK of Rhodococcus opacus, and systematically explains its amino acid sequence characteristics, three-dimensional spatial conformation, phylogenetic relationship, and post-translational modification potential by bioinformatics analysis. Through multi-tool joint analysis, the conserved functional domain and key active residues of the enzyme were accurately identified, and the distribution of phosphorylation modification sites was predicted, and then the possible activity regulation network was evaluated. This study not only provides a molecular basis for the rational design of this key enzyme but also provides an innovative idea for reshaping the carbon metabolic flow of Rhodococcus through metabolic engineering, which is of great value for expanding its application prospects in the fields of bio-manufacturing and environmental remediation.

2. Materials and Methods

2.1. Experimental Materials

This study downloaded the target sequence WP_005244435.1 of PyK from the NCBI database. In addition, the multi-species homologous sequences used for subsequent comparison and evolutionary analysis are derived from UniProt database, including M1 and M2 isozymes of Homo sapiens (P14618-2, P14618), as well as rabbits (P11974), Aedes aegypti (Q16F38), Saccharomyces cerevisiae (1A3W), Leishmania (3HQP), Cryptosporidium (Q5CSM7), Mycobacterium tuberculosis (5W5B) and PyK sequences of Escherichia coli (1PKY).

2.2. Experimental Methods

The information on databases, servers, and online software tools used in bioinformatics analysis in this study is presented in Table 1.

2.2.1. Domain and Secondary and Tertiary Structure Analysis of Rhodococcus PyK

Firstly, the domain prediction analysis of the Rhodococcus PyK protein sequence was carried out using the SMART online tool, and the FASTA format sequence of the target protein was submitted to the SMART platform. In the Normal mode, the parameter options, such as Outlier homologues and homologues of known structure, Pfam domains, signal

peptides, and internal repeats, were checked. After the system completes the analysis, the prediction result with the highest confidence is selected for recording and saving. In addition, the domain of its protein sequence was further predicted by the Pfam online prediction website.

In this study, the secondary and tertiary structures of Rhodococcus PyK were predicted based on the SWISS-MODEL online platform. In the prediction of monomer structure in protein, the AlphaFold prediction model with the highest sequence consistency (gene identification: A0A149ZLB6_9NOCA) was selected as the template to construct the three-dimensional structure, and the PDB format file was downloaded for multi-sequence alignment visualization of ESPript 3.0. For the simulation of the tetramer structure, the template of mycobacterial PyK with the best QSQE score is selected for homologous modeling, and on this basis, the specific amino acid residues interacting with ATP molecules in Rhodococcus PyK are further identified and labeled.

Table 1. Bioinformatics Analysis Softwares and Websites Used

Use	Software/Website Name	Source Website Information
Protein sequence download	NCBI	https://www.ncbi.nlm.nih.gov/protein/
Protein sequence download	Uniprot	https://www.uniprot.org/
Protein domain prediction	SMART	https://smart.embl-heidelberg.de/
Protein domain prediction	Pfam	https://pfam-legacy.xfam.org/
Spatial structure prediction of protein	SWISS-MODEL	https://swissmodel.expasy.org/
The comparative analysis of multiple sequences	DNAman 6.0	https://www.lynnon.com/dnaman.htmL/
The visualization of multi-sequence comparison	ESPript 3.0	https://espript.ibcp.fr/ESPript/ESPript/
Construction of phylogenetic tree	MEGA 6.0	https://www.megasoftware.net/
Prediction of Phosphorylation site	NetPhos 3.1	https://services.healthtech.dtu.dk/services/NetPhos-3.1/

2.2.2. The Comparative Analysis of Multi-sequence and Phylogenetic Tree Construction of Rhodococcus PyK Homologous Genes

In order to analyze the evolutionary relationship and conservation of Rhodococcus PyK, multiple-sequence comparison and phylogenetic analysis were carried out in this study. At first, DNAman 6.0 was used to compare the homologous sequences in FASTA format from NCBI and UniProt databases, and a result file in ALN format was generated. This file and the structural model predicted by SWISS-MODEL (PDB format) were submitted to the ESPript 3.0 server, and the comparison map was generated by adjusting the visualization parameters. Finally, PowerPoint was used to graphically mark the key domains and ATP-binding residues. The Neighbor-Joining Tree in MEGA 6.0 software is used to construct the phylogenetic tree. All the FASTA sequences mentioned above are imported into the software, the default parameters are selected, and the generated tree diagram is also annotated and beautified using PowerPoint.

2.2.3. Prediction and Analysis of Phosphorylation Sites of Rhodococcus PyK

The amino acid sequence of the Rhodococcus PyK protein was submitted to the NetPhos 3.1 online website tool in FASTA format, and three sites, serine, threonine, and tyrosine, were selected. The threshold was set to 0.5, and the task was submitted. After saving the image, PowerPoint was used to modify the font and annotate the information.

3. Results and Analysis

3.1. Protein Structure Characteristics of Rhodococcus PyK

The results showed that the sequence of Rhodococcus PyK protein was composed of 511 amino acids, and it also had PK-like superfamily domain (Figure 1a), including PyK barrel domain located at amino acids 22-362 and PyK C-terminal located at amino acids 373-485. E-valve is $4.80e^{-122}$ and $2.80e^{-24,}$ respectively (Figure 1a, b).

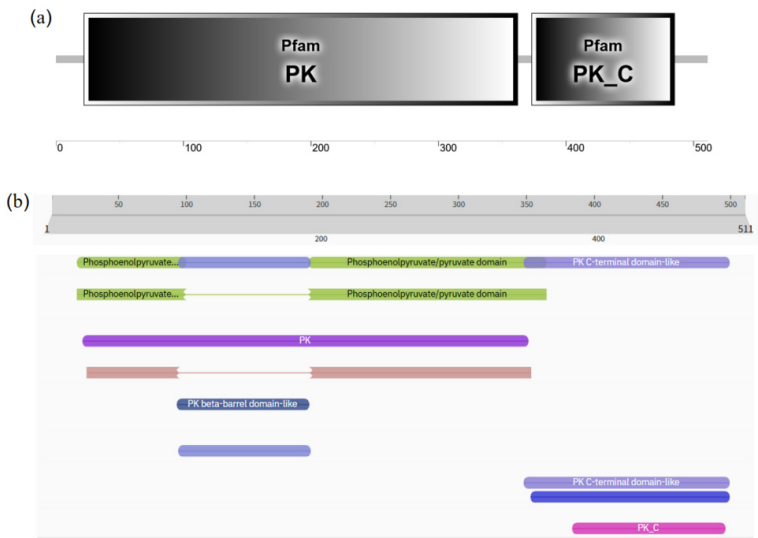


Figure 1. Sequence Domain Characteristics of the PyK Protein from Rhodococcus

The monomer subunit of PyK mainly includes four different domains: A, B, C, and N [7]. The monomer structure of Rhodococcus PyK constructed by the Swiss-Model website has 98.43% homology with the AlphaFold template sequence, and the GMQE score is 0.91. According to the spatial structure prediction of protein, the monomer structure of Rhodococcus PyK can be divided into four distinct parts: a cylindrical A-domain with a large upper middle, a flaky B-domain with a small top, a spiral C-domain with a small lower middle, and an N-domain tail at the bottom (Figure 2a).

PyK generally exists in the form of a homotetramer in organisms, and its activation generally requires potassium ions, dimethyl metal ions (magnesium or manganese), and ADP [7]. The homology between the tetramer structure of Rhodococcus PyK constructed by the Swiss-Model website and the template sequence of binding Mycobacterium PyK is 75.32%, and the GMQE score is 0.91. According to the spatial structure prediction of protein, Rhodococcus PyK tetramer can interact with multiple potassium ions and ADP, and has multiple amino acid active sites associated with it (Figure 2b).

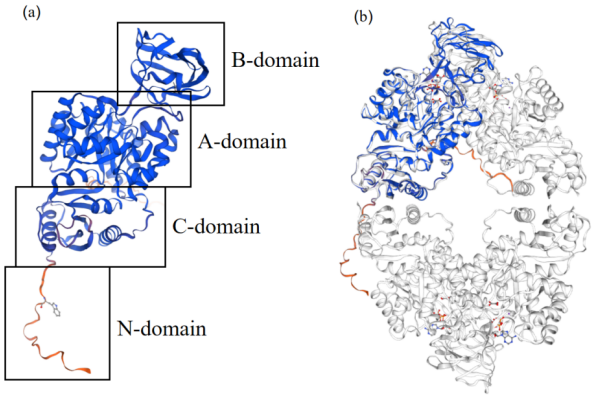


Figure 2. Spatial Structure Characteristics of Rhodococcus PyK

Note: (a) Spatial structure characteristics of Rhodococcus PyK monomer, and the box is used to mark the location of different domains. (b) Spatial structure characteristics of Rhodococcus PyK tetramer, and different colors are used to highlight the position of the pyruvate protein monomer in the tetramer.

3.2. Conservation and Phylogenetic Characteristics of PyK Homologous Genes in Rhodococcus

Based on the results of the three-dimensional structure model and multi-sequence comparison, the secondary structure composition and conserved characteristics of Rhodococcus PyK were further analyzed in this study. The protein consists of 16 α helices and 21 β folds, and its A domain consists of two segments: 21–91 (A1) and 191–359 (A2), forming a typical α/β barrel topology. The B domain is rich in β -β-sheet and located between β3 and α4 in the A domain. C domain (after amino acid 360) is also composed of α helix and β sheet, while the N-terminal region is located before residue 20 (Figure 3). The comparative results show that the key secondary structural units of PyK (such as α helix and β sheet) are highly conserved among different species, and its ADP and metal ion binding sites show strict sequence conservation from Rhodococcus to humans (Figure 3).

In the phylogenetic tree based on KDM5 sequences of 10 representative species, it was found that the enzyme protein can be clearly divided into four groups according to the evolutionary relationship: animal kingdom, fungus kingdom, protist kingdom, and monera kingdom (Figure 4). Among them, the homologous sequences of Rhodococcus PyK and Mycobacterium tuberculosis first clustered into one branch, and then together with other groups formed a complete evolutionary topology (Figure 4). In addition, phylogenetic analysis shows that PyK has high sequence conservation in the evolution process (Figure 4).

3.3. Distribution of Phosphorylation Sites of Rhodococcus PyK

It has been confirmed that protein phosphorylation can regulate the homotetramer assembly of PyK [8]. In this study, 44 potential modification sites, including 18 serine, 25 threonine, and one tyrosine, were identified in Rhodococcus PyK through the prediction analysis of phosphorylation sites (Figure 5). Further analysis shows that the overall prediction reliability of the serine site is high, while the score of the tyrosine site is close to the threshold value of 0.5, and its modification possibility is low.

4. Discussion

Rhodococcus plays an important role in industrial biotechnology, especially in the degradation of organic pollutants and microbial lipid synthesis [3]. PyK is the key rate-limiting enzyme in the glycolytic pathway, and its catalytic activity directly regulates the flux and direction of intracellular carbon metabolic flow [4]. In this study, the structural and functional characteristics of PyK from Rhodococcus opacus were systematically analyzed by comprehensive bioinformatics analysis, which provided a theoretical basis for optimizing the industrial performance of Rhodococcus opacus by metabolic engineering.

Structural analysis showed that Rhodococcus PyK showed high evolutionary conservation in monomer folding mode, tetramer assembly, and ADP, potassium ion, and divalent metal ion binding sites. However, the phylogenetic position of the enzyme does not follow the typical eukaryotic-prokaryotic dichotomy pattern, but belongs to two different protein clusters [9, 10]. Studies show that cluster I enzyme has conserved glutamic acid residues at the junction of A and B domains, and its activity depends entirely on potassium ions. However, the cluster II enzyme is lysine residue at the corresponding site, and its activity is less dependent on potassium ions [11,12]. This study found that Rhodococcus PyK is lysine at this key site, which indicates that it may belong to cluster II type. This feature provides a new idea for designing metabolic regulation strategies that are not completely dependent on potassium ions. Post-translational modification analysis shows that Rhodococcus PyK has 44 potential phosphorylation sites, of which serine and threonine sites are dominant, suggesting that its activity may be finely regulated by the typical serine/threonine protein kinase system in prokaryotes. Studies in eukaryotic systems have confirmed that phosphorylation can regulate PyK function by affecting tetramer assembly, allosteric regulation, and substrate binding, and similar regulatory mechanisms may also exist in Rhodococcus [8,13]. This provides a new idea for regulating the sugar metabolism pathway of strains through post-translational modification at the transcription level.

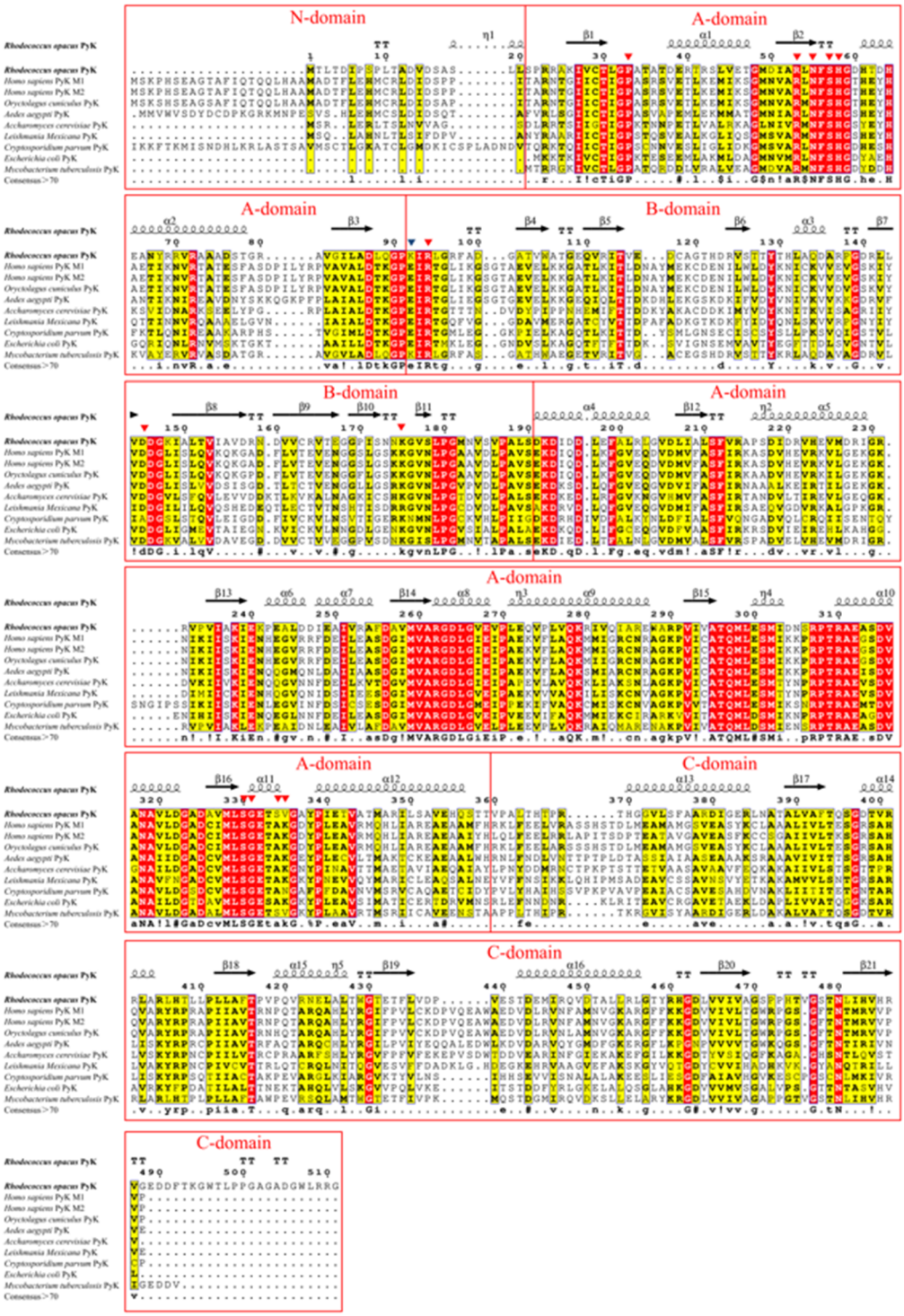


Figure 3. Multi-sequence Comparison of the PyK Homologous Gene Protein of Rhodococcus

Note: The red box is used to represent four different domains of PyK A, B, C, and N. Red ▼ is used to represent amino acid residues interacting with ADP; Blue ▼ is used to label amino acid residues that can distinguish different clusters and interact with potassium ions. White font and red fill color are used to represent amino acid residues with the same sequence; Bold fonts and yellow fill colors are used to represent amino acid residues with similarity greater than 70%. See section 2.1 for details of the protein sequence used in this analysis.

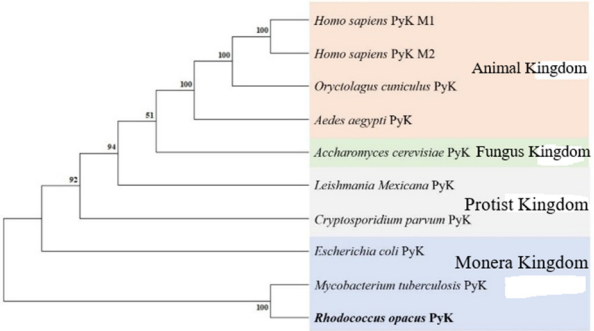


Figure 4. Phylogenetic Tree Analysis of the PyK Homologous Gene Protein of Rhodococcus

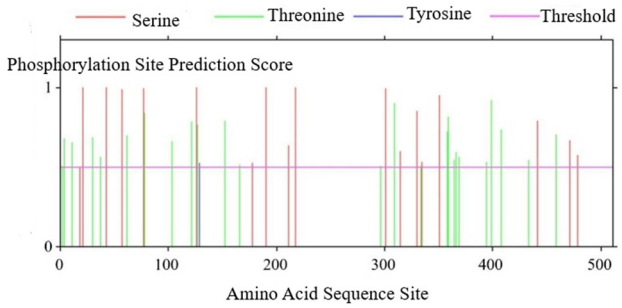


Figure 5. Prediction of the Phosphorylation Site of the PyK Protein in Rhodococcus

To sum up, this study systematically revealed the conservative characteristics of PyK from the perspective of bioinformatics in the catalytic center, spatial structure, and cofactor binding site, and found that it may be finely regulated by a complex phosphorylation regulatory network. These findings not only deepen the understanding of the regulation mechanism of carbon metabolism of Rhodococcus, but also provide multi-level theoretical support for optimizing the industrial application performance of Rhodococcus through metabolic engineering strategy.

References

[1] Wu Jinfang, Zhang Yuxiao, Peng Ren. Progress in the research on the omics, genetic transformation, and application of Rhodococcus [J]. *Chinese Journal of Applied and Environmental Biology*, 2022,28(01):239-246.

[2] Du Y, Wang Y, Peng G, et al. Reducing COD and BOD, as well as producing triacylglycerol by LDS5 grown in CTMP effluent [J]. *Bioresources*, 2011, 6(3)：3505-3514.

[3] Liu TT, Xu Y, Liu H, et al. Functional characterization of a gene cluster involved in gentisate catabolism in Rhodococcus sp. strain NCIMB 12038[J]. *Appl Microbiol Biotechnol*, 2011, 90 (2):671-678.

[4] Wulfert S, Schilasky S, Krueger S. Transcriptional and Biochemical Characterization of Cytosolic Pyruvate Kinases in Arabidopsis thaliana[J]. *Plants*, 2020,9 (3):353.

[5] Zhong W, Cui L, Goh BC, et al. Allosteric pyruvate kinase-based “logic gate” synergistically senses energy and sugar levels in Mycobacterium tuberculosis [J]. *Nat Commun*, 2017,8 (1): 1986.

[6] Xu X, Li X, Liu Y, et al. Pyruvate-responsive genetic circuits for dynamic control of central metabolism [J]. *Nat Chem Biol*. 2020; 16 (11):1261-1268.

[7] Schormann N, Hayden KL, Lee P, et al. An overview of the structure, function, and regulation of pyruvate kinases[J]. *Protein Sci*, 2019,28 (10):1771-1784.

[8] Huang Xiaolan, Li Chen, Ren Chi, et al. Research progress of pyruvate kinase post-translational modification on meat quality [J]. *Food Science*, 2024,45(16):303-310.

[9] Schramm A, Siebers B, Tjaden B, et al. Pyruvate kinase of the hyperthermophilic crenarchaeote Thermoproteus tenax: physiological role and phylogenetic aspects[J]. *J Bacteriol*, 2000, 182 (7):2001-2009.

[10] Johnsen U, Hansen T, Schonheit P. Comparative analysis of pyruvate kinases from the hyperthermophilic archaea Archaeoglobus fulgidus, Aeropyrum pernix, and Pyrobaculum aerophilum and the hyperthermophilic bacterium Thermotoga maritima: unusual regulatory properties in hyperthermophilic archaea[J]. *J Biol Chem*, 2003,278 (28):25417-25427.

[11] Oria-Hernández J, Riveros-Rosas H, Ramírez-Sílva L. Dichotomic phylogenetic tree of the pyruvate kinase family: K+ -dependent and -independent enzymes[J]. *J Biol Chem*, 2006,281 (41):30717-30724.

[12] Laughlin LT, Reed GH. The monovalent cation requirement of rabbit muscle pyruvate kinase is eliminated by substitution of lysine for glutamate 117[J]. *Arch Biochem Biophys*, 1997,348 (2):262-267.

[13] Zheng S, Liu Q, Liu T, et al. Post-translational modification of pyruvate kinase type M2 (PKM2): novel regulation of its biological roles to be further discovered[J]. *J Physiol Biochem*,2021,77 (3):355-363.