



## Research article

# Pyruvate kinase deficiency and *PKLR* gene mutations: Insights from molecular dynamics simulation analysis

Yang Wang<sup>a,1</sup>, Jiaqi Liu<sup>b,1</sup>, Tao Liu<sup>a</sup>, Xizhou An<sup>a</sup>, Lan Huang<sup>a</sup>, Jiacheng Li<sup>a</sup>, Yongjie Zhang<sup>a</sup>, Yan Xiang<sup>a</sup>, Li Xiao<sup>a</sup>, Weijia Yi<sup>a</sup>, Jiebin Qin<sup>a</sup>, Lili Liu<sup>c</sup>, Cuilan Wang<sup>d</sup>, Jie Yu<sup>a,\*</sup>

<sup>a</sup> Department of Hematology and Oncology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing Key Laboratory of Pediatrics, 136 Zhongshan er lu, Yu zhong district, Chongqing 400014, China

<sup>b</sup> Shanghai Cinopath Medical Testing Co Ltd, Shanghai 200000, China

<sup>c</sup> Department of Cardiovascular Medicine, Affiliated Hospital of North China University of Science and Technology, Tangshan 063000, China

<sup>d</sup> Department of Neurology, Affiliated Hospital of North China University of Science and Technology, Tangshan 063000, China

## ARTICLE INFO

## Keywords:

Mutation

Pyruvate kinase deficiency

Pyruvate kinase

Molecular dynamics simulation

## ABSTRACT

Pyruvate kinase deficiency is a rare hereditary erythrocyte enzyme disease caused by mutations in the pyruvate kinase liver and red blood cell gene. The clinical presentations of pyruvate kinase deficiency are significantly heterogeneous, ranging from just mild anemia to hemolytic crisis or even death. The proband in our study was a 2-year-old girl for severe skin and scleral icterus with progressive aggravation. We collected the family's data for further analysis. Whole exome genome sequencing of the pedigree revealed a novel compound heterozygous mutation, *c.1097del (p.P366Lfs\*12)* and *c.1493G > A (p.R498H)*, in the *pyruvate kinase liver and red blood cell* gene. Furthermore, molecular dynamics simulations were employed to uncover differences between the wild type and mutant pyruvate kinase liver and red blood cell proteins, focusing on structural stability, protein flexibility, secondary structure, and overall conformation. The combined bioinformatic tools were also utilised to assess the effects of the missense mutation on protein function. Thereafter, wild type and mutant plasmids were constructed and transfected into 293T cells, and Western blot assay was conducted to validate the impact of the mutations on the expression of pyruvate kinase liver and red blood cell protein. The data presented in our study enriches the genotype database and provides evidence for genetic counseling and molecular diagnosis of pyruvate kinase deficiency.

## 1. Introduction

Pyruvate kinase deficiency (PKD) (OMIM#609712) is a rare autosomal recessive disease caused by defective functional pyruvate kinase (PK), which is a rate-limiting enzyme in the glycolytic pathway [1]. Defective PK can negatively impact the metabolism of

\* Corresponding author. Department of hematology and oncology, Children's Hospital of Chongqing Medical University, 136 Zhongshan er lu, Yuzhong District, Chongqing, 400014, China.

E-mail address: [1808106652@qq.com](mailto:1808106652@qq.com) (J. Yu).

<sup>1</sup> Yang Wang and Jiaqi Liu contributed equally to this study.

<https://doi.org/10.1016/j.heliyon.2024.e26368>

Received 7 April 2022; Received in revised form 25 December 2023; Accepted 12 February 2024

Available online 16 February 2024

2405-8440/© 2024 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

erythrocytes, resulting in decreased intracellular adenosine triphosphate production and ion pump dysfunction. This further causes a significant loss of potassium and water from erythrocytes, leading to hemolysis [2]. The clinical manifestations of PKD include varying degrees of anemia, yellow-stained skin and sclera, splenomegaly, and gallstones. However, both clinical manifestations and hematological features of PKD lack specificity and cannot be used for PKD diagnosis [3]. Gene sequencing can provide more accurate and intuitive evidence for the diagnosis of rare disorders. Currently, the gold standard for diagnosing PKD is either enzyme activity or genetic sequencing [1]. Given the potential limitations of both PK enzyme activity and genetic testing, the diagnosis of PKD is made through a combination of both tests, if available [4]. With the popularization of sequencing, a substantial number of variants have been identified, many of which are novel. It is of utmost importance to document these. Furthermore, interpreting the pathogenicity of novel variants presents an even greater challenge [5].

Molecular dynamics (MD) simulations use quantum and classical mechanics to simulate molecular motion numerically, clarifying structures, properties, and explaining experimental phenomena [6]. MD simulations have become the third most used method for microstructure and macroscopic property determination [7]. We used it to illustrate the distinctions between the proteins encoded by the wildtype (WT) and mutant pyruvate kinase liver and red blood cell (*PKLR*) gene. Additionally, we validated these results by *in vitro* experiments.

In summary, we reported a pedigree of congenital hemolytic anemia. The proband was ultimately diagnosed with PKD by gene sequencing. We conducted MD simulations, bioinformatic tools and *in vitro* experiments to further determine the effect of the mutations. We identified the novel compound heterozygous mutation, exon7 *c.1097del (p.P366Lfs\*12)/exon10 c.1493G > A (p.R498H)*, which we believe would broaden the spectrum of mutations known to cause PKD. Our findings provide valuable insights into genetic counseling and the interpretation of mutations.

## 2. Materials and methods

### 2.1. Study participants and clinical investigations

The participant was a 2-year-old girl admitted to the Department of Hematological Oncology, Children's Hospital of Chongqing Medical University for severe skin and scleral icterus with progressive aggravation. She was the second child of non-consanguineous parents and was born in the Sichuan Province of China. We performed a detailed genealogical study on the proband and recorded the age, gender, and medical status of the family members to prepare a genealogical map.

### 2.2. Ethical approval

This study was conducted in accordance with the World Medical Association Declaration of Helsinki and the International Ethical Guidelines for Human Biomedical Research. Genetic screening on the entire family was performed after obtaining their signed informed consent form. The research was approved by the Medical Ethics Committee of the Children's Hospital of Chongqing Medical University (approval file number: 2021.456).

### 2.3. Genomic DNA extraction and whole exome sequencing

Genomic DNA (gDNA) was extracted from peripheral blood using centrifugal column-type blood gDNA extraction kits (Tiangen, Beijing, China), and its concentration was evaluated using the Sim-100 ultra-micro spectrophotometer (Xunjie, Hangzhou, China). The gDNA (2 µg) of the proband was next fragmented by sonication with a Qsonica Q800R and amplified to construct next-generation sequencing libraries with a GenNext NGS library quantification kit (TOYOBO, Japan) and an HTP library Preparation kit (Kapa Biosystems). Next, the Illumina Nextseq 500 sequencing platform was used to generate a high-throughput sequencing library. The mean sequencing depth for the targeted sequencing regions was approximately 500–1000 × . The target region capture sequencing was provided by AmCare Genomic Laboratory. The sequencing data were assessed with the adapted Illumina Sequence Control Software and only the qualified data reads were subjected to bioinformatics analyses [8]. Relevant published articles on mutation-related information were obtained from the PubMed database, and the pathogenicity of the variant locus was annotated by combining the previously reported results and the classification criteria and guidelines for genetic variants (American College of Medical Genetics and Genomics) [9]. The potential disease-related variant sites were then validated by Sanger sequencing [10]. The analysis method was as before [11].

### 2.4. Homology modeling

We performed homology screening using the Protein Data Bank (PDB; PDB ID:6NN7) and used SWISS-MODEL to render the 3D structures of WT and mutant PLKR proteins (R498H, and P366Lfs\*12) [12,13]. Thereafter, we assessed the quality of the optimized protein model using PROCHECK and prepared the corresponding Ramachandran plots [14,15].

### 2.5. Bioinformatic analysis of the missense mutation

An expanding array of online tools is currently under development to predict the pathogenicity of missense mutations. DynaMut webserver was used to visualize and assess the protein stability and interactions [16]. Novopro (<https://www.novopro.cn/tools/>

[protein-hydrophilicity-plot.html/](http://protein-hydrophilicity-plot.html/)) was used to evaluate protein hydrophobicity [17]. PONDR (<http://www.pondr.com/>) was used to predict inherently disordered regions (IDRs) [18]. To generate a prediction, the query sequences are provided as input and fed into the individual programs as described. After waiting for a while, the prediction results will be displayed on the web page. The nucleotide sequences and amino acid sequences can be found in the supplementary material.

## 2.6. MD simulations

We performed MD simulations for the WT, R498H, and P366Lfs\*12 proteins using the GROMACS [19]. MD simulations were performed under constant temperature and pressure, as well as periodic boundary conditions, using the Amber14SB all-atom force field and transferable intermolecular potential 3 points water model [20]. The LINCS algorithm was employed to constrain all the bonds associated with hydrogen atoms, and the time step was set to 2 fs [20]. The Particle Mesh Ewald was used to reveal electrostatic interactions [21]. The non-bonded interaction cutoff was set to 10 Å and updated every 10 steps. The V-rescale temperature coupling method and the Parrinello-Rahman method were used to control temperature (300 K) and pressure (1 bar), respectively, during the simulation [21]. Initially, the system energy was minimized using the steepest descent method to eliminate any close contacts between the atoms. Subsequently, 1 ns of multistep isothermal–isovolumetric and isothermal–isobaric conditions equilibrium simulations were performed at 300 K, respectively. Finally, the system underwent 50 ns of simulation, with the conformations saved every 10 ps. The resulting trajectories were analyzed to observe and compare the structural differences between the WT and mutant proteins. The root-mean-square deviation (RMSD), root-mean-square fluctuations (RMSF), radius of gyration (Rg), and solvent accessible surface area (SASA) were computed using the GROMACS and the Origin software. Thereafter, secondary structures were calculated and simulated, and the resulting structures were analyzed along with the global average structures of native and mutant proteins. Visual Molecular Dynamics program was used to visualize the simulation results [22,23].

## 2.7. Preparation of mutant constructs

A plasmid containing full-length *PKLR* cDNA was purchased from Youbao Company (Changsha, China). The WT *PKLR* gene was cloned into the pCDNA3.1 (+)-3XFlag-C expression vector and 5'-ATGGACTACAAGGACGACGATGACAAG-3' (encoding FLAG) was inserted before the initiation codon. The mutant *PKLR* genes were obtained using the KOD-Plus Mutagenesis Kit (TOYOBO, Osaka, Japan). The plasmid map can be found in [Supplementary Fig. 1](#). The primer sequences can be found in [Supplementary Table 1](#). Sanger sequencing was used for the validation of all the expression constructs.

## 2.8. Cell culture and transfection

The 293T cells donated by Li Jiang, were grown in 10% fetal bovine serum/Dulbecco's Modified Eagle Medium containing 1% penicillin streptomycin, at 37 °C and 5% CO<sub>2</sub>. The cell density was adjusted to  $5 \times 10^5$  using the Automated Cell Counter IM1200 (Countstar, Shanghai, China), and the cells were seeded in culture dishes. When the cell confluency reached 70–80%, we transfected the cells with both WT and mutant plasmids using Lipofectamine 2000 Transfection Reagent (Invitrogen, #11668019) according to the manufacturer's instructions. After 48 h, the transfected cells were harvested for further analysis.

## 2.9. Western blot

The experiment samples were divided into 5 groups: NC, WT, P366Lfs\*12, R498H, and P366Lfs\*12 + R498H group. The total protein was extracted using a total protein extraction kit (BBproExtra, China) and quantified using a bicinchoninic acid assay kit (Biyuntian, China). The protein samples were then subjected to 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred onto a polyvinylidene fluoride membrane (Millipore, USA). Thereafter, the membrane was blocked using a blocking solution (YWB0501, Yoche Biotechnology, Shanghai, China) and incubated with anti-FLAG (20543-1-AP, Proteintech, Wuhan, China) and anti-GAPDH (390035, ZEN-BIOSCIENCE, Chengdu, China) primary antibodies (1:2000), 4 °C, overnight. Subsequently, the membrane was incubated with secondary anti-rabbit antibodies (TA373083, OriGene, Wuxi, China), at room temperature for 1.5 h. Lastly, the images were obtained using the gel imaging systems (Gene, USA).

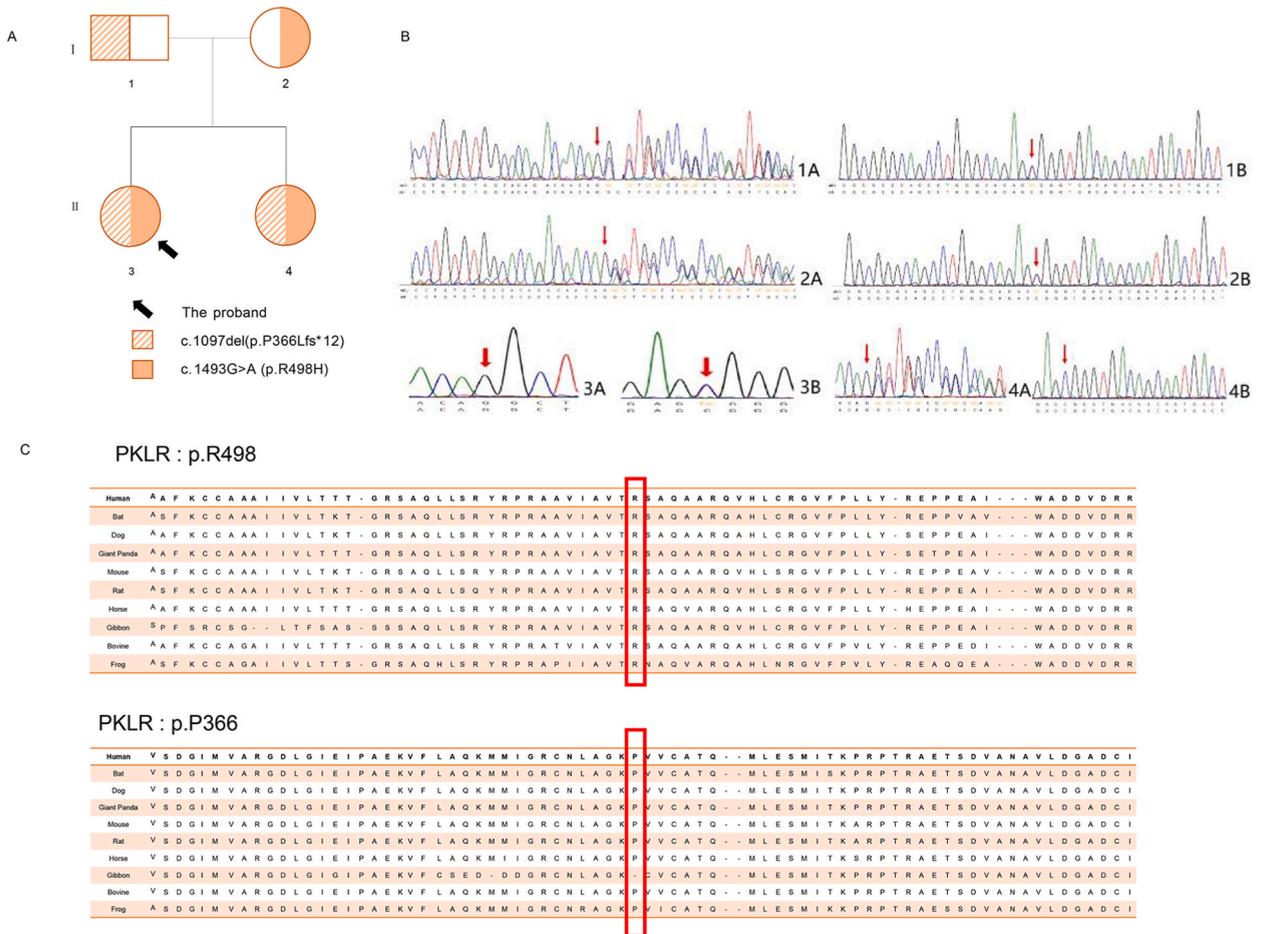
## 3. Results

### 3.1. Clinical data and mutation analysis

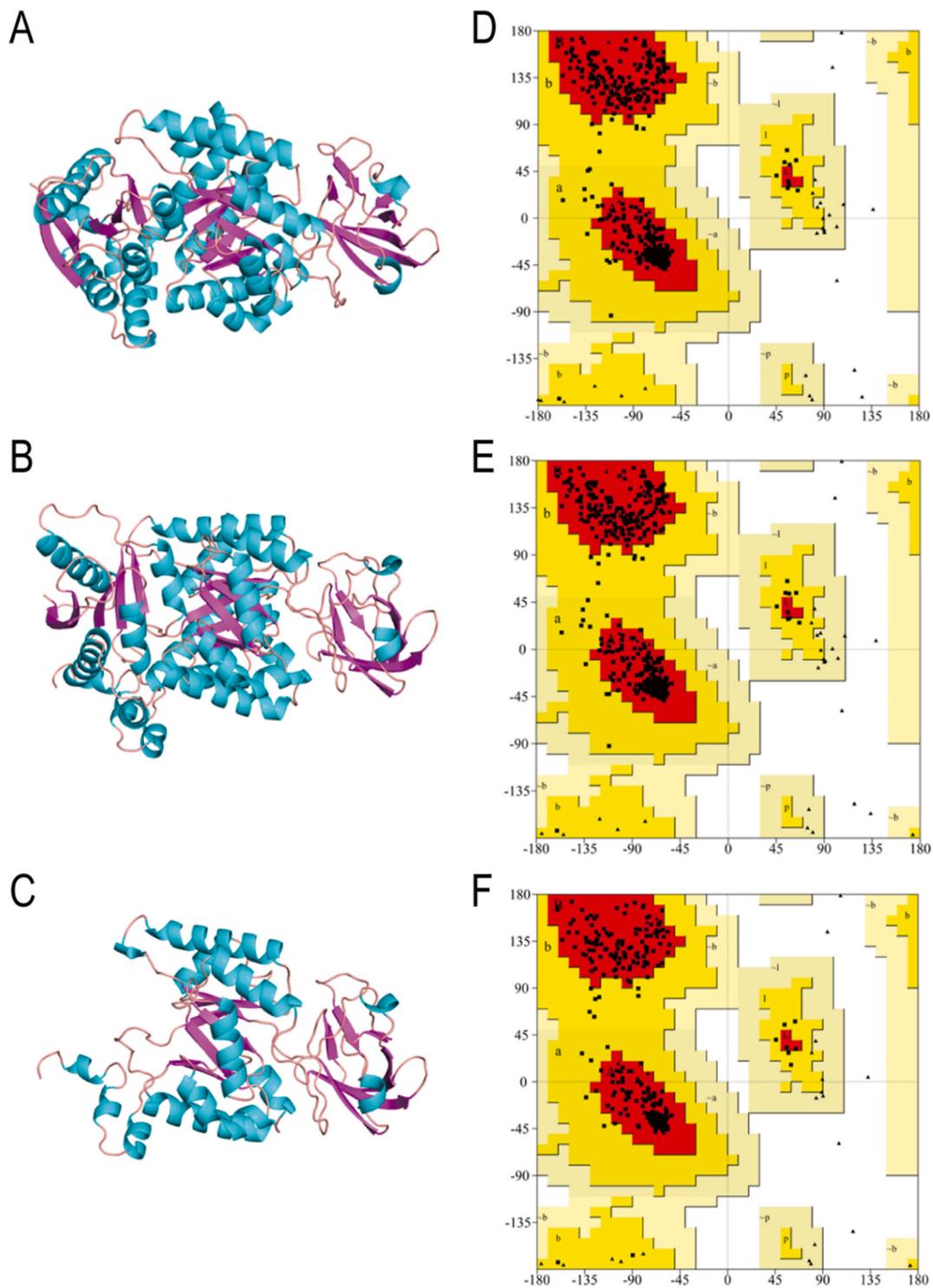
The proband was the second child of non-consanguineous parents and had extremely severe skin and scleral icterus with progressive aggravation. A physical examination of the patient revealed poor nutritional status and mental distress. The proband's body mass index was lower than that of other children of the same age, and she also experienced easy fatigability. Her extremely severe anemic condition was manifested by pallor, scleral jaundice, pale lips, mild cyanosis, and pale nail beds; however, her distal extremities remained warm. The abdomen was soft to palpation, and the liver and spleen were noted below the costal margin. The masses palpated were soft and had a smooth surface. Laboratory examination of the patient's blood sample revealed the following values red blood cells (RBCs):  $1.00 \times 10^{12}/L$ , hemoglobin: 21 g/L, reticulocyte percent: 13.1% (0.5–1.5%), reticulocyte count:  $0.2760 (0.024\text{--}0.084 \times 10^{12}/L)$ , white blood cells:  $11.96 \times 10^9/L$ , and platelets:  $222 \times 10^9/L$ . Further analysis of some biochemical

indicators revealed abnormally high values of total bilirubin (43.4  $\mu\text{mol/L}$ ; 0–21  $\mu\text{mol/L}$ ), unconjugated bilirubin (42.2  $\mu\text{mol/L}$ ; 0–13.68  $\mu\text{mol/L}$ ), glutamate transaminase (85.3 U/L; 4–40 U/L), and lactate dehydrogenase (1255.3 U/L; 109–245 U/L); the other residual indicators were normal. The liver and renal functions of the patient were within normal limits. Additionally, no abnormalities were detected in the proband in the erythrocyte permeability osmotic fragility test, glucose-6-phosphate dehydrogenase activity, and copper cyanine tests. All the members of the patient’s family were negative for the direct coombs test. Furthermore, there was no clearly positive evidence of bacterial and viral infections. The morphology of the erythrocytes was analyzed multiple times by blood smear test and no spherical, oval, or other aberrantly shaped erythrocytes were observed. The abdominal ultrasound revealed that the right hepatic midclavicular line was located 4.3 cm below the costal margin, and the oblique diameter of the right hepatic lobe was 10.3 cm. The spleen was situated 2.7 cm below the left costal margin and had a diameter and thickness of approximately 7.2 and 2.3 cm, respectively. The pale and yellow discoloration of the skin and sclera, hemoglobinuria, and hepatosplenomegaly strongly suggest hemolysis.

We found that the proband’s family had a history of hereditary hemolytic anemia, wherein the proband’s father suffered from  $\alpha$ -thalassemia, while the sibling suffered from  $\alpha$ -thalassemia and PKD. However, genetic testing of the proband revealed no relevant mutations for thalassemia. In the PKD family, the proband and both her sister and father had manifestations of anemia, while the mother showed no abnormal clinical presentation. Previous results of examination revealed that both I-1 and II-4 carried copy number heterozygous deletion variants of at least 4.5 kb in the 16p13.3 gene region. Thus, the father and sister were clearly diagnosed with  $\alpha$ -thalassemia. Gene sequencing indicated that the proband carried frameshift mutation (*c.1097del*) and missense mutation (*c.1493G > A*) occurring in exon 7 and 10, respectively, in the *PKLR* gene. The proband’s sibling also carried the same compound heterozygous mutation. The detected mutations were verified by sanger sequencing, indicating that *c.1097del* and *c.1493G > A* was inherited from the father and mother, respectively. The genetic mutation profiles of the proband’s family were shown in Fig. 1A–B. The observed



**Fig. 1.** Mutations in the PKD family (A) Pedigrees of the proband’s family based on genetic testing data and clinical presentation. (B) 1A, 1B: The proband carries the *c.1097del* and *c.1493G > A* compound heterozygous variant. 2A, 2B: the sibling carries the same mutation as the proband. 3A, 3B: the mother is purely heterozygous for the natural *c.1097del* locus and *c.1493G > A* heterozygous variant. 4A, 4B: the father is heterozygous for the *c.1097del* variant and the natural *c.1493G > A* locus. (C) Evolutionary conservation of residue 366 and 498. Amino acid (aa) positions are denoted by the red box. PKD: Pyruvate Kinase Deficiency.

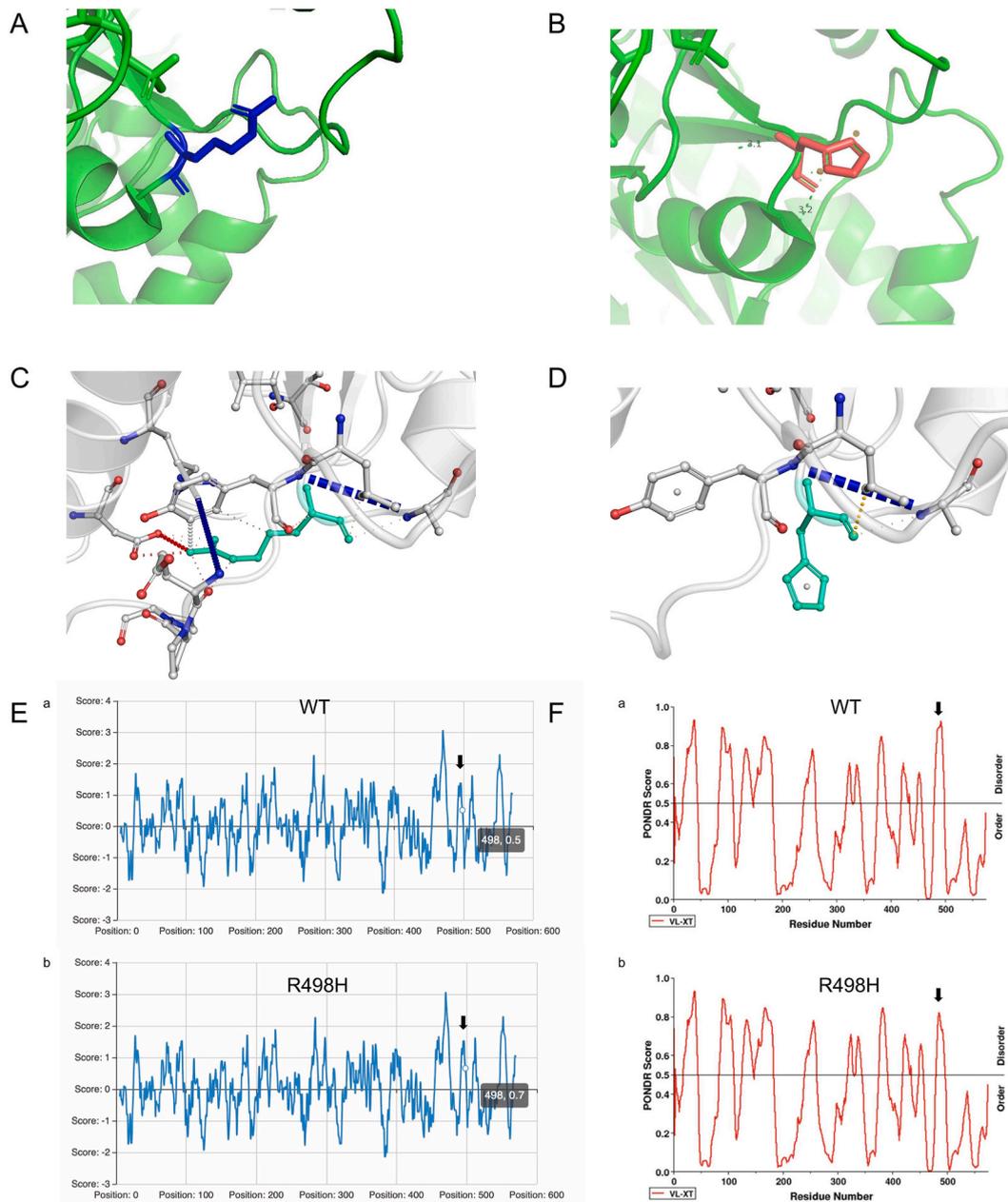


**Fig. 2.** Protein model construction and evaluation. (A–C) Individual crystal structures of WT, R498H and P366Lfs\*12 according to the model template (PDB ID: 6NN7). (D–F) Ramachandran plots of WT, R498H and P366Lfs\*12. The red color indicates the most favored regions, yellow indicates additional allowed regions, and white indicates the forbidden regions. WT: wildtype.

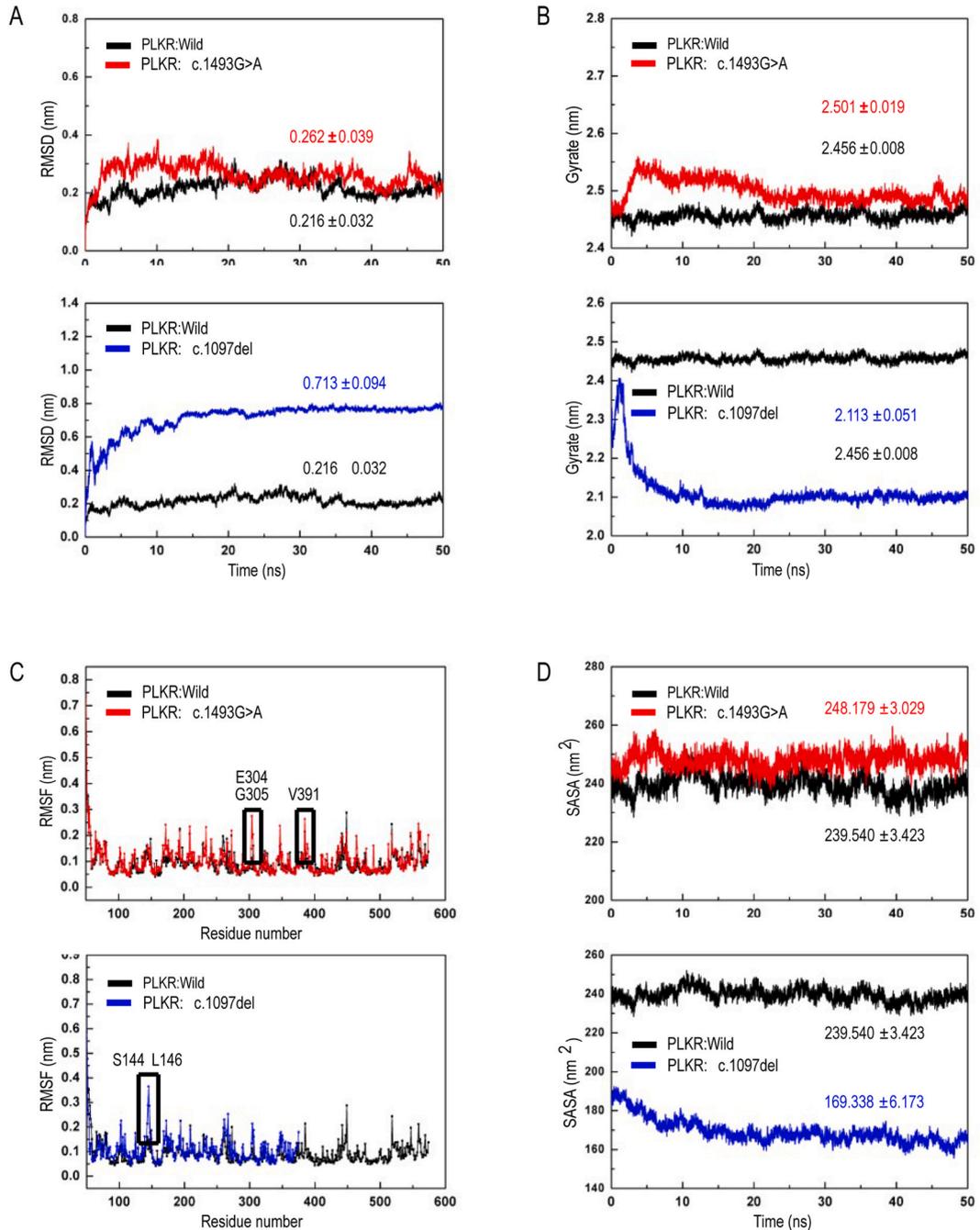
missense and frameshift mutant is conserved (Fig. 1C).

### 3.2. Homologous modeling and model quality evaluation

The protein structure with the code 6NN7 was accessible in the PDB database and was used for constructing models (Fig. 2A–C). Ramachandran plots were used to assess the sensitivity and specificity of the forecasted structure and visualize the dihedral angles  $\psi$  and  $\phi$  of the aa residues. Ramachandran plot analysis (Fig. 2D–F) revealed that 93.8%, 93.8%, and 92.8% of the aa residues of the WT,



**Fig. 3.** Bioinformatic analysis of c.1493G > A (p.R498H). (A–B) Three-dimension structure of the PKLR protein. Residue 498 is marked in blue and the R498H is marked in red. (C–D) Predictive interactions of the WT and R498H protein. Residues in the WT and R498H are coloured light green and shown as sticks. The respective chemical interactions are labelled as dotted lines and coloured as follows: hydrogen bonds—red; weak hydrogen bonds—orange; hydrophobic contacts—green; amide-amide contacts—blue and ionic interactions—gold. Amino acid residues are also coloured according to type, namely nitrogen (blue), oxygen (red), and sulphur (yellow). (E) Hydrophobicity analysis indicates heightened hydrophobicity in the R498H. (F) Protein disorder prediction indicates a decrease the residues located in IDRs. PKLR: pyruvate kinase liver and red blood cell; WT: wildtype; IDRs: inherently disordered regions.

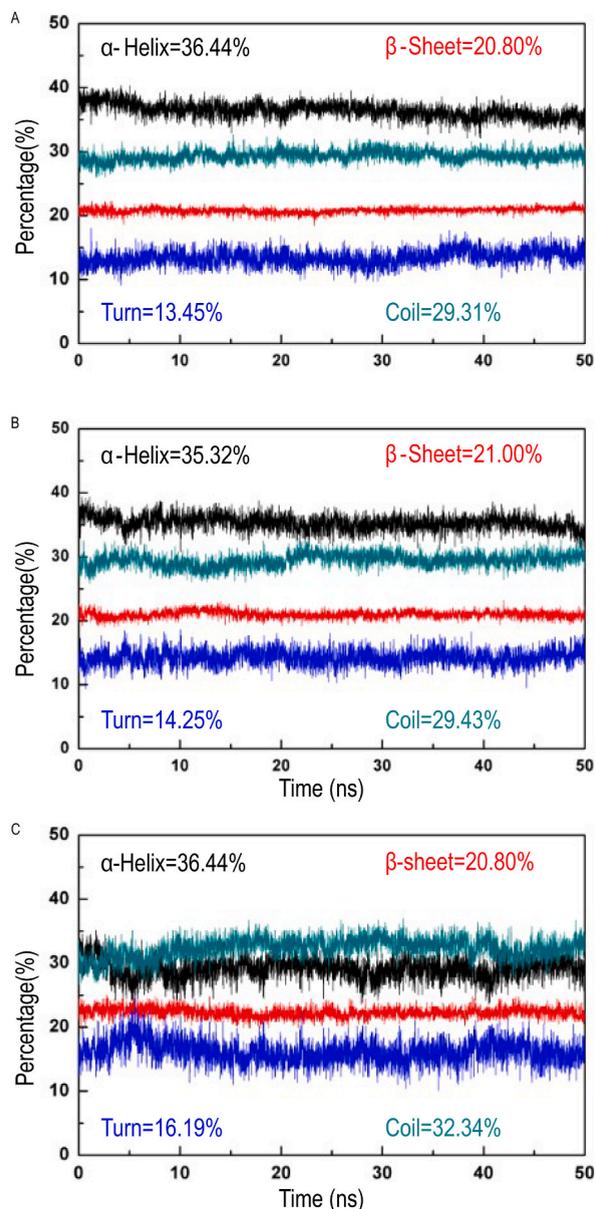


**Fig. 4.** The RMSD, RMSF, Rg, and SASA values of the WT, R498H, and P366Lfs\*12 proteins. (A) The RMSD values of the WT, R498H, and P366Lfs\*12 models were found to be  $0.216 \pm 0.032$ ,  $0.262 \pm 0.039$ , and  $0.713 \pm 0.094$  nm, respectively; (B) the Rg values of the WT, R498H, and P366Lfs\*12 models were determined as  $2.456 \pm 0.008$ ,  $2.501 \pm 0.019$ , and  $2.113 \pm 0.051$  nm, respectively; (C) the R498H mutation increased the overall flexibility of the mutant protein, compared to the WT, with the highest increase in flexibility a E304, G305, and V391. The P366Lfs\*12 mutant, with deleted C-terminal structural domain, showed a small margin of overall structural flexibility in the overlapping region before the deletion, specifically in the S144-L146 region; (D) the SASA values of the WT, R498H, and P366Lfs\*12 models were  $239.540 \pm 3.423$ ,  $248.179 \pm 3.029$ , and  $169.338 \pm 6.173$  nm<sup>2</sup>, respectively. RMSD: root-mean-square deviation; RMSF: root-mean-square fluctuations; Rg: radius of gyration; SASA: solvent accessible surface area; WT: wild type.

R498H, and P366Lfs\*12 proteins, respectively, were located in the favored regions, and none of the models had aa residues located in the forbidden regions. Therefore, the models were all reasonably consistent with the rules of stereochemistry.

### 3.3. Bioinformatic analysis of *c.1493G > A (p.R498H)*

DynaMut was employed to visualize the structure of the PKLR protein. The missense mutation, R498H, resulted in aa residue change from arginine to histidine (Fig. 3A and B). Compared to the WT, some interactions (hydrophobic, amideamide, and ionic bonds) were observed to decrease in the R498H (Fig. 3C and D). Furthermore, R498H was more hydrophobic, with a score of 0.7 (Fig. 3E). The R498H variant results in a decrease the residues located in IDRs (Fig. 3F).



**Fig. 5.** Analysis of the secondary structural components ( $\alpha$ -helix,  $\beta$ -sheet, turn, and coil) of the WT, R498H, and P366Lfs\*12 proteins. (A) The secondary structural components of WT proteins; (B) the secondary structural components of R498H proteins; (C) the secondary structural components of P366Lfs\*12 proteins. WT: wild type.

### 3.4. Results of MD simulations

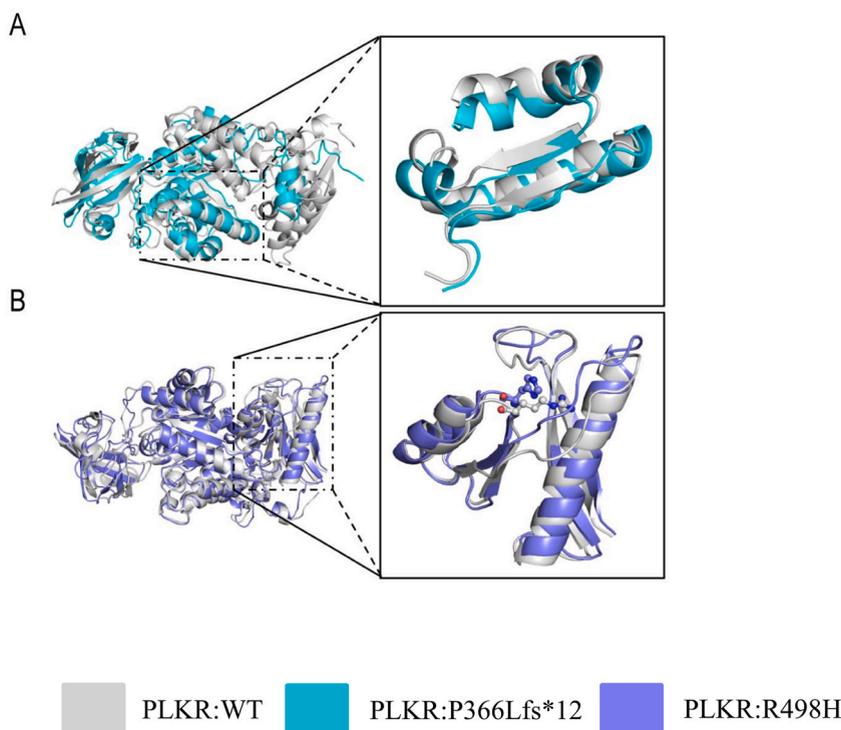
#### 3.4.1. MD simulation parameters

RMSD is an essential parameter for evaluating protein stability. It reflects the deviation of the target conformation from the initial structure, thus a smaller RMSD value indicates a higher similarity between the target conformation and the initial structure [24]. The system stabilized at approximately 10 ns. The RMSD values of the WT, R498H, and P366Lfs\*12 models were found to be  $0.216 \pm 0.032$ ,  $0.262 \pm 0.039$ , and  $0.713 \pm 0.094$  nm, respectively (Fig. 4A). The increased RMSD values of the two mutant proteins compared to that of the WT protein, suggest that the corresponding gene mutations affect the stability of the protein. In particular, the *c.1097del* frameshift mutation, resulting in C-terminal truncation of the structural domain, had a more severe effect on protein stability. Furthermore, the mutant systems had a higher RMSD range. Therefore, the mutant systems showed larger RMSD values and greater fluctuations.

The Rg value is used to analyze the compactness of a protein structure, which may affect its ligand-binding property [25]. A higher Rg value indicates lower compactness of the system (Fig. 4 B). The Rg values of the WT, R498H, and P366Lfs\*12 models were determined as  $2.456 \pm 0.008$ ,  $2.501 \pm 0.019$ , and  $2.113 \pm 0.051$  nm, respectively. These results suggest that the R498H mutation decreased the compactness of the protein, indicating that it might be accompanied by local structural unwinding. Compared with the WT, the P366Lfs\*12 mutation, causing a large C-terminal truncation, led to an increase in the compactness of protein.

The RMSF value of each C  $\alpha$  atom, relative to its average position, provides insights into the fluctuation and flexibility of the aa residue (Fig. 4C). The R498H mutation increased the overall flexibility of the mutant protein, compared to the WT, with the highest increase in flexibility a E304, G305, and V39. Comparative structural analysis revealed that the aa residues with large fluctuations in flexibility were all close to the R498H site. This demonstrates that mutations in this site could affect protein flexibility, which could further affect protein function and stability in the physiological environment, ultimately causing diseases. Interestingly, the P366Lfs\*12 mutant, with deleted C-terminal structural domain, showed a small margin of overall structural flexibility in the overlapping region before the deletion, specifically in the S144-L146 region. These results indicate that the mutation-induced changes in protein flexibility have a large impact on the biological functions of the protein.

The SASA is the area of a protein that is directly accessible to the surrounding solvents. In general, the active site of enzyme located in the hydrophobic cavity of the protein and the substrate has stronger interactions with the catalytic groups of the active site present in a non-polar environment. The SASA values of the WT, R498H, and P366Lfs\*12 models were  $239.540 \pm 3.423$ ,  $248.179 \pm 3.029$ , and  $169.338 \pm 6.173$  nm<sup>2</sup>, respectively (Fig. 4D).



**Fig. 6.** Superimposed structures of the WT, P366Lfs\*12, and R498H proteins during the most stable period (40–50 ns) of the MD simulations. (A) Structural comparisons of the N316-E375 and N316-R375 regions in the WT and P366Lfs\*12 protein. (B) Structural comparison of the D451-G536 regions in the WT and R498H protein. WT: wild type; MD simulations: molecular dynamics simulations.

### 3.4.2. Analysis of protein secondary structure

The secondary structures of the proteins include  $\alpha$ -helix,  $\beta$ -sheet, turn, and coil. A higher percentage of  $\alpha$ -helix indicates increased stability of the overall domain structure [26]. In this study, we calculated the changes in secondary structural elements of the WT and mutant proteins, using simulations (Fig. 5). We found that, compared to the WT (Fig. 5A), R498H showed a 1.12% reduction in the  $\alpha$ -helix elements, which were mainly converted to turns (Fig. 5B). Moreover, the number and proportion of  $\alpha$ -helix elements decreased significantly in the P366Lfs\*12 (Fig. 5C), which may seriously affect its biological function. In summary, these mutations lead to weakened protein structure stability, which may further affect its binding and activity.

### 3.4.3. Structural superposition of the global average structures

The effect of dimer deletion on the total structure was observed by a comparative analysis of the average structure during the most stable period (40–50 ns) of the simulation (Fig. 6). The differences between the overall structures of the WT and R498H were mostly insignificant; however, a segment of the loop next to the R498H mutation loci was significantly structurally displaced after the mutation, moving closer to the R498H (Fig. 6B). In contrast, there was a distinct segment of structural domain variation between the WT and P366Lfs\*12 (Fig. 6A). The *c.1097del* mutation caused premature termination of the mutant protein at residue at 366, such that the WT and P366Lfs\*12 sequences were PVVCATQMLE and LLSVPHRCWR, respectively. This change in the P366Lfs\*12 sequence led to a change in the secondary structure of the PKLR protein from  $\beta$ -sheet to coil, which could affect its biological function.

### 3.5. P366Lfs\*12 and R498H variants lead to loss of function of PKLR gene

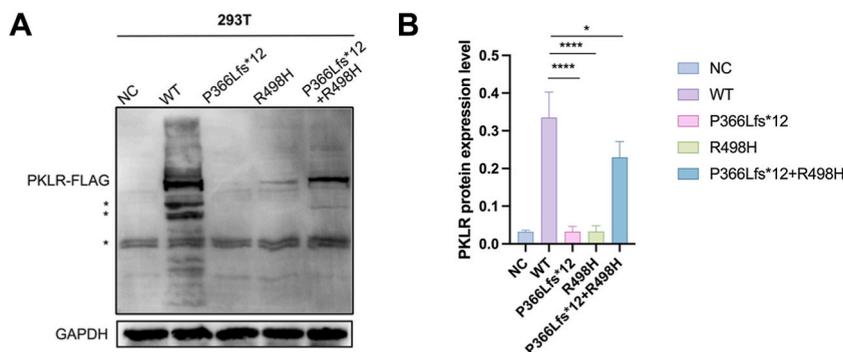
We determined the pathogenicity of PKLR mutations by overexpressing the WT and mutant proteins in 293T cells. Western blot analyses revealed that the mutant proteins showed reduced expression compared to the WT and that truncated proteins were not expressed (Fig. 7A and B). The original images of WB stain were given in Supplementary Figs. 2–3.

## 4. Discussion

PKD is an autosomal recessive congenital disorder that was first reported in 1961 [27]. The incidence of PKD is variable in different regions and ethnic groups [28]. PK activity assay and genetic testing are important for the diagnosis of PKD [29,30]. In our study, we identified the novel compound heterozygous loss-of-function PKLR mutation (*c.1097del: p. P366Lfs\*12* and *c.1493G > A: p. R498H*), which was responsible for PKD. We further investigated the molecular mechanisms of mutant proteins by using advanced computational methods to predict their structural perturbations and functional implications. These results were then corroborated by experimental data.

The human PKLR gene, located on chromosome 1q21, consists of 12 exons spanning 9.5 kb. Mutations in the PKLR gene leading to a total loss or severe reduction of the enzyme activity can lead to hemolysis [31]. Clinical symptoms arise when an individual inherits homozygous or heterozygous mutations of the PKLR gene from both parents [32]. Thus far, over 300 PKLR gene mutations have been reported globally and the majority of these are point mutations, followed by insertion–deletion mutations [33]. Approximately 70% of the PKLR mutations are compound heterozygous mutations, among which one is a missense mutation, while the other is a frameshift or nonsense mutation. *c.1529G > A* and *c.1456C > T* are the most common missense mutations. Patients with the *c.1529G > A* mutation are mainly distributed in the United States (42%) and central and northern Europe (41%), while patients with the *c.1456C > T* mutation are mainly distributed in southern Europe (Spain: 32%, Portugal: 35%, and Italy: 29%). In addition, patients with the *c.1468C > T* mutation are primarily distributed in Asia [34].

PK (200–240 kDa) is an allosteric enzyme that catalyzes the conversion of phosphoenolpyruvate and adenosine diphosphate to pyruvate and adenosine triphosphate, respectively, in the glycolytic pathway [33]. It is highly conserved protein consisting of 4 subunits/domains: A, B, C, and Nter domain. The conformational space possesses mirror symmetric geometry [35]. In general, a highly



**Fig. 7.** Western blot analysis of the WT and mutant proteins. (A) Protein bands of western blotting. The asterisks indicate non-specific or protein degradation products. (B) Quantification of protein bands. \* $P < 0.05$  and, \*\*\*\* $P < 0.0001$ . Western blot assay was repeated for three times. WT: wild type.

elaborate functional unit is dependent upon the molecular structure of the protein. Aa mutations may significantly alter the conformation of the PKLR protein. Different mutations have different effects on the PK conformational transition mechanism, leading to significant differences in clinical manifestations among PKD patients. MD simulations act as a computational “microscope”, which are valuable for exploring the effects of mutation on protein structure [36]. For instance, MD simulations on WT and mutant ND4L-ND6 subunits revealed that the mutations limit the passage of water molecules leading to type 2 diabetes mellitus and cataracts [37]. Additionally, MD simulations on a mutation in the arylsulfatase A protein altered its structural and functional behavior, ultimately led to metachromatic leukodystrophy [38]. Identifying the effects of mutations on protein structure is critical for exploring disease pathogenesis at the molecular level.

The missense mutation, R498H, resulted in residue change from arginine to histidine and occurred in the core hydrophobic region of the A domain, which is involved in the formation of the highly conserved active site. It is speculated that R498H may lead to functional alteration. Our in-depth study found that R498H is more hydrophobic and results in a decrease in the residues located in IDRs. In addition, R498H decreased intermolecular forces. Moreover, mutations affect the IDRs and its abundant multivalent weak interactions, which can induce liquid-liquid phase separation, resulting in the formation of liquid condensate phases [39]. The above predicted results all suggest that R498H works through loss-of-function mechanisms. The frameshift mutation, P366Lfs\*12, resulted in early termination of the translation product of the *PKLR* gene and occurred in the structural subunit of the C domain, which plays a crucial role in stabilizing tetramerization and variant regulation. MD simulations showed the differences in the stability, compactness, and flexibility of the overall protein structures of the WT and mutant proteins. The results revealed that the global secondary structure was significantly altered by these mutations. Furthermore, *in vivo* experiments revealed that the expression of P366Lfs\*12 and R498H variants was significantly lower than that of the WT, suggesting that the mutant proteins may be subject to nonsense-mediated decay. The functional effect trends with protein expression level [40]. Since, protein expression level is associated with its function, we hypothesized that these mutations may affect the catalytic activity of PK. Changes in the aa sequence and physicochemical properties of the protein lead to variations in the crystal conformation of the enzyme as well as in the spatial conformation of the active site, ultimately leading to hemolysis [41]. The compound heterozygous mutation, identified in this study, can provide new evidence for genetic diagnoses of PKD.

Currently, PKD is an incurable disorder. PKD patients are primarily treated to reduce symptom presentation, and RBC transfusion is used to improve their general state. Additionally, resection of the enlarged spleen, in PKD patients, is conducted by clinicians after deliberation of its pros and cons. Allogeneic hematopoietic stem cell transplantation has shown excellent treatment efficacy in patients with severe PKD [42]. Some PKD studies revealed that lentiviral vectors, carrying hPGK-coRPK-targeting hematopoietic stem cells (HSCs), effectively cured PKD in mouse models without causing genotoxicity [43]. In addition, new drugs developments have brought new hopes for patients with PKD. Therapeutic options for PKD are rapidly expanding. Modification of PK may be a future strategy for PKD treatment. Mitapivat (AG-348, Agios Pharmaceuticals) is an oral small molecule that allosterically activates erythrocyte PK [44]. Mitapivat can increase WT and variant PKR enzyme activity and have the advantages of high safety and favorable tolerability [45]. Studies have revealed that the majority of patients harboring at least one missense variant, whereas those with two non-missense mutations (a homozygous or compound heterozygous state) exhibited minimal or absent responses [46]. Also, animal models demonstrate the effectiveness of gene therapy for PK deficiency [47]. For instance, Fañanas-Baquero et al. restored normal adenosine triphosphate levels in PKD-HSCs-derived erythrocytes by using clustered regularly interspaced short palindromic repeats and recombinant adeno-associated vector, and this gene-editing strategy may serve as a lifelong therapy to correct the function of PK in RBCs of PKD patients [48]. Although these results are promising, the long-term efficacy and safety of gene therapy need to be further validated by large-scale studies [49]. Furthermore, vectors used in gene therapy pose additional challenges. For instance, viral vectors have safety concerns, while non-viral vectors have low delivery efficiency [50]. Therefore, research and development of novel vectors are crucial for gene therapy. H.N. Abdelhamid et al. developed carbonized chitosan-encapsulated hierarchical porous zeolitic imidazolate frameworks, which greatly improved gene delivery [51]. The gene therapy clinical trials of PKD are currently undergoing [52]. Therefore, continuous developments in precision medicine and gene therapy are essential for PKD treatment.

In this study, the proband's sister had  $\alpha$ -thalassemia and PKD, and genetic testing revealed that her father carried the  $\alpha$ -thalassemia gene, while her mother carried a heterozygous pathogenic variant of PKD, which is linked to a higher risk of hereditary disease. The clinically relevant interaction between  $\alpha$ -thalassemia predisposition and PKD may be a possible research direction in the future. Notably, genetic counseling has significantly reduced the incidence of congenital diseases [53]. However, less than 10% of clinicians provide important information to patients suspected of having genetic disorders [54]. Therefore, preconception counseling should be actively offered to all the parents of patients with rare diseases, to avoid further cases.

## 5. Conclusions

Genetic testing is important for the diagnosis of diseases. MD simulations are used to validate and demonstrate the effect of deleterious mutations on the protein's structure. Genetic counseling and proper engagement between the patients and clinicians can facilitate in early detection of hereditary diseases, thus reducing the morbidity and mortality associated with these diseases. With the emergence of new drugs and therapies, we believe that a feasible cure for PKD will be available in the future.

## Funding

This work was supported by 2022 Research Projects of Chongqing Municipal Health and Health Commission (No. 2022WSJK005) and the Intelligence Medicine Project of Chongqing Medical University (YJSZHYX202103).

## Data availability statement

The data analyzed in the current study are available in the GenBank database (Gene ID: 5313). The cDNA and gDNA sequence numbers are NM\_000298.6 and NG\_011677.1, respectively. More detailed patient data or sequencing data is available by contacting Dr. Jie Yu ([1808106652@qq.com](mailto:1808106652@qq.com)) or Dr. Yang Wang ([15732899563@163.com](mailto:15732899563@163.com)).

## CRediT authorship contribution statement

**Yang Wang:** Writing – original draft. **Jiaqi Liu:** Writing – review & editing. **Tao Liu:** Validation, Writing – review & editing. **Xizhou An:** Conceptualization, Visualization. **Lan Huang:** Investigation. **Jiacheng Li:** Investigation. **Yongjie Zhang:** Investigation, Project administration. **Yan Xiang:** Conceptualization, Resources. **Li Xiao:** Funding acquisition, Resources. **Weijia Yi:** Resources. **Jiebin Qin:** Resources. **Lili Liu:** Methodology. **Cuilan Wang:** Methodology. **Jie Yu:** Methodology, Supervision.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Li Xiao reports financial support was provided by 2022 Research Projects of Chongqing Municipal Health and Health Commission (No. 2022WSJK005). Li Xiao reports financial support was provided by Intelligence Medicine Project of Chongqing Medical University (YJSZHYX202103). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We would like to thank the patient, her family, and the investigators for their contribution to this study. We would also like to thank Phadcalc ([www.phadcalc.com](http://www.phadcalc.com)) for the molecular docking simulations and the Beijing HitroBio Biotechnology Co. Ltd. for the reagents. We also extend our thanks to Gao Mengxue for his help with figure preparation. We thank Bullet Edits Ltd. for editing and proofreading the manuscript. We thank Home for Researchers ([www.home-for-researchers.com](http://www.home-for-researchers.com)) for providing the speech corpus.

## Abbreviations

PKD	Pyruvate Kinase Deficiency
PK	Pyruvate Kinase
PKLR	pyruvate kinase liver and red blood cell
WT	wildtype
RMSD	root-mean-square deviation
RMSF	root-mean-square fluctuations
Rg	radius of gyration
SASA	solvent accessible surface area
MD simulations	molecular dynamics simulations
IDRs	inherently disordered regions
PDB	Protein Data Bank
RBCs	red blood cells
gDNA	Genomic DNA
HSCs	hematopoietic stem cells

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2024.e26368>.

## References

- [1] C. Laas, et al., Improving the laboratory diagnosis of pyruvate kinase deficiency, *Br. J. Haematol.* 193 (5) (2021) 994–1000.
- [2] B. Van Dooijeweert, et al., Untargeted metabolic profiling in dried blood spots identifies disease fingerprint for pyruvate kinase deficiency, *Haematologica* 106 (10) (2021) 2720–2725.
- [3] S. Chonat, et al., Pyruvate kinase deficiency in children, *Pediatr. Blood Cancer* 68 (9) (2021) e29148.
- [4] N. Luke, et al., Updates and advances in pyruvate kinase deficiency, *Trends Mol. Med.* 29 (5) (2023) 406–418.
- [5] J. Jimenez-Sainz, R.B. Jensen, Imprecise medicine: BRCA2 variants of uncertain significance (VUS), the challenges and benefits to integrate a functional assay workflow with clinical decision rules, *Genes* 12 (5) (2021).
- [6] W. Wang, Recent advances in atomic molecular dynamics simulation of intrinsically disordered proteins, *Phys. Chem. Chem. Phys.* 23 (2) (2021) 777–784.
- [7] F. Sohraby, H. Aryapour, Rational drug repurposing for cancer by inclusion of the unbiased molecular dynamics simulation in the structure-based virtual screening approach: challenges and breakthroughs, *Semin. Cancer Biol.* 68 (2021) 249–257.

- [8] L. Fan, et al., A novel mutation c.3392G>T of COL2A1 causes spondyloepiphyseal dysplasia congenita by affecting pre-mRNA splicing, *Front. Genet.* 13 (2022) 827560.
- [9] M.M. Li, et al., Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the association for molecular pathology, American society of clinical Oncology, and College of American pathologists, *J. Mol. Diagn.* 19 (1) (2017) 4–23.
- [10] R.S. Daniels, et al., A Sanger sequencing protocol for SARS-CoV-2 S-gene, *Influenza Other Respir. Viruses* 15 (6) (2021) 707–710.
- [11] A.V. Harder, et al., Migraine genetics: status and road forward, *Cephalalgia* 43 (2) (2023) 3331024221145962.
- [12] J.S. McFarlane, et al., Changes in the allosteric site of human liver pyruvate kinase upon activator binding include the breakage of an intersubunit cation- $\pi$  bond, *Acta Crystallogr. F Struct. Biol. Commun.* 75 (Pt 6) (2019) 461–469.
- [13] A. Waterhouse, et al., SWISS-MODEL: homology modelling of protein structures and complexes, *Nucleic Acids Res.* 46 (W1) (2018) W296–w303.
- [14] A. Tirziu, V. Paunescu, Cytotoxic T-cell-based Vaccine against SARS-CoV-2: a hybrid immunoinformatic approach, *Vaccines* 10 (2) (2022).
- [15] R.A. Laskowski, et al., AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR* 8 (4) (1996) 477–486.
- [16] Y. Wang, et al., The novel compound heterozygous variants identified in a Chinese family with glucose phosphate isomerase deficiency and pathogenicity analysis, *BMC Med. Genom.* 16 (1) (2023) 162.
- [17] W. Xu, et al., A novel homozygous TUB mutation associated with autosomal recessive retinitis pigmentosa in a consanguineous Chinese family, *BMC Med. Genom.* 16 (1) (2023) 9.
- [18] R. Linding, et al., Protein disorder prediction: implications for structural proteomics, *Structure* 11 (11) (2003) 1453–1459.
- [19] D. Van Der Spoel, et al., GROMACS: fast, flexible, and free, *J. Comput. Chem.* 26 (16) (2005) 1701–1718.
- [20] T.S. Hofer, M.J. Wiedemair, Towards a dissociative SPC-like water model II. The impact of Lennard-Jones and Buckingham non-coulombic forces, *Phys. Chem. Chem. Phys.* 20 (45) (2018) 28523–28534.
- [21] D. Nguyen, P. Macchi, A. Volkov, Fast analytical evaluation of intermolecular electrostatic interaction energies using the pseudoatom representation of the electron density. III. Application to crystal structures via the Ewald and direct summation methods, *Acta Crystallogr., A Found Adv.* 76 (Pt 6) (2020) 630–651.
- [22] M. Amir, et al., Investigation of deleterious effects of nsSNPs in the POT1 gene: a structural genomics-based approach to understand the mechanism of cancer development, *J. Cell. Biochem.* 120 (6) (2019) 10281–10294.
- [23] F.I. Khan, et al., Molecular mechanism of Ras-related protein Rab-5A and effect of mutations in the catalytically active phosphate-binding loop, *J. Biomol. Struct. Dyn.* 35 (1) (2017) 105–118.
- [24] A. Zhuravleva, L.M. Gierasch, Substrate-binding domain conformational dynamics mediate Hsp70 allostery, *Proc. Natl. Acad. Sci. U. S. A.* 112 (22) (2015) E2865–E2873.
- [25] S. Kozaka, et al., Design of swollen lipidic cubic phase to increase transcutaneous penetration of biomacromolecules, *ACS Appl. Mater. Interfaces* 13 (46) (2021) 54753–54761.
- [26] M. Maj, et al., Site-specific detection of protein secondary structure using 2D IR dihedral indexing: a proposed assembly mechanism of oligomeric hIAPP, *Chem. Sci.* 9 (2) (2018) 463–474.
- [27] H. Al-Samkari, et al., Early-onset reduced bone mineral density in patients with pyruvate kinase deficiency, *Am. J. Hematol.* 98 (3) (2023) E57–e60.
- [28] R.F. Grace, W. Barcellini, Management of pyruvate kinase deficiency in children and adults, *Blood* 136 (11) (2020) 1241–1249.
- [29] H. Al-Samkari, et al., Health-related quality of life and fatigue in children and adults with pyruvate kinase deficiency, *Blood Adv* 6 (6) (2022) 1844–1853.
- [30] M. Morado, et al., Consensus document for the diagnosis and treatment of pyruvate kinase deficiency, *Med. Clin.* 157 (5) (2021) 253.e1–253.e8.
- [31] O.A. Enegele, F. Anjum, Pyruvate kinase deficiency, in: *StatPearls*, StatPearls Publishing Copyright © 2023, StatPearls Publishing LLC, Treasure Island (FL), 2023.
- [32] H.W. Chueh, N. Kim, Molecular genetic testing enabled the diagnosis of otherwise undiagnosable cases of pyruvate kinase deficiency, *Pediatr. Hematol. Oncol.* 39 (2) (2022) 166–173.
- [33] P. Bianchi, E. Fermo, Molecular heterogeneity of pyruvate kinase deficiency, *Haematologica* 105 (9) (2020) 2218–2228.
- [34] G. Canu, et al., Red blood cell PK deficiency: an update of PK-LR gene mutation database, *Blood Cells Mol. Dis.* 57 (2016) 100–109.
- [35] B. Milanese, et al., Six novel variants in the PKLR gene associated with pyruvate kinase deficiency in Argentinian patients, *Clin. Biochem.* 91 (2021) 26–30.
- [36] A. Nin-Hill, et al., Photopharmacology of ion channels through the light of the computational microscope, *Int. J. Mol. Sci.* 22 (21) (2021).
- [37] W. Destiarani, et al., Molecular dynamics simulation of T10609C and C10676G mutations of mitochondrial ND4L gene associated with proton translocation in type 2 diabetes mellitus and cataract patients, *Bioinf. Biol. Insights* 14 (2020) 1177932220978672.
- [38] Z. Ataei, et al., Novel in-frame duplication variant characterization in late infantile metachromatic leukodystrophy using whole-exome sequencing and molecular dynamics simulation, *PLoS One* 18 (2) (2023) e0282304.
- [39] H. Zhang, et al., Liquid-liquid phase separation in biology: mechanisms, physiological functions and human diseases, *Sci. China Life Sci.* 63 (7) (2020) 953–985.
- [40] K. Ma, et al., Loss of miR-638 in vitro promotes cell invasion and a mesenchymal-like transition by influencing SOX2 expression in colorectal carcinoma cells, *Mol. Cancer* 13 (2014) 118.
- [41] F. Climent, et al., Red cell glycolytic enzyme disorders caused by mutations: an update, *Cardiovasc. Hematol. Disord.: Drug Targets* 9 (2) (2009) 95–106.
- [42] E. Maisonneuve, et al., Prenatal management of fetal anemia due to pyruvate kinase deficiency: a case report, *Transfusion* 63 (1) (2023) 257–262.
- [43] M. Garcia-Gomez, et al., Safe and efficient gene therapy for pyruvate kinase deficiency, *Mol. Ther.* 24 (7) (2016) 1187–1198.
- [44] B. Fattizzo, I. Motta, Rise of the planet of rare anemias: an update on emerging treatment strategies, *Front. Med.* 9 (2022) 1097426.
- [45] H. Al-Samkari, et al., Mitapivat versus placebo for pyruvate kinase deficiency, *N. Engl. J. Med.* 386 (15) (2022) 1432–1442.
- [46] R.F. Grace, et al., Safety and efficacy of mitapivat in pyruvate kinase deficiency, *N. Engl. J. Med.* 381 (10) (2019) 933–944.
- [47] N. Luke, et al., Updates and advances in pyruvate kinase deficiency, *Trends Mol. Med.* 29 (5) (2023) 406–418.
- [48] S. Fañanas-Baquero, et al., Clinically relevant gene editing in hematopoietic stem cells for the treatment of pyruvate kinase deficiency, *Mol. Ther. Methods Clin. Dev.* 22 (2021) 237–248.
- [49] J.D. Schwartz, et al., Who should be eligible for gene therapy clinical trials in red blood cell pyruvate kinase deficiency (PKD)?: toward an expanded definition of severe PKD, *Am. J. Hematol.* 97 (3) (2022) E120–e125.
- [50] H.N. Abdelhamid, et al., Gene delivery using cell penetrating peptides-zeolitic imidazolate frameworks, *Microporous Mesoporous Mater.* (2020) 300.
- [51] H.N. Abdelhamid, M. Dowaidar, U. Langel, Arbonized chitosan encapsulated hierarchical porous zeolitic imidazolate frameworks nanoparticles for gene delivery, *Microporous Mesoporous Mater.* (2020) 302.
- [52] J.L. López Lorenzo, et al., Lentiviral mediated gene therapy for pyruvate kinase deficiency: a global phase 1 study for adult and pediatric patients, *Blood* 136 (Supplement 1) (2020), 47-47.
- [53] S. Richards, et al., Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of medical genetics and genomics and the association for molecular pathology, *Genet. Med.* 17 (5) (2015) 405–424.
- [54] N. Tanabe, et al., A survey of the practice patterns of gynecologic oncologists dealing with hereditary cancer patients in Japan, *Fam. Cancer* 13 (3) (2014) 489–498.